MTA DOKTORI ÉRTEKEZÉS

ADVANCES IN BEHAVIORAL INDUSTRIAL ORGANIZATION

Készítette:

Kőszegi Botond
Közép-európai Egyetem

Budapest, 2014. január

Contents

Chapter 1

Introduction to Behavioral Economics

1.1 The Goal of this Dissertation

This dissertation familiarizes the reader with some recent advances in applied behavioral economics, especially behavioral industrial organization. To set up the stage for these newest developments, this chapter gives a general background and perspective on behavioral economics, and introduces the models of individual decisionmaking that will be used by the applied models.

1.2 What Is Behavioral Economics?

Psychology and economics—often also called "behavioral economics"—is not an easy field to define. In my view, it is not even really a field of economics at all, but more like a way of thinking about and doing research in any field. It is a mindset: the belief that economists should aspire to making assumptions about humans that are as realistic as possible, and hence that we should develop methods and habits to learn what is psychologically realistic.

What does this mindset entail? In a large part, it is not at all different from the mindset of the overwhelming majority of modern economics. We conceptualize economic phenomena by starting from individual behavior that is goal-driven: people try to understand their environment and to achieve their goals to the extent possible within the constraints they have. As in almost any field of economics, our aim is to understand how the goal-driven individual behavior plays out in different economic environments and what the welfare consequences are. Some of the habits and criteria we use in our teaching and research are therefore identical to those used by most economists. We formalize ideas using mathematical models, in which decisionmakers are

often highly sophisticated. In our models and explanations, we highly value simplicity as well as generality, and in fact view this as a major advantage of many of the ideas we propose. And we recognize that markets and incentives play an important role in shaping behavior, that one of the main goals of economic analysis is to evaluate the performance of market institutions and policies, and that therefore it is important to test ideas using data on *market* behavior. All this means that psychology and economics is a field of *economics* rather than psychology.

But there is a part of the habit and mindset of psychology and economics that is more new. Namely, we study more carefully and in more detail than most neoclassical economists what motivates people and how they go about maximizing their well-being. Our interest in "unpacking" how people might be thinking about and making decisions in turn implies an attentiveness and open-mindedness toward exploring behavior through experiments and surveys, and especially toward research in psychology, the other social sciences, and the brain sciences. What emerges from this investigation is a more nuanced, more detailed, and more accurate picture of individual economic actors than is typical in economics. But once we incorporate the more detailed knowledge of decisionmaking into our hypotheses, our goal is to understand economic consequences, especially market outcomes and welfare effects.

In some ways, the developments that led to the current field of behavioral economics are no different from those in game theory and information economics several decades ago. In both of these fields, researchers replaced previous—useful but simplistic—assumptions with more realistic and more detailed premises, and tremendously improved our understanding of economics as a result. Before the introduction of game theory into the field of industrial organization, many researchers thought of firm behavior in the market either in terms of perfect competition or in terms of monopoly. With game theory, it became possible to analyze intermediate cases more fully, and to understand how a firm's strategic decisions depend on its information, resources, and the structure of the industry. Similarly, firm behavior has and still is often usefully conceptualized in terms of profit maximization (including in this dissertation). Information economics, however, recognizes and assumes that a firm involves multiple individuals with possibly different motives and information. This allows a more nuanced understanding of the firm that is useful in many circumstances, such as when trying to understand the boundaries of the firm and the labor market.

Psychology and economics does something similar, but where it expands or improves previous assumptions is in the realm of individual decisionmaking. As with the above developments, this is more useful in some settings than

in others. This dissertation covers models of individual decisionmaking and applications that I believe are very important economically.

The historical development of behavioral economics can be categorized into three overlapping waves. In the first wave (which happened mostly in the 1970's and 80's), researchers identified systematic and important ways in which economic theory had been wrong, and suggested alternative ways of understanding behavior based on simple psychological principles. In the second wave (most of which started in the 90's, and which is still going on today), economists formally modeled some of the alternatives, and established their empirical importance in the laboratory and the field. Finally, the third wave (most of which started in the 2000's) involves full integration of the new psychologically based models into economic analysis, to address the same questions economists have always been interested in: how individual behavior plays out in organizations and markets, what the welfare consequences are, and how policy should respond to market outcomes.

The applications in this dissertation belong solidly in third-wave behavioral economics. They are built on a simple common principle of an asymmetry between consumers and firms. On the one hand, individual consumers are likely to be subject to the psychological phenomena documented by psychologists and behavioral economists. Hence, we assume that consumers exhibit these phenomena when making decisions in the marketplace. On the other hand, firms face incentives to maximize profits and have substantial resources and can create complex systems to make this happen. To capture this in an extreme way, we assume that they do not exhibit the psychological phenomena in question at all, but instead act as classical profit-maximizing firms.

The next two sections introduce the models of consumer behavior that will be used in later chapters. I introduce the taste for immediate gratification, which will be used in Chapter 2, in the next section. Then, I introduce loss aversion, which will be used in Chapters 3 and 4.

1.3 The Taste for Immediate Gratification

This section formalizes the taste for immediate gratification, and possible naivete regarding this taste, as modeled using the β-δ approach by Laibson (1997) and O'Donoghue and Rabin (1999). While I present some evidence as a way of motivating the model, I do not attempt to cover the broad array of evidence that has been accumulated over the years in favor of the model. For excellent reviews of the evidence, see Rabin (1998) and DellaVigna (2009).

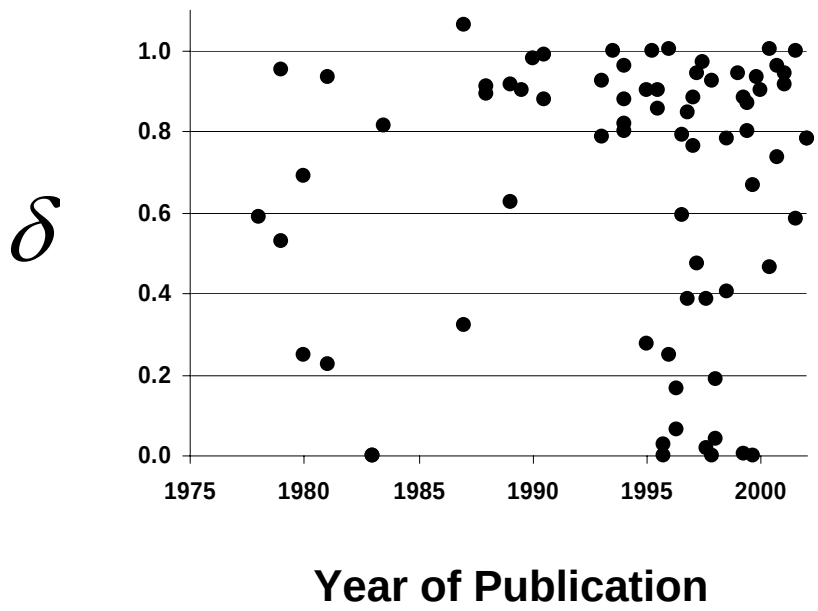

Figure 1.1: Estimated Annual δ's from Economics Research

1.3.1 Introduction: The Need to Move Beyond Exponential Discounting

The standard theory for modeling intertemporal choices in economics is Samuelson's exponential-discounting model. In a version of this model, an agent making a decision at time t aims to maximize

$$u_t + \delta u_{t+1} + \delta^2 u_{t+2} + \cdots = \sum_{\tau \geq t} \delta^{\tau - t} u_\tau,$$

where u_τ is the instantaneous utility at time τ and δ is the agent's *discount factor*. The discount factor measures the proportional discount the agent applies to any one-period delay: instantaneous utility at any time is worth δ times as much as instantaneous utility in the previous period. In this model, the single variable that captures intertemporal preferences is δ. So it is little wonder that economists have tried to estimate δ over the years. Taken from Frederick, Loewenstein and O'Donoghue (2002), Figure 1.1 illustrates estimates of annual δ's from economics research in the years 1975-2002. As is clear from the figure, there is tremendous variation in the estimates.

That economists' estimates of δ are all over the place could mean three things. It could mean that economists have not made much progress in converging on a good estimate of δ, and that we need much more research to figure out which estimates are right. It could also mean that the discount factor is wildly variable across different situations and populations in a way that follows no logical pattern, so that there is no hope in pinning down an even approximate value for it. I find it implausible that these explanations provide a full account for the dispersion in measured δ's, and hence will take a third perspective (one that I will argue helps explain the data): that δ simply cannot be measured accurately because the exponential discounting model tries to cram too much into this single parameter. Exactly as Samuelson intended, the model is an excellent way of conceptualizing and thinking about the fact that people make intertemporal tradeoffs, and the discount factor is a useful "summary statistic" for how much the decisionmaker cares about the future in the *particular* situation. But intertemporal choice is a reflection of many psychological processes, some of which are more conducive to patience and some of which are more conducive to impatience. In addition, these forces are relevant in different choice problems to different degrees, so an economist who ignores them and tries to fit a single parameter onto every situation will keep measuring a different δ each time. Hence, the exponential discounting model is not a good model to shed light on how behavior *varies across situations*, and understanding the underlying psychological processes better will help us understand the variation better. As the most important example, I study self-control problems in this dissertation.

1.3.2 Short-Run Desires Versus Long-Run Goals

This section discusses two types of motivation for the model of hyperbolic discounting I introduce in the next section, both covering the model's basic features and illustrating its economic importance. First, when it comes to tradeoffs between now and the near future, people are quite impatient. Second, when it comes to tradeoffs between nearby future dates, people are much more patient.

Evidence on Impatience in Short-Run Decisions

I begin this section by describing evidence from a range of decisionmaking situations showing that individuals can be quite impatient when it comes to choices that implicate immediate pleasures and pains. A common and economically relevant example is the so-called *payday effect*: individuals who live

from paycheck to paycheck spend a lot of their income immediately after getting paid. Huffman and Barenstein (2005), for instance, document that the consumption of working households in the United Kingdom is 18% higher the week after their payday than the week before. This indicates that when they get paid, they are eager to consume, and do not care so much that they will be suffering a few weeks down the line when they run out of money. Motivated by the same hypothesis, Shapiro (2005) finds that the caloric intake of food-stamp recipients declines by 13-14% over the month, while expenditure on food declines by about 20% over the month (indicating that households switch to less expensive foods over the course of the month). Shapiro argues that a high degree of short-run impatience is the most likely explanation for his findings, and rules out some alternative possible explanations.[1]

People's impatience regarding consumption is reflected not only in their spending a lot of their available funds immediately after they receive it, but also in their eagerness to borrow from future income. One expensive form of short-term borrowing is through credit cards, and in countries where credit cards are widely available, credit-card debt tends to be quite high. For example, in the United States in January 2012, the amount of outstanding revolving debt held by households (most of which is credit-card debt) was $800.9 billion. That is about $6,975 per household, including households that do not carry revolving debt. As of November 2011, the average interest rate on credit-card accounts with debt was 12.78%.[2] And credit cards are among the cheaper forms of short-term credit. One of the most expensive forms of (legal) borrowing in the United States is from one's upcoming paycheck. About 10 million people borrow money through payday loans, and do so at annualized compounded percentage rates often exceeding 1000%. Yet there are more payday-loan and check-cashing outlets in the United States than there are McDonald's and Starbucks combined.

[1] In an interesting—if somewhat sad—twist to the payday effect, Hastings and Washington (2010) document that store pricing responds to the effect: in areas with a lot of food-stamp recipients (but not in other areas), prices rise at the time beneficiaries receive their food stamps. This reaction is especially relevant for the current dissertation. Chapters 2-4 study how profit-maximizing firms react to consumers' psychological phenomena, and the paper by Hastings and Washington demonstrates in a compelling way that firms do this.

[2] Source for these statistics: `http://www.federalreserve.gov/releases/g19/current/`. It is worth mentioning that households use credit cards not only for immediate consumption, but also for purchases such as durables. One of the main results of Section 2 is showing that such purchases can also be understood in terms of the taste for immediate gratification.

y	z	median x	yearly δ
\$15	1 month	20	0.032
\$15	ten years	100	0.83
\$40	six months	50	0.64
\$40	four years	90	0.82

Table 1.1: Some Findings of Thaler (1981)

Patience in Decisions About the Future

All the situations discussed in the previous section implicate the possibility of immediate or near-immediate monetary receipts or consumption. I now argue that the high level of impatience people display is specific to such situations: for similar tradeoffs that are further in the future, individuals tend to be more patient, so that exponential discounting cannot capture their attitudes toward delay for both short-run and long-run decisions.

The simplest way to see this is to carry out some thought experiments on the implications of applying the above type of impatience to intertemporal tradeoffs in general. Specifically, If a level of impatience that seems reasonable for short-run delays is applied to any delay of equal length, the implied level of impatience over long-run delays would be unreasonable. To illustrate, consider some simple arithmetic:

$$0.99^{365\times 10} \approx \frac{1}{8541609622012070} \quad \text{and} \quad 0.999^{3\times 365\times 10} \approx \frac{1}{57266}.$$

If a person values tomorrow just one percent less than today—so that she ever-so-slightly prefers to put off an unpleasant task until tomorrow, even if that means the task becomes ever-so-slightly more difficult to perform—and applies the same preference to any one-day delay, she must value anything that happens in 10 years as totally insignificant relative to what happens today. Similarly, if she ever-so-slightly prefers to put off doing something from the morning to the evening, she must value today 57,266 times as much as 10 years from now (so she would not invest \$100 to get \$5 million in ten years).

These simple calculations suggest that the kind of short-run impatience found in the previous section cannot possibly describe intertemporal tradeoffs over longer horizons. Indeed, there is also direct evidence indicating that people are more impatient over short-term decisions than over long-term decisions. For instance, Thaler (1981) asked questions of the following form: "What amount x makes you indifferent between \$$y$ today and \$$x$ in z time?"

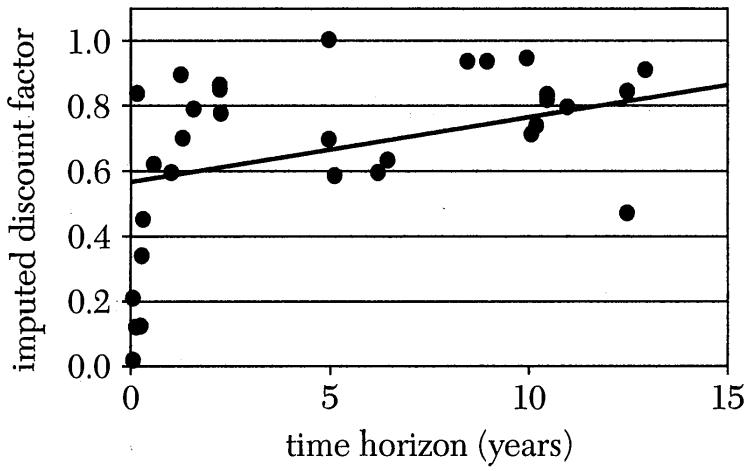

Figure 1.2: Estimated Discount Factors as a Function of the Time Delay in Frederick, Loewenstein, and O'Donoghue (2002)

For each possible time delay z, one can calculate the equivalent yearly discount factor δ, and see how δ depends on the time delay. Table 1.1 gives some of Thaler's findings. Clearly, the longer is the time delay, the larger is the implied yearly δ, indicating that individuals tend to become more patient on longer-run decisions.

Frederick and Loewenstein (1999) repeat the same exercise as Thaler, but do so using economists' estimates of yearly discount factors rather than hypothetical choices. Figure 1.2 graphs the estimated yearly discount factor δ as a function of the delay in the decisionmaking situation based on which δ was estimated. Just like in Thaler's case, the relationship is clearly and significantly positive. Hence, from analyses using different sources of data and different methods, the same pattern emerges over and over again: short-run discount factors are lower than long-run discount factors.

1.3.3 Modeling the Conflict Between Short-Run Desires and Long-Run Goals

Discounted Utility Function

As argued in detail in the previous section, evidence suggests that people discount nearby events quite heavily, but they are more patient for events

further away. Recent research has given rebirth to attempts to incorporate this pattern, and the interpersonal conflict it generates, formally into economics. In this section, I consider a modification of exponential discounting that captures these forces, and investigate some consequences of the new formulation.

Recall that Samuelson's exponentially discounted utility model says that utility at time t is

$$u_t + \delta u_{t+1} + \delta^2 u_{t+2} + \delta^3 u_{t+3} + \dots.$$

Instead of Samuelson's formulation, Laibson (1997) and O'Donoghue and Rabin (1999) assume that at time t, a person aims to maximize

$$u_t + \beta\delta u_{t+1} + \beta\delta^2 u_{t+2} + \beta\delta^3 u_{t+3} + \dots$$

with $0 < \beta \leq 1$. This is the *hyperbolic discounting* or (more precisely) *quasi-hyperbolic discounting* model. The extra parameter β that applies to all future periods captures the extra discounting people apply to the future relative to the present due to their taste for immediate gratification. Since they apply β to all future periods equally, this means that they are relatively impatient when it comes to tradeoffs between the present and the future, but they are relatively patient when it comes to tradeoffs that occur in the future. And since β captures the extra discounting that occurs between the present and the future, it can be thought of as the short-run discount factor. When $\beta = 1$, so that there is no extra discounting on the future, we get back to exponential discounting.

Analyzing the Model: Sophistication versus Naivete

To illustrate how to work with the hyperbolic discounting model, as well as to raise an additional issue, I demonstrate some of its implications in a simple decision. This analysis is based on O'Donoghue and Rabin (1999).

Consider a student's decision of when to do a problem set. She could do it right after lecture ($t = 0$), when she remembers the material very well, or tomorrow ($t = 1$), when she remembers the material less well, or the day after tomorrow ($t = 2$), when she has forgotten everything. Because of this, the cost of doing the problem set increases, the later she does it. At date 0, the instantaneous disutility (immediate cost) of doing the problem set is 1. At date 1, the instantaneous disutility of doing the problem set is 3/2. At date 2, the instantaneous disutility of doing the problem set is 5/2. The student is a hyperbolic discounter with $\beta = 1/2$ and $\delta = 1$.

To begin thinking about this example, first assume that the student can decide at time 0 when to do the problem set. This means, for example, that

she can decide at time 0 to do the problem set at time 1. Because she can decide what she will do in the future, we say that the student can *commit* or has access to *commitment*. For example, she could set up a study group that she cannot cancel later (or that is prohibitively costly to cancel later).

Since the student has different preferences at different points in time, we refer to her in multiple ways that reflect her changing preferences: we let self 0 be the period-0 incarnation of the student, self 1 her period-1 incarnation, and self 2 her period-2 incarnation. To decide when self 0 would like to do the problem set, we have to look at the discounted costs of doing it from the perspective of period 0:

- If the problem set is done in period 0, the discounted cost is 1.
- If the problem set is done in period 1, the discounted cost is $\frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}$.
- If the problem set is done at period 2, the discounted cost is $\frac{1}{2} \cdot \frac{5}{2} = \frac{5}{4}$.

Because it minimizes the discounted cost of doing the problem set, self 0 commits to doing it in period 1. Intuitively, because of her taste for immediate gratification, the student prefers to delay—because she cares much more about the present than about the future, she is willing to put off the task even though she knows it will become harder. But because she is more patient when it comes to future tradeoffs, she does not want to delay more than one period.

Now suppose that the student has no access to a commitment technology. So, in both period 0 and period 1, she freely decides whether she does the problem set then or later (at date 2, she has to do it if she has not already finished it). It would still be the case that self 0 would *want to* do the problem set at time 1. Would she actually do it at time 1? To answer this, we look at discounted costs from the perspective of period 1:

- If the problem set is done in period 1, the discounted cost is $\frac{3}{2}$.
- If the problem set is done in period 2, the discounted cost is $\frac{5}{4}$.

So she now prefers to do the problem set at time 2. Because the student's earlier preference for what she should do at a later point (her period-0 preference for doing the problem set at time 1) is different from what she wants to do when the time comes (her period-1 preference to delay doing the problem set), the student is dynamically *in*consistent. This intertemporal conflict leads to a *self-control problem*: the student wants to exercise self-control tomorrow, but once tomorrow come she may not want to.

But given this conflict, when does the student actually do the problem set? It turns out that with the information introduced so far, it is impossible to answer this question. To understand how the self-control problem plays out, we need to know whether the student is *aware* that she will change her mind in the future. There are two extreme assumptions one can make in this regard. *Naive* decisionmakers are not aware of their own self-control problem—they persist in happily thinking that whatever they plan will actually be carried out by their later selves. *Sophisticated* decisionmakers, on the other hand, perfectly anticipate their future behavior—they do not harbor illusions about their ability to carry through plans. This means that they will try to take actions to make sure they stick with current plans. These two possible assumptions about the student's self-awareness are very extreme, and most of us have features of both sophistication and naivete.

Now we are ready to analyze the student's behavior, separately by their degree of sophistication. We start with a naive student. We have already calculated that self 0 would prefer to do the problem set at time 1, and since (being naive) she assumes she will behave "correctly" in the future, that is what she assumes she will do. Hence, she does not do the problem set at time 0, thinking, "Well, this is no big deal, I will just do it tomorrow..." We have also calculated that she does not do the problem set at time 1, so a naive student does the problem set in period 2.

Intuitively, since she believes that she will do the problem set before this will be too hard, she *believes* she cannot lose much by delaying. Next period, she again perceives the cost of delaying to be small, so she delays again.

Note that from the point of view of time 0, the student incurs a discounted cost of 5/4, higher than if she did the problem set in period 0. That is, the student does something she considers unambiguously bad for herself. This could not happen with exponential discounting: with exponential discounting, whatever self 0 thinks is the best thing to do, later selves will be willing to do. Hence, the behavior of the naive student is not driven purely by impatience—time inconsistency is implicated as well. That hyperbolic discounters often do things that they perceive as unambiguously bad for themselves is an important property of these models, and plays a central role in applications of hyperbolic discounting, including that in Chapter 2.

We now turn to sophisticated students. We know that if the student does not do the problem set in period 0, she does not do it in period 1, either. A sophisticated student realizes this, so she knows that if she does not do the problem set in period 0, she will not do it until period 2. Since (according to the above calculation) she prefers to do it in period 0 rather than in period 2, a sophisticated student does problem set in period 0.

Intuitively, a sophisticated student recognizes that if she delays, she will delay more. Since she knows that would be too costly, she reluctantly does the problem set at time 0.

As mentioned above, the most realistic assumption seems to be that most individuals are neither fully sophisticated, nor fully naive. How do we model such decisionmakers? O'Donoghue and Rabin (2001) assume that the agent believes with probability 1 that her future β will equal $\hat{\beta}$. In this formulation, $\hat{\beta} = \beta$ corresponds to full sophistication, and $\hat{\beta} = 1$ corresponds to full naivete. Eliaz and Spiegler (2006) and Asheim (2008) assume that the agent has beliefs that put some probability on $\hat{\beta} = \beta$, and the complementary probability on $\hat{\beta} = 1$. In some settings, Heidhues and Kőszegi (2010a) (Chapter 2) allow for a person's beliefs to be a full distribution.

1.4 Reference Dependence and Loss Aversion

This section introduces reference dependence and loss aversion, which will be used as the model of consumer behavior in Chapters 3 and 4. As in the case of hyperbolic discounting, I discuss some evidence as a way of motivating the model, but do not discuss the full array of evidence in favor of the model. For evidence, see Rabin (1998) and DellaVigna (2009).

1.4.1 An Illustration and an Introduction

Take a look at Figure 1.3 illustrating the Gradient Illusion. The central stripe in the illustration is actually uniform in color. Yet the part surrounded by a darker grey looks lighter than the part surrounded by a lighter grey. The illusion persists even after one has been told that the central stripe is of uniform color, after one has measured the color in Photoshop, and even if one has created the figure oneself.

This illusion illustrates the extent to which our brain tends to perceive things *relative* to other things. Just by putting these different background colors next to the central stripe, we can induce the brain to automatically make the comparison, and for that to dominate the judgment about the absolute color of the stripe. Even once you are fully convinced that this is an illusion, the central stripe still does not *seem* uniform in color. It is just hard to see it in any other way. More generally, in judging perceptive things such as brightness, loudness, or temperature, the stimuli are perceived in relation to some neutral point.

The tendency to compare stimuli to other stimuli extends to the economic

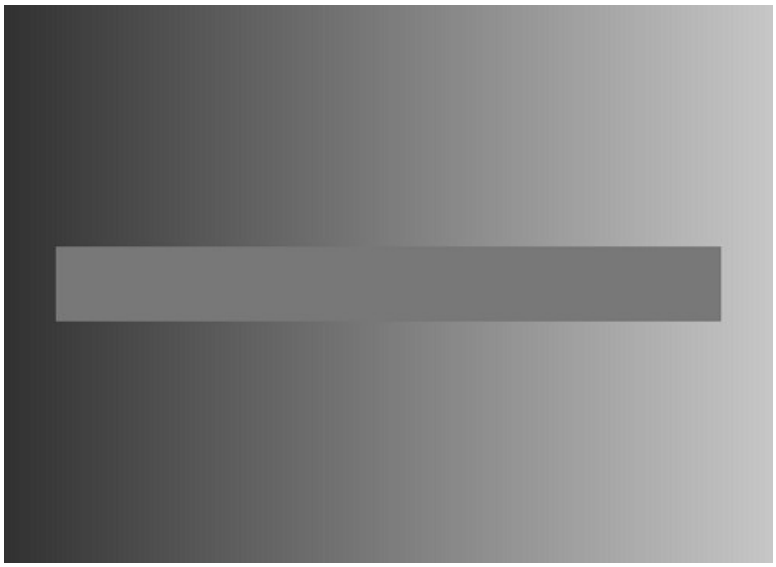

Figure 1.3: The Gradient Illusion. The central stripe has uniform color—yet it seems much lighter on the left side.

domain in a big-time way. We call this the phenomenon of *reference-dependent preferences*: that the utility level we derive from an outcome depends in a major way on comparisons to certain "benchmark" outcomes or *reference points*—not only on an absolute evaluation of the outcome itself. A large literature starting with Kahneman and Tversky (1979) is devoted to modeling reference-dependent preferences and its implications for economics.

1.4.2 Loss Aversion

The most important property of reference-dependent preferences is loss aversion—people dislike losses relative to the reference point more than they like same-sized gains. I illustrate two kinds of evidence on loss aversion, that based on people's willingness to trade their current position for another one, and that based on choices over risky gambles.

Loss aversion is manifested in the striking endowment effect documented first by Kahneman, Knetsch, and Thaler (1990, 1991) and subsequently by many other researchers: once a person comes to possess a good, she almost immediately values it more than before she possessed it. These experiments usually start by randomly giving half the subjects (often a class) mugs. These subjects become the "owners" or potential sellers, and the others are "non-

owners" or potential buyers. The owners are then asked to examine the mug and think about how useful it might be to them. They are also asked to pass their mug to the closest non-owner, so that they can examine it as well. This is an important part of the design, because it reduces the information asymmetry between owners and non-owners. Buying and selling prices are then elicited in an incentive-compatible way using the Becker-DeGroot-Marschak procedure (Becker, DeGroot and Marschak 1964). Prototypical experiments starting with Kahneman, Knetsch and Thaler (1990), have consistently found a major gap, with selling prices being about twice the buying prices.

The endowment effect—the fact that owners value a good more than otherwise identical non-owners—is usefully conceptualized as a case of loss aversion. Individuals who are randomly given mugs treat the mugs as part of their reference levels or endowments, and consider not having a mug to be a loss. Individuals without mugs consider not having a mug as remaining at their reference point, and getting a mug as a gain. Since people are more sensitive to losses than they are to same-sized gains, the sellers "value" the mug more: by keeping the mug, they avoid a loss, whereas buyers would merely make a gain if they got the mug.

Another important manifestation of loss aversion is in attitudes toward risky gambles. For instance, most people would turn down an immediate fifty-fifty gain $550 or lose $500 gamble. This kind of risk aversion seems such an intuitively obvious fact that for a long time researchers have not even bothered to check it. But recently, Barberis, Huang and Thaler (2006) offered the gamble for real to MBA students, financial analysts, and even very rich investors (with median financial wealth over $10 million!). A majority of all these people, including 71% of the investors, rejected the gamble.

The standard economic explanation for people's rejection of this gamble is risk aversion or (equivalently for our purposes) diminishing marginal utility of wealth. Indeed, diminishing marginal utility of wealth is an excellent assumption based on good psychology: people satisfy their most important needs and desires first and the less important ones only if they have something left over, so the first $1,000,000 in wealth generates more utility than the next $1,000,000. This is a great explanation for large-scale risk aversion, such as the decision to take $4 million for sure rather than $10 million with probability one-half.

But most of the risky decisions we face are not in the $1 million range or even the $100,000 range. They are much smaller. And in a key article, Rabin (2000a) argued that expected-utility-over-wealth maximizers—who care only about final wealth outcomes—should not reject such a gamble unless they turn down phenomenally favorable larger risks. Since most people do take

many risks, expected utility is not a reasonable explanation for rejecting the small-scale gamble. Rabin's mathematical argument centers around proving statements of the following form: "If an individual with expected utility over wealth turns down a fifty-fifty lose $\$l$ or gain $\$g$ gamble over a range of wealth levels, she also turns down a fifty-fifty lose $\$L$ or gain $\$G$ gamble," where G is huge relative to L and L is not that large (G is infinite in some examples). The argument proceeds by using that if a person turns down the g/l gamble for some wealth level, her marginal utility must diminish by some non-trivial amount over the range of the gamble. Using that this is the case for multiple wealth levels, we conclude that over the range of these wealth levels marginal utility diminishes quite a lot. But this implies extreme sensitivity to larger gambles.

Here is an illustration of the precise argument. Suppose Johnny is a classical utility maximizer with diminishing marginal utility of wealth who would turn down a fifty-fifty lose \$500 or gain \$550 bet for a non-trivial range of initial wealth levels. Let us take a concave, increasing utility function over wealth, $u(\cdot)$. Rejection of this bet means that

$$\frac{1}{2}u(w+550)+\frac{1}{2}u(w-500)<u(w),$$

which implies

$$u(w+550)-u(w)<u(w)-u(w-500).$$

But notice that by the concavity of $u(\cdot)$, $u(w)-u(w-500)<500\cdot u'(w-500)$, and $u(w+550)-u(w)>550\cdot u'(w+550)$. Therefore,

$$500\cdot u'(w-500)>550\cdot u'(w+550),$$

or

$$u'(w-500)>\frac{11}{10}u'(w+550).$$

Now suppose Johnny was \$1,050 poorer in lifetime terms. This is a *very* small change in lifetime wealth, equivalent to something less than \$50 per year. It is implausible that risk aversion would diminish significantly with such small changes in initial wealth, especially for decreases in wealth. If so, then by the same argument as above but now applied to a wealth level of $w-1050$,

$$u'(w-1550)>\frac{11}{10}u'(w-500).$$

Combining the two

$$u'(w-1550)>\left(\frac{11}{10}\right)^2u'(w+550),$$

and by the same reasoning

$$u'(w - 2100) > \left(\frac{11}{10}\right)^2 u'(w).$$

But this implies that marginal utility for wealth *skyrockets* for larger decreases in wealth unless there are dramatic shifts in risk attitudes over larger changes in wealth: for every decrease of \$1,050 in Johnny's wealth, his marginal utility of wealth increases by a factor of 11/10. Doing this fifty times... If Johnny became \$52,500 poorer in lifetime wealth—which is something less than \$2,500 in pre-tax income per year, say—then he would value income at least 117 times ($\approx \left(\frac{11}{10}\right)^{50}$) as much as he currently does. While none of us know Johnny, we know this is a false fact about Johnny.

Furthermore, such a plummeting marginal utility of money leads to wild risk aversion over large stakes: if Johnny's marginal utility of wealth increases by a factor of 117 if he were \$52,500 poorer, for instance, then—*even if he were risk neutral above his current wealth level*—then Johnny would turn down a fifty-fifty lose \$110,000 or gain \$6.4 million bet at his current wealth level.[3]

By a similar calculation, if Johnny were risk neutral above his current wealth level but averse to 50/50 lose \$10 / gain \$11 bets below his current wealth level, then he would turn down a 50/50 lose \$22,000 / gain \$100 billion bet. Rabin gives many further numerical examples.

Since this kind of risk aversion is inconceivable (how many would turn down this last bet?), we can conclude that diminishing marginal utility of wealth cannot reconcile risk aversion over modest stakes with reasonable risk aversion over large stakes. And these results are just bounds, and vastly understate the severity of large-scale risk aversion implied by small-scale risk aversion.

So why do people reject a fifty-fifty lose \$500 or gain \$550 risk? Most likely because of loss aversion. They dislike the prospect of an unpleasant loss of \$500 much more than they like the prospect of a gain of \$500. Loss aversion is not subject to the same critique as diminishing marginal utility over wealth because it does not assume that risk preferences over any level of wealth are determined by a single function. It could be that at any wealth level, a person dislikes a \$500 loss much more than she likes a \$550 gain—if her reference point is her current wealth. But this does not mean that her utility function is

[3] To get this number, I used that Johnny's marginal utility at wealth levels below the current wealth minus \$52,500 is at least 117 times that at his current wealth level. A \$110,000 loss is a loss of more than \$55,000 extra. He cares about this extra loss at least 117 times as much than about gains from the current wealth level. So even a gain of $6,400,000 < 117 \times 55,000$ would not be enough to compensate him.

very concave overall, because it does not imply that her utility function must at the same time curve at each of these wealth levels.

In other words, loss aversion gets around the Johnny logic by assigning a special role to current wealth (or another reference point), and making a strong distinction between gains and losses. Because losses are much more painful than equal-sized gains are pleasant, it may well be that a gain of $550 is not as attractive as a loss of $500 is scary. But with loss aversion, it is not necessarily the case that a loss of an extra $500 is worse than the loss of the first $500—since both of these are losses. So the above logic breaks down.

1.4.3 The Reference Point

Predictions of reference-dependent preferences and loss aversion of course depend crucially on what we assume the reference point is. While many theories of reference-point determination have been proposed, the most frequently used theory in recent years is that of Kőszegi and Rabin (2006). To motivate the key assumption in our model, I begin with an experiment. Abeler, Falk, Götte and Huffman (2011) gave students a menial task (entering data into the computer) to perform for a piece-rate. Students could work as long as they wanted. The twist in the authors' experiment was in how students were paid. After a student finished working, a random draw was made: with probability one-half, the student received what she earned in the task, and with probability one-half, she received a predetermined amount. For a randomly chosen half of the subjects, the predetermined amount was €3.50, and for the other half, it was €7.00. Subjects knew all these details of the experiment, including their own predetermined amount, in advance.

Abeler et al. (2011) found a striking difference in how much the two groups of students worked: the group whose predetermined amount was €3.50 tended to stop more when they earned €3.50, and the group whose predetermined amount was €7.00 tended to stop more when they have earned €7.00. In a sense, the predetermined amount became a target for how much to earn. This indicates that a subject's *recent expectations* (i.e., probabilistic beliefs) about how much she might earn determine her reference point for earnings. Chapters 3 and 4 of the dissertation build on this assumption, first formalized by Kőszegi and Rabin (2006). Other evidence also lends support to the expectations-based model. 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.[4] Crawford and Meng (2011) 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.

[4] 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. In a similar experiment, however, Smith (2008) does not find the same effect.

Chapter 2

Exploiting Naivete about Self-Control in the Credit Market[1]

2.1 Introduction

Researchers as well as policymakers have expressed concerns that some contract features in the credit-card and subprime mortgage markets may induce consumers to borrow too much and to make suboptimal contract and repayment choices.[2] These concerns are motivated in part by intuition and evidence on savings and credit suggesting that consumers have a time-inconsistent taste for immediate gratification, and often naively underestimate the extent of this taste.[3] Yet the formal relationship between a taste for immediate gratifica-

[1] This chapter is coauthored with Paul Heidhues, and appeared in the American Economic Review (2010), 100(5), pp. 2279-2303. Some features of real-life credit contracts we discuss in the paper no longer exist in the United States due to recently enacted regulations—also mentioned in the paper—that are very similar in spirit to the welfare-improving interventions we propose. Hence, our discussion of contract terms is most appropriate for the period ending around 2010.

[2] See, for instance, Ausubel (1997), Durkin (2000), Engel and McCoy (2002), Bar-Gill (2004), Warren (2007), and Bar-Gill (2008).

[3] Laibson, Repetto and Tobacman (2007) estimate that to explain a typical household's simultaneous holdings of substantial illiquid wealth and credit-card debt, the household's short-term discount rate must be higher than its long-term discount rate. Complementing this finding, Meier and Sprenger (2010) document that low and middle-income individuals who exhibit a taste for immediate gratification in experimental choices over monetary payments have more outstanding credit-card debt. Laibson et al. (2007) calculate that many households are made worse off by owning credit cards, so the fact that they get those cards suggests some degree of naivete about future use. Consistent with this idea, consumers overrespond to the introductory "teaser" rates in credit-card solicitations relative to the length of the introductory period (Shui and Ausubel 2004) and the post-introductory inter-

tion and consumer behavior and welfare in the credit market remains largely unexplored and unclear. Existing work on contracting with time inconsistency (DellaVigna and Malmendier 2004, Kőszegi 2005, Eliaz and Spiegler 2006) does not investigate credit contracts and especially welfare and possible welfare-improving interventions in credit markets in detail. Furthermore, because borrowing on a mortgage or to purchase a durable good typically involves up-front effort costs with mostly delayed benefits, models of a taste for immediate gratification do not seem to predict much of the overextension that has worried researchers and policymakers.

In this paper, we provide a formal economic analysis of the features and welfare effects of credit contracts when some consumers have a time-inconsistent taste for immediate gratification that they may only partially understand. Consistent with real-life credit-card and subprime mortgage contracts but (we argue) inconsistent with natural specifications of rational time-consistent theories, in the competitive equilibrium of our model firms offer seemingly cheap credit to be repaid quickly, but introduce large penalties for falling behind this front-loaded repayment schedule. The contracts are designed so that borrowers who underestimate their taste for immediate gratification both pay the penalties and repay in an ex-ante suboptimal back-loaded manner more often than they predict or prefer. To make matters worse, the same misprediction leads non-sophisticated consumers to underestimate the cost of credit and borrow too much—despite borrowing being for future consumption. And because the penalties whose relevance borrowers mispredict are large, these welfare implications are typically large even if borrowers mispredict their taste for immediate gratification by only a little bit and firms observe neither borrowers' preferences nor their beliefs. Accordingly, for any positive proportion of non-sophisticated borrowers in the population, a policy of disallowing large penalties for deferring small amounts of repayment—akin to recent new US regulations limiting prepayment penalties on mortgages and certain interest charges and fees on credit cards—can raise welfare.

Section 2.2 presents our model. There are three periods, 0, 1, and 2. If the consumer borrows an amount c in period 0 and repays amounts q and r in periods 1 and 2, respectively, self 0, her period-0 incarnation, has utility $c - k(q) - k(r)$, where $k(\cdot)$ represents the cost of repayment. Self 1 maximizes

est rate (Ausubel 1999), suggesting that they end up borrowing more than they intended or expected. Skiba and Tobacman (2008) find that the majority of payday borrowers default on a loan, yet do so only after paying significant costs to service their debt. Calibrations indicate that such costly delay in default is only consistent with partially naive time inconsistency. For further discussions as well as evidence for a taste for immediate gratification in other domains, see DellaVigna (2009).

$-k(q) - \beta k(r)$ for some $0 < \beta \leq 1$, so that for $\beta < 1$ the consumer has a time-inconsistent taste for immediate gratification: in period 1, she puts lower relative weight on the period-2 cost of repayment—that is, has less self-control—than she would have preferred earlier. Since much of the borrowing motivating our analysis is for future consumption, self 0 does not similarly discount the cost of repayment relative to the utility from consumption c. Consistent with much of the literature, we take the long-term perspective and equate the consumer's welfare with self 0's utility, but the overborrowing we find means that self 1 and self 2 are also hurt by a non-sophisticated borrower's contract choice. To allow for self 0 to be overoptimistic regarding her future self-control, we follow O'Donoghue and Rabin (2001) and assume that she believes she will maximize $-k(q) - \hat{\beta}k(r)$ in period 1, so that $\hat{\beta}$ satisfying $\beta \leq \hat{\beta} \leq 1$ represents her beliefs about β.

The consumers introduced above can sign exclusive non-linear contracts in period 0 with competitive profit-maximizing suppliers of credit, agreeing to a consumption level c as well as a menu of installment plans (q, r) from which self 1 will choose. Both for theoretical comparison and as a possible policy intervention, we also consider competitive markets in which disproportionately large penalties for deferring small amounts of repayment are forbidden. Formally, in a *restricted market* contracts must be linear—a borrower can shift repayment between periods 1 and 2 according to a single interest rate set by the contract—although as we discuss, there are other ways of eliminating disproportionately large penalties that have a similar welfare effect.

Section 2.3 establishes our main results in a basic model in which β and $\hat{\beta}$ are known to firms. Since a sophisticated borrower—for whom $\hat{\beta} = \beta$—correctly predicts her own behavior, she accepts a contract that maximizes her ex-ante utility. In contrast, a non-sophisticated borrower—for whom $\hat{\beta} > \beta$—accepts a contract with which she mispredicts her own behavior: she believes she will choose a cheap front-loaded repayment schedule (making the contract attractive), but she actually chooses an expensive back-loaded repayment schedule (allowing firms to break even). Worse, because the consumer fails to see that she will pay a large penalty and back-load repayment—and *not* because she has a taste for immediate gratification with respect to consumption—she underestimates the cost of credit and borrows too much. Due to this combination of decisions, a non-sophisticated consumer, no matter how close to sophisticated, has discontinuously lower welfare than a sophisticated consumer. This discontinuity demonstrates in an extreme form our main point regarding contracts and welfare in the credit market: that because the credit contracts firms design in response postulate large penalties for deferring repayment, even relatively minor mispredictions of preferences by borrowers can

have large welfare effects.

Given the low welfare of non-sophisticated borrowers in the unrestricted market, we turn to identifying welfare-improving interventions. Because in a restricted market borrowers have the option of paying a small fee for deferring a small amount of repayment, non-sophisticated but not-too-naive borrowers do not drastically mispredict their future behavior, and hence have higher utility than in the unrestricted market. Since sophisticated borrowers achieve the highest possible utility in both markets, this means that a restricted market often Pareto-dominates the unrestricted one. If many borrowers are very naive, a restricted market can be combined with an interest-rate cap to try to limit borrowers' misprediction and achieve an increase in welfare.

The properties of non-sophisticated borrowers' competitive-equilibrium contracts, and the restriction disallowing disproportionately large penalties for deferring small amounts of repayment, have close parallels in real-life credit markets and their regulation. As has been noted by researchers, the baseline repayment terms in credit-card and subprime mortgage contracts are typically quite strict, and there are large penalties for deviating from these terms. For example, most subprime mortgages postulate drastically increased monthly payments shortly after the origination of the loan or a large "balloon" payment at the end of a short loan period, and failing to make these payments and refinancing triggers significant prepayment penalties. Similarly, most credit cards do not charge interest on any purchases if a borrower pays the entire balance due within a short one-month grace period, but do charge interest on all purchases if she revolves even $1. To protect borrowers, new regulations restrict these and other practices involving large penalties: in July 2008 the Federal Reserve Board severely limited the use of prepayment penalties, and the Credit CARD Act of 2009 prohibits the use of interest charges for partial balances the consumer has paid off, and restricts fees in other ways. Opponents have argued that these regulations will decrease the amount of credit available to borrowers and exclude some borrowers from the market. Our model predicts the same thing, but also says that this will benefit rather than hurt consumers—who have been borrowing too much and will now borrow less because they better understand the cost of credit.

In Section 2.4, we consider equilibria when β is unknown to firms, and show that with two important qualifications the key results above survive. First, since sophisticated and non-sophisticated borrowers with the same $\hat{\beta}$ are now indistinguishable to firms, the two types sign the same contract in period 0. This contract has a low-cost front-loaded repayment schedule that a sophisticated borrower chooses, and a high-cost back-loaded repayment schedule that a non-sophisticated borrower chooses. As before, even if a non-sophisticated

borrower is close to sophisticated, the only way she can deviate from the front-loaded repayment schedule is by paying a large fee. Furthermore, we identify reasonable conditions under which consumers self-select in period 0 into these same contracts even if β and $\hat{\beta}$ are both unknown to firms. Second, while the restricted market does not Pareto-dominate the unrestricted one, we establish that for *any* proportion of sophisticated and non-sophisticated borrowers, if non-sophisticated borrowers are not too naive, then the restricted market has higher total welfare.

In Section 2.5, we generalize our basic model—in which a non-sophisticated borrower believes with certainty that her taste for immediate gratification is above β—as well as other existing models of partial naivete and allow borrower beliefs to be a full distribution $F(\hat{\beta})$. We show that whether or not borrower beliefs are known, the qualitative predictions we have emphasized for non-sophisticated borrowers—overborrowing, often paying large penalties, and getting discretely lower welfare than sophisticated borrowers—depend not on $F(\beta) = 0$, but on $F(\beta)$ being bounded away from 1. Since this condition is likely to hold for many or most forms of near-sophisticated borrower beliefs, our observation that small mispredictions have large welfare effects is quite general. For example, even if the borrower has extremely tightly and continuously distributed beliefs centered around her true β, her welfare is not close to that of the sophisticated borrower. We also highlight an important asymmetry: while overestimating one's self-control, even probabilistically and by a small amount, has significant welfare implications, underestimating it has no welfare consequences whatsoever.

In Section 2.6, we discuss how our theory contributes to the literature on contracting with time-inconsistent or irrational consumers and relates to neoclassical screening. We are not aware of a theory with rational time-consistent borrowers that explains the key contract features predicted by our model, and we argue that natural specifications do not do so. Because the main predictions of our model are about repayment terms, the most likely neoclassical screening explanation would revolve around heterogeneity in borrowers' ability to repay the loan early. If borrowers know at the time of contracting whether they can repay fast, a lender will offer an expensive loan with back-loaded repayment intended for those who cannot, but achieving this using a prepayment penalty and going through the costs refinancing is inefficient. If borrowers do not know at the time of contracting whether they can repay fast, a model of sequential screening (Courty and Li 2000) or post-contractual hidden knowledge predicts that—analogously to business travelers' expensive but flexible airline tickets—the optimal loan is expensive if repaid quickly but allows borrowers to cheaply change the repayment schedule. This is of course exactly the opposite pattern

of what we find and what is the case in reality.

In Section 2.7, we conclude the paper by emphasizing some shortcomings of our framework, especially the importance of studying two major questions raised by our results: what regulations non-sophisticated borrowers will accept, and whether and how borrowers might learn about their time inconsistency. Proofs are in the Web Appendix.

2.2 A Model of the Credit Market

2.2.1 Setup

In this section, we introduce our model of the credit market, beginning with borrower behavior. There are three periods, $t = 0, 1, 2$. Self 0's utility is $c - k(q) - k(r)$, where $c \geq 0$ is the amount the consumer borrows in period 0 and $q \geq 0$ and $r \geq 0$ are the amounts she repays in periods 1 and 2, respectively.[4] Self 1 maximizes $-k(q) - \beta k(r)$, where β satisfying $0 < \beta \leq 1$ parameterizes the time-inconsistent taste for immediate gratification (as in Laibson 1997). Note that while self 1 discounts the future cost of repayment by a factor of β, because much of the borrowing motivating our analysis is for future consumption,[5] self 0—from whose perspective c, q, r are all in the future—does not discount the cost of repayment relative to the utility from consumption. The cost function $k(\cdot)$ is twice continuously differentiable with $k(0) = 0, \beta > k'(0) > 0$, $k''(x) > 0$ for all $x \geq 0$, and $\lim_{x\to\infty} k'(x) = \infty$. Our results would not fundamentally change if the utility from consumption c was concave instead of linear. Moreover, since self 1 makes no decision regarding c, under separability from the cost of repayment our analysis would be unaffected if—as is reasonable for mortgages and durable goods—the utility from consumption was decomposed into a stream of instantaneous utilities and added to self 1's utility function.

Following O'Donoghue and Rabin's (2001) formulation of partial naivete, we assume that self 0 believes with certainty that self 1 will maximize $-k(q) -$

[4] The bounds on q and r are necessary for a competitive equilibrium to exist when β and $\hat{\beta}$ defined below are known. In this case, the model yields a corner solution for the amount the borrower expects to pay in period 2. Any finite lower bound, including a negative one, yields the same qualitative results. Section 2.4 demonstrates that when β is unknown and $k'(0)$ is sufficiently low, the bounds are not binding.

[5] Most mortgages require substantial time and effort during the application process, and yield mostly delayed benefits of enjoying the new or repaired home. Similarly, a significant amount of credit-card spending seems to be on durables and other future-oriented goods (Hayhoe, Leach, Turner, Bruin and Lawrence 2000, Reda 2003).

$\hat{\beta}k(r)$, where $\beta \leq \hat{\beta} \leq 1$. The parameter $\hat{\beta}$ reflects self 0's beliefs about β, so that $\hat{\beta} = \beta$ corresponds to perfect sophistication regarding future preferences, $\hat{\beta} = 1$ corresponds to complete naivete about the time inconsistency, and more generally $\hat{\beta}$ is a measure of sophistication. Because the O'Donoghue-Rabin specification of partial naivete using degenerate beliefs is special, in Section 2.5 we allow borrower beliefs to be any distribution, and show that so long as a non-sophisticated borrower attaches non-trivial probability to her time inconsistency being above β, most of our qualitative results survive. In addition, although evidence indicates that people are more likely to have overly optimistic beliefs ($\hat{\beta} > \beta$), in Section 2.5 we consider the possibility of overly pessimistic beliefs ($\hat{\beta} < \beta$), and show that—unlike overoptimism—this mistake has no consequences in equilibrium.

We think of a group of consumers who are indistinguishable by firms as a separate market, and will define competitive equilibrium for a single separate such market. We assume that the possible β's in a market are $\beta_1 < \beta_2 < \cdots < \beta_I$, and $\hat{\beta} \in \{\beta_2, \ldots \beta_I\}$. For any given $\hat{\beta} = \beta_i$, the borrower has $\beta = \beta_i$ with probability p_i and $\beta = \beta_{i-1}$ with probability $1 - p_i$. If firms observe $\hat{\beta}$, then $I = 2$; and if they also observe β, then in addition $p_2 = 0$ or $p_2 = 1$.

Since the credit market seems relatively competitive—at least at the initial stage of contracting—we assume that the borrowers introduced above interact with competitive, risk-neutral, profit-maximizing lenders.[6] For simplicity, we assume that firms face an interest rate of zero, although this does not affect any of our qualitative results. Borrowers can sign non-linear contracts in period 0 regarding consumption and the repayment schedule, and these contracts are exclusive: once a consumer signs with a firm, she cannot interact with other firms.[7] An *unrestricted credit contract* is a consumption level c along with a

[6] By standard indicators of competitiveness, the subprime loan origination market seems quite competitive: no participant has more than 13% market share (Bar-Gill 2008). By similar indicators, the credit-card market is even more competitive. For the subprime mortgage market, however, observers have argued that because borrowers find contract terms confusing, they do not do much comparison shopping, so the market is de facto not very competitive. Our analysis will make clear that when $\hat{\beta}$ is known, the features and welfare properties of contracts are the same in a less competitive market. But Section 2.4.2.4.2's and Section 2.5's results on the sorting of consumers according to their beliefs in period 0 do take advantage of our competitiveness assumption.

[7] While the effects of relaxing exclusivity warrants further research, in general it would not eliminate our main points regarding non-sophisticated borrowers. Even if borrowers had access to a competitive market in period 1, our results remain unchanged so long as the original firm can include in the contract a fee—such as the prepayment penalties in subprime mortgages—for refinancing with *any* firm in the market. If firms cannot postulate such a fee for refinancing on the competitive market, then in our three-period setting a borrower will always avoid repaying more than expected. But as predicted by O'Donoghue and Rabin

finite menu $\mathcal{C} = \{(q_s, r_s)\}_{s \in S}$ of repayment options, and is denoted by $(c, \mathcal{C})$. To focus on the role of borrower mispredictions regarding repayment, we suppose that there is no possibility of default. Note that this specification allows the set of repayment options to be a singleton $\{(q, r)\}$, committing the borrower's future behavior and fully solving her self-control problem.

To enable us to focus on the contracts accepted by consumers, we suppress the strategic interaction between firms and define equilibrium directly in terms of the contracts that survive competitive pressure.[8] Since a borrower's behavior in period 0 can depend only on $\hat{\beta}$, the competitive equilibrium will be a set of contracts $\{(c_i, \mathcal{C}_i)\}_{i \in \{2,\ldots,I\}}$ for the possible $\hat{\beta}$ types β_2 through β_I.[9] For a firm to calculate the expected profits from a contract, and for a borrower to decide which of the contracts available on the market to choose, market participants must predict how a borrower will behave if she chooses a given contract. They do this through an incentive-compatible map:

Definition 2.1. The maps $q_i, r_i : \{\beta_1, \ldots, \beta_I\} \to \mathbb{R}_+$ are jointly *incentive compatible* for $\mathcal{C}_i$ if $(q_i(\beta), r_i(\beta)) \in \mathcal{C}_i$ for each $\beta \in \{\beta_1, \ldots, \beta_I\}$, and

$$-k(q_i(\beta)) - \beta k(r_i(\beta)) \geq -k(q) - \beta k(r) \text{ for all } (q, r) \in \mathcal{C}_i.$$

A consumer of type $(\hat{\beta}, \beta)$ believes in period 0 that she will choose $(q_i(\hat{\beta}), r_i(\hat{\beta}))$ from $\mathcal{C}_i$, whereas in reality she chooses $(q_i(\beta), r_i(\beta))$ if confronted with $\mathcal{C}_i$. Based on the notion of incentive compatibility, we define:

Definition 2.2. A *competitive equilibrium* is a set of contracts $\{(c_i, \mathcal{C}_i)\}_{i \in \{2,\ldots,I\}}$ and incentive-compatible maps $(q_i(\cdot), r_i(\cdot))$ for each $\mathcal{C}_i$ with the following properties:

(2001) and is consistent with evidence in Shui and Ausubel (2004), in a more realistic, long-horizon, setting non-sophisticated borrowers may procrastinate for a long time before finding or taking advantage of favorable refinancing opportunities. And even if a non-sophisticated borrower refinances, she might perpetually do so using contracts of the type we predict, and eventually repay according to such a contract. Indeed, Engel and McCoy (2002) document that subprime mortgages are often refinanced with similarly structured loans, and credit-card balance-transfer deals and teaser rates also draw consumers into contracts similar to those they had before.

[8] This approach is similar in spirit to Rothschild and Stiglitz's (1976) definition of competitive equilibrium with insurance contracts. By thinking of borrowers as sellers of repayment schedules $\mathcal{C}$, lenders as buyers of these schedules, and c as the price of a schedule $\mathcal{C}$, we can modify Dubey and Geanakoplos's (2002) competitive-equilibrium framework for our setting in a way that yields the same contracts as Definition 2.2.

[9] Although in principle different borrowers with the same $\hat{\beta}$ may choose different contracts, by assuming that there is exactly one contract for one $\hat{\beta}$ type, this approach for simplicity imposes that they do not.

1. [Borrower optimization] For any $\hat{\beta} = \beta_i \in \{\beta_2, \ldots, \beta_I\}$ and $j \in \{2, \ldots, I\}$, one has $c_i - q_i(\hat{\beta}) - r_i(\hat{\beta}) \geq c_j - q_j(\hat{\beta}) - r_j(\hat{\beta})$.

2. [Competitive market] Each $(c_i, \mathcal{C}_i)$ yields zero expected profits.

3. [No profitable deviation] There exists no contract $(c', \mathcal{C}')$ with jointly incentive-compatible maps $(q'(\cdot), r'(\cdot))$ such that (i) for some $\hat{\beta} = \beta_i$, $c' - q'(\hat{\beta}) - r'(\hat{\beta}) > c_i - q_i(\hat{\beta}) - r_i(\hat{\beta})$; and (ii) given the types for whom (i) holds, $(c', \mathcal{C}')$ yields positive expected profits.

4. [Non-redundancy] For each $(c_i, \mathcal{C}_i)$ and each installment plan $(q_j, r_j) \in \mathcal{C}_i$, there is a type $(\hat{\beta}, \beta)$ with $\hat{\beta} = \beta_i$ such that either $(q_j, r_j) = (q_i(\hat{\beta}), r_i(\hat{\beta}))$ or $(q_j, r_j) = (q_i(\beta), r_i(\beta))$.

Our first requirement for competitive equilibrium is that of borrower optimization: given a type's predictions about how she would behave with each contract, she chooses her favorite one from the perspective of period 0. Our next two conditions are typical for competitive situations, saying that firms earn zero profits by offering these contracts, and that firms can do no better.[10]

The last, non-redundancy, condition says that all repayment options in a contract are relevant in that they affect the expectations or behavior of the consumer accepting the contract. This assumption simplifies statements regarding the uniqueness of competitive equilibrium, but does not affect any of our predictions regarding outcomes and welfare.[11] Due to the non-redundancy condition, the competitive-equilibrium contracts we derive exclude most options by assumption; in particular, non-sophisticated borrowers' only option to change the repayment schedule will be to change it by a lot for a large fee. As is usually the case in models of non-linear pricing, the same outcomes can also be implemented by allowing other choices, but making them so expensive that the borrower does not want to choose them. In fact, this is how it works in the real-life examples discussed below, where deferring even small amounts of repayment carries disproportionately large fees.

[10] We could have required a competitive equilibrium to be robust to deviations involving multiple contracts, rather than the single-contract deviations above. In our specific setting, this makes no difference to the results. This is easiest to see when $\hat{\beta}$ is known: then, offering multiple contracts instead of one cannot help a firm separate different consumers, so it cannot increase profits.

[11] For general distributions of β and $\hat{\beta}$, our definition of non-redundancy would have to be more inclusive. Specifically, it would have to allow for a repayment schedule in $\mathcal{C}_i$ to be the expected choice from $\mathcal{C}_i$ of a consumer type *not* choosing $(c_i, \mathcal{C}_i)$—because such an option could play a role in *preventing* the consumer from choosing $(c_i, \mathcal{C}_i)$. Clearly, this consideration is unimportant if $\hat{\beta}$ is known. Given our assumptions, it is also unimportant if $\hat{\beta}$ is unknown, because the competitive equilibrium in Section 2.4.2.4.2 already fully sorts consumers according to $\hat{\beta}$.

One of our main interests in this paper is to study borrower welfare in the above market, and to find welfare-improving interventions. While using self 1's or self 2's utility as our welfare measure will often yield similar insights (because the overborrowing our model predicts implies that in the unrestricted market selves 1 and 2 are stuck having to repay large amounts), we follow much of the literature on time inconsistency (DellaVigna and Malmendier 2004, Gruber and Kőszegi 2004, O'Donoghue and Rabin 2006, for example) and identify welfare with long-run, period-0, preferences.[12] In our stylized setting, there are then many ways of increasing welfare. Notably, since the optimal outcome c, q, r is known and easy to describe—equating the marginal cost of repayment in each period with the marginal utility of consumption, $k'(q) = k'(r) = 1$, and $c = q + r$—a policy just mandating this allocation is an optimal policy. But we are interested in more plausible policies, ones that do not cause harm because of features of the credit market missing from our model—which such a mandate clearly does if the social planner does not know an individual borrower's preferences.[13] Hence, we will focus on interventions that leave substantial flexibility in market participants' hands, and that target the central contract feature generating low welfare: that non-sophisticated borrowers' only way to reschedule repayment is to pay a large penalty. We propose to restrict contracts by requiring them to allow the deferral of small amounts of repayment, and—more importantly—prohibiting disproportionately large penalties for deferring small amounts. Since (as we argue in Section 2.6) the large penalties are unlikely to be serving a neoclassical purpose, and we are also unaware of unmodelled "behavioral" reasons for them, such a policy is unlikely to do harm. Indeed, we discuss parallels between our restriction and recent new regulations in the credit-card and mortgage markets.

Formally, in a *restricted market* the permissible repayment options must form a linear set: the contract specifies some R and L, and the set of permissible repayment schedules is $\{(q, r)|q + r/R = L \text{ and } q, r \leq M\}$, where M is an

[12] Although we simplify things by considering a three-period model, in reality time inconsistency seems to be mostly about very immediate gratification that plays out over many short periods. Hence, arguments by O'Donoghue and Rabin (2006) in favor of a long-run perspective apply: in deciding how to weight any particular week of a person's life relative to future weeks, it is reasonable to snub that single week's self—who prefers to greatly downweight the future—in favor of the many earlier selves—who prefer more equal weighting. In addition, the models in Bernheim and Rangel (2004a, 2004b) can be interpreted as saying that a taste for immediate gratification is often a mistake not reflecting true welfare.

[13] Because in our model all consumers know their future circumstances in period 0, another optimal policy is to require borrowers to commit fully to a repayment schedule. As Amador, Angeletos, and Werning (2006) show, however, this intervention is suboptimal if consumers are subject to ex-post shocks in their financial circumstances.

exogenous bound on q and r that can be arbitrarily large and that we impose as a technical condition to ensure the existence of competitive equilibrium, and for which we require $k'(M) > 1/\beta$.[14] As we note below, many other ways of eliminating disproportionately large penalties have the same or similar welfare effect.

2.2.2 A Preliminary Step: Restating the Problem

As a preliminary step in our analysis, we restate in contract-theoretic terms the requirements of a competitive equilibrium when $\hat{\beta}$ is known and the consumer may be non-sophisticated ($I = 2, p_2 < 1$). To help understand our restatement, imagine a firm trying to maximize profits from a borrower who has an outside option with perceived utility $\underline{u}$ for self 0. Restricting attention to non-redundant contracts, we can think of the firm as selecting consumption c along with a "baseline" repayment schedule $(q_2(\beta_2), r_2(\beta_2))$ the borrower expects to choose in period 0 and that a sophisticated type (if present) actually chooses in period 1, and an alternative repayment schedule $(q_2(\beta_1), r_2(\beta_1))$ a non-sophisticated borrower actually chooses in period 1. In designing its contract, the firm faces the following constraints. First, for the borrower to be willing to accept the firm's offer, self 0's utility with the baseline schedule must be at least $\underline{u}$. This is a version of the standard participation constraint (PC), except that self 0 may make her participation decision based on incorrectly forecasted future behavior. Second, if self 0 is to think that she will choose the baseline option, then given her beliefs $\hat{\beta}$ she must think she will prefer it to the alternative option. We call this constraint a perceived-choice constraint (PCC). Third, if a non-sophisticated consumer is to actually choose the alternative repayment schedule, she has to prefer it to the baseline. This is analogous to a standard incentive-compatibility constraint (IC) for self 1.

It is clear that a competitive-equilibrium contract must be a solution to the above maximization problem with $\underline{u}$ defined as self 0's perceived utility from accepting this contract: given that a competitive-equilibrium contract earns zero profits, if this was not the case, a firm could solve for the optimal contract and increase c slightly, attracting all consumers and making strictly positive expected profits. In addition, for the solution to the above maximization problem to be a competitive equilibrium, $\underline{u}$ must be such that the highest achievable expected profit is zero. In fact, this is also sufficient:

[14] Strictly speaking, we have defined a competitive equilibrium only for the case of unrestricted contracts. When considering the restricted market, one needs to replace the finite set of repayment options $\mathcal{C}_i$ with an infinite but linear set.

Lemma 2.1. *Suppose $\hat{\beta}$ is known ($I = 2$), the possible β's are $\beta_1 < \hat{\beta}$ and $\beta_2 = \hat{\beta}$, and $p_2 < 1$. The contract with consumption c and repayment options $\{(q_2(\beta_1), r_2(\beta_1)), (q_2(\beta_2), r_2(\beta_2))\}$ is a competitive equilibrium if and only if there is a $\underline{u}$ such that the contract maximizes expected profits subject to a PC with perceived outside option $\underline{u}$, PCC, and IC, and the profit level when maximizing profits subject to these constraints is zero.*

2.3 Non-Linear Contracting with Known β and $\hat{\beta}$

We begin our analysis of non-linear contracting with the case when both β and $\hat{\beta}$ are known. We show that non-sophisticated borrowers get a very different contract from sophisticated ones, and because they mispredict whether they will pay the large penalty their contract postulates for changing the repayment schedule, they have discontinuously lower welfare. We establish that prohibiting such large penalties for deferring small amounts of repayment can raise welfare. Finally, we show that the misprediction of time-consistent preferences has no implications for outcomes, indicating that time inconsistency is necessary for our results.

2.3.1 Competitive Equilibrium with Unrestricted Contracts

We start with the remark that if borrowers are time consistent and rational, the organization of the credit market does not matter:

Fact 2.1. *If $\beta = \hat{\beta} = 1$, the competitive-equilibrium consumption and repayment outcomes are the same in the restricted and unrestricted markets, and both maximize welfare.*

For the rest of the paper (with the exception of Section 2.3.2.3.3), we assume that $\beta < 1$. First, we consider the case of a perfectly sophisticated borrower, for whom $\hat{\beta} = \beta$. By the same logic as in DellaVigna and Malmendier (2004), since a sophisticated borrower correctly predicts her own behavior, it is profit-maximizing to offer her a contract that maximizes her utility:

Proposition 2.1. *Suppose β and $\hat{\beta}$ are known, and $\hat{\beta} = \beta$. Then, the competitive-equilibrium contract has a single repayment option satisfying $k'(q) = k'(r) = 1$, and $c = q + r$.*

The situation is entirely different for a non-sophisticated borrower, for whom $\hat{\beta} > \beta$. Applying Lemma 2.1, the competitive-equilibrium contract consists of a consumption level c, a repayment schedule (q, r) self 1 actually

chooses, and a possibly different baseline repayment schedule $(\hat{q}, \hat{r})$ self 0 expects to choose, that solve

$$\max_{c,q,r,\hat{q},\hat{r}} \quad q + r - c \tag{2.1}$$

$$\text{s.t.} \quad c - k(\hat{q}) - k(\hat{r}) \geq \underline{u}, \tag{PC}$$

$$-k(\hat{q}) - \hat{\beta}k(\hat{r}) \geq -k(q) - \hat{\beta}k(r), \tag{PCC}$$

$$-k(q) - \beta k(r) \geq -k(\hat{q}) - \beta k(\hat{r}), \tag{IC}$$

PC binds because otherwise the firm could increase profits by reducing c. In addition, IC binds because otherwise the firm could increase profits by increasing q. Given that IC binds and $\hat{\beta} > \beta$, PCC is equivalent to $q \leq \hat{q}$: if self 1 is in reality indifferent between two repayment options, then self 0—who overestimates her future self-control by at least a little bit—predicts she will prefer the more front-loaded option. Conjecturing that $q \leq \hat{q}$ is optimal even without PCC, we ignore this constraint, and confirm our conjecture in the solution to the relaxed problem below.

Given the above considerations, the problem becomes

$$\max_{c,q,r,\hat{q},\hat{r}} \quad q + r - c$$

$$\text{s.t.} \quad c - k(\hat{q}) - k(\hat{r}) = \underline{u}, \tag{PC}$$

$$-k(q) - \beta k(r) = -k(\hat{q}) - \beta k(\hat{r}). \tag{IC}$$

Notice that in the optimal solution, $\hat{r} = 0$: otherwise, the firm could decrease $k(\hat{r})$ and increase $k(\hat{q})$ by the same amount, leaving PC unaffected and creating slack in IC, allowing it to increase q. Using this, we can express $k(\hat{q})$ from IC and plug it into PC to get $c = k(q) + \beta k(r) + \underline{u}$. Plugging c into the firm's maximand yields the unconstrained problem

$$\max_{q,r} \quad q + r - k(q) - \beta k(r) - \underline{u},$$

and gives the following proposition:

Proposition 2.2. *Suppose β and $\hat{\beta} > \beta$ are known. Then, the competitive-equilibrium contract has a baseline repayment schedule $(\hat{q}, \hat{r})$ satisfying $\hat{q} > 0, \hat{r} = 0$ that the borrower expects to choose and an alternative schedule (q, r) satisfying $k'(q) = 1, k'(r) = 1/\beta$ that she actually chooses. Consumption is $c = q + r > \hat{q}$, and is higher than that of a sophisticated borrower. The borrower has strictly lower welfare than a sophisticated borrower.*

The first important feature of the equilibrium contract is that it is flexible in a way that induces the borrower to unexpectedly change her mind regarding how she repays. To see why this is the case, consider why the sophisticated borrower's contract—which is also the non-sophisticated borrower's favorite *among* fully committed contracts—is not a competitive equilibrium. The reason is that a firm can deviate by offering slightly higher consumption and still allow the same repayment terms, but introduce an alternative option to defer part of the first installment for a fee. Thinking that she will not use the alternative option, the consumer likes the deal. But since she does use the option, the firm earns higher profits than with a committed contract.

Beyond showing that the equilibrium contract is flexible in a deceptive way, Proposition 2.2 says that $k'(q) = \beta k'(r)$, so that self 1's preferences fully determine the allocation of actual repayment across periods 1 and 2. Hence, the ability to commit perfectly to a repayment schedule does not mitigate the consumer's time inconsistency regarding repayment at all. Intuitively, once a firm designs the contract to induce repayment behavior self 0 does not expect, its goal with the chosen option is to maximize the gains from trade with the self that makes the repayment decision, so it caters fully to self 1's taste for immediate gratification.

To make matters worse, the competitive-equilibrium contract induces overborrowing in two senses: the non-sophisticated consumer borrows more than the sophisticated one, and she borrows more than is optimal *given that* repayment is allocated according to self 1's preferences.[15] Unlike in existing models of time inconsistency, self 0 overborrows not because she undervalues the cost of repayment relative to consumption, but because she mispredicts how she will repay her loan, in effect leading her to underestimate its cost. To see how the exact level of c is determined, recall that the contract is designed so that self 0 expects to finish her repayment obligations in period 1 ($\hat{r} = 0$). Hence, when deciding whether to participate, self 0 trades off c with $k(\hat{q})$. But from the firm's perspective, $k(\hat{q})$ is just the highest actual total cost of repayment that can be imposed on self 1 so that she is still willing to choose the alternative installment plan. This means that the tradeoff determining the profit-maximizing level of borrowing is between c and *self 1's* cost of repayment, which discounts the second installment by β.

Notice that due to the excessive borrowing in period 0, the non-sophisticated

[15] The prediction regarding the amount of borrowing contrasts with predictions of hyperbolic discounting in standard consumption-savings problems, such as Laibson (1997). In those problems, whether more naive decisionmakers borrow more or less than sophisticated ones depends on the per-period utility function. In our setting, non-sophisticated consumers borrow more for any $k(\cdot)$.

borrower is worse off than the sophisticated one not only from the perspective of period 0, but also from the perspective of period 1—repaying the same amount in period 1 and more in period 2. Hence, the fact that the borrower is fooled into changing her mind and allocating repayment according to self 1's preferences is ultimately worse for self 1 as well.

All of the above holds for *any* $\hat{\beta} > \beta$, so that all non-sophisticated borrowers, even near-sophisticated ones, receive discretely different outcomes from and discretely lower welfare than sophisticated borrowers. The discontinuity is an extreme form of one of our main points in the paper: that due to the credit contracts profit-maximizing firms design in response, even small mispredictions of preferences by borrowers often have large welfare effects. The welfare effects are large because a borrower is allowed to change her repayment schedule only by paying a large fee, and the fee is designed so that she mispredicts whether she will pay it.[16] Hence, even if self 0 mispredicts her future *utility* by only a little bit, she mispredicts her future *outcomes* by a lot, and because she is time-inconsistent this means she mispredicts her *welfare* by a lot—repaying her loan in a much more costly way than she expects.

While our main interest is in the implemented repayment schedule (q, r), the structure of the baseline schedule $(\hat{q}, \hat{r})$ is also intriguing: the firm asks the borrower to carry out all repayment in period 1, even if the marginal cost of repaying a little bit in period 2 is very low. Intuitively, because the baseline terms are never implemented, the firm's goal is not to design them efficiently. Instead, its goal is to attract the consumer in period 0 without reducing the total amount she is willing to pay through the installment plan she actually chooses in period 1. Front-loading the baseline repayment schedule achieves this purpose by making the schedule relatively more attractive to self 0 than to self 1.

Finally, the above analysis makes it clear how competition matters: through $\underline{u}$. For a monopolist, $\underline{u}$ is a borrower's perceived outside option when not taking a loan. In a perfectly competitive market, $\underline{u}$ is set endogenously such that profits are zero. Since the repayment options in the optimal contract are independent of $\underline{u}$, whether the market is perfectly competitive or monopolistic matters only for determining the consumption level c.[17]

[16] As we have mentioned above, the fact that a borrower literally has no other option but to pay a large fee and defer a large amount of repayment follows from the non-redundancy condition in Definition 2.2. The same outcome can also be implemented by allowing the deferral of small amounts of repayment, but charging disproportionately large fees for this—as the real-life contracts we discuss do.

[17] In a Hotelling-type model of imperfect competition in contract offers, an intermediate level of competition generates a contract identical to that implied by the above analysis for a

The properties of the non-sophisticated borrower's competitive-equilibrium contract—a relatively low-cost front-loaded repayment schedule with a large penalty to switch out of it—arguably closely resemble some features of real-life credit arrangements.[18] Loaded with cash-back bonuses, free rental-car insurance, and other perks, the typical credit-card deal is extremely favorable—so long as the consumer repays all of her debt within the one-month grace period. If she revolves even $1, she is charged interest on *all* purchases, and all of a sudden credit-card use becomes quite expensive. Similarly, in-store financing and credit-card balance-transfer deals often involve no interest for a few months, but if a consumer does not repay fully within the allotted time, she is charged interest from the time of purchase. Most credit cards also charge late-payment, over-the-limit, and other fees that are large even for small violations of terms. In the subprime market, the most common, "hybrid," form of mortgage starts with low payments, but after a short period resets to high monthly payments that will be difficult for most borrowers to meet. Even more extreme is the "balloon" mortgage, which requires the borrower to pay off the entire remaining balance in a large payment at the end of a relatively short loan period. In addition, these types of mortgages typically include hefty prepayment penalties.[19] As emphasized by Hill and Kozup (2007) and especially Renuart (2004) and as the logic of our model suggests, the high monthly payments or the balloon payment drive borrowers to refinance, and the high prepayment penalty—folded into the principal and financed—serves to make

level of $\underline{u}$ that is in-between the competition and monopoly extremes, with the appropriate $\underline{u}$ increasing monotonically as competition increases and approaching that in the competitive market above. Formally, suppose there are two firms A and B located at the endpoints of the unit interval, and there is a mass one of borrowers uniformly distributed along this interval. The period-0 self of a borrower located at χ derives utility $c^A - k(q^A) - k(r^A) - d\chi$ from firm A's contract, where c^A is the consumption level offered by firm A and q^A and r^A are the repayments made to firm A. The period-0 self of the same borrower derives utility $c^B - k(q^B) - k(r^B) - d(1-\chi)$ from firm B's contract, and 0 when rejecting both firms' contract offers. To find the equilibrium contract offers, think of firm A as first maximizing its profits for any perceived utility $\underline{u} = c^A - k(\hat{q}^A) - k(\hat{r}^A)$ it chooses to offer to the borrower located at $\chi = 0$, and then selecting the optimal perceived utility level for this borrower. The first step is identical to the problem above, so the repayment options are also identical to those found above. Optimizing over c gives that if d is sufficiently low, the market is covered in equilibrium and $c = q + r - d$, generating a $\underline{u}$ that increases with an increase in competition as captured by a decrease in d.

[18] We focus on the non-sophisticated borrower's contract because (as we show in Section 2.4) when β is unknown sophisticated and non-sophisticated borrowers accept the same contract, and this contract much resembles the above contract for non-sophisticated borrowers.

[19] Demyanyk and Van Hemert (2008) report that 54.5 percent of US subprime mortgages originated in 2006 were of the hybrid type, 25.2 percent were of the balloon type, and 71 percent postulated a prepayment penalty.

this profitable to the lender. In a practice known as "loan flipping," creditors sometimes refinance repeatedly (Engel and McCoy 2002). Indeed, Demyanyk and Van Hemert (2008) find that the majority of subprime mortgages is obtained for refinancing into a larger new loan for the purposes of extracting cash.[20]

2.3.2 A Welfare-Increasing Intervention

Given non-sophisticated borrowers' suboptimal welfare, it is natural to ask whether there are welfare-improving interventions. If borrowers are sufficiently sophisticated, there is a simple one:

Proposition 2.3. *A sophisticated borrower ($\hat{\beta} = \beta$) is equally well off in the restricted and unrestricted markets. If a non-sophisticated borrower ($\hat{\beta} > \beta$) is sufficiently sophisticated ($\hat{\beta}$ is sufficiently close to β), she is strictly better off in the restricted than in the unrestricted market.*

By counteracting her tendency for immediate gratification as given by β, a restricted contract with an interest rate $R = 1/\beta$ aligns self 1's behavior with the borrower's long-run welfare. And since sophisticated borrowers understand their own behavior perfectly, it is profit-maximizing to offer such a contract to them. Hence, for sophisticated borrowers the restricted and unrestricted markets both generate the highest possible level of utility.

More interestingly, restricting contracts to have a linear structure prevents firms from fooling non-sophisticated but not-too-naive borrowers into discretely mispredicting their behavior, and hence raises these borrowers' welfare. For any interest rate R, a slightly naive borrower mispredicts her future behavior by only a small amount, which leads her to make only a small mistake in how much she wants to borrow. This means that her behavior is very close to that of a sophisticated borrower, so that she gets a contract very close to that offered to a sophisticated borrower. As a result, her utility is close to optimal.

[20] A weakness of our theory is that it does not convincingly explain why contracts look so different in the prime and subprime mortgage markets. Many prime contracts feature very simple installment plans (for example, the same nominal payment every month for 30 years), and have little or no prepayment penalties. Although this is consistent with our theory if borrowers in the prime market are time-consistent, we find this explanation implausible. A simple plausible explanation (but one completely outside our theory) is that unlike borrowers in the subprime market, borrowers in the prime market have access to plenty of other sources of credit that would make refinancing their mortgage an unattractive way to make funds available for short-term consumption, substantively violating our exclusivity assumption.

In the case of observable β and $\hat{\beta}$ and sufficiently sophisticated borrowers, therefore, our intervention satisfies the most stringent criteria of "cautious" or "asymmetric" paternalism (Camerer, Issacharoff, Loewenstein, O'Donoghue and Rabin 2003): it greatly benefits non-sophisticated borrowers, while it does not hurt sophisticated borrowers. Furthermore, if everyone in the population is rational (sophisticated), the intervention has no effect on outcomes at all.

The linearity of the allowable set of repayment options is not fundamental for the intervention to be welfare-improving. What is important is to rule out disproportionately large penalties for deferring small amounts of repayment, preventing borrowers from discretely mispredicting their behavior. Any contract in which r is a convex function of q has this property. For instance, Proposition 2.3 still holds if we allow contracts with a "focal" installment plan $\bar{q}, \bar{r}$ and a higher interest rate when repaying less than $\bar{q}$ in period 1 than when repaying more. Similarly, we could allow linear contracts with meaningful bounds on how much can be repaid in period 1.

Some recently enacted regulations aimed at protecting borrowers in the mortgage and credit-card markets in the US are interpretable in terms of Proposition 2.3's message to prohibit large penalties for small deviations from contract terms. In July 2008, the Federal Reserve Board amended Regulation Z (implementation of the Truth in Lending Act) to severely restrict the use of prepayment penalties for high-interest-rate mortgages. By 12 C.F.R. §226.35(b)(2), a prepayment penalty can only apply for two years following the commencement of the loan, and only if the monthly payment does not change in the first four years. This regulation will prevent lenders from collecting a prepayment penalty by requiring a high payment in the near future that induces borrowers to refinance. Title I, Section 102.(a)-(b) of the Credit Card Accountability, Responsibility, and Disclosure (Credit CARD) Act of 2009 prohibits the use of interest charges for partial balances the consumer pays off within the grace period, and Section 101.(b) prohibits applying post-introductory interest rates to the introductory period, ruling out exactly the kinds of large penalties we have discussed above. The act also limits late-payment, over-the-limit, and other fees to be "reasonable and proportional to" the consumer's omission or violation.

Note that the restricted market mitigates non-sophisticated but not-too-naive consumers' overborrowing, so if there is a non-trivial proportion of these consumers in the population, lenders extend less total credit in the restricted market than in the unrestricted market. This insight is relevant for a central controversy surrounding the above regulations of the credit market. Opponents have repeatedly argued that the new regulations will decrease the amount of credit available to borrowers and exclude some borrowers from the market,

intimating that this will be bad for consumers.[21] Our model predicts that these opponents may well be right in predicting a decreased amount of credit, but also says that in as much as this happens, it will benefit rather than hurt consumers—because consumers were borrowing too much to start with.[22]

Proposition 2.3 holds in general only for sufficiently sophisticated borrowers because both restricted and unrestricted contracts can lead a very naive borrower to severely overestimate how much she will be willing to pay back in period 1. If many consumers are very naive and as a result establishing the restricted market is not in itself an effective intervention, this can be combined with other regulations to limit borrowers' misprediction of their own behavior. One simple regulation is to restrict the amount of repayment that can be shifted to period 2, mechanically limiting borrowers' mispredictions. Another possible regulation is to set an interest-rate cap. For some commonly used utility functions, in fact, non-sophisticated borrowers are better off in a restricted market with an interest-rate cap of even zero than in an unrestricted market:[23]

Proposition 2.4. *Suppose $k(x) = x^\rho$ for some $\rho > 1$ or $k(x) = (y-x)^{-\rho} - y^{-\rho}$ for some $y > 0, \rho > 0$. Then, for any $\hat{\beta} > \beta$, a borrower has higher utility in a restricted market with $R = 1$ than in an unrestricted market.*

Intuitively, in both the unrestricted market and in the restricted market with an interest-rate cap of zero (which will clearly bind), repayment is allocated across periods 1 and 2 according to self 1's preferences ($k'(q) = \beta k'(r)$). But because contracts are more restricted in the latter market, non-sophisticated borrowers mispredict their behavior by less, and hence do not overborrow as much. Of course, allowing at least a small positive interest rate leads to even higher welfare for non-sophisticated borrowers, because it

[21] See, for instance, "Senate Passes Credit-Card Reform Bill by Vote of 90-5," FOXBusiness, May 19, 2009, `http://www.foxbusiness.com/story/markets/senate-passes-credit-card-reform-bill-vote/`; and "How the Banks Plan to Limit Credit-Card Protections," Time, April 27, 2009 `http://www.time.com/time/politics/article/0,8599,1894041,00.html`.

[22] If we relax the simplifying assumption that $k'(0) < \beta$, the exclusion from the market mentioned above occurs in our model for a non-sophisticated but not-too-naive borrower with $1/\beta > k'(0) > 1$. Such a borrower participates in the unrestricted market, but will stay out of a restricted market—and because her marginal cost of repayment is greater than the benefit of consumption, staying out is the better outcome.

[23] These utility functions guarantee that with linear contracts, non-sophisticated consumers borrow more than sophisticated ones, and this and further overborrowing lowers ex-ante utility. Our proof makes use of these features, but no other feature of the utility functions in Proposition 2.4.

induces them to repay more of their loan earlier. Despite these advantages, an interest-rate cap is more problematic than other policies we suggest in this paper because it harms sophisticated borrowers with a low β by preventing them from getting the ex-ante optimal high-interest-rate contract. Hence, an interest-rate cap is welfare improving only if we are confident that there is a sizable portion of non-sophisticated borrowers in the population.

2.3.3 The Role of Time Inconsistency

The theory in this paper makes two major assumptions that deviate from most classical theories of the credit market: that borrowers have a time-inconsistent taste for immediate gratification, and that they might mispredict this taste. Since (as we have shown above) sophisticated consumers receive the maximum achievable level of utility, the misprediction of preferences is necessary for our central welfare results regarding overborrowing and suboptimal repayment. In this section, we show that the misprediction of time-*consistent* preferences has no welfare consequences for the borrower, establishing that time inconsistency is also necessary for our central results.

Suppose that the borrower's true period-1 utility is given by $-k(q) - k(r)$ (that is, $\beta = 1$), and she is time-consistent: self 0 weights the repayment costs the same way that she believes self 1 does. But self 0 might mispredict self 1's preferences, believing that self 1's utility will be $-k(q) - \hat{\beta}k(r)$ for some $\hat{\beta} \geq 1$. Hence, although true ex-ante utility is $c - k(q) - k(r)$, self 0 believes it to be $c - k(q) - \hat{\beta}k(r)$. This situation is conceivable, for instance, if self 0 mispredicts how painful it will be to make a loan payment in period 1 relative to period 2, but thinks that the decision to allocate repayment across the two periods should be made according to this pain. With these changes to the model, PCC in Problem 2.1 above does not change, while PC changes to $c - k(\hat{q}) - \hat{\beta}k(\hat{r}) \geq \underline{u}$ and IC changes to $-k(q) - k(r) \geq -k(\hat{q}) - k(\hat{r})$. Analyzing the resulting problem yields:

Proposition 2.5. *In the time-consistent model, for any $\hat{\beta} \geq \beta = 1$ the repayment schedule chosen by the borrower in a competitive equilibrium satisfies $k'(q) = k'(r) = 1$, and the borrowed amount is $c = q + r$.*

Proposition 2.5 says that the competitive-equilibrium contract maximizes the borrower's utility for any period-0 beliefs. As in the time-inconsistent case, for $\hat{\beta} > \beta$ the borrower is induced to unexpectedly change her mind and repay according to self 1's preferences—but this is the welfare-maximizing repayment schedule in the time-consistent case. In addition, because preferences are time-consistent—and hence the repayment schedule self 1 chooses is not more

costly from the ex-ante point of view than what self 0 expects—mispredicting repayment behavior does not lead the borrower to underestimate the cost of credit, so she does not overborrow.[24]

Although a non-sophisticated time-consistent borrower ends up maximizing ex-ante utility just like a sophisticated borrower, her contract is different in that it includes a very front-loaded repayment option $(\hat{q}, \hat{r})$ satisfying $\hat{q} > 0, \hat{r} = 0$ that she expects to choose. This is an artifact of the assumption that β and $\hat{\beta}$ are known: unlike in the time-inconsistent case we analyze in Section 2.4, under time-consistent preferences with β unknown a near-sophisticated borrower mispredicts her repayment behavior by only a little bit. Intuitively, fooling a borrower regarding her repayment schedule is profitable because it makes the lender's offer seem cheaper, and hence makes it easier to attract the borrower. With a near-sophisticated time-consistent borrower, however, a lender cannot make the loan seem much cheaper than it actually is. At the same time, because a sophisticated borrower will actually follow the ex-ante expected repayment schedule, if the firm does not know which type it is facing, fooling the near-sophisticated borrower by distorting the ex-ante expected repayment terms is costly. As a result, it is not optimal to fool her by more than a little.

2.4 Non-Linear Contracting with Unknown Types

This section investigates competitive equilibria when either β, or both β and $\hat{\beta}$, are unknown to firms. Beginning with the former case, we show that with two important qualifications, our key insights from Section 2.3 survive. First, because sophisticated and non-sophisticated consumers with the same beliefs cannot be distinguished by firms, these two types sign the same contract—although they still choose very different repayment schedules from that contract and have very different welfare levels. Second, a restricted market no longer Pareto dominates the unrestricted market—although it still has higher total welfare for any proportion of sophisticated and near-sophisticated borrowers. We then assume that both β and $\hat{\beta}$ are unknown, and identify conditions under which the competitive equilibrium remains the same as when $\hat{\beta}$ is known.

[24] That borrowers are completely unaffected by mispredicting time-consistent preferences relies on the market being competitive. Although allocations would still be efficient, a monopolist would use the borrower's misprediction to extract more rent. As in Laibson and Yariv (2007), in a competitive market firms give all of this rent back to borrowers in an effort to attract them.

2.4.1 Known $\hat{\beta}$, Unknown β

Suppose that a borrower's $\hat{\beta}$ is known ($I = 2$), and she has $\beta_1 < \hat{\beta}$ with probability p_1 and $\beta_2 = \hat{\beta}$ with probability p_2. For technical convenience, we assume that $k'(0) < p_1, 1-p_1, \beta_1$, which guarantees that first-order conditions throughout the section describe optimal choices.

Because sophisticated and non-sophisticated borrowers have the same beliefs in period 0, they accept the same contract. The following proposition identifies key features of this contract.

Proposition 2.6 (Period-1 Screening)**.** *Suppose $\hat{\beta}$ is known, and β takes the values $\beta_1 < \hat{\beta}$ and $\beta_2 = \hat{\beta}$ with probabilities p_1 and $p_2 = 1 - p_1$, respectively. The unique competitive-equilibrium contract, accepted by both types, has two installment plans (q_1, r_1) and (q_2, r_2), which are chosen in period 1 by types β_1 and β_2, respectively. These satisfy $q_1 < r_1$, $q_2 > r_2$, $q_1 + r_1 > q_2 + r_2$, and*

$$\frac{k'(q_2)}{k'(r_2)} - 1 = (1-\beta_1) \cdot \frac{k'(q_2)}{k'(q_1)} \cdot \frac{p_1}{p_2}, \tag{2.2}$$

$$\frac{k'(q_1)}{k'(r_1)} - \beta_1 = 0. \tag{2.3}$$

Furthermore, consumers overborrow on average: $p_1 k'(q_1) + p_2 k'(q_2), p_1 k'(r_1) + p_2 k'(r_2) > 1$.

By Equation 2.2, the sophisticated borrower's repayment schedule calls for a first installment that is too high even from the long-term perspective of period 0. And by Equation 2.3, the non-sophisticated borrower's repayment schedule caters fully to self 1's preferences. These results are closely related to those in standard screening problems in which the tradeoff between increasing efficiency for the less profitable type and decreasing the information rent paid to the more profitable type leads to a distorted outcome for the less profitable type and an efficient outcome for the more profitable type. In our model, however, the relevant preferences in this tradeoff exist at different times. Since a sophisticated borrower sticks to her ex-ante preferred installment plan, the profit the firm can extract from her depends on period-0 preferences, so this side of the tradeoff takes the period-0 perspective. But since a non-sophisticated borrower abandons her ex-ante preferred installment plan, the profit the firm can extract from her depends partly on period-1 preferences, so this side of the tradeoff takes the period-1 perspective.

The difference between the sophisticated and non-sophisticated borrowers' first-order conditions implies a generalization of our insight above that there is

a discontinuity in outcomes and welfare at full sophistication, with the discontinuity now generated by the large penalties for deferring repayment stipulated in the contract that both sophisticated and non-sophisticated borrowers sign. As β_1 approaches β_2 from below, q_1 approaches a number strictly smaller than q_2 does. In other words, a non-sophisticated borrower, even if she is arbitrarily close to sophisticated, repays in a discontinuously different way from a sophisticated borrower, and is discontinuously worse off as a result.

We now show that if non-sophisticated borrowers are not too naive, eliminating disproportionately large penalties for deferring small amounts of repayment is still welfare-improving:

Proposition 2.7. *Suppose $\hat{\beta}$ is known, and β takes each of two values, $\beta_1 < \hat{\beta}$ and $\beta_2 = \hat{\beta}$, with positive probability. Borrowers strictly prefer the competitive-equilibrium contract in the unrestricted market over that in the restricted market, and a sophisticated borrower is indeed better off in the unrestricted market. If the non-sophisticated borrower is sufficiently sophisticated (β_1 is sufficiently close to $\hat{\beta}$), her welfare, as well as the population-weighted sum of type 1's and type 2's welfare, is greater in the restricted market than in the unrestricted market.*

As is the case when β is known, if non-sophisticated borrowers are not too naive, their welfare is higher in the restricted market than in the unrestricted one. The basic reason is also the same as before: because in the restricted market non-sophisticated borrowers have the option of deferring a small amount of repayment for a proportionally smaller fee, they do not drastically mispredict their own behavior. In the current setting, however, sophisticated borrowers are worse off in the restricted than in the unrestricted market, so the restricted market does not Pareto-dominate the unrestricted one; and since all borrowers think they are sophisticated, they all prefer the unrestricted market. The intuition for this result is related to a point first emphasized by Gabaix and Laibson (2006): because non-sophisticated borrowers are more profitable, in a competitive equilibrium it must be that firms make money on non-sophisticated borrowers and lose money on sophisticated borrowers. This cross-subsidy, and consequently the utility of sophisticated borrowers, is lower in the restricted market than in the unrestricted one. When β is unknown, therefore, our intervention does not satisfy the stringent requirement of asymmetric paternalism to avoid hurting fully rational consumers. Nevertheless, for any p_1 and p_2 the restricted market is still socially superior by the measure typically used in public economics: the population-weighted sum of individuals' welfare. Hence, this intervention is "robust" in that it is likely to be

welfare-improving even if we do not know much or do not agree about the prevalence of non-sophisticated types in the population.

2.4.2 Unknown β and $\hat{\beta}$

We now consider competitive equilibria when β and $\hat{\beta}$ are both unobservable to firms, providing a condition under which the contracts we have derived in Section 2.4.2.4.1 sort borrowers according to $\hat{\beta}$ in period 0. This means that even when firms observe neither consumers' preferences nor their degree of sophistication, any non-sophisticated consumer endogenously selects a contract with which she changes her mind regarding repayment, making her strictly worse off than a sophisticated consumer with the same time-preference parameter β.

We build our analysis on that of Section 2.4.2.4.1, where $\hat{\beta}$ is known. Let $\underline{u}_i$ be the sophisticated borrower's utility from the competitive-equilibrium contract when $\hat{\beta} = \beta_i$ is known, with probability p_i a borrower is sophisticated, and with probability $(1-p_i)$ she is type β_{i-1}. Our key condition is the following:

Condition 2.1. $\underline{u}_i$ is increasing in β_i.

Condition 2.1 states that if $\hat{\beta}$ was observable, the sophisticated borrower's utility from the equilibrium contract would be increasing in $\hat{\beta}$. That is, the closer a sophisticated borrower is to being time consistent, the higher is her utility. While this is an endogenous condition, it is intuitively plausible: it requires roughly that borrowers who are more optimistic about their future behavior tend to be more naive about it. Since firms compete more fiercely for such profitable borrowers, they drive up the utility of sophisticated borrowers.[25]

We argue that under Condition 2.1, there is a competitive equilibrium in which consumers sign the same contracts as when $\hat{\beta}$ is observed. The crucial part is that from such a set of contracts, consumers self-select according to $\hat{\beta}$ in period 0; then, since there would be no profitable deviation even if firms knew $\hat{\beta}$, there is certainly none when they do not know $\hat{\beta}$. By Condition 2.1, the credit contract intended for a borrower with higher $\hat{\beta}$ offers a better deal if the borrower can stick to the more favorable repayment schedule, but requires greater self-control to stick to that schedule. Hence, because a

[25] Condition 2.1 is clearly non-empty. Consider, for instance, a setting with two possible $\hat{\beta}$'s. If the lower $\hat{\beta}$ type is almost certain to be sophisticated while the higher $\hat{\beta}$ type has a non-trivial probability of being non-sophisticated, Condition 2.1 holds. More generally, in the current setting with two types of β for each $\hat{\beta}$, we require that consumers who believe themselves to be less time-inconsistent are non-sophisticated with sufficiently higher probability.

consumer takes the most favorable credit contract with which she believes she can repay according to the ex-ante preferred schedule, she chooses the contract corresponding exactly to her $\hat{\beta}$.

To illustrate the logic of this self-selection through an example, consider a consumer looking to buy a TV on sale financed using store credit that does not accrue interest for six months. The nicer the TV, the sweeter is the deal both because the sale is steeper and because the six-month interest-free period is more valuable. At the same time, it is more difficult to pay back a larger loan in six months. Hence, the consumer chooses the TV which she believes she can *just* pay off in time. But if she is even slightly naive, this TV will be too nice, and she will fail to pay it off.

In fact, the above competitive equilibrium is the unique one:

Proposition 2.8 (Period-0 Screening)**.** *Suppose Condition 2.1 holds. Then, in the unique competitive equilibrium with $\hat{\beta}$ unobserved, each consumer accepts the same contract as when $\hat{\beta}$ is observed.*

2.5 General Borrower Beliefs

In the basic model used throughout the paper, a borrower believes with certainty that her taste for immediate gratification will be $\hat{\beta}$ (as in O'Donoghue and Rabin 2001). While this assumption is analytically convenient, it is also very special. In this section, we investigate outcomes for a general specification of borrower beliefs that incorporates existing formulations of partial naivete as special cases. We clarify when a discontinuity in outcomes and welfare at full sophistication occurs, and identify an important asymmetry: while overestimating one's self-control has drastic welfare consequences, underestimating it has none.

Let the cumulative distribution function $F(\hat{\beta})$ with support in $[0, 1]$ represent a borrower's beliefs about her taste for immediate gratification β. Because we cannot solve a model with fully general beliefs and preferences both unobserved, we suppose that firms know borrowers' β. Since firms have a lot of information about consumers and spend a lot on researching their behavior, we find this scenario plausible for many borrowers.

It is straight-forward to extend the definition of competitive equilibrium to allow for a borrower to be uncertain about what she will choose in period 1. Our key result is the following:

Proposition 2.9. *Both when firms know borrowers' beliefs and when they do not, in a competitive equilibrium the repayment schedule a borrower with beliefs*

$F(\cdot)$ actually chooses satisfies

$$k'(q) = 1; \quad k'(r) = \frac{1}{F(\beta) + (1 - F(\beta))\beta}. \tag{2.4}$$

The borrowed amount is $c = q + r$. If $F(\beta) = 1$, the borrower believes in period 0 that she will choose (q, r) with probability 1. If $F(\beta) < 1$, then there is a unique other repayment schedule $(\hat{q}, \hat{r})$ such that the borrower believes in period 0 that she will choose (q, r) with probability $F(\beta)$ and $(\hat{q}, \hat{r})$ with probability $1 - F(\beta)$. This other schedule satisfies $\hat{q} > 0, \hat{r} = 0$ and $q < \hat{q} < q + r$.

Proposition 2.9 generalizes many of the central points regarding outcomes and welfare we have made in this paper. In particular, non-sophisticated consumers with $F(\beta) < 1$ delay repayment more often than they expect, and they borrow more and have lower welfare than sophisticated consumers. In addition, the fact that firms cannot observe consumers' beliefs does not affect the competitive equilibrium at all.[26]

Equation 2.4 in the proposition also clarifies that the extent to which a non-sophisticated consumer overborrows, repays in a back-loaded way, and has lower welfare than a sophisticated consumer, depends on $1 - F(\beta)$, the probability she attaches to unrealistically high levels of self-control. As a result, whether a borrower with beliefs close to sophisticated has discontinuously lower welfare than a sophisticated borrower depends on whether $F(\beta)$ is close to 1. We argue that for most natural senses in which beliefs can approach sophistication, $F(\beta)$ does not approach 1, so that near-sophisticated borrowers will typically have discretely lower welfare than sophisticated borrowers. Consider a sequence F_n of distributions, and let F^* be the distribution (corresponding to perfect sophistication) that assigns probability 1 to the true β. As a possible example of an increase in sophistication, if each F_{n+1} is obtained by shifting F_n to the left, with the mean of F_n approaching β, then $F_n(\beta)$ does not approach 1, and this is the case even if the support of each F_n is extremely tight. Alternatively, if the F_n are symmetric continuous distributions with mean β whose variance approaches zero as n approaches infinity, $F_n(\beta)$ does not change at all (and is equal to one-half). Combining these two possibilities,

[26] To see why borrowers self-select, notice that a borrower's competitive-equilibrium contract when beliefs are known maximizes her perceived expected utility subject to a zero-profit condition determined by the borrower's actual behavior. Since given the contract the borrower signs her behavior is independent of her beliefs, the zero-profit condition is independent of borrower beliefs. This implies that each borrower prefers the competitive-equilibrium contract she gets with her beliefs known to contracts borrowers with other beliefs get.

if the F_n are symmetric continuous distributions whose mean approaches β from above and whose variance approaches zero, then $F_n(\beta) \leq 1/2$ for all n. More generally, a natural formulation of convergence to sophistication with general beliefs is that $F_n \to F^*$ in distribution (or, equivalently, $F_n \to F^*$ in probability), and this statement does *not* imply that $F_n(\beta) \to F^*(\beta) = 1$. In fact, this implication seems extremely special, especially for sequences that approach F^* from the direction of overoptimistic beliefs.

Intuitively, a non-sophisticated borrower has much lower utility than a sophisticated borrower if she assigns a non-trivial probability to unrealistically high levels of self-control. Knowing that these beliefs are wrong, firms offer a contract that requires such unrealistic levels of self-control to repay in an advantageous way, thereby making credit seem cheap and fooling the consumer into overborrowing and paying a large fee for back-loading repayment. Note that although we have assumed that β is known to firms, this intuition suggests that the basic mechanism operates more generally—whenever there is a β such that borrowers attach unrealistically high probability on average to $\hat{\beta} > \beta$, and firms know this.

Proposition 2.9 and the above intuition make clear that in our setting, previous formalizations of near sophistication can be seen as opposite extremes. Translated into our model, Eliaz and Spiegler (2006) and Asheim (2008) assume that $F(\cdot)$ is binary, assigning probability p to being time-consistent ($\beta = 1$) and probability $1 - p$ to the true β. In this model of partial naivete, a near-sophisticated borrower puts a high probability on her actual taste—$1 - p = F(\beta) \approx 1$—so she cannot be fooled much regarding how she will repay. In the O'Donoghue and Rabin (2001) model of partial naivete, a near-sophisticated consumer puts zero weight on her actual taste or lower—$F(\beta) = 0$—so she can be completely fooled. For many or most notions of near sophistication, $F(\beta)$ is neither close to zero nor close to one, so the borrower can be partially fooled. This means that welfare is discretely lower than for sophisticated consumers, although by less than with the O'Donoghue-Rabin specification.

Proposition 2.9 also indicates that in a market situation, there is a fundamental asymmetry between overly optimistic and overly pessimistic beliefs about time inconsistency. This is true at the individual level: the weight a person puts on too high levels of $\hat{\beta}$ has significant welfare implications, but the weight she puts on too low levels of $\hat{\beta}$ has no implications in that it is as if she put the same weight on her true β. And a similar conclusion holds when comparing individuals with different beliefs: whereas a small amount of confident overoptimism (e.g. a degenerate $\hat{\beta} > \beta$) leads to a discontinuous drop in welfare, a small amount of overpessimism ($\hat{\beta} < \beta$) leads to no welfare

loss at all. The intuition derives from which kind of misprediction firms can profitably take advantage of. As we have emphasized throughout the paper, a firm can attract an overly optimistic borrower by leading her to think she will repay more of her loan early than she actually will, making credit seem cheap and generating overborrowing and a change of mind regarding repayment. In contrast, the only way a firm could mislead a pessimistic borrower is by making her think that she will repay *less* of her loan early than she actually will. Since the borrower considers her future self too present-oriented to start with, she would dislike this possibility, so she would be reluctant to sign such a contract. Hence, there is no point in misleading her in this direction.[27]

Similarly to the predictions on contract terms and welfare in the unrestricted market, our conclusion that the restricted market can yield higher welfare also extends, with minor qualifications, to the more general formulation of borrower beliefs. By the same argument as in Sections 2.3 and 2.4, such an intervention benefits near-sophisticated borrowers with $F(\beta)$ non-trivially different from 1. Since a borrower with $F(\beta) \approx 1$ gets utility close to that of a sophisticated borrower anyway, the same intervention cannot benefit her by much. And since an overly pessimistic borrower gets the same utility as a sophisticated borrower, she can only be made worse off by the intervention. But while it will not help much, neither does the intervention hurt the latter two types of borrowers by much. Since the welfare gain for the former types of borrowers is discrete, therefore, if there is even a very small fraction of these borrowers in the population, a restricted market may have higher social welfare than an unrestricted market. For the same reason, our model implies that the restricted market can generate substantially higher welfare even if borrowers are not only all close to sophisticated, but also on average correct about their future preferences—with some overestimating β and some underestimating it.[28]

[27] The above logic also explains why for any borrower beliefs there are at most two (relevant) repayment options in the competitive-equilibrium contract. To the extent that the borrower puts weight on unrealistically high levels of self-control ($\hat{\beta} > \beta$), she can be fooled into believing she will choose a cheap front-loaded repayment schedule, so a lender offers a single repayment schedule that will make credit seem cheapest. To the extent that the borrower puts weight on unrealistically low levels of self-control ($\hat{\beta} < \beta$), it is unprofitable to fool her, so a lender offers the repayment option she will actually choose.

[28] As we have discussed in Section 2.3, if many consumers are very naive it is unclear whether the restricted market yields higher welfare than the unrestricted one. But even in that case, a restricted market combined with an interest-rate cap is often better than an unrestricted market.

2.6 Related Literature

2.6.1 Related Psychology-and-Economics Literature

Our model builds on several recent papers on contracting with time-inconsistent or boundedly rational consumers. While we discuss other differences between these theories and ours below, the most important difference is that we consider a richer set of welfare implications, and also analyze welfare-increasing interventions.

Our paper belongs to the small literature on contracting with time inconsistency, including DellaVigna and Malmendier (2004), Gottlieb (2008), and Hafalir (2008) on specific contract forms and Kőszegi (2005) and Eliaz and Spiegler (2006) on general non-linear contracts. Closest to our work, DellaVigna and Malmendier (2004) develop a model in which firms sell to time-inconsistent individuals using two-part tariffs consisting of an initial lump-sum transfer and a later price for consuming. Analogously to our prediction that deferring repayment is costly, they show that for a product with immediate benefits and delayed costs, the price is above cost. Although this has no welfare effect in their setting, in an extension they also show that firms choose renewal fees so that all non-sophisticated consumers mispredict whether they will renew. But because their model exogenously imposes the contract forms, and because it is not specifically written for the credit market, it does not make many of our finer predictions on contract features and outcomes (such as the overborrowing by non-sophisticated consumers, the excessively front-loaded baseline repayment schedules, and the disproportionately large fees for deferring small amounts of repayment).

Eliaz and Spiegler (2006) develop a two-period model in which a monopolist offers contracts in the first period to a population of consumers who have homogeneous time-inconsistent preferences about an action to be taken in the second period, but attach heterogeneous prior probabilities to the change in preferences. We modify Eliaz and Spiegler (2006) by assuming a different form of naivete about preferences and by focusing on perfect competition, and as a result get a discontinuity in outcomes and welfare at full sophistication that is not present in their model. By extending their and our model to allow for any borrower beliefs, we show that the discontinuity holds for many or most forms of these beliefs. We also extend their theory by considering heterogeneity in preferences in addition to beliefs. And we specialize their model to a credit market in which time inconsistency derives from a taste for immediate gratification, yielding specific predictions that would not make immediate sense in their setting.

Modeling a phenomenon that is clearly very important in credit markets, Gabaix and Laibson (2006) assume that there is an exogenously given costly add-on (e.g. a printer's cartridge costs or a credit card's fees) that naive consumers might partially or fully ignore when making purchase decisions, and that sophisticated consumers take steps to avoid. Gabaix and Laibson's main finding is that because competitive firms lose money on sophisticated consumers and make money on naive consumers, they may not have an incentive to debias the latter ones. While both forms of naivete are clearly relevant, our focus is on what happens when consumers might misunderstand their *reaction* to a contract rather than the *terms* of the contract. This has the advantage that we can derive borrowers' misprediction of the cost of credit from a general model of consumer preferences and beliefs interacting with profit-maximizing firms—rather than take this misprediction as exogenous—allowing us to endogenize more features of credit contracts (e.g., a low-cost overly front-loaded baseline repayment schedule along with a large penalty to switch out of it) and propose plausible interventions. There is also a major difference between the two models in the source of inefficiency: whereas in Gabaix and Laibson's model the welfare loss comes from sophisticated consumers' costly effort to avoid the add-on, in ours it derives largely from the suboptimal contracts non-sophisticated borrowers receive—an aspect that seems very realistic for credit markets.

Grubb (2009) considers contracting with consumers who overestimate the extent to which they can predict their demand for a product (e.g., their cell-phone usage). To exploit consumers' misprediction, firms convexify the price schedule by selling a number of units at zero marginal price and further units at a positive marginal price. The high marginal price for high amounts of consumption is similar to our basic prediction that deferring repayment is expensive. Unlike in Grubb (2009), however, in our setting the price of deferring repayment is imposed as a large fee, and beyond this fee the marginal price can be low to encourage self 1 to defer more of her repayment. This feature seems consistent with credit markets; for instance, although a subprime mortgage typically carries a large prepayment penalty, once a borrower pays that penalty there is little extra cost in refinancing more of the mortgage.

2.6.2 Predictions of Neoclassical Models

We are not aware of neoclassical theories that explain the contract features we have derived. Beyond this observation, we argue in this section that natural versions of neoclassical models do not generate qualitatively similar features.

Since the main predictions of our model concern a contract's repayment

terms and how a borrower chooses from these terms, we begin with discussing situations in which there is heterogeneity in borrowers' ability or willingness to repay the loan fast—to which screening using repayment terms would seem to be the natural response. If borrowers know at the time of contracting whether they will be able to repay fast, it is optimal for lenders to offer an expensive loan aimed at late payers that allows back-loaded repayment. But a contract with a prepayment penalty is a very inefficient way of achieving this—it would be better to simply offer an expensive mortgage that postulates later repayment to start with, avoiding the costs of refinancing. Similarly, a credit-card contract intended for a late payer could simply be more expensive and have a longer grace period, rather than require fast repayment and feature a large penalty for deviations.

If borrowers do not know at the time of contracting whether they will be able to repay fast, but are rational regarding this uncertainty and are time consistent, we get a situation of classical sequential screening (Courty and Li 2000, for example) or post-contractual hidden knowledge (Laffont and Martimort 2001, Section 2.11, for example). But specifying such a model in a natural way for our setting yields essentially the opposite qualitative contract features than what we have found. As a simple example in the context of hidden knowledge, suppose that each borrower is interested in buying a product for a price of 1, and she has the option of paying for the product out of pocket in period 1. She can, however, also obtain a loan for buying the product from a single lender. If the borrower obtains a loan, she pays back an amount q in period 1 and an amount r in period 2, with costs $\theta k(q)$ and r, respectively. The variable θ, with support equal to some positive interval $[\underline{\theta}, \overline{\theta}]$, captures differences in the cost of repaying early. Neither party knows θ at the time of contracting, but the borrower learns it before choosing q in period 1. Then, it is easy to show that the lender's optimal contract involves a loan that is expensive if repaid early—if θ is low, the borrower wishes she had paid out of pocket—but whose repayment schedule is free to change. In contrast, our model predicts loans that are cheap if repaid early but whose repayment terms are expensive to change.[29]

[29] The formal derivation of the optimal contract in the case of hidden knowledge, as well as some discussion of the above assumptions, is available from the authors upon request. By the same basic logic, sequential screening seems to yield similarly different contracts from those predicted by our theory. In the main example given by Courty and Li (2000), there is a business traveler with highly uncertain valuation for an airplane ticket and a leisure traveler with less uncertain valuation, and the airline screens these travelers by offering an expensive refundable ticket to the business traveler and a cheap non-refundable ticket to the leisure traveler. Analogously, a lender should offer an expensive flexible mortgage to

While our model assumes no default and therefore ignores issues of credit risk, it is unlikely that the contract features we have found could be explained by this consideration. As shown in classical contributions by Stiglitz and Weiss (1981) and Bester (1985, 1987), the primary screening tools lenders would use when facing heterogeneity in credit risk are credit rationing and collateral requirements. If there is a negative correlation between credit risk and the ability to repay early, then screening in part using repayment terms might be an optimal response. But if this was the case, borrowers who repay quickly would be the most profitable—a prediction that is empirically false. Credit-card companies only appear to break even on consumers who repay their full bill every month, and make the bulk of their ex-post profits on consumers who carry a balance (Ausubel 1991, Chakravorti and Emmons 2003). In fact, consumers who regularly pay off their balances are sometimes referred to in the industry as "deadbeats" or "freeloaders" (Chakravorti and Shah 2001). Similarly, as mentioned above, subprime mortgage lenders seem to have generated a significant portion of their profits from prepayment penalties and refinancing fees.

The contract features we have derived also do not seem consistent with a screening model in which rational time-consistent borrowers differ in their need for credit. If this was the case, the primary screening tool lenders would likely use is the amount of credit rather than the time structure of repayment.

Finally, the large penalties predicted by our theory are at first glance similar to penalties used by principals in moral-hazard and screening models to prevent an agent from taking actions the principal does not want.[30] In contrast to these penalties that serve only a preventive role and that agents rarely or never pay in equilibrium, in our model non-sophisticated borrowers do pay the penalties. In fact, the penalties are a central source of firm profits and designing them is a central part of a firm's contract-design problem.

2.7 Conclusion

While it captures some salient features of real-world credit markets and identifies simple welfare-improving interventions, our setting leaves unanswered important questions about whether and in what way partial naivete justifies intervention. Although the intervention we propose is welfare-improving in the

borrowers who face uncertainty regarding their ability to repay early—one that is expensive if repaid early but has a lot of flexibility on how to pay back.

[30] In the classical case of moral hazard, see Mirrless (1999) and Bolton and Dewatripont (2005, page 140).

sense typically used in economics (social welfare), in the spirit of libertarian paternalism's (Sunstein and Thaler 2003) respect for individual liberty, we can formulate another criterion for interventions: that they should be accepted by consumers. In our theory, all borrowers believe they are rational, so if they correctly predicted what contracts they would receive in a restricted market, they would be against intervention. Investigating whether this generalizes to settings where firms do not redistribute all of their profits to sophisticated borrowers, and whether there are modifications of our intervention that consumers would accept, is left for future work.

Another important issue we have completely ignored in this paper is the source of consumer beliefs. Consumers may learn about their preferences from their own behavior and that of the firms, and they often seem to have a generic skepticism regarding contract offers even if they do not know how exactly the contract is looking to exploit them. Since our model (like most models of naivete with which we are familiar) starts from exogenously given beliefs, it cannot easily accommodate such learning and meta-sophistication.[31] Nevertheless, our results suggest that learning can sometimes lower welfare. So long as a borrower does not become *fully* sophisticated, she might switch away from her preferred repayment schedule ex post, so that her increased sophistication does not help in achieving full self-control in repayment. In addition, her pessimism might mean that—in a futile attempt at achieving self-control—she chooses a worse deal up front, lowering her utility.

Appendix: Proofs

Proof of Lemma 2.1. ($\Rightarrow$) Suppose $(c, \mathcal{C})$ satisfies the condition of the lemma. Since only this contract is offered and it satisfies the borrower's PC, it is optimal for her to accept the contract and her choice between contracts is trivial. Thus Condition 1 of Definition 2.2 is satisfied. Conditions 2 and 4 hold by construction. The key part is to check Condition 3. Consider a contract $(c', \mathcal{C}')$ with incentive compatible repayment terms that the borrower strictly prefers. Incentive compatibility guarantees that the contract satisfies IC and PCC, and the fact that the borrower strictly prefers it implies that PC

[31] The only paper we know that systematically studies whether individuals will learn their taste for immediate gratification is Ali (2011). In the model, a decisionmaker who is too optimistic about her self-control does not restrict her choices, and hence keeps learning about her self-control from her own behavior. As a result, overoptimism about self-control tends to be eliminated by learning. Given the evidence that many people are overoptimistic, we view Ali's (2011) theory as deepening the puzzle of how learning affects the behavior of time-inconsistent individuals.

is satisfied when the outside option is $\underline{u}$. Hence, because $(c', \mathcal{C}')$ satisfies all constraints that $(c, \mathcal{C})$ does, and $(c, \mathcal{C})$ is optimal given these constraints and yields zero profits, $(c', \mathcal{C}')$ cannot yield positive expected profits.

($\Leftarrow$) Since there is only one $\hat{\beta}$ type, there can only be one contract. Let $(c, \mathcal{C})$ be that competitive-equilibrium contract. Condition 4 (non-redundancy) implies that there are only two repayment options in the contract, one for β_1 and one for β_2. Incentive compatibility implies that $(c, \mathcal{C})$ satisfies IC and PCC, and it trivially satisfies PC with $\underline{u}$ defined as the perceived utility from $(c, \mathcal{C})$. Now suppose by contradiction that $(c, \mathcal{C})$ does not maximize profits given these constraints. Then, there is a contract $(c', \mathcal{C}')$ that satisfies the same constraints and yields strictly positive profits. This means that for a sufficiently small $\epsilon > 0$, $(c' + \epsilon, \mathcal{C}')$ attracts all borrowers and yields strictly positive profits, violating Condition 3 of Definition 2.2.

Proof of Fact 2.1. It follows from Proposition 2.1 that she borrows $c = 2(k')^{-1}(1)$ and repays $(k')^{-1}(1)$ in each period in the unrestricted market, and from the proof of Proposition 2.3 that she borrows and repays the same amounts in long-term restricted market.

Proof of Proposition 2.1. A sophisticated borrower correctly foresees the repayment option she eventually chooses. Thus, a non-redundant contract (i.e. one that satisfies Condition 4 of Definition 2.2) has a single repayment option (q, r). Using this fact, Conditions 1 and 3 of Definition 2.2 imply that any competitive contract $(c, \mathcal{C})$ must solve

$$\begin{aligned} \max_{c,q,r} \quad & q + r - c \\ \text{s.t.} \quad & c - k(q) - k(r) \geq \underline{u}, \end{aligned} \tag{PC}$$

where $\underline{u}$ is the perceived utility from accepting the competitive contract $(c, \mathcal{C})$. It is clear that in the maximization problem above PC is satisfied with equality; otherwise, the firm could increase profits by lowering c. Plugging PC into the maximand, we can rewrite the firm's problem as

$$\max_{q,r} \quad q + r - k(q) - k(r).$$

Solving this maximization yields $k'(q) = k'(r) = 1$ in any competitive contract. Furthermore, the zero-profit condition (Condition 2) implies that $c = q + r$, and this completely characterizes the unique competitive-equilibrium contract.

Proof of Proposition 2.2. We have established in the text that $\hat{q} > 0, \hat{r} = 0, k'(q) = 1, k'(r) = 1/\beta$, and Lemma 2.1 implies that $c = q + r$. Using Proposition 2.1, the sophisticated and non-sophisticated borrowers repay the same amount in period 1, but the non-sophisticated borrower repays more

in period 2. Hence, the non-sophisticated consumer borrows more than the sophisticated one.

To show that $q + r > \hat{q}$, suppose by contradiction that $\hat{q} \geq q + r$. Then, notice that for a sufficiently small $\epsilon > 0$, self 0 strictly prefers the repayment schedule $(\hat{q}/2 + \epsilon, \hat{q}/2 + \epsilon)$ to $(\hat{q}, 0)$, the terms she thinks she is going to choose with the competitive-equilibrium contract. Hence, the firm could increase profits by offering a single repayment schedule $(\hat{q}/2+\epsilon, \hat{q}/2+\epsilon)$, a contradiction.

Finally, from the proof of Proposition 2.1 it is clear that the contract offered to a sophisticated borrower is the unique contract that maximizes period-0 welfare among all contracts that break even ($c = q + r$). Since the borrower's contract also breaks even and differs from the sophisticated one, the borrower is strictly worse off than a sophisticated borrower.

Proof of Proposition 2.3. Let a restricted contract be described by the triplet (c, R, L), where c is consumption, R is the interest rate, and L is the present discounted value of total repayment from the perspective of period 1, using the interest rate R.

Consider sophisticated borrowers first. Notice that a contract with $R = 1/\beta$ will induce the borrower to repay in equal installments. This means that a contract that combines $R = 1/\beta$ with the ex-ante optimal consumption level c^* and the competitive L^* maximizes the borrower's utility subject to the constraint that consumption is equal to total repayment. Conversely, no other contract with which a firm breaks even maximizes the borrower's utility: for the borrower to repay according to $k'(q) = k'(r) = 1$, the contract must have $R = 1/\beta, L = L^*$, and then for the firm to break even consumption must be c^*. Hence, if this contract was not offered but firms made zero profits, for a sufficiently small $\epsilon > 0$ the contract $(c^* - \epsilon, 1/\beta, L^*)$ could be profitably introduced. Hence, $(c^*, 1/\beta, L^*)$ is the unique competitive-equilibrium contract.

Now we consider non-sophisticated borrowers. For any R, L, there is a unique repayment schedule (q, r) the borrower follows, and hence a unique $c(R, L) = q + r$ with which a firm breaks even. Let $\mathcal{B}$ be the set of contracts $(c(R, L), R, L)$; this is the set of contracts that if accepted break even given the borrower's actual behavior, and is independent of $\hat{\beta}$. Furthermore, consider the borrower's perceived utility $\underline{U}_{\hat{\beta}}(c, R, L)$ as a function of (c, R, L) over $\mathcal{B}$; this is a function of $\hat{\beta}$. Notice that a competitive-equilibrium contract maximizes $\underline{U}_{\hat{\beta}}$ over $\mathcal{B}$; otherwise, a firm could find a contract that breaks even and gives the borrower higher perceived utility, and starting from this contract could decrease c slightly, attracting the borrower and earning positive profits. To see that competitive equilibrium exists, we first show that without loss of generality we can assume that $R \in [k'(0)/(\hat{\beta}k'(M)), k'(M)/(\beta k'(0))]$, and $L \in$

$[0, M + M\hat{\beta}k'(M)/k'(0)]$. The borrower believes she will choose $\hat{q}$ to solve

$$\min_{\hat{q}} k(\hat{q}) + \hat{\beta}k(R(L-\hat{q})) \text{ s.t. } 0 \leq \hat{q} \leq M \text{ and } 0 \leq R(L-\hat{q}) \leq M, \tag{2.5}$$

and she actually chooses q to solve the above problem with q and β replacing $\hat{q}$ and $\hat{\beta}$. Hence, for any $R \geq k'(M)/(\beta k'(0))$ we have a corner solution in which $q = \hat{q} = M$ and hence the second-period repayment amounts are $\hat{r} = r = R(L-M)$. The firm can thus replicate the outcome of any contract (c, R, L) in which $R > k'(M)/(\beta k'(0))$ by one in which $R = k'(M)/(\beta k'(0))$ and L is appropriately adjusted. Similarly, if $R \leq k'(0)/(\hat{\beta}k'(M))$, then $q = \hat{q} = 0$, so that we can replace any contract featuring $R < k'(0)/(\hat{\beta}k'(M))$ with a contract featuring $R = k'(0)/(\hat{\beta}k'(M))$. Hence, without loss of generality we can restrict attention to contracts in which $R \in [k'(0)/(\hat{\beta}k'(M)), k'(M)/(\beta k'(0))]$. Since repayment amounts in each period are bounded from above by M and the interest rate from below by $k'(0)/(\hat{\beta}k'(M))$, we can furthermore restrict attention to $L \in [0, M + M\hat{\beta}k'(M)/k'(0)]$. Now since q, r (and hence $c = q + r$) and $\hat{q}, \hat{r}$ are continuous in R, L and R, L are chosen from compact sets, it follows that a contract exists that maximizes $\underline{U}_{\hat{\beta}}(c, R, L)$ over $\mathcal{B}$.

Now notice that given a contract (c, R, L), the borrower's perceived repayment behavior is continuous in $\hat{\beta}, R, L$, which in turn implies that $\underline{U}_{\hat{\beta}}(c, R, L)$ is continuous in $\hat{\beta}, c, R, L$. For $\hat{\beta} = \beta$, we have shown above that $\underline{U}_{\hat{\beta}}$ has a unique maximum at $(c^*, 1/\beta, L^*)$. We complete the proof by showing that as a result, if $\hat{\beta} \to \beta$, any selection of maximizers $(c(\hat{\beta}), R(\hat{\beta}), L(\hat{\beta}))$ of $\underline{U}_{\hat{\beta}}$ over $\mathcal{B}$ must approach $(c^*, 1/\beta, L^*)$. This means that in the restricted market the welfare of a non-sophisticated borrower approaches that of a sophisticated borrower as $\hat{\beta} \to \beta$. In contrast, by Propositions 2.1 and 2.2, in the unrestricted market the welfare of a non-sophisticated borrower does not approach that of a sophisticated borrower as $\hat{\beta} \to \beta$, so for $\hat{\beta}$ sufficiently close to β the restricted market yields higher welfare.

Suppose by contradiction that there is some selection of maximizers $(c(\hat{\beta}), R(\hat{\beta}), L(\hat{\beta}))$ of $\underline{U}_{\hat{\beta}}$ over $\mathcal{B}$ that does not converge to $(c^*, 1/\beta, L^*)$ as $\hat{\beta} \to \beta$. Since the $(c(\hat{\beta}), R(\hat{\beta}), L(\hat{\beta}))$ are within a compact set, there must be a convergent subsequence with limit $(c, R, L) \neq (c^*, 1/\beta, L^*)$. Since $\mathcal{B}$ is closed, $(c, R, L) \in \mathcal{B}$. We know that $U_{\hat{\beta}}(c(\hat{\beta}), R(\hat{\beta}), L(\hat{\beta})) \geq U_{\hat{\beta}}(c^*, 1/\beta, L^*)$, so by continuity $U_\beta(c, R, L) \geq U_\beta(c^*, 1/\beta, L^*)$, contradicting that U_β has a unique maximum over $\mathcal{B}$ at $(c^*, 1/\beta, L^*)$.

Proof of Proposition 2.4. Let us call the restricted market in which the interest rate is zero (i.e. $R = 1$) the capped market. We begin by showing that the borrower's consumption is lower in the capped market than in the unrestricted market. Since self 0 thinks self 1's cost of repayment is $k(q) +$

$\hat{\beta}k(r)$, she believes that for any L, self 1 will choose the repayment schedule by minimizing $k(q) + \hat{\beta}k(L-q)$ subject to $q, L-q \leq M$; let the solution be $\hat{q}$, and set $\hat{r} = L - \hat{q}$. In the competitive equilibrium of the capped market, the amount of credit c maximizes the borrower's perceived utility subject to $c = L$; otherwise, the firm could offer a contract that both has higher perceived utility and has $c < L$, attracting the borrower and making positive profits. We first observe that the competitive-equilibrium c is such that $\hat{q}, \hat{r} < M$. Suppose by contradiction that $\hat{q} \geq M$ or $\hat{r} \geq M$. Then, because $\hat{\beta} \leq 1$ implies $\hat{r} \geq \hat{q}$, we must have $\hat{r} = M$. Hence $k'(\hat{r}) = k'(M) \geq 1/\beta$, and using the perceived cost minimization of the borrower, $k'(\hat{q}) \geq \hat{\beta}k'(\hat{r}) \geq \hat{\beta}/\beta > 1$. Therefore, because the perceived marginal cost of repayment in both periods is strictly greater than the marginal utility of consumption, decreasing c and $L = c$ by a small amount increases the borrower's perceived utility independently of how she believes she will allocate the decreased L across periods 1 and 2, a contradiction. By a similar argument, we can show that competitive-equilibrium c is such that $\hat{q}, \hat{r} > 0$. Suppose by contradiction that this is not the case. Since $\hat{r} \geq \hat{q}$, this means that $\hat{q} = 0$. Then $k'(\hat{q}) = k'(0) < \beta$, and therefore $k'(\hat{r}) \leq k'(\hat{q})/\hat{\beta} < \beta/\hat{\beta} < 1$. Hence, because the perceived marginal cost of repayment in both periods is strictly lower than the marginal utility of consumption, increasing c and $L = c$ by a small amount increases the borrower's perceived utility independently of how she believes she will allocate the increased L across periods 1 and 2, a contradiction.

Because in a competitive equilibrium $0 < \hat{q}, \hat{r} < M$, the solution to the borrower's perceived repayment-cost minimization problem is described by the first-order condition $k'(\hat{q}) = \hat{\beta}k'(L - \hat{q})$. Let $\hat{q}(L)$ denote the unique solution to this first-order condition; this is the amount self 0 thinks self 1 will repay in period 1 if she owes L. Note that $\hat{q}(L)$ is a continuously differentiable function of L, with a derivative strictly between zero and one.

Again using that the competitive-equilibrium c maximizes the borrower's perceived utility subject to $L = c$, the competitive-equilibrium c solves

$$\max_c \quad c - k(\hat{q}(c)) - k(c - \hat{q}(c)),$$

yielding the first-order condition

$$1 = k'(\hat{q}(c))\hat{q}'(c) + k'(\hat{r}(c))(1 - \hat{q}'(c)).$$

Plugging in $k'(\hat{r}(c)) = k'(\hat{q}(c))/\hat{\beta}$ gives

$$1 = k'(\hat{q}(c))[\hat{q}'(c) + (1 - \hat{q}'(c))/\hat{\beta}].$$

Since the term in square brackets is greater than 1, $k'(\hat{q}(c)) \leq 1$, which implies that $k'(\hat{r}(c)) \leq 1/\hat{\beta} < 1/\beta$. Because $\hat{q}(c) + \hat{r}(c) = L = c$, we thus have $c < (k')^{-1}(1) + (k')^{-1}(1/\beta)$, which establishes that consumption is less than in the unrestricted market.

Now we use the fact that the borrower consumes more in the unrestricted market than in the capped market to show that she has lower welfare than in the capped market. Simple arithmetic yields the following lemma:

Lemma 2.2. *Suppose either (i) $k(x) = x^\rho$ for some $\rho > 1$; or (ii) $k(x) = (y-x)^{-\rho} - y^{-\rho}$ for some $y > 0, \rho > 0$. Then, in the capped market c is increasing in $\hat{\beta}$.*

Proof. We begin by establishing this for case (i). The borrower expects to repay c in a way such that $k'(\hat{q}) = \hat{\beta}k'(c - \hat{q})$, which in case (i) simplifies to

$$\hat{q}(\hat{\beta}, c) = \underbrace{\frac{\hat{\beta}^{\frac{1}{\rho-1}}}{1 + \hat{\beta}^{\frac{1}{\rho-1}}}}_{\equiv b(\hat{\beta})} c. \tag{2.6}$$

Thus, her perceived-period-zero utility is $c - (b(\hat{\beta})c)^\rho - ((1 - b(\hat{\beta}))c)^\rho$, which can be rewritten as $c - c^\rho \left[b(\hat{\beta})^\rho + (1 - b(\hat{\beta}))^\rho\right]$. The borrower chooses c to maximize her perceived utility so that $1 = \rho c^{\rho-1} \left[b(\hat{\beta})^\rho + (1 - b(\hat{\beta}))^\rho\right]$. Since $b(\hat{\beta})$ is increasing and less than 1/2, the term in square brackets is decreasing in $\hat{\beta}$, and thus c is increasing in $\hat{\beta}$.

In case (ii), let $W \equiv 2y - c$, $s \equiv y - \hat{q}$, and $t \equiv y - \hat{r}$. Hence in the capped market $t = W - s$. Rewriting $k'(\hat{q}) = \hat{\beta}k'(c - \hat{q})$, yields

$$s(\hat{\beta}, W) = \underbrace{\frac{\hat{\beta}^{\frac{-1}{1+\rho}}}{1 + \hat{\beta}^{\frac{-1}{1+\rho}}}}_{\equiv b(\hat{\beta})} W. \tag{2.7}$$

Observe that $b(\hat{\beta})$ is decreasing and greater than 1/2. The borrower's perceived period-zero utility is $c - \left(b(\hat{\beta})W(c)\right)^{-\rho} - \left((1 - b(\hat{\beta}))W(c)\right)^{-\rho} - 2y^{-\rho}$, which can be rewritten as $c - W(c)^{-\rho}\left[b(\hat{\beta})^{-\rho} + (1 - b(\hat{\beta}))^{-\rho}\right] - 2y^{-\rho}$. Since the power function with the exponent $-\rho$ is convex, and $b(\hat{\beta})$ decreasing and greater than 1/2, an increase in $\hat{\beta}$ decreases the term in square brackets. Since at the perceived optimal c, $1 = \rho W(c)^{-(\rho+1)}[b(\hat{\beta})^{-\rho} + (1 - b(\hat{\beta}))^{-\rho}]$, an increase in $\hat{\beta}$ must lead to a decrease of $W(c)$ or—in other words—an increase in c. □

To complete the proof, consider contracts in the capped market and restrict attention to contracts for which consumption is equal to total repayment ($c = L$). We show that for any $\beta, \hat{\beta}$, the actual repayment amounts satisfy $0 < q(c) \leq r(c) < M$. The part $r(c) \geq q(c)$ is obvious. For $\hat{\beta} = \beta$, we have already established that $\hat{q}(c) > 0$ and thus $q(c) > 0$. Because by Lemma 2.2 c is increasing in $\hat{\beta}$, we also have $q(c) > 0$ for all $\hat{\beta} \geq \beta$. For $\hat{\beta} = 1$, $k'(\hat{q}) = k'(\hat{r}) = 1$. Since $q(c) > 0$ implies $k'(q(c)) \geq \beta k'(r(c))$, we must have $k'(r(c)) < 1/\beta$, so that $r(c) < M$. Again using Lemma 2.2, since c is increasing in $\hat{\beta}$, for any $\hat{\beta} \leq 1$ we must have $r(c) < M$.

Since $0 < q(c), r(c) < M$, replacing $\hat{\beta}$ by β in Equations 2.6 and 2.7 shows that the repayment amounts $q(c), r(c)$ increase linearly in c. Hence in the capped market the borrower's welfare is $c - k(a_1 + bc) - k(a_2 + (1-b)c)$ for some constants $a_1, a_2 \in \mathbb{R}$, and $b \in (0,1)$. Twice differentiating with respect to c shows that for the utility functions in the proposition, among contracts where $R = 1$ and $c = L$ the borrower's welfare is single-peaked in consumption. By revealed preference, the maximum occurs at the consumption level that the sophisticated borrower chooses in the capped market. Lemma 2.2 implies that a non-sophisticated borrower consumes more in the capped market than the sophisticated borrower, and we established above that she consumes even more than that in the unrestricted market. This implies that she has lower welfare in the unrestricted than in the capped market.

Proof of Proposition 2.5. The firm's problem is

$$\begin{aligned} \max_{c,q,r,\hat{q},\hat{r}} \quad & q + r - c \\ \text{s.t.} \quad & c - k(\hat{q}) - \hat{\beta}k(\hat{r}) \geq \underline{u}, && \text{(PC)} \\ & -k(\hat{q}) - \hat{\beta}k(\hat{r}) \geq -k(q) - \hat{\beta}k(r), && \text{(PCC)} \\ & -k(q) - k(r) \geq -k(\hat{q}) - k(\hat{r}). && \text{(IC)} \end{aligned}$$

The steps in the analysis are very similar to those in the time-inconsistent case. PC binds because otherwise the firm could increase profits by reducing c. In addition, IC binds because otherwise the firm could increase profits by increasing q. Given that IC binds and $\hat{\beta} > 1$, PCC is equivalent to $q \leq \hat{q}$, so conjecturing that $q \leq \hat{q}$ is optimal even without PCC, we ignore this constraint, and confirm our conjecture in the solution to the relaxed problem below.

The relaxed problem is

$$\begin{aligned} \max_{c,q,r,\hat{q},\hat{r}} \quad & q + r - c \\ \text{s.t.} \quad & c - k(\hat{q}) - \hat{\beta}k(\hat{r}) = \underline{u}, && \text{(PC)} \\ & -k(q) - k(r) = -k(\hat{q}) - k(\hat{r}). && \text{(IC)} \end{aligned}$$

Notice that in the optimal solution, $\hat{r} = 0$: otherwise, the firm could decrease $k(\hat{r})$ and increase $k(\hat{q})$ by $\hat{\beta}$ times the same amount, leaving PC unaffected and creating slack in IC, allowing it to increase q. Using this, we can express $k(q)$ from IC and plug it into PC to get

$$c = k(q) + k(r) + \underline{u}.$$

Plugging c into the firm's maximand and solving yields all the statements in the proposition. Finally, using $\hat{r} = 0$ it follows from IC that $\hat{q} > q$, and thus the solution to the relaxed problem indeed satisfies PCC.

Proof of Proposition 2.6. Applying Lemma 2.1, we set up a firm's problem as choosing a type-independent consumption c and a menu of type-dependent repayment options $\{(q_1, r_1), (q_2, r_2)\}$ subject to participation, incentive, and perceived-choice constraints. Notice that because both types initially believe they are the sophisticated type β_2 and the sophisticated borrower chooses the baseline repayment schedule, the non-sophisticated borrower's perceived-choice constraint is identical to the sophisticated borrower's incentive constraint. As in textbook models of screening (e.g. Bolton and Dewatripont 2005, Chapter 2), we solve a relaxed problem with only type 1's incentive constraint, and verify ex-post that the solution satisfies type 2's incentive constraint. Given these considerations, the firm's relaxed problem is

$$\max_{c,q_1,r_1,q_2,r_2} \quad p_1(q_1 + r_1) + p_2(q_2 + r_2) - c \tag{2.8}$$

$$\text{s.t.} \quad c - k(q_2) - k(r_2) \geq \underline{u}, \tag{PC}$$

$$-k(q_1) - \beta_1 k(r_1) \geq -k(q_2) - \beta_1 k(r_2). \tag{IC}$$

In the optimal solution, IC binds; otherwise, the firm could increase q_1 without violating IC or PC, increasing profits. In addition, PC binds; otherwise, the firm could decrease c and thereby increase profits. From the binding constraints, we get $k(q_2) = c - k(r_2) - \underline{u}$ and $k(q_1) = k(q_2) + \beta_1(k(r_2) - k(r_1))$.

We first establish uniqueness of the competitive equilibrium. Based on the above arguments, the firm's problem reduces to

$$\max_{c,q_1,r_1,q_2,r_2} p_1(q_1 + r_1) + p_2(q_2 + r_2) - c$$

$$c - k(q_2) - k(r_2) = \underline{u} \tag{PC}$$

$$k(q_2) + \beta_1 k(r_2) = k(q_1) + \beta_1 k(r_1). \tag{IC}$$

We prove that $r_1 < r_2$ is suboptimal. Supposing by contradiction that $r_1 < r_2$, using IC we have $k(q_2) + k(r_2) = k(q_1) + \beta_1 k(r_1) + (1 - \beta_1)k(r_2) > k(q_1) + k(r_1)$. Then, if $q_1 + r_1 \geq q_2 + r_2$, the firm could eliminate the repayment option

(q_2, r_2) without decreasing profits, creating slack in PC and thereby allowing it to decrease c. And if $q_1 + r_1 < q_2 + r_2$, the firm would be strictly better off not offering (q_1, r_1), yielding the desired contradiction.

Now, substituting PC into the maximand gives

$$\begin{aligned} \max \quad & p_1(q_1 + r_1) + p_2(q_2 + r_2) - k(q_2) - k(r_2) \\ k(q_2) + \beta_1 k(r_2) \quad = \quad & k(q_1) + \beta_1 k(r_1) \qquad \text{(IC)}. \end{aligned}$$

Let $A = k(q_2), B = k(r_2), D = k(r_1) - k(r_2)$. Then, $k(r_1) = B + D$ and using the IC constraint $k(q_1) = A - \beta_1 D$. Let $f = k^{-1}$. Since k is strictly increasing and strictly convex, f is strictly increasing and strictly concave, and our assumptions on k furthermore ensure that $\lim_{x\to\infty} f'(x) = 0$. Then, the firm's maximization problem can be written as

$$\max_{A\geq 0, B\geq 0, 0\leq D\leq A/\beta_1} \quad p_1(f(A-\beta_1 D)+f(B+D))+(1-p_1)(f(A)+f(B))-A-B \tag{2.9}$$

with no constraints. The first-order conditions are:

$$p_1 f'(A - \beta_1 D) + (1 - p_1) f'(A) = 1, \qquad (\text{FOC}_A)$$
$$p_1 f'(B + D) + (1 - p_1) f'(B) = 1, \qquad (\text{FOC}_B)$$
$$f'(B + D) - \beta_1 f'(A - \beta_1 D) = 0. \qquad (\text{FOC}_D)$$

Notice that there is a lower bound T such that if $A, B \geq T$, then $p_1(f(A - \beta_1 D) + f(B + D)) + (1 - p_1)(f(A) + f(B)) - A - B \leq 0$ for any permissible D. Since the maximand is strictly positive if the firm offers the optimal committed contract (for which $D = 0$ and $A = B = A - \beta_1 D = k[(k')^{-1}(1)]$), this means that there is a global maximum that either satisfies the above first-order conditions or is at a corner. We show that for $k'(0) < 1 - p_1, \beta_1$, or equivalently $f'(0) > 1/(1 - p_1), 1/\beta_1$, the global maximum is not at a corner. It is clear from the derivatives of the maximand with respect to A and B that the firm's objective function does not obtain a maximum when $A = 0$ or $B = 0$. If $D = A/\beta_1$, either FOC_B does not hold, in which case the maximum is not attained, or FOC_B holds, in which case $f'(B + D) < 1$ and thus $f'(0) > 1/\beta_1$ implies that the derivative of the maximand with respect to D is negative, ruling out such a corner solution. For $D = 0$, either FOC_A and FOC_B do not both hold, in which case the maximum is not attained, or FOC_A and FOC_B both hold, in which case $f'(A) = f'(B) = 1$ and hence the derivative of the maximand with respect to D is positive, ruling out such a corner solution as well.

We have established that a global maximum must satisfy the system of first-order conditions. To prove that the competitive equilibrium is unique, we

next show that the solution to the system of first-order conditions is unique. Because $k'(0) < p_1$ and hence $f'(0) > 1/p_1$, for any $D \geq 0$ there is a unique $A > \beta_1 D$ satisfying FOC_A; call this $\alpha^A(D)$. Since $\alpha^A(D)$ is strictly increasing in D, $\alpha^A(D) - \beta_1 D$ must be strictly decreasing in D. Also, notice that if $B \geq 0$ is fixed, then for any $D \geq 0$ there is either a unique $A \geq \beta_1 D$ satisfying FOC_D or—in case $f'(B + D) > \beta_1 f'(0)$—there exists no solution to this first-order condition; if a solution exists for some B and D, one also exists for higher B and D. If the solution exists, we refer to it as $\alpha_B^D(D)$ and otherwise we set $\alpha_B^D(D) = \beta_1 D$. Note also that if $\alpha_B^D(D) > \beta_1 D$, $\alpha_B^D(D) - \beta_1 D$ is strictly increasing in D.

Since f is strictly concave, f' and f'^{-1} are strictly decreasing. Consider the range of B given by $B \leq f'^{-1}(\beta_1)$, or equivalently $f'(B) \geq \beta_1$. If for fixed B and $D = 0$ there is an A satisfying FOC_D, then $\alpha_B^D(0) = f'^{-1}(f'(B)/\beta_1)$; and otherwise $\alpha_B^D(0) = 0$. In either case, $\alpha_B^D(0) \leq f'^{-1}(1) = \alpha^A(0)$. Using the implicit function theorem,

$$\frac{d\alpha^A(D)}{dD} = \beta_1 \frac{p_1 f''(\alpha^A(D) - \beta_1 D)}{p_1 f''(\alpha^A(D) - \beta_1 D) + (1 - p_1) f''(\alpha^A(D))} < \beta_1,$$

and whenever $\alpha_B^D(D) > \beta_1 D$,

$$\frac{d\alpha_B^D(D)}{dD} = \frac{f''(B + D) + \beta_1^2 f''(\alpha_B^D(D) - \beta_1 D)}{\beta_1 f''(\alpha_B^D(D) - \beta_1 D)} > \beta_1.$$

Since at any crossing point of the two curves $\alpha^A(D) = \alpha_B^D(D) > \beta_1 D$, this means that at any crossing point α_B^D is steeper. In addition, since $\lim_{y\to\infty} f'(y) = 0$, it follows from FOC_D that as $D \to \infty$, $\alpha_B^D(D) > \beta_1 D$ and $f'(\alpha_B^D(D) - \beta_1 D) \to 0$ while FOC_A implies that $f'(\alpha^A(D) - \beta_1 D) > 1$ for any $D > 0$. Hence $\alpha_B^D(D) > \alpha^A(D)$ for sufficiently large D. Summarizing, since $\alpha_B^D(0) \leq \alpha^A(0)$, $\alpha_B^D(D)$ is steeper than $\alpha^A(D)$ at any crossing point, both curves are continuous, and for a sufficiently high D we have $\alpha_B^D(D) > \alpha^A(D)$, for this range of B there is a unique A and D satisfying first-order conditions FOC_A and FOC_D. Call these solutions $A^*(B)$ and $D^*(B)$, respectively. If $B > f'^{-1}(\beta_1)$ then $\alpha_B^D(0) > \alpha^A(0) > \beta_1 D$ and since $\alpha_B^D(D)$ is steeper than $\alpha^A(D)$ at any crossing point no solution to the first-order conditions FOC_A and FOC_D exists in this range of B.

To complete the proof, notice that since $\alpha^A(D)$ is independent of B and $\alpha_B^D(D)$ is increasing in B, $A^*(B)$ and $D^*(B)$ are decreasing in B; by FOC_A, this means that $A^*(B) - \beta_1 D^*(B)$ is increasing in B, which by FOC_D means that $B + D^*(B)$ is increasing in B. Hence, the function $p_1 f'(B + D^*(B)) + (1 - p_1) f'(B)$, which is continuous in B, is strictly decreasing in B. Furthermore,

because $k'(0) < 1-p_1$, $f'(0) > 1/(1-p_1)$, so $p_1 f'(0+D^*(0))+(1-p_1)f'(0) > 1$. Since for $B = f'^{-1}(\beta_1)$, $\alpha_B^D(0) = \alpha_A(0)$, one has $\beta_1 = f'(B) = f'(B + D^*(B))$ for this value of B. Hence for $B = f'^{-1}(\beta_1)$, one has $p_1 f'(B + D^*(B)) + (1 - p_1)f'(B) < 1$. Since $p_1 f'(B + D^*(B)) + (1 - p_1)f'(B)$ is strictly decreasing in B, this implies there exists a unique $B \in (0, f'^{-1}(\beta_1))$ for which $B, D^*(B)$ satisfies FOC_B. Because for $B \leq f'^{-1}(\beta_1)$, $A^*(B), D^*(B)$ characterize a solution to FOC_A and FOC_B, we have shown that $B, A^*(B), D^*(B)$ is the unique solution to the system of first-order conditions. Thus we have shown that the competitive equilibrium is unique.

To characterize the optimal installment plan, we invert the expressions for $k(q_1)$ and $k(q_2)$ found above and plug them into the principal's objective function, yielding

$$\begin{aligned} \max_{c,r_1,r_2} \quad & p_1 \left[k^{-1}\big(c - k(r_2) - \underline{u} + \beta_1 \left(k(r_2) - k(r_1)\right)\big) + r_1\right] \\ + \quad & p_2 \left[k^{-1}\big(c - k(r_2) - \underline{u}\big) + r_2\right] - c. \end{aligned} \tag{2.10}$$

The first-order-conditions with respect to r_1 and r_2 are:

$$\begin{aligned} p_1 \left[1 - \beta_1 \frac{k'(r_1)}{k'(q_1)}\right] &= 0, \\ p_2 \left[1 - \frac{k'(r_2)}{k'(q_2)}\right] - p_1(1-\beta_1)\frac{k'(r_2)}{k'(q_1)} &= 0. \end{aligned}$$

Rewriting these first-order conditions gives the equations in the proposition, which in turn imply that $q_1 < r_1$ and $q_2 > r_2$. It remains to establish that $q_1 + r_1 > q_2 + r_2$. Suppose by contradiction that $q_1 + r_1 \leq q_2 + r_2$. Then the firm would be at least as well off offering a single repayment option (q_2, r_2): the resulting contract satisfies PC and, since there is no choice in period 1, it also satisfies PCC and IC, and yields at least as high profits. This, however, contradicts the fact that in any optimal contract $q_1 < r_1$ and $q_2 > r_2$.

Finally, we show that borrowers overborrow on average. Taking the first-order condition of the maximization problem 2.10 with respect to c gives

$$p_1 \frac{1}{k'(q_1)} + p_2 \frac{1}{k'(q_2)} = 1.$$

By Jensen's inequality, the left-hand side is greater than

$$\frac{1}{p_1 k'(q_1) + p_2 k'(q_2)},$$

which gives $p_1 k'(q_1) + p_2 k'(q_2) > 1$.

To show the analogous inequality for r_1 and r_2, we solve for $k(r_1)$ and $k(r_2)$ from the binding constraints (instead of solving for $k(q_1)$ and $k(q_2)$), invert these, and plug them into the principal's objective function to get

$$\begin{aligned}\max_{c,q_1,q_2} \quad & p_1\left[q_1 + k^{-1}\big(c - k(q_2) - \underline{u} + (k(q_2) - k(q_1))/\beta_1\big)\right] \\ + \quad & p_2\left[q_2 + k^{-1}\big(c - k(q_2) - \underline{u}\big)\right] - c.\end{aligned}$$

Again taking the first-order condition with respect to c and using Jensen's inequality completes the proof.

Proof of Proposition 2.7. First, we show that the borrower strictly prefers the unrestricted market over the restricted one by showing that the perceived utility $\underline{u}$ generated by the competitive-equilibrium contract in the unrestricted market is higher than the borrower's perceived utility in the restricted market. Suppose by contradiction that this is not the case. Then, a contract with the consumption and repayment terms the two types of borrowers choose in the restricted market satisfies the constraints PC, IC, and PCC in Lemma 2.1, and breaks even, and is therefore a competitive-equilibrium contract. But this is impossible since a competitive equilibrium identified in Proposition 2.6 does not replicate outcomes in the restricted market: for the condition $k'(q_1) = \beta_1 k'(r_1)$ to hold, the firm needs to set $R = 1$, and at this interest rate sophisticated borrowers will not repay more in period 1 than 2.

Since sophisticated borrowers understand their behavior, the fact that their perceived utility is higher than in the restricted market implies that their actual welfare is also higher.

We next consider social welfare. The same steps as in Proposition 2.3 establish that as $\beta_1 \to \beta_2$, the competitive-equilibrium contract approaches $(c^*, 1/\beta_2, L^*)$, so that both types' outcomes approach the welfare-maximizing outcome (the only difference in the argument is that the break-even $c(R, L)$ must be defined in expectation). Since in the unrestricted market $k'(q_1) = \beta_1 k'(r_1)$ for any $\beta_1 < \beta_2$, total welfare remains bounded away from the welfare-maximizing level as $\beta_1 \to \beta_2$. Hence, for β_1 sufficiently close to β_2 the restricted market yields higher social welfare. Finally, since a non-sophisticated borrower has lower welfare than a sophisticated borrower, the fact that total welfare remains bounded away from optimal as $\beta_1 \to \beta_2$ implies that the non-sophisticated borrower's welfare also does. Since her welfare in the restricted market approaches the optimum as $\beta_1 \to \beta_2$, for β_1 sufficiently close to β_2 the restricted market yields higher welfare for her.

Proof of Proposition 2.8. We begin by establishing that there is a competitive equilibrium in which the same contracts are offered as when $\hat{\beta}$ is known and each borrower selects the contract designed for her belief $\hat{\beta}$. We first

show borrower optimality (Condition 1 of Definition 2.2). Since a borrower of type $\hat{\beta}$ expects to choose the baseline repayment option in a contract intended for any $\hat{\beta}' \leq \hat{\beta}$, among these contracts she prefers the one intended for her because by Condition 2.1 it gives her the highest perceived period-0 utility. Second, while from a period-0 perspective the borrower prefers the baseline option in the contract for $\hat{\beta}' > \hat{\beta}$ to the baseline option in the contract for her own type, she also believes that she will switch away from this option ex post. Once she takes this into account, the period-0 utility from the contract designed for $\hat{\beta}' > \hat{\beta}$ is lower. To see this last point, suppose by contradiction that a type $\hat{\beta}$ preferred to select the contract designed for $\hat{\beta}' > \hat{\beta}$. Then, the contract for $\hat{\beta}$ is suboptimal when $\hat{\beta}$ is known: the contract designed for $\hat{\beta}'$ both attracts $\hat{\beta}$ types and induces all of them to choose the non-sophisticated repayment option, which by Proposition 2.6 is strictly profitable, and thus this contract guarantees positive profits when $\hat{\beta}$ is known. Since the contracts are identical to the ones in which $\hat{\beta}$ is observable they satisfy the zero-profit and non-redundancy requirements (Conditions 2 and 4). Furthermore, Condition 3 is satisfied because any other contract that gives a borrower of type $\hat{\beta}$ a higher perceived utility makes losses on this type of borrowers, since otherwise this contract could also be profitably introduced when $\hat{\beta}$ is observable.

Now we argue that this competitive equilibrium is unique. Consider any purported equilibrium in which not all $\hat{\beta}$ types are offered the competitive-equilibrium contract for the case in which $\hat{\beta}$ is known. Let $\underline{u}_i'$ be the perceived utility of $\hat{\beta}_i$ in this situation. First, we show that there is some i such that $\underline{u}_i' < \underline{u}_i$. Suppose by contradiction that $\underline{u}_i' \geq \underline{u}_i$ for all i. Then, even if $\hat{\beta}$ was observable, a firm could only break even on each type, and do so only using the competitive-equilibrium contract for each type—contradicting that not all $\hat{\beta}$ types get the same contract as when $\hat{\beta}$ is known.

Now consider the highest i such that $\underline{u}_i' < \underline{u}_i$. For a sufficiently small $\epsilon > 0$, a contract that is optimal for type $\hat{\beta}_i$ with the outside option $\underline{u}_i' + \epsilon$ attracts $\hat{\beta}_i$ and makes positive expected profits on this type. Furthermore, since for any $j > i$, $\underline{u}_i' < \underline{u}_i \leq \underline{u}_j \leq \underline{u}_j'$, the contract does not attract higher $\hat{\beta}_j$. If it attracts $\hat{\beta}_j$ for some $j < i$, it makes strictly positive profits on these borrowers, since they all select the non-sophisticated repayment option in the contract. Hence, the contract makes positive expected profits.

Proof of Proposition 2.9. As in the case of degenerate borrower beliefs, the notion of competitive equilibrium is based on the notion of incentive compatible maps determining what a borrower expects to choose for each possible $\hat{\beta}$ and what she actually chooses (similarly to Definition 2.1). Accordingly, we think of a firm's problem as selecting $(\hat{q}(\hat{\beta}), \hat{r}(\hat{\beta}))$ the borrower thinks she

will choose for each possible $\hat{\beta}$, as well as a $(q, r) = (\hat{q}(\beta), \hat{r}(\beta))$ the borrower actually chooses, where $\hat{q}(\cdot)$ and $\hat{r}(\cdot)$ must be incentive compatible.

First suppose that firms know $F(\cdot)$. Denote the support of F by $\overline{F}$. Rewriting 2.1, the firm's problem is

$$\max_{c,q,r,\hat{q}(\hat{\beta}),\hat{r}(\hat{\beta})} \quad q + r - c$$

$$\text{s.t.} \quad \int \left[c - k(\hat{q}(\hat{\beta})) - k(\hat{r}(\hat{\beta}))\right] dF(\hat{\beta}) \geq \underline{u}, \tag{PC}$$

$$-k(\hat{q}(\hat{\beta})) - \hat{\beta}k(\hat{r}(\hat{\beta})) \geq -k(\hat{q}(\hat{\beta}')) - \hat{\beta}k(\hat{r}(\hat{\beta}')) \text{ for any } \hat{\beta} \in \overline{F}, \hat{\beta}' \in \overline{F} \cup \{\beta\}, \tag{PCC}$$

$$-k(q) - \beta k(r) \geq -k(\hat{q}(\hat{\beta})) - \beta k(\hat{r}(\hat{\beta})) \text{ for any } \hat{\beta} \in \overline{F}. \tag{IC}$$

As before, PC binds because otherwise a firm could raise profits by decreasing c. Notice that for any $\hat{\beta} \leq \beta$, for PCC and IC to both hold we must have $\hat{q}(\hat{\beta}) \leq q$ and $\hat{r}(\hat{\beta}) \geq r$. Hence, the IC constraint $k(q) + \beta k(r) \leq k(\hat{q}(\hat{\beta})) + \beta k(\hat{r}(\hat{\beta}))$ implies that $k(q)+k(r) \leq k(\hat{q}(\hat{\beta}))+k(\hat{r}(\hat{\beta}))$, with a strict inequality if $(q, r) \neq (\hat{q}(\hat{\beta}), \hat{r}(\hat{\beta}))$. Hence, given PC it is optimal to set $(\hat{q}(\hat{\beta}), \hat{r}(\hat{\beta})) = (q, r)$ for all $\hat{\beta} \leq \beta$, and in any optimal contract the set of $\hat{\beta} \leq \beta$ for which this equality does not hold must have measure zero under the agent's beliefs $F(\cdot)$.

Next consider $\hat{\beta} > \beta$. We ignore PCC for these $\hat{\beta}$; it is obvious to check that the resulting contract satisfies it. It is optimal to set $\hat{r}(\hat{\beta}) = 0$ for all $\hat{\beta} > \beta$: for any $\hat{\beta}$ with $\hat{r}(\hat{\beta}) > 0$, we can decrease $k(\hat{r}(\hat{\beta}))$ by some amount and increase $k(\hat{q}(\hat{\beta}))$ by β times the same amount, leaving IC unaffected and weakly increasing the left-hand side of PC. Furthermore, in any optimal contract the set of $\hat{\beta} > \beta$ for which $\hat{r}(\hat{\beta}) > 0$ must have measure zero; otherwise, these steps would create a slack in PC, allowing the firm to decrease c. With $\hat{r}(\hat{\beta}) = 0$ for all $\hat{\beta} > \beta$ (other than a measure zero set under $F(\cdot)$), it is optimal to set $\hat{q}(\hat{\beta}) = \hat{q}$ at the level such that IC binds, and the set of $\hat{\beta} > \beta$ for which this is not the case must have measure zero under $F(\cdot)$.

Given these simplifications, the firm's problem becomes

$$\max_{c,q,r,\hat{q},\hat{r}} \quad q + r - c$$

$$\text{s.t.} \quad F(\beta)\left[c - k(q) - k(r)\right] + (1 - F(\beta))\left[c - k(\hat{q})\right] = \underline{u}, \tag{PC}$$

$$-k(q) - \beta k(r) = -k(\hat{q}). \tag{IC}$$

Expressing $k(\hat{q})$ from IC, plugging it into PC, and solving for c and plugging it into the maximand yields that the firm wants to maximize

$$q + r - k(q) - \left[F(\beta) + (1 - F(\beta))\beta\right] k(r).$$

Solving this yields Equation 2.4. That $q < \hat{q} < q + r$ follows from the fact that IC binds.

Finally, we argue that the above (essentially unique) contract is the competitive-equilibrium contract for a borrower with beliefs $F(\cdot)$ even if firms do not observe borrowers' beliefs. The argument is in two parts.

I. Offering these contracts is a competitive equilibrium. To see this, notice first that the profits a firm earns from an accepted contract are independent of the borrowers' beliefs. Suppose by contradiction that a borrower with beliefs $F(\hat{\beta})$ strictly prefers the contract $(c', \mathcal{C}')$ to a contract $(c, \mathcal{C})$ we have solved for above. Then, the firm could offer a contract $(c' - \epsilon, \mathcal{C}')$ for some $\epsilon > 0$ when $F(\hat{\beta})$ is known and earn strictly positive profits, contradicting the no-profitable-deviation condition of competitive equilibrium.

II. There is no other competitive equilibrium. Let $(c, \mathcal{C}) = (q + r, \{(\hat{q}, 0), (q, r)\})$ be the (essentially unique) competitive-equilibrium contract when $F(\cdot)$ is known (for which we have solved above). Suppose by contradiction that there is a competitive equilibrium in which a borrower with beliefs $F(\cdot)$ accepts a contract $(c', \mathcal{C}')$ that does not satisfy the conditions specified in the proposition. Let $\underline{u}'$ be her perceived utility from $(c', \mathcal{C}')$, and let $\underline{u}$ be her perceived utility from $(c, \mathcal{C})$. Notice that $\underline{u}$ maximizes the borrower's perceived utility among contracts that earn zero profits given the borrower's actual behavior. Since $(c', \mathcal{C}')$ is not a competitive equilibrium when $F(\cdot)$ is known but earns zero profits, this implies that $\underline{u} > \underline{u}'$. Therefore, a firm can offer $(c - \epsilon, \mathcal{C})$, and for a sufficiently small $\epsilon > 0$ both attract the borrower with beliefs $F(\cdot)$ and make positive profits from her. Since a borrower's behavior is independent of her beliefs, a firm still makes positive profits if it also attracts other borrowers, contradicting the no-profitable-deviation condition.

Chapter 3

Competition and Price Variation when Consumers are Loss Averse[1]

3.1 Introduction

Menu costs, tacit collusion, search costs, kinked demand curves, and many other theories have been invoked to explain a widespread empirical fact: that prices in imperfectly competitive industries often do not change when costs or demand change.[2] Existing theories, however, cannot convincingly account for another way in which prices vary surprisingly little: they are often identical across differentiated products. As is part of industrial-organization folklore starting from Hall and Hitch (1939) and Sweezy (1939) and is documented by the Competition Commission of the United Kingdom (1994), Beck (2004), and Heidhues and Kőszegi (2007), many non-identical competitors charge identical, "focal," prices for their differentiated products. In addition, as documented for instance by McMillan (2004) and Einav and Orbach (2007) and familiar to anyone who has bought clothes, books, or movie tickets, many retailers selling multiple products with different cost and/or demand characteristics charge the same, "uniform," price for them.

[1] This chapter is coauthored with Paul Heidhues, and appeared in the American Economic Review (2008), 98(4), pp. 1245-1268.

[2] Indirect evidence for this fact—often referred to as price stickiness—is provided by Kashyap (1995), Slade (1999), and Chevalier, Kashyap and Rossi (2003), who document in various retail industries that regular prices typically do not change for months at a time. There is also direct evidence that marginal-cost changes are sometimes fully absorbed by retailers (Competition Commission of the United Kingdom 1994, page 150, Section 7.41). And evidence on countercyclical markups reviewed in Rotemberg and Woodford (1999) suggests that even when prices do adjust to circumstances, they move less than marginal costs.

To explain the above tendencies toward reduced price variation, we develop a model of price competition between profit-maximizing firms selling to loss-averse consumers. Because consumers are especially averse to paying a price when it exceeds their expectation of the purchase price, the price responsiveness of demand—and hence the intensity of competition—is greater at higher than at lower market prices, reducing or eliminating price variation. Unlike in most previous theories, this logic applies both to different possible prices of the same product and to prices of different products, so that our theory not only explains the unresponsiveness of prices to changing circumstances, it also often predicts focal and uniform pricing as the *unique* outcome even for asymmetric firms and products. And because a change in the responsiveness of demand affects competition more when the value of an extra consumer is high, we predict that these tendencies are stronger in more concentrated industries.

Section 3.2 presents our model and illustrates our solution concepts and some key results using a two-firm example. Building on Salop (1979), a consumer's "taste" is drawn uniformly from the circumference of a circle, and she is looking to buy exactly one of n products located equidistant from each other on the same circle. Her utility from or "satisfaction with" a product is decreasing in the product's distance from her taste, and she also suffers additive disutility from paying the product's price. But (applying Kőszegi and Rabin 2006, 2007) we posit that in addition to this *intrinsic utility*, a consumer derives *gain-loss utility* from comparing outcomes in money and product satisfaction to her lagged rational expectations about those outcomes, with losses being more painful than equal-sized gains are pleasant. For example, if she had been expecting to spend \$14.99 on a Britney Spears CD—her favorite music—she experiences a sensation of loss if she buys that CD for \$18.98, and also if she instead buys a—less agreeable—Madness CD for \$14.99.[3] And if she expected to pay either \$14.99 or \$19.99 for something, paying \$18.98 for it generates a mixture of two feelings, a loss of \$3.99 and a gain of \$1.01, with the weight on the loss equal to the probability with which she expected to pay \$14.99.

The firms, none of whom owns two neighboring products, are standard: they face uncertain privately observed costs of production and simultaneously set prices to maximize expected profits given other firms' behavior and consumer expectations. We begin the analysis in Section 3.3 by showing that the necessary and sufficient condition for a focal-price equilibrium—an equilibrium in which all firms always charge the same focal price p^*—to exist is

[3] Actual prices are taken on September 4th, 2005, from www.amazon.com. \$14.99 is the retail price of both CD's (and numerous others), while \$18.98 is a typical list price.

that any two cost realizations of any two firms are within a given constant. This condition allows for, say, one firm to have higher costs than another in all states of the world. If consumers had expected to pay p^* with probability one, they assess buying at a price greater than p^* as a loss in money and buying at a price lower than p^* merely as a gain in money, so that demand is more responsive to unilateral price increases than decreases from p^*. Due to this asymmetry, for a range of cost levels p^* is the optimal price to charge.

We next establish two properties of focal-price equilibria. First, a focal-price equilibrium is more likely to exist in more concentrated industries. Since the profits from a consumer are higher in this environment, the asymmetric demand responsiveness at p^* creates a greater difference in marginal profits from price increases versus price decreases from p^*, and hence yields a greater range of costs for which p^* is the optimal price to charge. Second, loss aversion increases prices. Since a consumer is more sensitive to a loss from getting a surprisingly low product satisfaction than to a gain from paying surprisingly little, attracting her from a competitor is difficult, decreasing competition.

In Section 3.4, we derive a sufficient condition that guarantees that *any* equilibrium is a focal-price equilibrium, even though it does not require firms to have the same cost distribution. The key is to argue that a firm sets a deterministic price in any equilibrium; then, if the supports of firms' cost distributions have even a single common point, firms cannot charge different deterministic prices, so any equilibrium is a focal-price one. If the consumer expected a firm's prices to be stochastic, the sense of loss from comparing the realized price to lower possible ones would make her demand more responsive at higher than at lower prices in the firm's anticipated distribution. Then, if the firm's costs do not vary much, in contradiction to equilibrium it could increase profits either by decreasing high prices (attracting a lot of extra demand) or by increasing low prices (not losing much demand).

In Section 3.5, we characterize all equilibria with common stochastic marginal costs and symmetric pricing strategies. This leads to a tractable model for studying price variation when conditions for focal pricing are not necessarily met, and for analyzing firms' responses to industry-wide cost shocks. As above, if a consumer had expected stochastic prices, her demand is more responsive at higher than at lower prices within the anticipated price distribution. Hence, competition is fiercer at higher prices, leading to markups that strictly decrease in cost. Using the empirical observation that costs are strongly procyclical, this means that markups are countercyclical. In addition, in some regions of cost it may be that competition at higher prices is tougher to an extent that is inconsistent with firms raising their price in response to cost increases at all. In such regions, the price must be constant in cost.

In Section 3.6, we argue that our results are robust to a number of modifications of our model, including dynamics, demand asymmetries and shocks, heterogeneity in consumer preferences, and the endogenous determination of the number of firms. In Section 3.7, we discuss theories of pricing most closely related to our model. By dint of predicting equal and sticky prices even in a one-shot setting, our theory cautions against the common interpretation of these patterns as signs of collusion. In fact, we can go further: because playing an equilibrium with *equal* in addition to *sticky* prices does not help firms in detecting each other's deviations, models of collusion do not provide a compelling reason for ex-ante asymmetric firms to set equal prices, whereas our theory does. In the same vein, other explanations of price stickiness are not intended to, and largely do not, predict focal and uniform pricing. For instance, menu costs may explain sticky pricing and can perhaps contribute to uniform pricing, but they cannot address equal pricing across firms. And if unexpected price increases trigger costly search by consumers but unexpected price decreases do not, sticky pricing can result, but there is no reason for differentiated products to have the same price.

3.2 Setup and Illustration

This section introduces our theory and explains the solution concepts and some key results through simple examples and heuristic calculations. We incorporate consumer loss aversion into the model of Salop (1979) using a disciplined approach introduced by Kőszegi and Rabin (2006): we base the reference-dependent "gain-loss utility" on classical "intrinsic utility" taken straight from Salop (1979) and fully endogenize the reference point as lagged rational expectations. If there is no loss aversion, our theory reduces to Salop's.

3.2.1 Reference-Dependent Utility

A mass of one of consumers have tastes $\chi \in [0, 1]$ distributed uniformly on the circumference of a circle with perimeter one. There are $n \geq 2$ products denoted $y_1, \ldots, y_n$ on the same circle equidistant from each other. A consumer can buy at most one product, and to avoid unenlightening extra notation, we assume her utility from not consuming is negative infinity, so that she always does buy a product.[4] Letting $d(\chi, y)$ denote the distance of χ and y on the

[4] Our results would be identical if consumers had an option of not buying, but v below was sufficiently high (or costs and product differentiation sufficiently low) that no consumer took advantage of this option in equilibrium. And in Section 3.6, we argue that our qualitative

circle, the intrinsic utility of consumer χ from buying product y at price p is $v - t \cdot d(\chi, y) - p$, where $k_1 = v - t \cdot d(\chi, y)$ is her intrinsic utility from or "satisfaction" with the good and $k_2 = -p$ is her intrinsic utility from paying its price. Like previous authors, we interpret χ as the consumer's "ideal variety," and $t \cdot d(\chi, y)$ as her disutility from consuming a product y different from her ideal. The constant t is a measure of the (intrinsic) differentiation between products.

For a riskless consumption outcome $k = (k_1, k_2)$ and riskless reference point $r = (r_1, r_2)$ defined over product satisfaction and money, total utility $u(k|r)$ is composed of two additive terms: intrinsic utility introduced above, and reference-dependent "gain-loss utility" equal to $\mu(k_1 - r_1) + \mu(k_2 - r_2)$. To capture loss aversion, we assume that μ is two-piece linear with a slope of 1 for gains and a slope of $\lambda > 1$ for losses.[5] This specification incorporates three key assumptions. First, the consumer assesses gains and losses in the two dimensions, satisfaction and money, separately. Hence, she evaluates a good that costs more but is closer to her ideal than the reference point as a loss in money and a gain in satisfaction—and not, for example, as a single gain or loss depending on total intrinsic utility relative to the reference point. This is consistent with much experimental evidence commonly interpreted in terms of loss aversion.[6] Second, while money is on a different psychological dimension from any of the n products, our model also says that the n products are on the *same* dimension. This assumption reflects our impression that goods that compete most strongly with each other are typically hedonically substitutable; indeed, that is partly why they compete. Third, since the gain-loss utility function μ is the same in both dimensions, the consumer's sense of gain or loss is directly related to the intrinsic value of the changes in question. While we find this assumption psychologically realistic, we point out in Section 3.6 that

results on reduced price variation would survive even if consumers made a relevant decision of whether to buy.

[5] In order not to clutter our formulas with extra notation, we do not introduce a weight on gain-loss utility relative to intrinsic utility. This does not qualitatively affect any of our results, and we have confirmed that our calibration at the end of Section 3.3 also remains unchanged.

[6] Specifically, it is key to explaining the endowment effect and other observed regularities in riskless trades. The common and intuitive interpretation of the endowment effect—that randomly assigned "owners" of an object value it more highly than "non-owners"—is that owners construe giving up an object as a painful loss that counts more than the money they receive in exchange, so they attach a high monetary value to the object. But if gains and losses were defined over the value of an 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 only necessary for one of our results, Proposition 3.2.

Since we assume below that the reference point is expectations, we extend the above utility function to allow for the reference point to be a probability measure Γ over $\mathbb{R}^2$:

$$U(k|\Gamma) = \int_r u(k|r)d\Gamma(r). \tag{3.1}$$

In evaluating k, the consumer compares it to each possibility in the reference lottery. For example, if she had been expecting to pay either \$15 or \$20 for her favorite CD, paying \$17 for it feels like a loss of \$2 relative the possibility of paying \$15, and like a gain of \$3 relative to the possibility of paying \$20. In addition, the higher the probability with which she expected to pay \$15, the more important is the loss in the overall experience.

We assume that consumers' prior on χ is identical to the population distribution, $U[0,1]$. Since it gives rise to the same distribution of satisfaction from each product, an equivalent model is one in which consumers know their ideal variety, but are uncertain about the positioning of products.[7] A situation where consumers have a very good idea about their ideal variety as well as the products offered corresponds to a narrow or even degenerate prior distribution on χ, and yields a different model. In Section 3.6, we argue that our results in Sections 3.3 and 3.5 carry over to this case unchanged, and that reasonable specifications also yield our results in Section 3.4.

3.2.2 Concepts and Results: Illustration

While Section 3.2.1 defined how the consumer's utility depends on her reference point, we must also specify what the reference point is and how firms behave when selling to loss-averse consumers. In this section we illustrate some of our definitions and results in a two-firm example and without the full formal detail of later sections.

To both motivate our model of reference-point determination and explain a key result, suppose the two firms in the market are expected to set deterministic prices p_1 and $p_2 > p_1$ for products 1 and 2. In the face of these prices, what is the consumer's reference point for evaluating her purchase? We posit that it is her *lagged rational expectations* about outcomes. But since these depend on her own behavior, our assumption requires elaboration. To illustrate, suppose that the consumer had planned to buy the cheaper product if her taste is within distance $\alpha \in (1/4, 1/2)$ of firm 1, and to buy the expensive product otherwise,

[7] More precisely, our model is identical to one in which each consumer knows her ideal variety, and product 1's location is drawn from a uniform distribution on the circle, with the n products still equidistant from each other.

as shown on the left-hand side of Figure 3.1. This plan induces an expected purchase-price distribution F with mass 2α on p_1 (the probability that χ falls within α of $y_1 = 0$) and mass $1 - 2\alpha$ on p_2, as well as an expected distribution of the purchased product's distance from ideal that is shown on the right-hand side of Figure 3.1. Hence, the consumer's reference point—and so her utility at the time of purchase—is affected by the plans she had formed earlier: if α is higher, she expected to pay less and get lower product satisfaction with higher probability, which makes paying a high price more painful and getting a less satisfying product less painful. To close the model, we follow Kőszegi and Rabin (2006) in requiring a consistency condition called personal equilibrium: the consumer can only make plans she knows she will follow through. In the current setting, this means that given the expectations above, when she has taste α the consumer must be indifferent between purchasing from firms 1 and 2.

While our model of consumer behavior is new, firm behavior is more or less standard: each firm maximizes expected profits given other firms' behavior and the consumer's reference point.[8] As in a standard model, a major factor in determining equilibrium is the consumer's reaction to price changes. To understand a key part of this reaction, we focus on the money dimension. If the consumer above (unexpectedly) pays a price p satisfying $p_1 < p < p_2$, her reference-dependent utility in money is

$$-p - \lambda(2\alpha)(p - p_1) + (1 - 2\alpha)(p_2 - p).$$

The first term is intrinsic utility. The second term represents a sense of loss from comparing p to the lower expected purchase price p_1—a loss of $p - p_1$ weighted by the probability with which she expected to pay p_1, 2α. And the third term represents a gain from comparing p to the higher expected purchase price p_2—a gain of $p_2 - p$ weighted by the probability with which she expected to pay p_2, $1-2\alpha$. Hence, a small price increase decreases the consumer's utility in the money dimension by $1 + \lambda(2\alpha) + (1 - 2\alpha)$. More generally, at any price p that is not a mass point of the expected purchase-price distribution F, the utility impact of a marginal price change is equal to $1+\lambda \cdot F(p)+1 \cdot (1-F(p))$. The intuition is simple. Paying p is experienced as a loss relative to lower prices in the expected purchase-price distribution, and as a gain relative to higher prices in that distribution. Due to this "comparison effect," a change in p is counted as a change in loss with weight $F(p)$—the probability with which the

[8] We assume profit maximization to capture our impression that firms display reference-dependent preferences far less than consumers do, and to isolate the effect of *consumer* loss aversion on market outcomes.

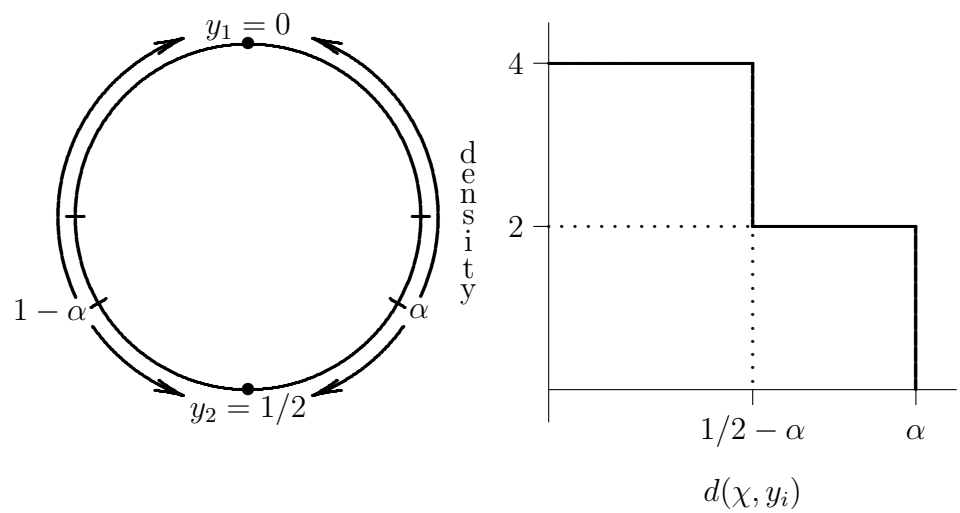

Figure 3.1: Illustration of a Consumer's Plans with Two Firms Charging Deterministic Prices

The figure on the left illustrates the consumer's strategy: if her taste is within distance $\alpha \in (1/4, 1/2)$ of firm 1, she buys the cheaper product 1, otherwise she buys the more expensive product 2. The figure on the right illustrates the density of the expected distribution of the purchased product's distance from ideal that is induced by this plan. If the consumer's taste is very close to a product, she buys that product, so the density is 4 for small distances. For larger distances, the consumer is willing to buy a product that far from her taste only if it is the cheaper product, so the density shrinks to 2. Given her plans, the consumer does not expect to purchase a product that is further than α from her ideal variety.

consumer expected to pay lower prices—and as a change in gain with weight $1 - F(p)$—the probability with which she expected to pay higher prices.

Based on the above considerations, at any price p_i that is not a mass point of F, the partial derivative of firm i's demand with respect to its price p_i is

$$-\frac{[1 + \lambda F(p_i) + (1 - F(p_i))]}{t \cdot z}, \tag{3.2}$$

where z reflects the consumer's gain-loss utility in product satisfaction. When p_i is a mass point of F, demand is continuous and left and right differentiable with the derivatives given by the left and right limits of Expression (3.2). Although z depends on the consumer's reference point and the prices set by the firm's neighbors, for the purposes of this section we assume heuristically that it is an exogenous constant.

Now we can illustrate results about possible equilibria in the market. Proposition 3.3 says (more generally and precisely) that so long as there exist realizations of marginal costs c_1 and c_2 with $c_1 \geq c_2$, an equilibrium in which the firms charge deterministic prices p_1 and $p_2 > p_1$ does not exist. With these prices, when $c_1 \geq c_2$ firm 2 has a higher markup than firm 1, so it benefits more from one extra consumer than firm 1 suffers from one less consumer. Furthermore, since firm 2 has lower infra-marginal demand than firm 1, its infra-marginal losses from lowering its price are lower than firm 1's infra-marginal gains from raising its price by the same amount. And because by Expression (3.2) the responsiveness of demand at prices just below p_2 is the same as the responsiveness just above p_1, either firm 1 wants to raise its price or firm 2 wants to lower its price.

In contrast, Proposition 3.1 says (generally and precisely) that even if the firms have different marginal-cost distributions, an equilibrium in which they charge the *same* deterministic price p^* often exists. If consumers expect to pay p^* with certainty, Expression (3.2) implies that the price responsiveness of a firm's demand when it unilaterally raises its price is $-(1 + \lambda)/(tz)$, while the responsiveness when it lowers its price is only $-2/(tz)$. Intuitively, a price decrease of a given amount expands demand less than a price increase of the same amount reduces demand because consumers are not as attracted by a gain in money as they dislike a loss in money. Since the effect of these price changes on profits from inframarginal consumers is symmetric, for a range of cost levels neither deviation can increase profits.

To conclude this section, we illustrate the reasoning behind our trickiest result, Proposition 3.4, which provides conditions for all firms to set a deterministic price. Combined with conditions above ruling out different deterministic prices for different firms, this leads to conditions under which any equilibrium

is a focal-price equilibrium. The essence of the argument can be seen most simply by assuming that in equilibrium firm 1's cost and price are continuously distributed on $(\underline{c}_1, \overline{c}_1)$ and $(\underline{p}_1, \overline{p}_1)$, respectively, and firm 2's price is also continuously distributed. We show a condition under which a contradiction results. If the marginal costs of the two firms are sufficiently similar and densely distributed, firm 1's expected demand is about 1/2. Using this and Expression (3.2), the fact that at cost $\overline{c}_1$ firm 1 does not want to set a price lower than $\overline{p}_1$ implies $\overline{p}_1 - \overline{c}_1 \lessapprox (tz)/[4+2(\lambda-1)F(\overline{p}_1)]$. Similarly, that at cost $\underline{c}_1$ firm 1 does not want to set a price above $\underline{p}_1$ implies $\underline{p}_1 - \underline{c}_1 \gtrapprox (tz)/[4+2(\lambda-1)F(\underline{p}_1)]$. Subtracting the latter inequality from the former one and using that the consumer must have expected to buy from firm 1 with probability of about one-half, so that $F(\overline{p}_1) - F(\underline{p}_1) \gtrapprox 1/2$, we get

$$\begin{aligned}
\overline{c}_1 - \underline{c}_1 &\gtrapprox \frac{tz}{2}\left[\frac{(\lambda-1)(F(\overline{p}_1) - F(\underline{p}_1))}{[2+(\lambda-1)F(\underline{p}_1)][2+(\lambda-1)F(\overline{p}_1)]}\right] + \overline{p}_1 - \underline{p}_1 \\
&\gtrapprox \frac{tz}{2}\frac{(\lambda-1)\frac{1}{2}}{(1+\lambda)^2} = \frac{tz}{4}\frac{\lambda-1}{(\lambda+1)^2}.
\end{aligned}$$

Hence, if $\overline{c}_1 - \underline{c}_1$ is small, firm 1's incentives are incompatible with the above purported equilibrium. Intuitively, if firm i chooses a stochastic price, then—no matter how close are its highest and lowest prices—expectations-based loss aversion dictates an amount by which the consumer is more price responsive at the firm's high price than at its low price. This in turn implies an amount by which the markup at the high price must be lower than at the low price. But if the firm's highest and lowest costs do not differ by that amount, such a situation is impossible.

3.2.3 Personal Equilibrium and Market Equilibrium

This section formally specifies consumer behavior and defines market equilibrium. We begin by defining the concept of personal equilibrium motivated above with two firms setting deterministic prices for more firms and arbitrary price distributions. Suppose that the consumer has a prior $H \in \Delta(\mathbb{R}^n_+)$ on the non-negative price vectors she might face. Her decision of which good to buy is made after observing the realized χ and the realized price vector, and is described by the strategy $\sigma : [0,1] \times \mathbb{R}^n_+ \to \{1, \ldots, n\}$. As emphasized above, her reference point for evaluating outcomes is her lagged rational expectations about outcomes. This is the distribution $\Gamma_{\sigma,H}$ induced by σ, H, and her uniform prior over χ, over vectors (k_1, k_2) of product satisfaction and expenditure.

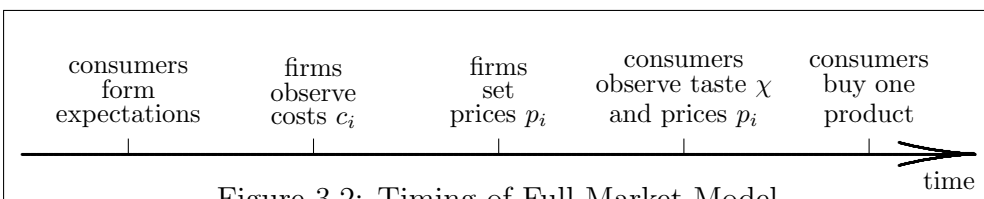

Figure 3.2: Timing of Full Market Model

To deal with the resulting interdependence between behavior (σ) and expectations ($\Gamma_{\sigma,H}$), personal equilibrium (Kőszegi and Rabin 2006) requires the behavior generating expectations to be optimal given the expectations:

Definition 3.1. σ is a personal equilibrium for the price distribution H if

$$\sigma(\chi, p) \in \text{argmax}_{i\in\{1,\ldots,n\}} \; U\left(v - t \cdot d(\chi, y_i)\;,\; p_i \mid \Gamma_{\sigma,H}\right)$$
$$\text{for all } \chi \in [0,1] \text{ and } p \in \mathbb{R}^n.$$

We now integrate consumer behavior into a notion of market equilibrium. The timing of our full market model is illustrated in Figure 3.2. Consumers first form the expectations regarding consumption outcomes that later determine their reference point. Next, firms observe their cost realizations and simultaneously set prices. Finally, each consumer observes her ideal variety and the realized market prices, and purchases a good.

For expositional simplicity, we assume that the n available products are produced by n different firms, with firm i producing good i, y_i. In Section 3.6, we argue that as long as each product is owned by exactly one firm and no firm owns neighboring products, our results on how products are priced extend unchanged to situations where some or all firms produce multiple products. Hence, results on focal pricing below also imply uniform pricing for multiproduct firms.

Firms' costs are jointly distributed according to Θ on the set $\prod_{i=1}^{n}[\underline{c}_i, \overline{c}_i]$, where $[\underline{c}_i, \overline{c}_i]$ is the smallest closed interval containing the support of firm i's cost distribution. Let $\underline{c} = \min\{\underline{c}_i\}$ and $\overline{c} = \max\{\overline{c}_i\}$. Denote firm i's pricing function by $P_i : [\underline{c}_i, \overline{c}_i] \to \mathbb{R}_+$. Let $P = (P_1, \ldots, P_n)$ be the vector of pricing strategies, H_P the market price distribution induced by P and Θ, and $P_{-i}(c_{-i}) = (P_j(c_j))_{j\neq i}$ the price vector of firms other than i.

Definition 3.2. The strategy profile $\{P, \sigma\}$ is a market equilibrium if

1. σ is a personal equilibrium for the price distribution H_P.

2. For each i and $c_i \in [\underline{c}_i, \bar{c}_i]$,

$$P_i(c_i) \in \text{argmax}_{p_i \in \mathbb{R}_+} \quad (p_i - c_i) \cdot \text{Prob}\left[\sigma(p_i, P_{-i}(c_{-i}), \chi) = i \;\mid\; c_i\right].$$

Our definition of market equilibrium extends Bayesian Nash equilibrium to allow for reference effects in consumer behavior. A market equilibrium needs to satisfy two conditions. First, consumers play a personal equilibrium given the correctly forecasted price distribution.[9] Second, each firm at each cost realization plays a best response to other firms' pricing strategies, taking consumers' reference point as given.[10]

One of our goals in this paper is to investigate circumstances under which all firms charge the same price irrespective of their cost positions. We introduce a term for such a situation.

Definition 3.3. A market equilibrium is a focal-price equilibrium if there is a price p^* that all firms charge with probability one. In a focal-price equilibrium, p^* is the focal price.

To address what we believe is mostly a technical issue that arises in our model as well as the standard Salop model with cost asymmetry, we introduce a restricted class of market equilibria.

Definition 3.4. A market equilibrium is an interior equilibrium if for any equilibrium price vector, all firms sell to a positive measure of consumers on each side of their location.

[9] For notational simplicity, our definition implicitly imposes that there is a single representative consumer, or all consumers play the same personal equilibrium. We argue in Section 3.6 that this does not affect our results.

[10] Hence, at the stage when firms' prices are chosen, these prices do not influence a consumer's reference point. This specification reflects our assumption that the reference point is *lagged* rather than contemporaneous expectations. If, for instance, a consumer had been confidently expecting to pay \$14.99 for a CD that she now finds costs \$18.98, she presumably adjusts her beliefs about the price distribution immediately, but she would still experience paying \$18.98 as a loss.

While the expectations that are relevant for specifying the reference point are clearly lagged, the fact that we do not specify when exactly these expectations are formed is a weakness of our approach. In addition, we exclude any direct influence of firm behavior on what expectations consumers form. As we show below, for instance, there are typically a continuum of focal prices, so firms have a strong incentive to manage consumers' price expectations. While firms might be able to do so through public commitments to prices, advertising, or other marketing activities, analyzing these motives is beyond the scope of this paper.

In principle, it is possible that a firm prices so low relative to a neighbor that it attracts all consumers between them. In the standard Salop model, at such a price level there is a discontinuity in the firm's demand, as it suddenly captures *all* consumers on the other side of its neighbor. In our model, both convex kinks and discontinuities in demand are possible. Since we do not know how to handle equilibria with such situations, we follow previous models with cost uncertainty (e.g. Aghion and Schankerman 2004) and focus on (what we call) interior equilibria.

In the following three sections, we analyze the above model. As a foundation for this analysis, Lemma 3.1 derives a key part of firms' incentives, the price responsiveness of demand. Suppose the consumer had expected the distribution of her purchase price and the distribution of her acquired product's distance from ideal to be $F(\cdot)$ and $G(\cdot)$, respectively. These distributions are determined as part of a personal equilibrium by the consumer's strategy σ and the distribution of market prices H in the following way:

$$\begin{aligned} F(p') &= \text{Prob}_{p\sim H,\chi\sim U[0,1]}[p_{\sigma(p,\chi)} \leq p'] \text{ for any } p' \geq 0, \text{ and} \\ G(s) &= \text{Prob}_{p\sim H,\chi\sim U[0,1]}[d(\chi, y_{\sigma(p,\chi)}) \leq s] \text{ for any } s \geq 0. \end{aligned}$$

Denoting right and left limits by subscripts $\downarrow$ and $\uparrow$, $F(\cdot)$ and $G(\cdot)$ can be used to express how demand depends on prices:

Lemma 3.1. *Suppose that given the consumer's lagged expectations $F(\cdot)$ and $G(\cdot)$ and the realized prices, the realized tastes that make the consumer indifferent between purchasing y_i and the respective neighboring products are at distances $x^+, x^- \in (0, 1/n)$ from y_i on its two sides. Then, the right derivative of firm i's residual demand, $D_i(p_i, p_{-i})$, with respect to its price p_i is*

$$\begin{aligned} D_{i\downarrow}(p_i, p_{-i}) = -\frac{1}{2t} \cdot \Bigg[& \frac{2 + (\lambda - 1)F(p_i)}{2 + \frac{\lambda-1}{2}(G(x^+) + G(1/n - x^+))} \\ & + \frac{2 + (\lambda - 1)F(p_i)}{2 + \frac{\lambda-1}{2}(G(x^-) + G(1/n - x^-))} \Bigg], \end{aligned} \tag{3.3}$$

and the left derivative $D_{i\uparrow}(p_i, p_{-i})$ is given by the expression in which $F_\uparrow(p_i)$ replaces $F(p_i)$ above.

Whenever firm i's two neighbors set the same price p, we will denote the indifferent consumer on each side by $x = x^+ = x^-$, and firm i's demand as a function of its price by $D_i(\cdot, p)$.

The price responsiveness of a firm's residual demand D_i derives partly from the comparison effect explained above: as reflected in the numerators in

Equation (3.3), the higher the probability with which the consumer expected lower prices, the more she experiences paying a given price as a loss, and hence the more responsive she is to price changes. The comparison effect has two important implications we will use repeatedly in the paper. First, the residual demand curve is kinked at p_i if $F(\cdot)$ has an atom at p_i, and it is differentiable at p_i if $F(\cdot)$ has no atom at p_i. Second, the price responsiveness of demand is greater at higher prices in the purchase-price distribution.

More subtle than the effect of utility from money itself is the effect of product satisfaction on the price responsiveness of demand. A small price change can affect a consumer's choice if she is approximately indifferent between firm i's product at a distance x^+ from ideal and the neighbor's product at a distance $1/n - x^+$ from ideal. If any of these options is evaluated as a loss to a greater extent—that is, if the expected probability of a better product, $G(x^+)$ or $G(1/n - x^+)$, is higher—then a change in the consumer's realized taste has a greater effect on which option she prefers. This means that a given price change reverses the consumer's decision for a smaller range of taste realizations, lowering the price responsiveness of firm i's demand.

3.3 Existence and Properties of Focal-Price Equilibria

As our first step in analyzing the model, we establish a necessary and sufficient condition for a focal-price equilibrium to exist, and explore the condition's implications for the price level and the effect of industry concentration on pricing. The condition allows for stochastic costs, and even for commonly known *differences in* (stochastic or deterministic) costs.

To derive our condition, we solve for the cost levels c_1 for which firm 1 does not want to deviate from a focal price of p^*. Market equilibrium requires that the consumer anticipated all prices to be p^*, so that she expected to spend p^* with probability 1. In addition, since (having expected equal prices) she expected to buy the product closest to her taste, $G(\cdot)$ is the uniform distribution on $[0, 1/(2n)]$. In addition, the consumers who are indifferent between a firm and its neighbor are at distances $x^+ = x^- = x = 1/(2n)$ from the firm, so $G(x) = G(1/n - x) = 1$.

Given these considerations, Equation (3.3) implies that $D_{1\downarrow}(p^*, p^*) = -1/t$. Using that $D_1(p^*, p^*) = 1/n$, so long as $(p^* - c_1)/t \geq 1/n$ firm 1 cannot benefit from locally raising its price. Similarly, since $D_{1\uparrow}(p^*, p^*) = -2/(t(1+\lambda))$, so long as $2(p^* - c_1)/(t(1+\lambda)) \leq 1/n$ firm 1 cannot benefit from locally lowering its price. Combining and rearranging these conditions, charging p^* is locally

optimal if and only if

$$p^* - \frac{t}{n} \cdot \frac{1+\lambda}{2} \leq c_1 \leq p^* - \frac{t}{n}. \tag{3.4}$$

In the appendix, we show that when local deviations are unprofitable, non-local deviations are also unprofitable. Therefore:

Proposition 3.1. *A focal-price equilibrium exists if and only if*

$$\overline{c} - \underline{c} \leq \frac{\lambda - 1}{2} \cdot \frac{t}{n}.$$

When there is no loss aversion ($\lambda = 1$), a focal-price equilibrium exists only if $\overline{c} - \underline{c} = 0$—if all firms have the same deterministic cost. As explained above, however, if consumers are loss averse and expect all firms to charge the same price p^*, there is a kink in residual demand at p^*, so for a range of cost levels p^* is the optimal price to charge. Hence, with loss aversion a focal-price equilibrium can exist despite cost differences and variation.

Proposition 3.1 has a number of important comparative-statics implications for when a focal-price equilibrium exists. Naturally, a focal price is easier to sustain when the range of marginal costs $\overline{c} - \underline{c}$ is smaller. Also, a focal-price equilibrium is more likely to exist when consumer loss aversion is greater. The greater is λ, the greater is the difference between a consumer's sensitivity to price increases from p^* and price decreases from p^*. Hence, the greater is the difference between the effects on profits of price increases and decreases, and the greater is the range of cost levels for which p^* is the optimal price. One implication of this comparative static and our model more generally may be that ceteris paribus, prices are less variable in consumer markets than in transactions between (presumably less loss averse) firms. Evidence in Blinder, Canetti, Lebow and Rudd (1998) is broadly consistent with this prediction.

Most interestingly, a focal-price equilibrium is more likely to exist when market power as measured by product differentiation relative to the number of firms (t/n) is greater. For an intuition, consider the price p^* at which a firm with cost $\overline{c}$ is just indifferent to raising its price. Then, due to a kink in demand, for a range of cost decreases it strictly prefers not to decrease its price. This range—and hence the allowed cost variation for a focal-price equilibrium to exist—is increasing in the markup $p^* - \overline{c}$, so that it is larger in less competitive industries. With a higher markup, the value of a marginal consumer is higher, so a change in the responsiveness of demand has a greater effect on the firm's incentives to change its price. Hence, the low responsiveness

of demand to price decreases makes the firm more reluctant to cut its price, and it will not want to do so for a greater range of cost decreases.

In addition to identifying conditions under which a focal-price equilibrium exists, Inequality (3.4) determines what the focal price level can be:

Proposition 3.2. *There is a focal-price equilibrium with focal price p^* if and only if*

$$\overline{c} + \frac{t}{n} \leq p^* \leq \underline{c} + \frac{t}{n} \cdot \frac{1+\lambda}{2}.$$

In the corresponding Salop model without loss aversion, the support of a firm's interior-Bayesian-Nash-equilibrium prices is bounded above by $\overline{c}+t/n$, and this bound can only be attained if the firm has realized cost $\overline{c}$.

Proposition 3.2 says that in a focal-price equilibrium, consumer loss aversion leads to increased prices: even at the *lowest* possible cost, a firm charges a higher price than it would in the standard model at the *highest* possible cost. Intuitively, if a firm unilaterally lowers its price, it attracts some consumers of the neighboring firms, who (unexpectedly) choose a good that both costs and matches their taste less than expected. Since consumers are more sensitive to the loss in satisfaction than to the gain in money, the firm attracts fewer of them than without loss aversion. But if the firm unilaterally raises its price, its consumers must either pay a higher price or get a less satisfactory product than they expected was possible, so—as either choice involves a loss—the firm loses the same number of consumers as without loss aversion. Since loss aversion decreases a firm's incentive to lower its price and leaves a firm's incentive to raise its price unchanged, it increases equilibrium prices.

Proposition 3.2 implies that if there is a focal-price equilibrium, there are generically multiple ones, with the set of possible focal prices being a closed interval. If consumers' expectation of the price increases from p to $p' > p$, the difference between paying p' and paying p turns from a loss to a foregone gain. Because this makes demand less responsive, firms are more willing to increase prices, within limits exactly matching the increased expectations.

Beyond a theoretical possibility, our model predicts that focal-price equilibria can exist for calibrationally non-trivial amounts of cost variation. Assuming $\lambda = 3$, which corresponds to the conventional assumption of about two-to-one loss aversion in observable choices (Tversky and Kahneman 1992, for example), a focal-price equilibrium exists for cost variation $\overline{c} - \underline{c}$ up to t/n. Since by Proposition 3.2 the equilibrium markup lies in the interval $[t/n, 2t/n]$, the allowed cost variation is between 50 percent and 100 percent of firms' markups.

3.4 Conditions for All Equilibria to be Focal

In this section, we identify sufficient conditions under which firms with possibly different cost distributions suppress cost shocks and adhere to focal pricing of differentiated products in *any* interior market equilibrium. To our knowledge, no price-setting model predicts focal prices so robustly. We first establish that if the intervals containing the supports of firms' cost distributions overlap, there cannot be an equilibrium with stable but different prices—if each firm sets a deterministic price, they set the same one. Then, we show that if the density of each firm's cost distribution is sufficiently large everywhere on its connected support, prices are stable—each firm sets a deterministic price. Then, when both conditions hold, any market equilibrium is a focal-price equilibrium. We also give examples illustrating that if firms' cost distributions do not overlap, equilibria with different deterministic prices can exist.

The following proposition is the first part of our argument:

Proposition 3.3. *Suppose* $\cap_{i\in N}[\underline{c}_i, \overline{c}_i] \neq \varnothing$. *If all firms set a deterministic price and either*

$$\lambda \leq 1 + \frac{2}{n-1}\left(1 + \sqrt{1 + 2n(n-1)}\right) \tag{3.5}$$

or $n = 2$, *the market equilibrium is a focal-price equilibrium.*

The intuition for Proposition 3.3 is the same as in the two-firm example of Section 3.2: if firms do not charge the same price, a highest-price firm has a higher markup and a lower inframarginal demand than a lowest-price firm, and because by the comparison effect it tends to face a greater responsiveness of demand, either it or the lowest-price firm wants to deviate. This intuition, however, ignores an effect that (for $n > 2$) makes it necessary to impose Condition (3.5). A change in a firm's price changes the distribution of marginal consumers in its two markets. By Lemma 3.1, this typically changes the price responsiveness of its residual demand. If demand responsiveness changed too fast, the firm's profit-maximization problem might not be single-peaked, and this would generate many technical difficulties. To rule out such possibilities, Proposition 3.3 above and Proposition 3.4 below impose restrictions on λ.

But Condition (3.5) is relatively weak. It only applies when $n > 2$, and it is satisfied for any number of firms whenever $\lambda \leq 1 + 2\sqrt{2} \approx 3.8$. Since the conventional assumption of two-to-one loss aversion is equivalent to $\lambda = 3$, the condition does not seem very problematic.

As a second ingredient for the main result of this section, we give conditions such that all firms charge a deterministic price. Because analyzing a

more general model is technically very difficult, we restrict attention to independent (idiosyncratic) cost shocks, still allowing for asymmetries in firms' cost distributions.[11]

Proposition 3.4. *Suppose costs are independently distributed with* $c_i \sim \Theta_i[\underline{c}_i, \overline{c}_i]$ *and corresponding densities* θ_i*. If* $38 > \lambda > 1$ *and* $(\overline{c}-\underline{c}) < (t/n)\cdot(3+\lambda)/(2(1+\lambda))$*, there is a real number* $\rho(\lambda, t, n, \overline{c}-\underline{c}) > 0$ *such that if*

$$\theta_i(c) > \rho(\lambda, t, n, \overline{c}-\underline{c})$$

for all $c \in [\underline{c}_i, \overline{c}_i]$*, then firm* i *sets a deterministic price in any interior equilibrium.*

Combining Propositions 3.3 and 3.4:

Corollary 3.1. *If the conditions of Propositions 3.3 and 3.4 hold, any interior market equilibrium is a focal-price equilibrium.*

It is worth noting that the function $\rho(\lambda, t, n, \overline{c}-\underline{c})$ that naturally drops out of our approximations underlying the proof of Proposition 3.4 is decreasing in t and increasing in n, and approaches zero as $t \to \infty$ and infinity as $n \to \infty$. Our sufficient conditions for all equilibria to be focal are therefore more likely to be met in less competitive industries.

To conclude this section, we provide some examples where the conditions of Proposition 3.3 do not hold but those of Proposition 3.4 may, illustrating the logic of market equilibrium with unequal prices and discussing further issues.

Example 3.1. Suppose $n = 2$, $\lambda = 5$, and $t = 1$. As we verify in the appendix, there is a market equilibrium in which firm 1 always charges price $p_1 = 2$, firm 2 always charges price $p_2 = 9/4$, and the consumer buys from firm 1 with probability 3/4, if and only if $c_1 \in [1/8, 5/4]$ and $c_2 \in [2, 49/24]$.

The above conditions for the existence of a market equilibrium with prices $p_1 = 2$ and $p_2 = 9/4$ allow for several possibilities. If costs are deterministic with $c_1 = 9/8$ and $c_2 = 97/48$, for instance, there is a market equilibrium with deterministic prices $p_1 = 2$ and $p_2 = 9/4$, and by Proposition 3.1 a focal-price equilibrium also exists. The intuition for why both types of equilibria can exist is the following. If consumers had expected the two firms to charge

[11] If costs are not independent, a change in c_i changes the distribution of competitors' prices conditional on c_i and hence also the distribution of marginal consumers for a given p_i. By Lemma 3.1, this typically changes the price responsiveness of residual demand. While we believe this consideration would not substantially modify the comparison effect, the main force driving our result, we cannot formally analyze this more general case.

the same price, demand will be very responsive to increases from this price and not very responsive to decreases from this price, so that it is optimal for both firms to charge this price. But if consumers had expected different prices, the responsiveness of demand in-between the two expected prices is at an intermediate level, so that it is optimal for the low-cost firm to charge the lower of the prices and for the high-cost firm to charge the higher of the prices.

In contrast, if costs are deterministic with $c_1 = 1$ and $c_2 = 97/48$, charging deterministic prices $p_1 = 2$ and $p_2 = 9/4$ is still a market equilibrium, but in this case a focal-price equilibrium does not exist. More generally, if firm 1's and firm 2's marginal costs are independently and narrowly distributed around 1 and 97/48, respectively, Proposition 3.4 implies that each firm charges a deterministic price in any market equilibrium, and Proposition 3.1 implies that these prices are different. Hence, sticky pricing—the unresponsiveness of prices to cost circumstances—does not necessarily go hand in hand with focal pricing—equal pricing across firms. Intuitively, if each individual firm's cost distribution is sufficiently narrow, the firm's price will be invariant to its cost realization. But if one firm is at the same time much more efficient than the other, the deterministic prices of the two firms must be different.

While not generating focal pricing, in some ways the above example still illustrates how loss aversion can lead to reduced price variation and lower competition—two important themes in the paper. It is easy to check that in a standard Salop model, an equilibrium with a price difference of 1/4 requires a cost difference of 3/4. In our example, a cost difference of up to 23/12 can support the same price difference, showing that with loss aversion prices can be much closer to each other. Indeed, unlike in our setting, in the standard model any cost difference above 3/2 would lead the low-cost firm to price the high-cost firm out of the market. Loss aversion therefore reduces competition and allows both firms to make positive profits.

Although our example does not speak directly to situations with more than two goods, its logic also suggests that in many situations uniform pricing is more likely to happen than focal pricing. If a single firm's cost distributions for its different products are narrow and overlapping, the firm will often set the same deterministic price for all its products. But again, if one firm has much lower costs overall than the other, the uniform prices of the two firms will have to differ.

3.5 Industry-Wide Cost Shocks

In this section, we fully characterize symmetric equilibria when firms always have identical marginal costs—that is, when they are subject only to industry-wide cost shocks. This allows us to study, in a tractable model, the responsiveness of price to cost when conditions for a focal-price equilibrium are not necessarily met. We find that markups strictly decrease with cost in any market equilibrium, and that the price may be sticky (unchanging in cost) in some regions. Furthermore, markups decline faster with cost, and prices tend to be more sticky, in more concentrated industries.

Suppose firms' common marginal cost is continuously distributed according to Θ on $[\underline{c}, \overline{c}]$, with corresponding density θ. We first establish two basic properties of symmetric market equilibria:

Lemma 3.2. *Suppose firms have identical, continuously distributed marginal costs. In a symmetric market equilibrium, price is a continuous and non-decreasing function of marginal cost.*

To understand the lemma, take costs c and c' and corresponding equilibrium prices p and $p' > p$, and suppose that residual demand is differentiable at both p and p'. Because firms use symmetric strategies, inframarginal demand is the same at the two prices (and equal to $1/n$). In addition, due to the comparison effect, demand is (weakly) more responsive to unilateral price changes at the high price p' than at the low price p. In order for firms' first-order conditions to be satisfied at both costs, therefore, c' must be greater than c and not arbitrarily close to it.[12]

We now fully characterize the set of symmetric-equilibrium pricing functions $P : [\underline{c}, \overline{c}] \to \mathbb{R}$, and then turn to a detailed discussion of the implications of this characterization. As a step toward a full analysis, we posit that for a cost c, $P(c)$ is not an atom of the market price distribution, and derive $P(c)$. Since in a symmetric equilibrium firms set identical prices in all states of the world, consumers always choose the product closest to their taste. Hence, as in Section 3.3, $G(\cdot)$ is the uniform distribution on the interval $[0, 1/(2n)]$. Furthermore, Equation (3.3) implies that the derivative of firm 1's residual demand exists at $P(c)$ and is equal to

$$-\frac{1}{t} \cdot \frac{2 + (\lambda - 1)F(P(c))}{1 + \lambda} = -\frac{1}{t} \cdot \frac{2 + (\lambda - 1)\Theta(c)}{1 + \lambda}, \tag{3.6}$$

[12] If the price distribution has atoms at p or p', so that residual demand is not differentiable, the same argument still works by considering—instead of first-order conditions—incentives to lower one's price from p' as compared to incentives to raise one's price from p.

where $F(P(c)) = \Theta(c)$ because $P(\cdot)$ is non-decreasing and $P(c)$ is not a pricing atom. Substituting Equation (3.6) into the firm's first-order condition, using that $D_1(P(c), P(c)) = 1/n$, and rearranging yields

$$P(c) = c + \frac{t}{n} \cdot \frac{2 + (\lambda - 1)}{2 + (\lambda - 1)\Theta(c)} \equiv \Phi(c). \tag{3.7}$$

Expression (3.7) and Lemma 3.2 impose strong restrictions on a symmetric-market-equilibrium pricing function. For any $c \in [\underline{c}, \overline{c}]$ that is not on a flat part of $P(\cdot)$, $P(c)$ is not a pricing atom, so $P(c) = \Phi(c)$. In addition, arbitrarily close to an interior end of a flat part there are costs c for which $P(c)$ is not a pricing atom, where again $P(c) = \Phi(c)$. Hence, at interior ends a flat part of $P(\cdot)$ connects continuously to $\Phi(\cdot)$. Finally, because for $c = \underline{c}$ Equation (3.6) is the left derivative of demand whether or not $\underline{c}$ is a pricing atom, for price decreases from $\underline{c}$ to be unprofitable we must have $P(\underline{c}) \leq \Phi(\underline{c})$; and by a similar argument, $P(\overline{c}) \geq \Phi(\overline{c})$.

The above conditions are in fact not only necessary, but also sufficient for $P(\cdot)$ to be a symmetric-market-equilibrium pricing function:

Proposition 3.5. *Suppose firms have identical marginal costs distributed according to Θ on $[\underline{c}, \overline{c}]$. A pricing function $P : [\underline{c}, \overline{c}] \to \mathbb{R}$ is a symmetric-market-equilibrium pricing function if and only if all of the following are satisfied:*

1. $P(\cdot)$ is continuous and non-decreasing.

2. There are disjoint intervals $[f_1, f_1'], [f_2, f_2'], \cdots \subset [\underline{c}, \overline{c}]$ such that $P(\cdot)$ is constant on all $[f_i, f_i']$ and not constant on any interval not contained in any $[f_i, f_i']$.

3. $P(c) = \Phi(c)$ for any $c \notin \cup_i [f_i, f_i']$.

4. $P(\underline{c}) \leq \Phi(\underline{c})$ and $P(\overline{c}) \geq \Phi(\overline{c})$.

To start identifying the implications of Proposition 3.5 in specific cases, suppose that $\Phi(\cdot)$ is strictly increasing. In that case, $P(\cdot)$ cannot have a flat part: because $P(\underline{c}) \leq \Phi(\underline{c})$ and $P(\overline{c}) \geq \Phi(\overline{c})$, a flat part cannot start at either of these points and connect continuously to $\Phi(\cdot)$; and an interior flat part cannot connect continuously to $\Phi(\cdot)$ at both ends. Hence, there are no pricing atoms, and the unique symmetric market equilibrium has $P(c) = \Phi(c)$ everywhere:

Corollary 3.2. *Under the conditions of Proposition 3.5, if $\Phi(c)$ is strictly increasing, the unique symmetric market equilibrium has pricing strategies $P(c) = \Phi(c)$. Otherwise, a symmetric equilibrium with strictly increasing pricing strategies does not exist.*

But $\Phi(\cdot)$ is not necessarily strictly increasing. Differentiating Equation (3.7) with respect to c,

$$\Phi'(c) = 1 - \frac{t}{n} \cdot \frac{(1+\lambda)(\lambda-1)\theta(c)}{[2+(\lambda-1)\Theta(c)]^2}, \tag{3.8}$$

which is negative if $\theta(c)$ is very high. If $\Phi(\cdot)$ is non-increasing, then $P(\cdot)$ cannot have a strictly increasing part—where it would have to coincide with a non-increasing $\Phi(\cdot)$—so that it is constant. Hence, in these situations any symmetric market equilibrium is focal:

Corollary 3.3. *Under the conditions of Proposition 3.5, if $\Phi(c)$ is non-increasing, any symmetric market equilibrium is a focal-price equilibrium. Otherwise, symmetric equilibria other than focal-price equilibria exist.*

As with Proposition 3.4, the intuition for this result is most easily seen by first assuming that consumers expected firms' prices to be strictly increasing in cost. If the density of the cost distribution is high, a small increase in c implies a large increase in $F(P(c))$ and hence a large increase in the comparison effect and the corresponding price responsiveness of demand. This is inconsistent with equilibrium: a firm can increase profits either by decreasing prices at higher costs and attracting substantial extra demand, or by increasing prices at lower costs without losing many consumers. Since this is true for any strictly increasing pricing strategy, the equilibrium price must be constant.

If $\Phi(\cdot)$ is neither strictly increasing nor non-increasing, Proposition 3.5 implies that market-equilibrium pricing functions will generally consist of flat parts pasted together continuously with strictly increasing parts that coincide with $\Phi(\cdot)$. Figure 3.3 illustrates a non-monotonic $\Phi(\cdot)$ and possible market equilibria. For $c \in [\underline{c}, c']$ and $c \in [c'', \overline{c}]$ the pricing function cannot have a flat part, because that could not be pasted continuously with $\Phi(\cdot)$. Hence, in these regions $P(\cdot)$ is strictly increasing and therefore equal to $\Phi(\cdot)$. The non-decreasing $P(\cdot)$, however, must be "ironed out" over the range where $\Phi(\cdot)$ is decreasing. Furthermore, because at the ends of a flat interval $P(\cdot)$ connects continuously to increasing parts of $\Phi(\cdot)$, it has exactly one flat part. $P^1(\cdot)$ and $P^2(\cdot)$ are two possible market-equilibrium pricing functions.

In combination with Equation (3.7), Proposition 3.5 has a number of important implications for symmetric equilibria. Two implications are about the level and variation in markups in our model relative to the standard one (identical to $\lambda = 1$ here). In the standard Salop model, the markup is constant in cost and equal to t/n. As in focal-price equilibria (Proposition 3.2), one effect of loss aversion is to increase the price level: the markup is strictly greater than

t/n for $c < \overline{c}$, and greater than or equal to t/n for $c = \overline{c}$. The consumers that a firm attracts by lowering its price experience a pure loss in product satisfaction (from choosing a product unexpectedly far from ideal), and unless $c = \overline{c}$, only some combination of gain and avoided loss in money. Hence, they are more difficult to attract than in the standard setting, decreasing competition and increasing prices.

The other effect of loss aversion is to decrease price variation by making markups strictly decreasing in c:

Corollary 3.4. *Under the conditions of Proposition 3.5, in any symmetric market equilibrium $P(c) - c$ is strictly decreasing in c on the support of Θ.*

This prediction of our theory is potentially relevant for understanding macroeconomic fluctuations. Extensive evidence reviewed by Rotemberg and Woodford (1999) indicates that costs are strongly procyclical. Hence, our model implies markups are countercyclical.[13] Intuitively, recall that due to the comparison effect, consumers are more responsive to price changes at higher than at lower prices within the price distribution. Since inframarginal demand is constant across the price distribution, this means that firms compete more fiercely at higher prices, reducing markups.

Proposition 3.5 implies not only that price variation is lower than in the standard model, but also that it is systematically related to the competitiveness of the market. The more concentrated is the industry (the lower is n) and the greater is product differentiation (the greater is t), the lower is $\Phi'(c)$ at any c (Equation (3.8)). As a result, the more countercyclical are markups—the faster $P(c) - c$ decreases with c—when price is strictly increasing in cost, and the more likely it is that any symmetric equilibrium is a focal-price one. Intuitively, with the higher average markups firms enjoy in a less competitive industry, the increased ability to attract consumers at higher prices has a greater impact on firms' incentive to cut prices, generating markups that decrease faster in cost. If markups are very high, the impact of an increase in demand responsiveness on firms' incentive to cut prices is so great that firms are unwilling to raise their price at all—they charge a sticky (and focal) price.

In fact, Proposition 3.5 allows us to more fully describe pricing patterns for industries ranging from very competitive ($t/n \approx 0$) to very uncompetitive ($t/n \to \infty$). If competition is sufficiently strong, the unique symmetric market equilibrium features a strictly increasing pricing function, which is close to marginal-cost pricing if competition is very strong. At lower levels of competition, markups are higher and more countercyclical. At even lower levels of

[13] Of course, if one measures countercyclicality using the Lerner index $(p-c)/p$, the Salop model without loss aversion also features countercyclical markups.

competition, the price is constant in cost near regions where the cost distribution is relatively dense, but may remain strictly increasing in cost in other regions. At very low levels of competition, the price is sticky and focal.

It is important to note that in this section identical pricing across firms was assumed, not derived. The question arises whether such identical pricing would be the outcome in an environment where idiosyncratic cost shocks also exist. For cases in which the cost variation is sufficiently small, Section 3.3 has shown that a focal-price equilibrium exists even when there is both industry-wide and idiosyncratic cost uncertainty. More strongly, although (for reasons mentioned above) we cannot fully analyze a general model with both types of cost uncertainty, intuition developed in the last two sections suggests that in regions where both components vary little, the price will be focal in any equilibrium.[14] But in regions where the common cost shocks are not absorbed, residual demand will be smooth, so idiosyncratic cost shocks will also not be absorbed.

3.6 Robustness

In this section, we argue that our results are largely robust to natural variations of our model. In short, most of our results rely on the simple intuition that—due to loss aversion in money—a consumer's sensitivity to price changes is increasing in the probability with which she expected to pay lower prices, and this force is not eliminated by reasonable modifications of the model.

Because in many situations consumers are unsure either about what they want or about what is available, we have assumed a dispersed prior on χ. But most of our results do not depend on this assumption. Even if χ is known perfectly, our results in Sections 3.3 and 3.5 remain unchanged, with the same logic and essentially the same proofs.[15] Under reasonable refinements, such as that consumers play the ex-ante optimal personal equilibrium, a version of Proposition 3.3 also remains true—although for a completely different reason than above.[16] To illustrate, suppose $n = 2$ and firms 1 and 2 charge prices p_1

[14] Again, the intuition can be seen by assuming that cost shocks are not absorbed. If costs vary little, the price distribution will be dense, so that the comparison effect implies that a small increase in the price leads to a large increase in the price responsiveness of demand. Then, a firm can increase profits either by increasing lower prices (where demand is relatively inelastic) or by decreasing higher prices (where demand is relatively elastic).

[15] Technically, if different consumers have different information about their preferred locations, there cannot be a representative consumer. The definition of market equilibrium can be modified easily to account for such heterogeneity.

[16] In fact, the assumption that consumers play the ex-ante optimal personal equilibrium

and $p_2 > p_1$. If a consumer prefers product 1 ex ante and plans to buy it, to avoid a loss in money she prefers it even more ex post. If she prefers product 2 ex ante and plans to buy it, to avoid a loss in satisfaction she prefers it even more ex post. With all consumers "locked in," both firms want to raise their price, contradicting equilibrium.

Our methods in Proposition 3.4 do not extend to the case when χ is known with certainty. The logic of our proof, however, only seems to rely on sufficiently many *marginal* consumers being sufficiently uncertain about their relative preference for at least two neighboring products that they are unsure as to which one they will buy. These consumers exhibit a similar pattern of behavior to our consumers above, so they give a firm similar incentives.

As do most applications of the Salop model, our model assumes that firms' prices affect only the allocation of demand, not its level. One way to model a market-size effect is to assume that consumers have an outside option with a randomly determined level of utility. In this case, the comparison effect makes consumers on the margin between two firms, as well as on the margin between a firm and the outside option, more responsive to price changes at higher than at lower prices in the purchase-price distribution. Hence, our qualitative results on reduced price variation (but not necessarily our result that loss aversion increases prices) are likely to survive.

We assume above that each firm sells exactly one product. As long as no firm owns neighboring products, our results carry over unchanged to multiproduct firms, so that results on focal pricing above translate directly into uniform pricing.[17] In an interior equilibrium, the incentive for locally changing one product's price is unaffected by how many non-neighboring products a firm owns. But global deviations are weakly less profitable for a multiproduct firm because such a firm might be cannibalizing its own market.[18]

In our model, consumer demand is stable and symmetric across firms, but firms have possibly different and uncertain marginal costs. In most industries,

is much stronger than we need. It suffices to assume, for example, that consumers have an arbitrarily small amount of self-discipline in the sense that they can select ex-ante whether to impose an arbitrarily small ex-post cost on themselves if they choose a certain action deemed undesirable ex ante.

[17] In this case, the relevant measure of the competitiveness of the industry depends (in addition to t) on the number of products rather than on the number of firms.

[18] Even if firms can own two neighboring products, all the forces behind our results are still present, so that focal pricing will often be an equilibrium, and often the only type of equilibrium. Our conditions and proofs, however, would have to account for the decreased competition between firms and for the fact that products may differ in how many neighboring products they compete with. When a firm can own three neighboring products, the middle one faces no immediate external competition, so the firm always sets a higher price for it.

firms do differ in the features and popularity of their products, leading to different elasticities of residual demand even if all of them set the same price. In our setting, for instance, the intrinsic valuations of firms' products could be different and random instead of always taking the same value v. Because a change in its product's intrinsic value has similar implications for a firm's pricing incentives as a change in its marginal cost, results closely analogous to those above would likely hold in this alternative model.

While we have assumed that industry structure is exogenous, our model can be extended to allow for endogenous entry. Suppose industry concentration is determined by a fixed cost that firms must pay to enter the market, and post-entry products are located equidistant from each other. Since the fixed cost determines the number of firms but has no impact on market equilibria given the number of firms, our qualitative results on the effect of industry concentration on market equilibria survive.

Our results on focal pricing and reduced price variation more generally also hold in a model in which consumers are loss averse only in money. This assumption would, in fact, substantially simplify some of our formal statements and proofs (especially those of Propositions 3.3 and 3.4). We make the assumption that consumers are loss averse also in the product dimension both because it is far more realistic and experimentally and theoretically well-motivated, and because it shows the robustness of results to including loss aversion in things other than money.

Our definition of market equilibrium assumes that all consumers play the same personal equilibrium. Relaxing this assumption does not affect our results. In all situations in Sections 3.3 and 3.5, selection is a non-issue simply because the personal equilibrium is unique. Our proofs in Section 3.4 work by investigating how the responsiveness of a firm's residual demand changes across the price distribution. Since our bounds hold for any personal equilibrium a person might be playing, they also hold if consumers play different equilibria.

The results in this paper are also robust to heterogeneity in loss aversion among consumers. Our estimation methods would have to account for such heterogeneity, but as long as there is some loss aversion in the population, the results would survive in some form.

A more fundamentally different extension of our model than all the considerations above is the incorporation of dynamics, and we conclude this section by intuitively discussing how this may affect our results. Suppose the firms play the pricing game T times with costs independently redrawn in every period, and consumers' reference points depend on lagged rational expectations regarding the distribution of outcomes they are going to get in each period.

Clearly, since for a range of costs firms cannot increase even current profits by deviating from focal pricing in any single period, our existence results easily extend to this dynamic setting. But whether focal and sticky pricing emerges as the *unique* possible outcome is far trickier. A major complication is that past prices can in general affect a consumer's expectations and hence also her reference point for future outcomes. While a full and realistic analysis of this issue seems important—and could potentially lead to novel models of advertising and price leadership—it is beyond the scope of this paper.[19] Hence, to abstract from this consideration, we assume that the expectations determining a consumer's reference points are formed before she observes any prices.

With this assumption, the same logic that underlies our results in this paper seems to imply that if firms' cost distributions overlap and have sufficiently high density, then in each period prices will be focal. Because there can be a continuum of static focal-price equilibria, however, it is not necessarily the case that firms set the same price from period to period. Whether this is guaranteed depends on how exactly lagged expectations determine consumers' reference point. At one extreme, suppose that consumers compare their outcomes in a period only to their lagged expectations specific to that period, possibly because new consumers arrive in each period. Then, any sequence of static focal prices is an equilibrium; it could be that consumers expect the price of a CD to be \$15 in one week and \$20 in the next, and firms comply with both of these expectations. At the other extreme, suppose that a consumer forms expectations regarding all T periods, and her reference point for outcomes in each period is an average of these expectations. Then, the logic of our results seems to indicate that the unique equilibrium is to charge the same focal price in each period: just like it cannot be an equilibrium for firms to charge different deterministic prices, it cannot be an equilibrium for them to charge different prices in different periods. More generally, if consumers' average reference point changes slowly—for example because a small fraction of consumers is replaced in each period—there is a tendency for firms not to move away from prices set in previous periods.

3.7 Related Literature

Loss aversion features prominently in at least two somewhat separate literatures. In experimental and behavioral economics, loss aversion and reference

[19] Nevertheless, there is a way in which observing past prices might increase the tendency for sticky pricing: if consumers take past prices as salient indicators of the future play of firms, firms may have a strong incentive to comply with these expectations.

dependence explain a number of robust phenomena, including the endowment effect and small-scale risk aversion mentioned in Chapter 1. More closely related to our topic, empirical evidence in marketing indicates loss aversion in consumer behavior that is broadly consistent with the consumer model of this paper. Consumers seem to compare actual market prices to "reference prices" determined at least partly by "price beliefs" or expectations, and purchases are more sensitive to losses from the reference price than to gains relative to it (Erickson and Johansson 1985, Kalwani and Yim 1992, Winer 1986). Hardie, Johnson and Fader (1993) find loss aversion in evaluations of quality as well. Our paper develops a model of reference prices based on insights from behavioral economics, and—asking a question not formally addressed in either literature—examines ways this impacts the strategic interactions between firms.

There are several prominent theories of price stickiness, some of which also feature focal-price equilibria. The stickiness of prices is a robust feature of these theories, but we will argue that their aim is not to explain the *equality* of prices: if they are extended to allow for asymmetric firms and differentiated products, they either become inconsistent with focal pricing, or predict a large number of equilibria with no compelling reason to select the focal one. Furthermore, these theories do not ask how the competitiveness of the industry affects price variation, and do not address the issue of uniform pricing.

Perhaps the most commonly invoked theory of price stickiness is that of menu costs. Menu costs generate a disincentive to change prices, but not an incentive to set identical prices. Furthermore, in some situations prices tend to be sticky even though menu costs seem to be zero.[20]

Formalizing the casual view of many researchers and observers that price stickiness and focal pricing are due to collusive behavior,[21] Athey, Bagwell and Sanchirico (2004) show that in a repeated price-setting game, the optimal symmetric equilibrium is often a focal-price equilibrium. This equilibrium is enforced by the threat of price war in case of a price change, and is efficient because the price war is never triggered. In asymmetric environments, however, there is no reason for firms' sticky prices to coincide. Similarly, if each

[20] For example, Kashyap (1995) finds sticky pricing in retail catalogues even when new catalogues are printed anyhow. Genesove (2003) documents substantial rigidity in apartment rents, even though a new lease is filled out and signed every year for most apartments in his sample.

[21] This view is expressed, for instance, in Carlton (1989, pages 914-915) and Knittel and Stango (2003, pages 1704-1705). In addition, focal prices and reduced price variability seem to have raised suspicions of collusion in other cases, such as the recent Sony-BMG merger case in Europe.

colluding oligopolist sells multiple differentiated products, there is no reason to set the same price for all those products.

Rotemberg (2011) develops a monopoly model in which consumers both dislike price changes and are willing to punish the firm if they perceive it to be insufficiently altruistic. Even if selfish, the firm pretends to be sufficiently altruistic by setting the highest acceptable price. The model predicts that observable increases in input costs lead to price increases but increases in demand may not. While this captures an important aspect of price dynamics our model misses, Rotemberg's single-product monopoly setup cannot address focal or uniform pricing.[22]

An important class of models with implications for price variation assumes that consumers must pay search costs to sample firms' products and prices. These models, however, often generate excess rather than reduced price variation. If search costs are bounded away from zero and the first search is costly, there is no focal-price equilibrium even with deterministic identical costs: if a consumer expecting price p shows up at a firm, the firm knows she values the good above p, and can raise the price. If search costs are not bounded away from zero, the situation is more complicated. If consumers observe the price distribution, a price increase by a firm triggers search by some consumers arriving at the firm, and a price decrease triggers search by some consumers arriving at other firms. Stiglitz (1987) shows that as a result of these opposing forces, price stickiness obtains if search costs are convex in the number of searches, but excess price variation results if search costs are concave. Finally, if consumers do not observe the price distribution, a price increase triggers search by consumers arriving at the firm, but (because it is not observed) a price decrease does not trigger search by consumers arriving at other firms. This can lead to price stickiness.[23] But even in this case, equilibria with non-equal prices exist, and with asymmetry there is no compelling reason to select a focal-price one.

Our model is related to an older literature on kinked demand curves (Hall and Hitch 1939, Sweezy 1939). In these models, each firm believes that rivals will follow price decreases but not price increases—leading to a kinked demand curve. Maskin and Tirole (1988) provide a game-theoretic foundation

[22] Rotemberg (2005, 2010) investigates the implications of Rotemberg's fair-pricing model for monetary policy and the frequency and size of price changes.

[23] Nevertheless, for reasons similar to the logic in Varian (1980), if there is a mass of informed consumers—who find out all prices for some reason—a focal-price equilibrium can once again not exist. With search costs, any equilibrium must have positive expected profits. Then, if all firms were to charge the same price, undercutting other firms slightly would attract all informed consumers, increasing profits.

for these beliefs, but do not investigate the impact of cost shocks on pricing behavior. In addition, once we drop their assumption that all consumers buy from a lowest-price firm, there is no reason to presume that equilibria would necessarily be focal. More distantly, our paper is also related to the literature on switching costs, whereby consumers face an exogenously given cost when buying a product different from the one they purchased previously.[24] Consumers in our model face a kind of "psychological" cost when switching away from their expected outcomes: they dislike trading an unanticipated loss in one dimension for an unanticipated gain in the other dimension. Both kinds of switching costs predict increased prices. But whereas in typical switching-cost models the profits from increased prices are competed away in earlier periods when firms fight for unattached consumers, the same is not true in our model. More importantly, in our model the size of the switching cost is endogenous and situation-dependent. In particular, the key feature of our model is that consumers are less reluctant to switch in response to a price change if they construe it as more of a change in a loss rather than a gain. Since this feature is not present in classical switching-cost models, these models do not predict reduced price variation.

3.8 Conclusion

A basic premise of our model is that consumers have accurate expectations about prices. Marketing studies disagree whether consumers can even recall prices for recently purchased products accurately, with estimates ranging from 5% to 50% of consumers.[25] The surveys upon which these estimates are based typically focus on the knowledge of average consumers and do not study the accuracy of price expectations. While for simplicity we have assumed that all consumers have correct price expectations, our effects are driven by marginal consumers—consumers who will switch in response to some relevant price changes—and so require only (some of) these consumers to have correct expectations. Unfortunately, we are not aware of any empirical work on whether they do.

In contrast to our results, which predict reduced price variation in a number of senses, there often seems to be excess price variation between even identical products.[26] Such price variation seems to occur primarily in industries where

[24] For a recent survey, see Farrell and Klemperer (2007).

[25] For a recent meta-study, which includes an extensive literature survey, see Estelami and Lehmann (2001).

[26] Baye, Morgan and Scholten (2004), for example, document that for many products,

consumers cannot or do not compare prices across different sellers, partly because firms deliberately make comparisons difficult.[27] Our theory clearly applies better when prices and relevant features of products are transparent to consumers.

Appendix: Proofs

Proof of Lemma 3.1. To derive our expression, we solve for how the location of the indifferent consumer between firms 1 and 2, which is equal to firm 1's demand on the interval $[0, 1/n]$, changes with p_1 for a given p_2. Suppose a consumer's realized taste is $\chi \in [0, 1/n]$. The consumer's utility from buying good 1 at price p_1 is then

$$u_1 = v - \chi t - p_1 - \lambda \int_0^{p_1} (p_1 - p)\ dF(p) + \int_{p_1}^{\infty} (p - p_1)\ dF(p)$$
$$- \lambda t \int_0^{\chi} (\chi - s)\ dG(s) + t \int_{\chi}^{1/n} (s - \chi)\ dG(s). \tag{3.9}$$

Replacing p_1 with p_2 and χ with $1/n - \chi$ in Equation (3.9), we get utility from buying good 2 at price p_2:

$$u_2 = v - ((1/n) - \chi)t - p_2 - \lambda \int_0^{p_2} (p_2 - p)\ dF(p) + \int_{p_2}^{\infty} (p - p_2)\ dF(p)$$
$$- \lambda t \int_0^{1/n-\chi} ((1/n - \chi) - s)\ dG(s) + t \int_{1/n-\chi}^{1/n} (s - (1/n - \chi))\ dG(s). \tag{3.10}$$

Equations (3.9) and (3.10) are differentiable with respect to χ, and right and left differentiable with respect to p_1. Using this together with the fact that

different internet-based retailers charge very different prices.

[27] Ellison (2005) and Gabaix and Laibson (2006) identify competitive advantages of making price comparisons difficult. Gabaix and Laibson (2006) show that in the presence of some consumers who ignore add-on costs, firms have no incentive to make add-on prices transparent, even when it is very cheap to do so. Ellison (2005) gives natural conditions under which rational consumers who are unresponsive to an ex-ante hidden add-on price also decrease competition in transparent aspects of the product.

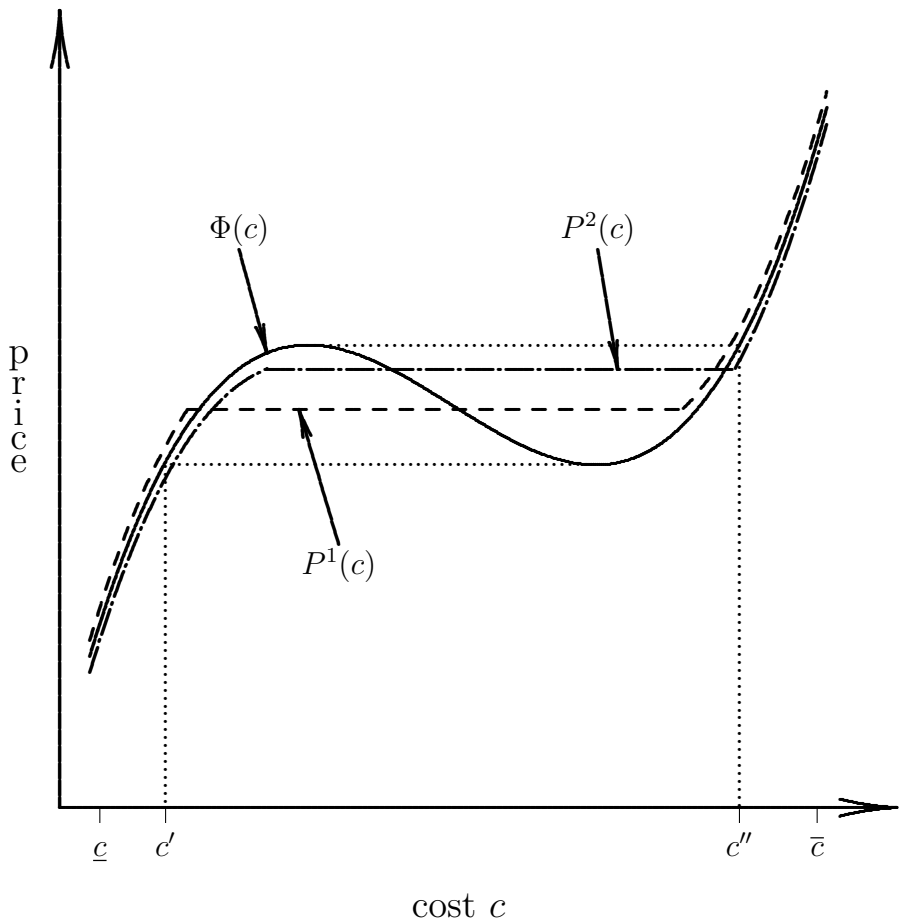

Figure 3.3: Determination of Equilibrium Pricing Functions with Industry-Wide Shocks

For better visibility, overlapping curves are drawn as close parallel curves. The solid curve $\Phi(c)$ is the solution to the first-order condition for optimal pricing assuming that $P(c)$ is not an atom of the price distribution. Since a market-equilibrium pricing function is non-decreasing and continuous, it consists of constant parts pasted together with strictly increasing parts that coincide with $\Phi(\cdot)$. Two market-equilibrium pricing functions are $P^1(c)$ and $P^2(c)$.

$u_1 = u_2$ for the indifferent consumer x^+ implies that

$$\left(\frac{dx^+}{dp_1}\right)_\downarrow = -\frac{\left(\frac{\partial u_1}{\partial p_1}\right)_\downarrow - \left(\frac{\partial u_2}{\partial p_1}\right)_\downarrow}{\frac{\partial u_1}{\partial \chi} - \frac{\partial u_2}{\partial \chi}} = -\frac{1}{2t} \cdot \left[\frac{2 + (\lambda - 1)F(p_1)}{2 + \frac{\lambda-1}{2}[G(x^+) + G(1/n - x^+)]}\right], \tag{3.11}$$

and that $(dx^*/dp_1)_\uparrow$ is given by the expression in which $F_\uparrow(p_1)$ replaces $F(p_1)$ above. Similar calculations give the responsiveness of demand on the other side of the firm. □

Proof of Proposition 3.1. If the condition in the proposition is satisfied, then there is a p^* satisfying $p^* - \frac{t}{n}\frac{1+\lambda}{2} \leq c \leq p^* - \frac{t}{n}$ for all $c \in [\underline{c}, \overline{c}]$. That this is a necessary and sufficient condition for local deviations to be unprofitable has been established in the text.

We now show that under the above condition, non-local deviations are also unprofitable. We start with increases in the price. First, note that the firm will never charge a price so high that it would be charging itself out of one market: if a deviating firm is charging itself out of one market, it is charging itself out of both, earning zero profits. Therefore, we only need to consider deviations for which $x \in \left(0, \frac{1}{2n}\right)$. Recall Equation 3.11:

$$\frac{dx}{dp_1} = -\frac{1}{2t}\left[\frac{2 + (\lambda - 1)F(p_1)}{\left\{2 + \frac{\lambda-1}{2}[G(x) + G((1/n) - x)]\right\}}\right]$$

Since $F(p_1) = G\left(\frac{1}{n} - x\right) = 1$ and $G(x)$ is increasing in x, in the range $x \in \left(0, \frac{1}{2n}\right)$, firm 1's demand (as a function of p_1) is concave. This implies that if local deviations are unprofitable, non-local increases in the price are also unprofitable.

Next, we rule out the possibility that firm 1 might like to charge a price so that $x \in \left[\frac{1}{2n}, \frac{1}{n}\right]$. In that case, Equations (3.9) and (3.10) imply that

$$\begin{aligned} &- xt - p_1 + (p^* - p_1) - \lambda t\left(x - \frac{1}{4n}\right) \\ &= -\left(\frac{1}{n} - x\right)t - p^* - \lambda t \cdot 2n \cdot \left(\frac{1}{n} - x\right)\frac{\frac{1}{n} - x}{2} + t \cdot 2n \cdot \left(x - \frac{1}{2n}\right)\frac{x - \frac{1}{2n}}{2}. \end{aligned}$$

Solving for p_1 gives

$$p_1 = p^* - \frac{1}{2}t\underbrace{\left[(\lambda + 1)\left(2x - \frac{1}{n}\right) + (\lambda - 1)\left(x - \frac{1}{4n} - nx^2\right)\right]}_{\equiv \kappa}.$$

To show that lowering the price to p_1 is not a profitable deviation, it is equivalent to show that

$$\frac{1}{n}(p^* - c) \geq 2x(p_1 - c) = 2x\left(p^* - c - \frac{1}{2}t\kappa\right).$$

Rearranging and using that $p^* - c \leq \frac{t(1+\lambda)}{2n}$ gives that it is sufficient to show that

$$\left(2x - \frac{1}{n}\right)\frac{1+\lambda}{n} \leq 2x\kappa, \tag{3.12}$$

or equivalently

$$\begin{aligned}(\lambda+1)\left(2x - \frac{1}{n}\right)^2 &\geq (\lambda-1)2x\left(nx^2 + \frac{1}{4n} - x\right)\\ &= (\lambda-1)2x\left(2x - \frac{1}{n}\right)\left(\frac{nx}{2} - \frac{1}{4}\right).\end{aligned}$$

This simplifies to

$$(\lambda+1)\left(2x - \frac{1}{n}\right) \geq (\lambda-1)2x\left(\frac{nx}{2} - \frac{1}{4}\right).$$

Notice that in the above inequality, the left-hand side is equal to the right-hand side for $x = \frac{1}{2n}$ and greater for $x = \frac{1}{n}$. Furthermore, the left-hand side is linear, while the right-hand side is quadratic and convex. This implies that the left-hand side is no less for all $\frac{1}{2n} \leq x \leq \frac{1}{n}$.

For $n > 2$, we are left to rule out that firm 1 undercuts its rival and steals more than the entire adjacent market. We begin by ruling out deviations in which the firm captures less than two adjacent markets on each side. Let p_1' be the price at which the consumer located at $\frac{1}{n}$ is indifferent between buying from firm 1 and buying from firm 2. This consumers utility of buying from firm 1 is

$$v - \frac{1}{n}t - p_1' + (p^* - p_1') - \lambda t\left[\frac{1}{n} - \frac{1}{4n}\right].$$

In case he buys from firm 2, her utility is

$$v - p^* + t\frac{1}{4n}.$$

Thus, if the consumer is indifferent

$$p^* - p_1' = \frac{t}{2n}\left[2 + \frac{3}{4}(\lambda - 1)\right].$$

Consider the maximum price at which a local deviation is unprofitable; for this price $p^* - \underline{c} = \frac{t}{2n}[2 + \lambda - 1]$ and in this case $p_1' - \underline{c} = \frac{t}{2n}\left[\frac{1}{4}(\lambda - 1)\right]$. Thus even if firm 1 would get the entire two adjacent markets when setting p_1', this is unprofitable as $\frac{1}{n}(p^* - \underline{c}) > \frac{4}{n}(p_1 - \underline{c})$.[28] Obviously undercutting is (weakly) less profitable for any lower focal price or any higher level of marginal cost.

We are left to consider the case in which $n > 4$, and firm 1 steals more than two adjacent markets on each side. We show that this is unprofitable because it requires firm 1 to price below marginal cost. For the consumer located at $\frac{2}{n}$ to weakly prefer buying from firm 1 rather than firm 3, it must be that

$$v - \frac{2}{n}t - p_1 + (p^* - p_1) - \lambda t\left[\frac{2}{n} - \frac{1}{4n}\right] \geq v - p^* + t\frac{1}{4n}.$$

Hence, in this case $p^* - p_1 \geq \frac{t}{2n}[4 + (\lambda - 1)\frac{7}{4}] > \frac{t}{2n}[2 + \lambda - 1] \geq p^* - c$, which completes the proof. □

Proof of Proposition 3.2. We have shown in the text that local deviations are unprofitable if and only if

$$p^* - \frac{t}{n} \cdot \frac{1 + \lambda}{2} \leq c_i \leq p^* - \frac{t}{n}$$

for all i and $c_i \in [\underline{c}_i, \overline{c}_i]$. It follows from the proof of Proposition 3.1 that if local deviations are unprofitable, so are global ones.

It remains to show the second part of the proposition. In the standard Salop model, for the consumer x between firms 1 and 2 who is indifferent between the two products,

$$x = \frac{\frac{t}{n} + p_2 - p_1}{2t}.$$

Hence, for realized cost c, firm 1's problem is

$$\max_{p_1} \quad \frac{p_1 - c}{2t} \cdot \left(\frac{2t}{n} - 2p_1 + E[p_2 + p_n|c]\right).$$

This implies that

$$P_1(c) = \frac{t}{2n} + \frac{E[p_2 + p_n|c]}{4} + \frac{c}{2}.$$

[28] Clearly if $n = 3$, the firm cannot attract two adjacent markets on each side, as there are only three local markets. Nevertheless, the upper bound on profitability we use is still valid.

Suppose that the supremum of prices charged by firms 1, 2, and n are $\overline{p}_1$, $\overline{p}_2$, and $\overline{p}_n$, respectively. Suppose without loss of generality that $\overline{p}_1$ is the supremum of market-equilibrium prices of all firms. Then for any $c \in [\underline{c}_1, \overline{c}_1]$,

$$P_1(c) \leq \frac{t}{2n} + \frac{\overline{p}_2 + \overline{p}_n}{4} + \frac{c}{2}. \tag{3.13}$$

Taking the supremum of both sides implies

$$\overline{p}_1 \leq \frac{t}{2n} + \frac{\overline{p}_1 + \overline{p}_1}{4} + \frac{\overline{c}}{2}.$$

Rearranging gives the upper bound in the proposition.

Finally, we show that this upper bound can only be attained at $\overline{c}$. If no firm's price attains $\overline{p}_1$, we are done. Next, suppose that for a price $c < \overline{c}$, $P_1(c) = \overline{p}_1$. By Inequality (3.13), again we are done. □

Proof of Proposition 3.3. Posit a candidate market equilibrium in which all firms set a deterministic price and in which the highest price p_H is strictly greater than the lowest price p_L. We prove that if the condition in the Proposition is satisfied, either (one of) the highest price firm(s) has a strict incentive to lower its price or (one of) the lowest price firm(s) has a strict incentive to raise its price, contradicting equilibrium.

We establish that the marginal profit of lowering the highest price is weakly greater than the marginal profit of raising the lowest price for all given cost realizations c. This is sufficient because it implies that the high-price firm has a strict incentive to lower its price when it has its lowest cost realization, or the low-price firm has a strict incentive to raise its price when it has its highest cost realization (which is higher than the high-price firm's lowest cost realization because the supports of the cost distributions overlap), contradicting equilibrium. Let x_H^+ and x_H^- be one of the highest cost firm's demands on its right and left, respectively. Define x_L^+ and x_L^- similarly. We want to establish that

$$(p_H - c)\Bigg[\underbrace{\frac{1}{2 + \frac{\lambda-1}{2}[G(x_H^+) + G(\frac{1}{n} - x_H^+)]}}_{\equiv \frac{1}{z_H^+}} + \underbrace{\frac{1}{2 + \frac{\lambda-1}{2}[G(x_H^-) + G(\frac{1}{n} - x_H^-)]}}_{\equiv \frac{1}{z_H^-}}\Bigg] \times [2 + F_\uparrow(p_H)(\lambda - 1)]$$

$$\geq (p_L - c)\left[\frac{1}{z_L^+} + \frac{1}{z_L^-}\right] \times [2 + F(p_L)(\lambda - 1)], \tag{3.14}$$

where z_L^+ and z_L^- are defined analogously to z_H^+ and z_H^-. For brevity, let $\eta_H \equiv [2 + F_\uparrow(p_H)(\lambda - 1)]$ and let $\eta_L \equiv [2 + F(p_L)(\lambda - 1)]$.

Notice that either $\left(z_L^+ \frac{1}{z_H^+} + z_L^- \frac{1}{z_H^-}\right) \leq \frac{1}{2}(z_L^+ + z_L^-)\left(\frac{1}{z_H^+} + \frac{1}{z_H^-}\right)$ or $\left(z_L^- \frac{1}{z_H^+} + z_L^+ \frac{1}{z_H^-}\right) \leq \frac{1}{2}(z_L^+ + z_L^-)\left(\frac{1}{z_H^+} + \frac{1}{z_H^-}\right)$. We distinguish two cases depending on whether the former (Case I) or the latter (Case II) holds.

Case I. We rewrite Equation 3.14 as

$$\eta_H \left(z_L^+ \frac{z_L^-}{z_H^+} + z_L^- \frac{z_L^+}{z_H^-}\right) \geq \left(1 - \frac{p_H - p_L}{p_H - c}\right) \eta_L \, (z_L^+ + z_L^-). \tag{3.15}$$

Equation 3.15 is equivalent to

$$\eta_H \left(z_L^+ \left(1 - \frac{z_H^+ - z_L^-}{z_H^+}\right) + z_L^- \left(1 - \frac{z_H^- - z_L^+}{z_H^-}\right)\right) \geq \left(1 - \frac{p_H - p_L}{p_H - c}\right) \eta_L \, (z_L^+ + z_L^-).$$

As $\eta_H > \eta_L$ a sufficient condition for Equation (3.14) to hold is that

$$\eta_H \left(z_L^+ \frac{z_H^+ - z_L^-}{z_H^+} + z_L^- \frac{z_H^- - z_L^+}{z_H^-}\right) \leq \frac{p_H - p_L}{p_H - c} \, \eta_L \, (z_L^+ + z_L^-). \tag{3.16}$$

Using that

$$|z_H^+ - z_L^-| = \frac{\lambda - 1}{2} \left| \left[G(x_H^+) - G(x_L^-)\right] - \left[G\left(\frac{1}{n} - x_L^-\right) - G\left(\frac{1}{n} - x_H^+\right)\right]\right|,$$

that $g(\cdot)$ is bounded by $2n$, and that for all $p < p_H$

$$\left|\frac{dx}{dp}\right|_\downarrow, \left|\frac{dx}{dp}\right|_\uparrow \leq \frac{1}{2t} \cdot \frac{2 + (\lambda - 1)F_\uparrow(p_H)}{2 + \frac{\lambda - 1}{2}},$$

we get that

$$|z_H^+ - z_L^-| \leq \frac{\lambda - 1}{2} 2n \, |x_H^+ - x_L^-| \leq \frac{\lambda - 1}{2} 2n \, (p_H - p_L) \left(\frac{2 + (\lambda - 1)F_\uparrow(p_H)}{2 + \frac{\lambda - 1}{2}} \, \frac{1}{2t}\right),$$

and by a similar logic $|z_H^- - z_L^+|$ has the same upper bound. Combining these with Equation 3.16 implies that it is sufficient to prove

$$\frac{1}{p_H - c} \eta_L (z_L^+ + z_L^-) \geq (\eta_H)^2 \frac{\lambda - 1}{2} 2n \, \frac{1}{2 + \frac{\lambda - 1}{2}} \frac{1}{2t} \left(z_L^+ \frac{1}{z_H^+} + z_L^- \frac{1}{z_H^-}\right).$$

Using that $\left(z_L^+ \frac{1}{z_H^+} + z_L^- \frac{1}{z_H^-}\right) \leq \frac{1}{2}(z_L^+ + z_L^-)\ (\frac{1}{z_H^+} + \frac{1}{z_H^-})$ it is sufficient to prove

$$\frac{1}{p_H - c}\eta_L \geq (\eta_H)^2 \frac{\frac{\lambda-1}{2}}{2 + \frac{\lambda-1}{2}} \frac{n}{2t}\left(\frac{1}{z_H^+} + \frac{1}{z_H^-}\right). \tag{3.17}$$

Since the high-price firm's demand is always less than or equal $\frac{1}{n}$, the fact that it does not want to lower its price implies

$$1 \geq \frac{n}{2t}(p_H - c)\eta_H \left(\frac{1}{z_H^+} + \frac{1}{z_H^-}\right).$$

Hence, a sufficient condition Equation 3.17 to hold is that

$$\eta_L \geq \eta_H \ \frac{\frac{\lambda-1}{2}}{2 + \frac{\lambda-1}{2}}.$$

For $n = 2$, $F(p_L) = F_\uparrow(p_H)$, so the above is satisfied for any $\lambda > 1$. For $n > 2$, using that $F(p_L) \geq 1/n$ and $\eta_H \leq 1 + \lambda$, a sufficient condition for the above inequality to hold is that

$$(4 + \lambda - 1)(2n + \lambda - 1) \geq n(2 + \lambda - 1)(\lambda - 1).$$

Setting $a = \lambda - 1$, this can be rewritten as

$$0 \geq (n-1)a^2 - 4a - 8n.$$

Since this quadratic has one positive and one negative root, if a is positive and

$$a \leq \frac{2}{n-1}\left(1 + \sqrt{1 + 2n(n-1)}\right),$$

the inequality is satisfied. This gives the bound in the proposition.

Case II. In this case, we rewrite Equation 3.14 as

$$\eta_H \left(z_L^- \frac{z_L^+}{z_H^+} + z_L^+ \frac{z_L^-}{z_H^-}\right) \geq \left(1 - \frac{p_H - p_L}{p_H - c}\right) \eta_L\ (z_L^+ + z_L^-).$$

The remaining steps are analogous to Case I and thus omitted. □

Proof of Proposition 3.4. We begin by proving that each firm's pricing function is continuous in cost. This fact follows from the following lemma.

Lemma 3.3. *Consider any interior equilibrium with $\lambda < 38$ and any cost realization c_i of firm i. Consider furthermore the range of prices $p_i \geq c_i$ such that for all equilibrium price vectors p_{-i}, the indifferent consumers $x^+(p_i, p_{-i})$ and $x^-(p_i, p_{-i})$ are located within distance $(0, 1/n)$ of firm i's ideal product. Over this range of prices, firm i's expected profits are single peaked.*

Proof. Since for all price vectors under consideration, there exists indifferent consumers within distance of $1/n$ of firm i, Equation 3.11 implies that profits are differentiable whereever the market price distribution does not have an atom—which is almost everywhere—and continuous. Furthermore, at prices where the profit function is not differentiable, demand has a concave kink, and hence (as long as $p_i \geq c_i$) so do profits.

Suppose by contradiction that the profit function is not single-peaked in the relevant price region. This implies that the profit function must have a trough. At this trough, it obviously cannot have a concave kink, so it is differentiable. To arrive at a contradiction, we prove that if firm 1's first-order condition is satisfied at some price p_1, profits are lower slightly to the right of p_1.

Let the subscript 1 denote partial derivative with respect to firm 1's price of x^+ and x^- respectively. Note that for each $x(p_1, p_{-1}) \in \{x^-(p_1, p_{-1}), x^+(p_1, p_{-1})\}$ one has

$$\limsup_{p_1' \searrow p_1} \frac{1}{p_1' - p_1} \cdot (x_1(p_1', p_{-1}) - x_1(p_1, p_{-1})) = -\frac{1}{2t} \frac{(\lambda - 1)F'(p_1)}{2 + \frac{\lambda-1}{2}(G(x(p_1, p_{-1})) + G(\frac{1}{n} - x(p_1, p_{-1})))}$$

$$-\frac{1}{2t} \frac{\frac{\lambda-1}{2}(2 + (\lambda - 1)F(p_1)) \cdot \limsup_{p_1' \searrow p_1} \frac{G(x(p_1,p_{-1})) - G(x(p_1',p_{-1})) + G(1/n - x(p_1,p_{-1})) - G(1/n - x(p_1',p_{-1}))}{p_1' - p_1}}{[2 + \frac{\lambda-1}{2}(G(x(p_1, p_{-1})) + G(\frac{1}{n} - x(p_1, p_{-1})))]^2}$$

$$\leq (x_1(p_1, p_{-1}))^2 \frac{\frac{\lambda-1}{2} 2n}{2 + \frac{\lambda-1}{2}(G(x(p_1, p_{-1})) + G(\frac{1}{n} - x(p_1, p_{-1})))} \leq (x_1(p_1, p_{-1}))^2 \frac{\frac{\lambda-1}{2} 2n}{2 + \frac{\lambda-1}{2}}. \tag{3.18}$$

Let $\pi(p) = (p - c)E[x^+(p, p_{-1}) + x^-(p, p_{-1})]$. We will prove that

$$\limsup_{p_1' \searrow p_1} \frac{\pi'(p_1') - \pi'(p_1)}{p_1' - p_1} < 0.$$

This is sufficient because it shows that the derivative of the profit function is negative to the right of and sufficiently close to p_1, so that profits are smaller there.

By Equation 3.18, it is sufficient to prove

$$(p_1 - c)\ \underbrace{E\left[\left(x_1^-(p_1,p_{-1})\right)^2 + \left(x_1^+(p_1,p_{-1})\right)^2\right]}_{I}\ \frac{\frac{\lambda-1}{2}2n}{2+\frac{\lambda-1}{2}}$$
$$+2\underbrace{E\left[x_1^-(p_1,p_{-1}) + x_1^+(p_1,p_{-1})\right]}_{II} < 0. \tag{3.19}$$

To bound the above, we begin showing that I divided by the square of II is less than or equal to $\frac{1}{2}\frac{(k+1)^2}{4k}$, where

$$k \equiv \frac{2+\lambda-1}{2+\frac{\lambda-1}{2}}.$$

Since $|x_1(p_1,p_{-1})| \geq \frac{1}{2t}\frac{2+(\lambda-1)F(p)}{2+(\lambda-1)}$ and $|x_1(p_1,p_{-1})| \leq \frac{1}{2t}\frac{2+(\lambda-1)F(p)}{2+\frac{\lambda-1}{2}}$, one has

$$\frac{\max\{|x_1^-(p_1,p_{-1})|, |x_1^+(p_1,p_{-1})|\}}{\min\{|x_1^-(p_1,p_{-1})|, |x_1^+(p_1,p_{-1})|\}} \leq \frac{2+\lambda-1}{2+\frac{\lambda-1}{2}} = k.$$

Now we use the following fact.

Fact 3.1. *Suppose $\tilde{a}_+$ and $\tilde{a}_-$ are positive random variables such that*

$$\frac{\sup\{\tilde{a}_+, \tilde{a}_-\}}{\inf\{\tilde{a}_+, \tilde{a}_-\}} \leq k$$

Then

$$\frac{E[\tilde{a}_+^2 + \tilde{a}_-^2]}{E[\tilde{a}_+ + \tilde{a}_-]^2} \leq \frac{1}{2}\frac{(k+1)^2}{4k}. \tag{3.20}$$

Proof. Suppose without loss of generality that $1 \leq \tilde{a}_+, \tilde{a}_- \leq k$. Since the quadratic function is convex, the ratio on the left-hand side of Inequality 3.20 is maximized if the support if $\tilde{a}_+, \tilde{a}_-$ consists of the extremal values $1, k$. Thus, the left-hand side of the inequality is less than or equal to

$$\max_{b^+,b^- \in [0,1]} \frac{b^+\left(k^2-1\right)+1+b^-\left(k^2-1\right)+1}{\left(b^+(k-1)+1+b^-(k-1)+1\right)^2},$$

which is equivalent to maximizing

$$\max_{b^+,b^-} \frac{1}{2}\left[\frac{\frac{b^++b^-}{2}\left(k^2-1\right)+1}{\left(\frac{b^++b^-}{2}(k-1)+1\right)^2}\right].$$

For brevity, let $b = \frac{b^+ + b^-}{2}$. Then the first-order condition for the above maximization is satisfied if and only if

$$(b(k-1)+1)^2\left(k^2-1\right) - \left(b\left(k^2-1\right)+1\right)2(k-1)(b(k-1)+1) = 0,$$

which yields $b = \frac{1}{k+1}$. Substituting this into the maximand and rewriting gives the desired inequality. □

Hence, for Inequality 3.19 to hold it is sufficient that

$$(p_1 - c)\left|E\left[x_1^-(p_1,p_{-1}) + x_1^+(p_1,p_{-1})\right]\right| \frac{1}{2}\frac{(k+1)^2}{4k}\frac{\frac{\lambda-1}{2}2n}{2+\frac{\lambda-1}{2}} < 2.$$

Since the firm prices its neighbors out of the market with probability zero, its first-order condition implies

$$(p_1 - c)\left|E\left[x_1^-(p_1,p_{-1}) + x_1^+(p_1,p_{-1})\right]\right| \leq \frac{2}{n}.$$

In this case, the above condition simplifies to

$$\frac{(k+1)^2}{4k}\frac{\frac{\lambda-1}{2}}{2+\frac{\lambda-1}{2}} < 1.$$

This condition holds for any $\lambda < 38$. □

Since in an interior equilibrium, prices are above marginal costs and there exists an indifferent consumer between any two products for any marginal cost realization, the above lemma implies the following corollary:

Corollary 3.5. *In an interior equilibrium with $\lambda < 38$, the pricing function is continuous in cost for each firm.*

We are now ready to prove the statement of the proposition. We prove by contradiction; suppose that there exists (at least one) firm that does not charge a deterministic price. Corollary 3.5 implies that there must exist a nontrivial interval of prices, each of which the firm charges for some cost. On this interval, consider a price p^0 and a sequence of prices $p^i \searrow p^0$ such that i.) F is differentiable at p^i, p^0; ii.) the pricing function $p(\cdot)$ is differentiable at p^0 with a strictly positive derivative.[29] Let the corresponding costs be c^0 and $c^i \searrow c^0$.

[29] Given our estimation in Lemma 3.3 (which we also use again below to bound the derivative of $p(\cdot)$), we can show that $p(\cdot)$ is Lipschitz continuous. Hence, we can apply the Fundamental Theorem of Calculus to conclude that its derivative must be strictly positive on a set of positive measure.

Taking the difference between the first order condition for p^i and p^0, dividing it by $p^i - p^0$, and taking the limit as $p^i \to p^0$ while making use of the same calculation as in the proof Lemma 3.3, establishes that

$$p'(c^0) = \frac{E[x_1^+(p^0, p_{-1}) + x_1^-(p^0, p_{-1})]}{(p^0 - c)\lim_{i\to\infty} \frac{E[x_1^+(p^i,p_{-1})+x_1^-(p^i,p_{-1})-x_1^+(p^0,p_{-1})-x_1^-(p^0,p_{-1})]}{p^i-p^0} + 2E[x_1^+(p^0, p_{-1}) + x_1^-(p^0, p_{-1})]}$$
$$< \frac{1}{2 - \frac{(k+1)^2}{2k}\frac{\frac{\lambda-1}{2}}{2+\frac{\lambda-1}{2}}}.$$

By the firm's maximization problem,

$$(p^i - c^i)E\left[x_1^-(p^i, p_{-1}) + x_1^+(p^i, p_{-1})\right] + E[x^+(p^i, p_{-1}) + x^-(p^i, p_{-1})] = 0 \quad (3.21)$$

for each i, and a similar condition holds at p^0.

Fix any p_{-1}. We find a condition under which for $x(\cdot,\cdot) \in \{x^+(\cdot,\cdot)x^-(\cdot,\cdot)\}$,

$$\limsup_{c^i\to c^0} \frac{[(p^i - c^i)x_1(p^i, p_{-1}) + x(p^i, p_{-1})] - [(p^0 - c^0)x_1(p^0, p_{-1}) + x(p^0, p_{-1})]}{c^i - c^0} < 0.$$

This is sufficient for a contradiction because it implies that the first-order condition 3.21 cannot hold for all p^i, p^0 (since for i sufficiently high, the difference between the left-hand sides of the first-order conditions is negative).

The above limsup is equal to

$$\begin{aligned}
&\lim_{c^i\to c^0} \frac{x(p^i, p_{-1}) - x(p^0, p_{-1})}{c^i - c^0} + \lim_{c^i\to c^0} \frac{[(p^i - p^0) - (c^i - c^0)]}{c^i - c^0} \cdot x_1(p^0, p_{-1}) \\
&+ \limsup_{c^i\to c^0} \ (p^i - c^i) \cdot \frac{x_1(p^i, p_{-1}) - x_1(p^0, p_{-1})}{c^i - c^0} \\
&= x_1(p^0, p_{-1}) \cdot (2p'(c^0) - 1) + (p^0 - c^0) \cdot \limsup_{c^i\to c^0} \ \frac{x_1(p^i, p_{-1}) - x_1(p^0, p_{-1})}{c^i - c^0}.
\end{aligned} \quad (3.22)$$

Now we work on the last term above, which is (with apologies for the small

font) equal to

$$
\begin{aligned}
&-\frac{p^0-c^0}{2t}\cdot\limsup_{c^i\to c^0}\frac{1}{c^i-c^0}\left\{\left[\frac{(\lambda-1)(F(p^i)-F(p^0))}{\left\{2+\frac{\lambda-1}{2}[G(x(p^0,p_{-1}))+G((1/n)-x(p^0,p_{-1}))]\right\}}\right]\right.\\
&+\left[\frac{(2+(\lambda-1)F(p^i))}{\left\{2+\frac{\lambda-1}{2}[G(x(p^i,p_{-1}))+G((1/n)-x(p^i,p_{-1}))]\right\}}\right.\\
&\left.\left.-\frac{(2+(\lambda-1)F(p^i))}{\left\{2+\frac{\lambda-1}{2}[G(x(p^0,p_{-1}))+G((1/n)-x(p^0,p_{-1}))]\right\}}\right]\right\}\\
&=-\frac{p^0-c^0}{2t}\cdot\frac{(\lambda-1)F'(p^0)\,p'(c^0)}{\left\{2+\frac{\lambda-1}{2}[G(x(p^0,p_{-1}))+G((1/n)-x(p^0,p_{-1}))]\right\}}-\frac{p^0-c^0}{2t}\cdot\limsup_{c^i\to c^0}\frac{2+(\lambda-1)F(p^i)}{c^i-c^0}\\
&\times\frac{\frac{\lambda-1}{2}\cdot\big(G(x(p^0,p_{-1}))-G(x(p^i,p_{-1}))+G((1/n)-x(p^0,p_{-1}))-G(1/n-x(p^i,p_{-1}))\big)}{\left\{2+\frac{\lambda-1}{2}[G(x(p^0,p_{-1}))+G((1/n)-x(p^0,p_{-1}))]\right\}\cdot\left\{2+\frac{\lambda-1}{2}[G(x(p^i,p_{-1}))+G((1/n)-x(p^i,p_{-1}))]\right\}}\\
&\le(p^0-c^0)\;x_1(p^0,p_{-1})\left[\frac{\lambda-1}{\lambda+1}p'(c^0)F'(p^0)\right.\\
&\left.+\limsup_{c^i\to c^0}\frac{1}{c^i-c^0}\frac{\frac{\lambda-1}{2}\cdot\big(G(x(p^0,p_{-1}))-G(x(p^i,p_{-1}))+G((1/n)-x(p^0,p_{-1}))-G(1/n-x(p^i,p_{-1}))\big)}{\left\{2+\frac{\lambda-1}{2}[G(x(p^0,p_{-1}))+G((1/n)-x(p^0,p_{-1}))]\right\}}\right].
\end{aligned}
$$

Now, notice that only either $G(x(p^0,p_{-1}))-G(x(p^i,p_{-1}))$ or $G(1/n-x(p^0,p_{-1}))-G(1/n-x(p^i,p_{-1}))$ can be strictly greater zero but not both, and since $G(s)-G(s')\le 2n(s-s')$ for any $s>s'$, the sum of these expressions is less than or equal to $2n\,|x(p^i,p_{-1})-x(p^0,p_{-1})|$. Using also $G(x(p^0,p_{-1}))+G(1/n-x(p^0,p_{-1}))\ge 1$, this implies that the above is less than or equal to

$$(p^0-c^0)p'(c^0)\left(x_1(p^0,p_{-1})\frac{\lambda-1}{\lambda+1}F'(p^0)+\big(x_1(p^0,p_{-1})\big)^2\frac{(\lambda-1)n}{2+\frac{\lambda-1}{2}}\right).$$

Substituting into Expression 3.22 and using that $|x_1(p^0,p_{-1})|\le\frac{1}{2t}\frac{1+\lambda}{2+\frac{\lambda-1}{2}}=\frac{k}{2t}$ implies that it is sufficient to prove

$$1+k^2\frac{\lambda-1}{\lambda+1}\frac{n}{2t}(p^0-c^0)p'(c^0)<2p'(c^0)+(p^0-c^0)p'(c^0)\frac{\lambda-1}{\lambda+1}F'(p^0).$$

Using that $F'(p^0)\ge D(p^0)\frac{\theta_1(c^0)}{p'(c^0)}$ and that

$$p^0-c^0=\frac{D(p^0)}{-D'(p^0)}\le\frac{\frac{2}{n}}{\frac{1}{2t}\frac{2}{1+\lambda}}=\frac{2t}{n}(1+\lambda),$$

the above becomes

$$1+k^2(\lambda-1)p'(c^0)<2p'(c^0)+(p^0-c^0)D(p^0)\frac{\lambda-1}{\lambda+1}\theta_1(c^0). \tag{3.23}$$

To finish our proof, we put a bound on the firm's profits $(p^0 - c^0)D(p^0)$. In a market equilibrium, no firm charges a price less than $\underline{c}$, so firm 1's profits are at least as much as it would make if both of its neighbors charge $\underline{c}$ with probability one. If firm 1 also charges $\underline{c}$, its demand in each of its two markets is $\frac{1}{2n}$. Now

$$|x_1(p_1, p_{-1})| \leq \frac{1}{2t}\frac{1+\lambda}{2+\frac{\lambda-1}{2}} = \frac{k}{2t}.$$

This implies that a sufficient condition for the firm to be able to sell profitably is

$$\overline{c} - \underline{c} < \frac{\frac{1}{2n}}{\frac{k}{2t}} = \frac{t}{n}\frac{1}{k} = \frac{t}{n}\frac{3+\lambda}{2(1+\lambda)}.$$

Furthermore, if this is the case, its profits are at least

$$2(p-c)\left(\frac{1}{2n} - \frac{k}{2t}(p-\underline{c})\right) = (p-c)\left(\frac{1}{n} - \frac{k}{t}(p-\underline{c})\right).$$

Maximizing this expression with respect to p and setting $c = \overline{c}$ gives

$$(p^0 - c^0)D(p^0) \geq \frac{k}{4t}\left(\frac{t}{n}\frac{1}{k} - (\overline{c} - \underline{c})\right)^2 = \frac{k}{4}\frac{t}{n^2}\left(\frac{1}{k} - \gamma\right)^2,$$

where $\gamma \equiv (\overline{c} - \underline{c})/(t/n)$.

Now we have two cases.

Case I: $k^2(\lambda - 1) \leq 2$. In this case, a sufficient condition for Inequality 3.23 to hold is

$$\frac{t}{n^2}\theta_1(c^0) > \frac{\lambda+1}{\lambda-1}\frac{4k}{(1-k\gamma)^2}.$$

Case II: $k^2(\lambda - 1) > 2$. Then, substituting our bound for $p'(c^0)$ into Inequality 3.23 and rearranging gives that a sufficient condition is

$$\frac{t}{n^2}\theta_1(c^0) > \frac{4k}{(1-k\gamma)^2} \cdot \frac{(1+\lambda)k^2 - \frac{(k+1)^2}{4}}{2 - \frac{\lambda-1}{\lambda+1}\frac{(k+1)^2}{4}}.$$

This completes our proof. □

Calculations for Example 3.1. For the low-price firm 1 to have market share 3/4, it must be that in personal equilibrium consumers who are within a distance $\alpha = 3/8$ of firm 1's location $y_1 = 0$ buy from firm 1. Personal

equilibrium requires that having expected this behavior, a consumer with realized taste $\chi = 3/8$ be indifferent between buying from the two firms. The above behavior induces expectations to pay p_1 with probability 3/4 and p_2 with probability 1/4, and (from Figure 3.1) also induces an expected distribution of the product's distance from ideal that is a step function with a density of 4 between 0 and 1/8, a density of 2 between 1/8 and 3/8, and a density of zero everywhere else. Given these expectations, a consumer's utility from buying good 1 at price p_1 if she has taste 3/8 is

$$u_1 = v - \frac{3}{8}t - p_1 + \frac{1}{4}(p_2 - p_1) - \lambda t \int_0^{1/8} \left(\frac{3}{8} - s\right) 4d(s) - \lambda t \int_{1/8}^{3/8} \left(\frac{3}{8} - s\right) 2d(s), \tag{3.24}$$

while the utility from buying good 2 at price p_2 is

$$u_2 = v - \frac{1}{8}t - p_2 - \lambda\frac{3}{4}(p_2 - p_1) - \lambda t \int_0^{1/8} \left(\frac{1}{8} - s\right) 4d(s) + t \int_{1/8}^{3/8} \left(s - \frac{1}{8}\right) 2d(s). \tag{3.25}$$

Setting $u_1 = u_2$ yields

$$\begin{aligned} & v - \frac{3}{8}t - p_1 + \frac{1}{4}(p_2 - p_1) - \frac{\lambda t}{2}\left(\frac{3}{8} - \frac{1}{16}\right) - \frac{\lambda t}{2}\left(\frac{3}{8} - \frac{2}{8}\right) \\ & = v - \frac{t}{8} - p_2 - \frac{3\lambda}{4}(p_2 - p_1) - \frac{\lambda t}{2}\left(\frac{1}{8} - \frac{1}{16}\right) + \frac{t}{2}\left(\frac{2}{8} - \frac{1}{8}\right). \end{aligned}$$

Rearranging gives

$$p_2 - p_1 = \frac{t\left(\frac{5}{16} + \lambda\frac{3}{16}\right)}{\frac{5}{4} + \lambda\frac{3}{4}} = \frac{1}{4}, \tag{3.26}$$

We now derive the range of marginal costs that can support the above prices, and the above personal equilibrium, as part of a market equilibrium. To do so, we take advantage of Lemma 3.1. For the indifferent consumer on either side of either firm, $G(x) + G(1/n - x) = G(3/8) + G(1/8) = 3/2$. Using that $\lambda = 5$, the responsiveness of demand to local deviations from the prices p_1, p_2 is

$$D_{2\downarrow}(p_2, p_1) = -\frac{6}{5}, \quad D_{2\uparrow}(p_2, p_1) = D_{1\downarrow}(p_1, p_2) = -1; \quad D_{1\uparrow}(p_1, p_2) = -\frac{2}{5}.$$

Hence, for firm 2 not to want to deviate locally from the proposed equilibrium,

p_2 must satisfy the following conditions:

$$\begin{aligned}(p_2 - c_2)D_{2\downarrow}(p_2, p_1) + D_2(p_2, p_1) = -(p_2 - c_2) \cdot \frac{6}{5} + \frac{1}{4} &\leq 0 \\ (p_2 - c_2)D_{2\uparrow}(p_2, p_1) + D_2(p_2, p_1) = -(p_2 - c_2) \cdot 1 + \frac{1}{4} &\geq 0.\end{aligned}$$

This implies:

$$p_2 - \frac{1}{4} \leq c_2 \leq p_2 - \frac{5}{24} \tag{3.27}$$

By similar calculations, for firm 1 not to want to deviate locally, p_1 and c_1 must satisfy

$$p_1 - \frac{15}{8} \leq c_1 \leq p_1 - \frac{3}{4}. \tag{3.28}$$

Lemma 3.3 implies that if a a local deviation is unprofitable, a non-local deviation to an "interior" price (a price such that the indifferent consumer x is within distance 1/2 of the firm) is also unprofitable. It is also clearly unprofitable to change one's price after capturing or losing the entire market, so that the above local conditions are sufficient for firms not to want to deviate. □

Proof of Lemma 3.2. We begin with proving continuity. Suppose by contradiction that $c^i \to c$ but $P(c^i) \nrightarrow P(c)$. Then, since the pricing function is obviously bounded, we can choose the sequence so that $P(c^i)$ converges; let $P(c^i) \to P' \neq P(c)$. Furthermore, suppose that $P' > P(c)$; the proof for the other case is analogous.

Since $P(c^i)$ is optimal when the marginal cost is c^i, a firm cannot benefit from marginally lowering its price. Using Equation 3.11 to express the firm's marginal profit from lowering its price, this implies that

$$\frac{1}{2n} - \left(P(c^i) - c^i\right) \cdot \frac{2 + (\lambda - 1)F_\uparrow\left(P(c^i)\right)}{1 + \lambda} \geq 0. \tag{3.29}$$

Similarly, since $P(c)$ is optimal when the marginal cost is c, a firm cannot benefit from marginally raising its price. Using Equation 3.11, this implies that

$$\frac{1}{2n} - \left(P(c) - c\right) \cdot \frac{2 + (\lambda - 1)F\left(P(c)\right)}{1 + \lambda} \leq 0. \tag{3.30}$$

Subtracting Inequality 3.29 from Inequality 3.30 gives

$$\left(P(c^i) - c^i\right) \cdot \frac{2 + (\lambda - 1)F_\uparrow\left(P(c^i)\right)}{1 + \lambda} - \left(P(c) - c\right) \cdot \frac{2 + (\lambda - 1)F\left(P(c)\right)}{1 + \lambda} \leq 0.$$

The limit of the left-hand side of this inequality as $i \to \infty$ is positive, a contradiction.

Next, we prove by contradiction that $P(c)$ is non-decreasing. Suppose that $c' > c$ and $P(c') < P(c)$. Since $P(c)$ is optimal when the marginal cost is c, a firm cannot benefit from marginally lowering its price. As above, this implies that

$$\frac{1}{2n} - (P(c) - c) \cdot \frac{2 + (\lambda - 1)F_{\uparrow}(P(c))}{1 + \lambda} \geq 0. \tag{3.31}$$

Similarly, since $P(c)$ is optimal when the marginal cost is c, a firm cannot benefit from marginally raising its price. Therefore,

$$\frac{1}{2n} - (P(c') - c') \cdot \frac{2 + (\lambda - 1)F(P(c'))}{1 + \lambda} \leq 0. \tag{3.32}$$

Subtracting Inequality 3.31 from Inequality 3.32 gives

$$(P(c) - c) \cdot \frac{2 + (\lambda - 1)F_{\uparrow}(P(c))}{1 + \lambda} - (P(c') - c') \cdot \frac{2 + (\lambda - 1)F(P(c'))}{1 + \lambda} \leq 0,$$

a contradiction. □

Proof of Proposition 3.5. We first show that any symmetric equilibrium pricing function satisfies the above properties. Property 1 follows from Lemma 3.2. Since $P(\cdot)$ is increasing and continuous, $P^{-1}(p)$ is a closed interval for any p on the range of $P(\cdot)$. Let $p_1, p_2, \ldots$ be the (at most countable) set of prices p_i such that $P^{-1}(p_i)$ is a non-trivial interval, and let $[f_i, f_i'] = P^{-1}(p_i)$. These $[f_i, f_i']$ satisfy Property 2 by construction. Also, for any $c \notin [f_i, f_i']$, $P(c)$ is not an atom of the pricing distribution, so a firm's demand is differentiable, and hence $P(c)$ must satisfy Equation 3.7. This implies that Property 3 holds. Notice that $D_{1\uparrow}(P(c), P_{-1}(c)) = -\frac{1}{t}\frac{2}{1+\lambda}$, so firm 1 does not want to decrease its price at $\underline{c}$ only if $(P(\underline{c}) - \underline{c})\frac{1}{t}\frac{2}{1+\lambda} \leq \frac{1}{n}$, which implies the first part of Property 4. Also, $D_{1\downarrow}(P(c), P_{-1}(c)) = -\frac{1}{t}$. So for raising the price marginally to be unprofitable, we must have $(P(\overline{c}) - \overline{c})\frac{1}{t} \geq \frac{1}{n}$, which implies the second part of Property 4.

We now argue that if $P(\cdot)$ satisfies the properties in the Proposition, it is an equilibrium pricing strategy. Notice that for any $c \in (\underline{c}, \overline{c}), c \notin [f_i, f_i')$, demand is differentiable from the right. Since $P(c) = \Phi(c)$ for all such c, our analysis in the text implies that there is no profitable local price increase. We are left to consider non-local price increases. Analogously to Proposition 3.1, since the demand curve is concave for price increases, the result is immediate.

Now for any $c \in (\underline{c}, \overline{c}), c \notin (f_i, f_i']$, demand is differentiable from the left. Furthermore, since $P(c) = \Phi(c)$ for all such c, our analysis in the text implies that local price decreases are unprofitable. We now consider non-local price decreases.

The proof mirrors the proof of Proposition 3.1. Suppose the realized cost is c, so that the firm's price in the posited equilibrium is $P(c)$. At this price, consumers' marginal utility in money from a price decrease is $2+(\lambda-1)F_{\uparrow}(P(c))$. We will use that as the price decreases, this marginal utility in money also decreases.

We first rule out the possibility that firm 1 might like to charge a price p_1 so that the indifferent consumer is $x \in \left[\frac{1}{2n}, \frac{1}{n}\right]$. Equating Expressions (3.9) and (3.10), setting $p_2 = P(c)$, and replacing the difference in money utilities,

$$\begin{aligned} P(c) - p_1 + &\left[-\lambda \int_0^{p_1} (p_1 - p)\ dF(p) + \int_{p_1}^{\infty} (p - p_1)\ dF(p)\right] \\ &- \left[-\lambda \int_0^{P(c)} (P(c) - p)\ dF(p) + \int_{P(c)}^{\infty} (p - P(c))\ dF(p)\right], \end{aligned}$$

with its upper bound $(2 + (\lambda - 1)F_{\uparrow}(P(c)))(P(c) - p_1)$, gives that for the indifferent consumer x

$$\begin{aligned} &-xt + (P(c) - p_1)(2 + (\lambda - 1)F_{\uparrow}(P(c))) - \lambda t \left(x - \frac{1}{4n}\right) \\ \geq\ & -\left(\frac{1}{n} - x\right) t - \lambda t \cdot 2n \cdot \left(\frac{1}{n} - x\right) \frac{\frac{1}{n} - x}{2} + t \cdot 2n \cdot \left(x - \frac{1}{2n}\right) \frac{x - \frac{1}{2n}}{2}, \end{aligned}$$

so that

$$P(c) - p_1 \geq \frac{t}{2 + (\lambda - 1)F_{\uparrow}(P(c))} \underbrace{\left[(\lambda + 1)\left(2x - \frac{1}{n}\right) + (\lambda - 1)\left(x - \frac{1}{4n} - nx^2\right)\right]}_{\equiv \kappa}. \tag{3.33}$$

To show that lowering the price to p_1 is not a profitable deviation, it is sufficient to show that

$$\frac{1}{n}(P(c) - c) \geq 2x(p_1 - c).$$

Using Inequality (3.33), it is sufficient to show that

$$\frac{1}{n}(P(c) - c) \geq 2x \left(P(c) - c - \frac{t}{2 + (\lambda - 1)F_{\uparrow}(P(c))}\kappa\right).$$

Rearranging and using that $P(c) - c = \frac{t(1+\lambda)}{n(2+(\lambda-1)F_{\uparrow}(P(c)))}$ gives

$$\left(2x - \frac{1}{n}\right)\frac{1+\lambda}{n} \leq 2x\kappa,$$

which is equivalent to Inequality (3.12), which we verified in the proof of Proposition 3.1.

For $n > 2$, we are left to rule out that firm 1 undercuts its rival and steals more than the entire adjacent market. We begin by ruling out deviations in which the firm captures less than two adjacent markets. Let p_1 be the price at which the consumer located at $\frac{1}{n}$ is indifferent between buying from firm 1 and buying from firm 2. Substituting $x = 1/n$ into Equation 3.33 gives

$$P(c) - p_1 \geq \frac{t}{(2+(\lambda-1)F_{\uparrow}(P(c)))n}\left[2 + \frac{3}{4}(\lambda-1)\right].$$

Using the expression for $P(c) - c$ we get $p_1 - c \leq \frac{t}{(2+(\lambda-1)F_{\uparrow}(P(c)))n}\left[\frac{1}{4}(\lambda-1)\right]$. Thus, even if firm 1 would get the entire two adjacent markets when setting p_1', this is unprofitable as $\frac{1}{n}(P(c) - c) > \frac{4}{n}(p_1 - c)$.

We are left to consider the case when $n > 4$ and firm 1 steals at least two adjacent markets on each side. We show that this is unprofitable because it requires firm 1 to price below marginal cost. For the consumer located at $\frac{2}{n}$ to weakly prefer buying from firm 1 rather than firm 3, it must be that

$$\begin{aligned} P(c) - p_1 &\geq \frac{t}{(2+(\lambda-1)F_{\uparrow}(P(c)))n}[4 + (\lambda-1)\frac{7}{4}] \\ &> \frac{t}{(2+(\lambda-1)F_{\uparrow}(P(c)))n}[2 + \lambda - 1] = P(c) - c. \end{aligned}$$

This completes the proof that non-local price decreases are unprofitable.

We have established that there is no profitable deviation for $c \in (\underline{c}, \overline{c}), c \notin [f_i, f_i']$. For any $c \in (\underline{c}, \overline{c}), c \in [f_i, f_i']$, we have $P(f_i) = P(c) = P(f_i')$. Since it is not profitable to lower the price at f_i, it is also not profitable to lower it for c, and since it is not profitable to raise the price for f_i', it is also not profitable to raise it for c.

We are left to prove that there are no profitable deviations for $\underline{c}$ and $\overline{c}$. Our analysis of non-local deviations above (which only used that $P(c) = \Phi(c)$) implies that for $P(\underline{c}) = \Phi(\underline{c})$, there is no profitable deviation. Now suppose that $P(\underline{c}) < \Phi(\underline{c})$. Demand responsiveness to price decreases from $P(\underline{c})$ is then the same as when $P(\underline{c}) = \Phi(\underline{c})$. Hence, with the markup being lower, the incentive to lower the price is smaller than for $P(\underline{c}) = \Phi(\underline{c})$, so there is

no profitable price decrease. Next, we deal with price increases from $\underline{c}$. Since $P(\underline{c}) < \Phi(\underline{c})$, we consider two cases. First, suppose that $P(c)$ is a constant p^*. Then, using that by Property 4 in the proposition $\Phi(\underline{c}) \geq p^* \geq \Phi(\overline{c})$, and Equation 3.7, the condition in Proposition 3.2 is satisfied. Hence, p^* is a market-equilibrium focal price. If $P(c)$ is not constant, there is a largest interval $[\underline{c}, f_1']$ for which it is constant, and where $f_1' < \overline{c}$. In this case, our argument in the previous paragraph applies. Finally, a similar argument works for price deviations from $\overline{c}$. □

Proof of Corollary 3.2. Suppose by contradiction that there is a constant interval $[f_1, f_1']$. By Conditions 3 and 4 of Proposition 3.5, we must have $P(f_1) \leq \Phi(f_1)$. But by the same two conditions, we must also have $P(f_1') \geq \Phi(f_1')$, which is impossible since $\Phi(\cdot)$ is strictly increasing on the interval while $P(\cdot)$ is constant. □

Proof of Corollary 3.3. We first prove by contradiction that if $\Phi(c)$ is weakly decreasing, then any symmetric equilibrium is a focal-price one. Suppose the price is not deterministic. Then, by the continuity of the pricing function, there are cost levels $c, c' > c$ such that $P(c)$ and $P(c')$ are not atoms of the price distribution. Thus, for these cost levels, the chosen price must satisfy Equation 3.7. Using that $\Phi(c)$ is strictly decreasing, this means that $P(c') < P(c)$, contradicting that the pricing function is non-decreasing.

If $\Phi(c)$ is not weakly decreasing, then there are obviously non-constant $P(\cdot)$ satisfying Proposition 3.5. □

Proof of Corollary 3.4. The statement is true on both the constant and strictly increasing parts of the pricing function. □

Chapter 4

Regular Prices and Sales[1]

4.1 Introduction

It is widely understood in the literature that loss aversion—whereby individuals dislike losses relative to a reference point more than they like same-sized gains—leads individuals to be very averse to small and modest-scale monetary risk, and some researchers believe that loss aversion is the primary explanation for aversion to such risk.[2] Many existing theories have exploited this basic implication of loss aversion to show that firms often have an incentive to shield loss-averse consumers or employees from uncertainty in the environment.[3] As a complement to previous work, in this paper we identify an economically central setting in which the opposite is the case: a firm selling to loss-averse consumers optimally introduces random "sales" into an otherwise riskless en-

[1] This chapter is coauthored with Paul Heidhues, and is forthcoming in Theoretical Economics.

[2] As explained in Chapter 1, Rabin (2000b) shows that in an expected-utility-over-wealth model non-trivial aversion to modest-scale risk must be associated with implausible and empirically unobserved extreme aversion to large-scale risk, so that expected utility over wealth cannot explain attitudes toward both modest-scale and large-scale risks. They argue that loss aversion is likely a better explanation for aversion to small and modest-scale risks. Benartzi and Thaler (1995) and Barberis, Huang and Santos (2001) demonstrate that investor loss aversion can help explain the equity premium puzzle. Sydnor (2010) documents that homeowners display extreme aversion to risk in their deductible choices for homeowners' insurance, and argues that loss aversion can contribute to explaining this behavior.

[3] For instance, Chapter 3 of this dissertation explains why non-identical competitors often sell differentiated goods at identical prices. Similarly, Herweg and Mierendorff (forthcoming) argue that the prevalence of flat-rate contracts can be due to consumer loss aversion. In models of moral hazard with loss-averse agents, Herweg, Müller and Weinschenk (2010) and Macera (2012) demonstrate that the optimal incentive contract features less variation in the wage than would be expected based on classical models.

vironment. The resulting distribution of prices is not only a theoretically novel implication of loss aversion, but it is consistent with some empirically documented patterns in retailer pricing summarized in Section 4.2. While we are unaware of compelling evidence on the importance of our mechanism relative to those in other models of sales, our theory seems like a promising explanation to consider for at least two reasons. First, it is based on loss aversion, one of the most well-documented phenomena in human behavior. Second, as demonstrated by the combination of previous work and ours, unlike most theories loss aversion is consistent with the puzzling combination of *flexibility* in observed consumer prices (reflected in frequent sales) and *stickiness* in observed consumer prices (reflected in the stickiness of the regular price). Furthermore, our theory makes additional predictions on the circumstances under which random sales are likely to be observed.

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. The monopolist announces a price distribution, and the consumer forms her expectations after observing the price distribution. Then, a price is drawn from the distribution, and the consumer decides whether to buy a single item of the good. Our main result establishes that the optimal price distribution consists of low and variable "sale" prices and a high and atomic "regular" price. 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. Furthermore, 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. We also show that market power is necessary for random pricing to be optimal: if two firms compete ex ante for consumers by announcing their price distributions, they choose deterministic prices in equilibrium.

Section 4.3 presents our basic model, which uses the framework of Kőszegi 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. Among credible plans, the consumer chooses one that maximizes her ex-ante expected utility, which we call a *preferred personal equilibrium* or *PPE*.

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 Δ approaches zero.

We analyze our basic model in Section 4.4. We show that for *any* loss-averse preferences by the consumer, the monopolist's limit-optimal price distribution looks qualitatively like that illustrated in Figure 4.1: it consists of a region of continuously distributed low sale prices and a single atomic high regular price. 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.[4] But consider instead the strategy of sometimes charging sale prices low enough to make not buying at these prices non-credible, and at other times charging

[4] 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.

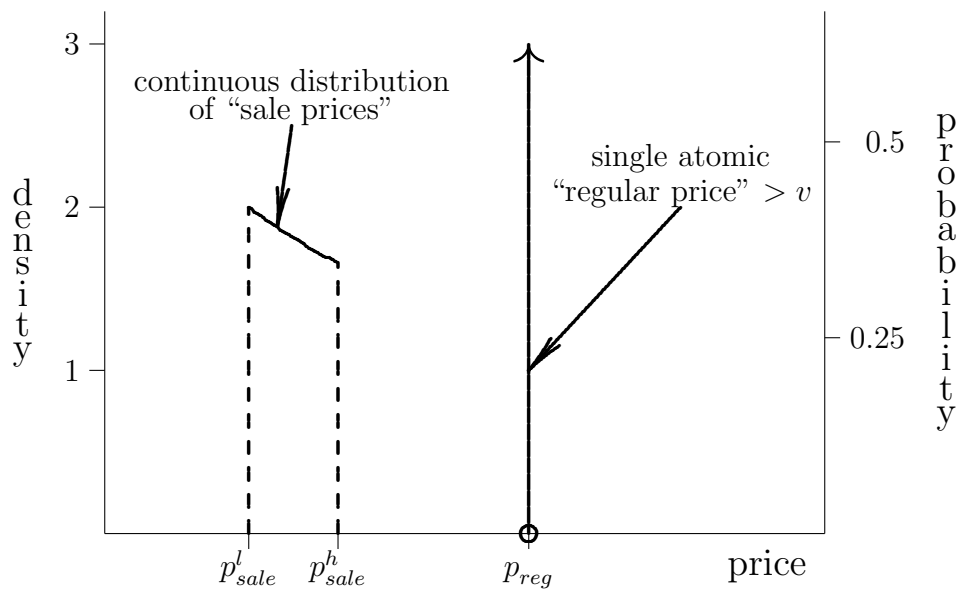

Figure 4.1: A Limit-Optimal Price Distribution

The figure graphs the limit-optimal price distribution when the monopolist sells to a single consumer with known consumption value v for the product, and—to be consistent with experimental evidence suggesting two-to-one loss aversion—loss-aversion parameters are $\lambda = 3$ and $\eta = 1$ (see Section 4.3 for a definition of these variables). The left axis shows the scale for the density of the sale prices, and the right axis shows the scale for the probability of the regular price atom. In this example, $p^l_{sale} = \underline{p} = 0.5 \cdot v, p^h_{sale} \approx 0.81 \cdot v$, and $p_{reg} \approx 1.47 \cdot v$. Although the location of the prices and the weight placed on the regular price is typically different, the limit-optimal price distribution has the same qualitative features (a region of continuously distributed low prices and an atomic high price) for any $\lambda > 1$ and $\eta > 0$.

a high regular price. The consumer's realization that she will buy at the sale prices engages an expectations-based variant of the endowment effect first discussed by Thaler (1980) and documented for instance by Kahneman et al. (1990): if she plans not to buy at the high regular price, she expects to get the product with an interior probability, so she feels an unpleasant loss if she does not get it. To avoid this sense of loss, she prefers to eliminate uncertainty in whether she will get the product, and is therefore willing to buy at a regular price that exceeds v somewhat. As with the endowment effect, the two-dimensional nature of loss aversion is crucial for this logic to hold: if the consumer experienced gain-loss utility over her total consumption utility rather than over the product and money dimensions separately, buying at a regular price exceeding v would neither be optimal ex post, nor reduce risk ex ante.

Going further, by exploiting a type of time inconsistency to push the consumer's expected utility below zero, the firm can lead her to pay not only a regular price exceeding v, but also 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 $\underline{p}$ chosen such that even if the consumer expected not to buy, she would buy at $\underline{p}$. Since the consumer realizes that she will buy at $\underline{p}$, 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, because the role of the regular price is completely different from that of sale prices, the monopolist chooses that price to be atomic. The regular price is chosen by the monopolist not to make a strategy of never buying non-credible, but to ensure that the consumer is willing to buy at all prices rather than just at the sale prices. Hence, there is no reason to make the regular price random—the monopolist just sets it at the consumer's endogenous willingness

to pay for the product.

In Section 4.5, we turn to various extensions and modifications of our framework, making several predictions on when random sales are less likely to be observed. In Section 4.5.1, 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. Because a price distribution that manipulates the consumer into suboptimal behavior would not attract her, in this case firms offer deterministic prices. This result shows that some market power is necessary for random pricing to obtain in equilibrium. In Section 4.5.2, we argue that—although the main point of this paper is to demonstrate the optimality of random prices—in a sense our paper does not contradict previous results that loss aversion leads to reduced price variation. In particular, our results are consistent with stickiness in the regular price, and with price stickiness in a competitive environment. Intuitively, if the regular price was uncertain, the consumer would experience a gain if it turned out relatively low and a loss if it turned out relatively high. Due to loss aversion, she would feel the loss more heavily, making her less willing to buy at an uncertain regular price. Hence, not only does a monopolist not need variation in the regular price (as we explained above), it has an incentive to keep the regular price sticky to induce the consumer to buy at the regular price in addition to the sale prices. Similarly, because a consumer dislikes uncertainty in the price, to attract her from a competitor a firm has an incentive to eliminate variation in the price, leading to sticky prices under ex-ante competition. Finally, in Section 4.5.3 we show that if it is sufficiently costly for consumers to observe the realized price before buying, then the monopolist chooses a deterministic price.

In Section 4.6, we summarize the behavioral-economics and pricing literatures most related to our paper. We conclude the paper in Section 4.7 by pointing out some pricing patterns that our model cannot explain. All proofs are in the Appendix.

4.2 Evidence on Pricing

The predictions of our model are qualitatively consistent with much of the evidence on supermarket pricing. Supermarket 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, Jaimovich and Rebelo 2009). Furthermore, most of these temporary price changes are sales (price decreases rather than in-

creases), with the mean deviation being -22% of the regular price (Kehoe and Midrigan 2008).

This price variability occurs despite considerable stickiness in the regular price, which change about once a year on average (Kehoe and Midrigan 2008, Eichenbaum et al. 2009). In addition, consumer retail prices tend to be sticky more broadly. 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 a classical reference-independent model, any change in the firm's cost or elasticity of demand creates an incentive to change prices. From this perspective, it is likely that changes in the economic environment are far too rapid to justify the above lags between price changes. As suggestive evidence for this observation, Eichenbaum et al. (2009) document that conditional on the weekly price being constant and equal to the regular price, the standard deviation of quantities sold is 42%.

Recent empirical research also shows that 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. 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.

It is important to note that 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. (2009) report that sale prices constitute about 30% of price observations at supermarkets, Klenow and Kryvtsov (2008) find that overall they constitute only 8% of non-food price observations. We are unaware of evidence on whether the pattern of sale frequencies across

different types of retailers is consistent with our model's predictions that sales are less likely to occur when prices are harder to observe or there is competition.

4.3 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. We suppose that c is sufficiently low for the monopolist to sell to the consumer; this will be the case whenever the revenue from the price distribution we identify below exceeds c. 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, makes a price-contingent purchase plan, and forms stochastic beliefs regarding her consumption outcomes. 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 $b \in \{0, 1\}$. For technical and expositional reasons, we assume that any indifference by the consumer in period 1 is broken in favor of buying.

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. One possible micro-foundation for this assumption is a model in which (based on Fudenberg and Levine 1989) the firm can develop a "reputation" for playing the optimal committed price distribution. More generally, it seems plausible to assume that over time consumers learn a firm's basic pricing strategy and incorporate it into their expectations, and that firms take this into account. This assumption is clearly crucial for our main result: once the consumer has formed expectations, the firm would prefer not to charge sale prices, so commitment is necessary for it to use sales as a way to induce an expectation to buy in the consumer.

Our model of consumer behavior follows the approaches of Kőszegi and Rabin (2006) and Heidhues and Kőszegi (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 any riskless consumption-utility outcome k^i and riskless reference point r^i in dimension i, we define total utility in dimension i as $u(k^i|r^i) = k^i + \mu(k^i - r^i)$. Hence, for any (k^v, k^p) and (r^v, r^p), total utility is

$$u(k^v|r^v) + u(k^p|r^p) = k^v + \mu(k^v - r^v) + k^p + \mu(k^p - r^p). \tag{4.1}$$

We assume that μ 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 η can be interpreted as the weight attached to gain-loss utility, and λ as the coefficient of loss aversion.[5]

Beyond loss aversion, our specification in Expression 4.1 incorporates the assumption that 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.[6] 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

[5] Consistent with most of the evidence and literature on loss aversion suggesting that individuals are loss averse even for small stakes, we assume a kink in gain-loss utility at zero. An alternative specification is one in which the marginal gain-loss utility changes quickly around zero, but there is no kink. The mechanism behind our results indicates that in such an alternative specification, charging random sale prices and separate regular prices would still be optimal. In a setting with cost uncertainty and downward-sloping demand, however, the regular prices would no longer be fully sticky, only compressed relative to what one would expect in a classical model.

[6] 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 the 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.

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. We will discuss how gain-loss utility and loss aversion in each of the two dimensions contributes to our results.

Since we assume below that expectations are rational, and in many situations such rational expectations are stochastic, we extend the utility function in Expression 4.1 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. For any consumption-utility outcome k^i and probability distribution over consumption utilities F^i in dimension i, we define

$$U(k^i|F^i) = \int_{r^i} u(k^i|r^i)dF^i(r^i), \tag{4.2}$$

and define total utility from outcome (k^v, k^p) as $U(k^v|F^v)+U(k^p|F^p)$. 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. Still applying Kőszegi and Rabin (2006), we assume that beliefs 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 (2010) 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 4.1. The cutoff price p^* is a personal equilibrium (PE) for price distribution Π 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 4.2. The cutoff price p^* is a preferred personal equilibrium (PPE) for price distribution Π if it is a PE, and for any PE cutoff price p^{**},

$$\begin{aligned} &E_{F^v(\Pi,p^*)}[U(k^v|F^v(\Pi, p^*))] + E_{F^p(\Pi,p^*)}[U(k^p|F^v(\Pi, p^*))] \\ \geq\ &E_{F^v(\Pi,p^{**})}[U(k^v|F^v(\Pi, p^{**}))] + E_{F^p(\Pi,p^{**})}[U(k^p|F^v(\Pi, p^{**}))]. \end{aligned} \tag{4.3}$$

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

$$\max_\Pi \left\{\Pi(p^*)E_P[p|p \leq p^*] - \Pi(p^*)c \mid p^* \text{ is the highest PPE for } \Pi(\cdot)\right\}. \tag{4.4}$$

To facilitate our statements and proofs, we make one more technical assumption: we suppose that the monopolist must choose a discrete price distribution in which neighboring atoms are at least $\Delta > 0$ apart. We think of Δ as being small. Together with the assumption that indifference by the consumer is broken in favor of buying, this ensures the existence of an optimal price distribution. In the Appendix, 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 Δ approaches zero:

Definition 4.3. 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 Δ_i such that $\Pi_i \to \Pi$ in distribution.

4.4 The Optimal Price Distribution

This section presents our main results on pricing with loss-averse consumers. We begin in Section 4.4.1 by illustrating the main idea behind random sales in a simplified model with no loss aversion in money. In Section 4.4.2, we identify the limit-optimal price distribution in our main model.

4.4.1 An Illustration: No Loss Aversion in Money

We illustrate the logic behind the role of randomization in the monopolist's optimal pricing strategy in a model with no loss aversion in money. This variant of our model simplifies many calculations, and is also relevant because, as argued for instance by Novemsky and Kahneman (2005) and Kőszegi and Rabin (2009), loss aversion may be weaker in the money than in the product dimension. We assume that for riskless reference points r^v, r^p, the consumer's gain-loss utility in the product dimension is $k^v + \mu(k^v - r^v)$ (as above), but in the money dimension it is simply $k^p + \eta \cdot (k^p - r^p)$. We define gain-loss utility for stochastic reference points analogously to above.

Because in this variant the consumer's disutility of paying a price p is $(1+\eta)p$ independently of r^p, her willingness to pay for the product depends only on r^v. To see the effect of this reference point, suppose that the consumer had expected to buy with probability q. Then, if she buys, her utility in the product dimension—consisting of consumption utility and a gain of v relative to the possibility of not buying she had expected with probability $1-q$—is $(1+(1-q)\eta)v$; and if she does not buy, her utility in the product dimension—consisting of a loss of v relative to the possibility of buying she had expected with probability q—is $-q\eta\lambda v$. Hence, the consumer buys if $(1+\eta)p \leq (1+\eta+\eta(\lambda-1)q)v$, or

$$p \leq \frac{(1+\eta+\eta(\lambda-1)q)v}{1+\eta}. \tag{4.5}$$

To see what this endogenous willingness to pay implies for pricing, first consider what the monopolist can do with a deterministic price p. In that case, the consumer faces a deterministic environment, so in any PE she gets what she expects.[7] This implies that her PE total utility is equal to her consumption utility, so that from an ex-ante perspective buying is optimal if and only if $p \leq v$. To conclude that buying is the PPE if and only if $p \leq v$, however, we must check that this constitutes a PE for any p. This is easy:

[7] This is true because—due to our assumption that indifference by the consumer in period 1 is broken in favor of buying—there are only pure-strategy PE.

for $p > v$, Inequality 4.5 implies that the consumer is willing to follow a plan to buy with probability zero ($q = 0$), and for $p \leq v$, the same inequality implies that the consumer is willing to follow a plan to buy with probability one ($q = 1$). Hence, the highest revenue the monopolist can earn with a deterministic price is v.

But the monopolist can do better with a stochastic price. Suppose that the monopolist charges $p = v$ with probability s_1, and higher prices (whose distribution we will determine momentarily) starting at $v+\Delta$ with probability $1 - s_1$. Then, in any PE the consumer buys at price $p = v$: by Inequality 4.5, even if she had expected to buy with probability zero, she would buy at price v. Given that the consumer therefore buys with probability of at least s_1 in any PE, Inequality 4.5 implies that it is not credible for her not to buy at price $v + \Delta$ if

$$v + \Delta \leq \frac{(1 + \eta + \eta(\lambda - 1)s_1)v}{1 + \eta} \Leftrightarrow s_1 \geq \frac{1 + \eta}{\eta(\lambda - 1)v}\Delta.$$

Intuitively, the consumer's realization that she will buy at price v raises her reference point in the product dimension and thereby creates a sense of loss if she does not buy. The motive to avoid this loss induces an "attachment effect" that raises her willingness to pay for the product.

Extending the above logic to any price distribution is straightforward. Suppose that the consumer faces the distribution $F(\cdot)$, and define $F_-(p) = \lim_{p' \nearrow p} F(p')$. Then, there is a unique PE in which the consumer buys with probability 1 if and only if

$$p \leq \frac{(1 + \eta + \eta(\lambda - 1)F_-(p))v}{1 + \eta} \Leftrightarrow F_-(p) \geq \frac{1 + \eta}{\eta(\lambda - 1)} \cdot \frac{p - v}{v}$$

holds for all p on the support of $F(\cdot)$. A "near-uniform" distribution that puts weights of $(1 + \eta)\Delta/(\eta(\lambda - 1)v)$ on each of the prices $v, v + \Delta, v + 2\Delta, \ldots,$ and the remaining weight on the highest price, satisfies this condition with equality at each of the atoms.

Notice that when facing the above price distribution, the consumer buys the product at an average price that exceeds v, so that she receives negative expected utility. In fact, for small Δ the consumer pays a price strictly above v with probability close to 1! Because the consumer has the option of making and following through a strategy of never buying—which would yield an expected utility of zero—this means that she behaves suboptimally among the strategies available to her. Intuitively, 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.[8] While the increase in the consumer's reference point due to the expectation to buy at low prices increases her willingness to pay, it also lowers her utility. When she makes her purchase decision in period 1, she takes the reference point (formed in period 0) as given, and therefore ignores this negative effect.

Adding loss aversion in money complicates the above logic underlying randomization for two reasons. First, because a consumer who did not expect to buy experiences a loss in money if she does buy, loss aversion in money reduces the highest price at which it is not credible not to buy. Second, once the consumer expects to buy at such a low price, she experiences paying higher prices as a loss, reducing the monopolist's ability to cash in on the attachment effect. Nevertheless, our main result below shows that for any loss-averse preferences by the consumer, a stochastic price remains optimal. Furthermore, our main result shows that the optimal price distribution features not only a densely packed region of sale prices similar to the uniform distribution above, but also a regular-price atom separated from the sale prices.

4.4.2 Main Result

Our main proposition identifies the features of the monopolist's limit-optimal pricing strategy when the consumer is loss averse in the money as well as the product dimension:

Proposition 4.1. *Fix any $\eta > 0$ and $\lambda > 1$. The profit-maximizing price distribution induces purchase with probability one. Furthermore, in that case any limit-optimal price distribution $\Pi(\cdot)$ has support $[p^l_{sale}, p^h_{sale}] \cup \{p_{reg}\}$, where $p^l_{sale} = \underline{p} = (1+\eta)v/(1+\eta\lambda) < p^h_{sale} < p_{reg}$ and $\Pi(\cdot)$ is continuously distributed on the interval $[p^l_{sale}, p^h_{sale}]$ with density $\pi(p) = (1+\eta\lambda)/[\eta(\lambda-1)(v+p)]$. The monopolist's expected revenue is strictly greater than v.*

Proposition 4.1 says that the limit-optimal price distribution has two parts (as illustrated in Figure 4.1): 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 4.2.

[8] That beliefs-based preferences can generate time-inconsistent behavior has been pointed out by Caplin and Leahy (2001) and Kőszegi (2010), and explored in more detail by Kőszegi (2010) and Kőszegi and Rabin (2009).

The proof of Proposition 4.1 has five main steps: (1) with a deterministic price, the monopolist cannot earn more than 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 make these "forcing" sale prices—i.e., the prices at which the consumer buys in any PE—stochastic; (4) it is suboptimal to rely solely on forcing sale prices, so that there is also a region of high regular prices separated from the sale prices; and (5) the high regular price is sticky. We have discussed the intuition for Steps (1), (2), (3), and (5) in the introduction, so here we discuss only Step (4).

To understand Step (4), suppose by contradiction that a forcing distribution—i.e., a price distribution that consists entirely of forcing prices and hence induces a unique PE in which the consumer buys at all prices—is optimal. By Steps 1 and 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 effect on utility when buying: besides generating gains in money, it eliminates losses in money and gains in the good, and since the average price is greater than v, the elimination of losses 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.

As the above intuition indicates, the prediction of our model that there is an atomic regular price separated from the sale prices does *not* rely on loss aversion in money: even if the consumer's disutility from monetary losses was equal to her utility from gains, the motive to eliminate these losses would lead her to buy at the higher price p'. This is straightforward to check in the version of our model in Section 4.4.1 with no loss aversion in money. But as the same intuition indicates, the regular price does rely on *gain-loss utility* in money: if the consumer did not derive disutility from paying a higher price than her reference point, she would not care about the above losses, and hence she would prefer not to buy at p'. To see this formally, we return to a subtle modification of our model in Section 4.4.1: we assume that the consumer's utility in money is not reference-dependent, but simply equal to $(1 + \eta)k_p$. This implies that for any price distribution the consumer's set of PE is the same as in Section 4.4.1, and the optimal deterministic price remains v. Nevertheless, absent

reference-dependence in money, we now show that the monopolist does not want to charge a regular price. Suppose that for some $p^* > v$ with $F(p^*) < 1$, there is a PE such that the consumer buys only for prices $p \leq p^*$. We argue that for any $p' > p^*$ with $F(p') > F(p^*)$, buying up to prices p' cannot be a PPE. For notational simplicity, we denote by $V(p'|p)$ the consumer's expected utility if she had formed expectations based on the plan to buy up to price p, and then follows a plan to buy up to price p'. In this notation, we have $V(p^*|p^*) \geq V(p'|p^*) > V(p'|p')$, where the first inequality follows from the definition of PE and the second follows from the fact that planning to buy for prices up to p' rather than p^* raises the consumer's reference point in the product dimension and hence lowers her utility. This means that with no gain-loss utility in money, the optimal price distribution induces a unique PE with purchase with probability 1 (much like the price distribution we identified in Section 4.4.1), and the firm's limit-optimal price distribution is the uniform distribution with support $[v, (1+\eta\lambda)v/(1+\eta)]$.

The qualitatively different nature of the optimal price distribution with and without gain-loss utility in money also reflects a subtle difference in how the monopolist exploits the consumer's attachment to the product to charge an expected price above v. In each case, the possibility of buying the product at a low price means that the consumer must expect to get the product with some probability. Without gain-loss utility in money, the monopolist exploits the consumer's *ex-post* (period-1) aversion to facing a sense of loss in the product dimension to charge higher prices. With gain-loss utility in money, the monopolist relies on this ex-post aversion to losses in the sales region, but the same aversion is insufficient to induce the consumer to buy at the regular price: if she expected not to buy at the regular price, she would strictly prefer not to do so ex post. Instead, the monopolist relies on the consumer's *ex-ante* (period-0) aversion to risk in whether she will get the product to induce her to plan on buying at the regular price.

An interesting possibility arises in our model if $c > \underline{p}$, yet the monopolist can profitably sell to the consumer. In this case, the monopolist's cost is higher than some of the prices it charges, providing a non-predatory rationale for potential below-marginal-cost pricing of a single-product firm. Going further, since the monopolist induces the consumer to buy at an average price exceeding v, it may sell the product even if $v < c$. In this case, loss aversion affects not only the monopolist's pricing strategy, but also its production decision, leading it to produce a socially wasteful product. Finally, note that if below-marginal-cost pricing is prohibited—as is the case in some countries—sales disappear altogether: since the firm cannot manipulate the consumer into buying against

her will, it chooses a sticky price if $v > c$, and chooses not to sell if $v < c$.[9]

Beyond the shape of the optimal price distribution, the observation that the consumer buys at an expected price exceeding v has an immediate welfare implication:

Proposition 4.2. *For any $\eta > 0, \lambda > 1$, and $\Delta < v - \underline{p}$, 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 4.2 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.[10] Two 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.[11] 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.

4.5 Extensions and Modifications

In this section we discuss a number of further predictions of our framework.

[9] Relatedly, with cost uncertainty the firm's opportunity cost of delivering the product could sometimes be greater than the highest possible price. 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 the monopolist may commit to selling even in situations in which it makes losses from doing so.

[10] Although we model neither multi-product retailers nor the wholesaler-retailer relationship, Proposition 4.2 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.

[11] Nevertheless, even with consumer heterogeneity, some (marginal) consumers who buy with positive probability would be better off making and following through a plan of never buying. See our working paper (Heidhues and Kőszegi 2010b) for details.

4.5.1 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 two simple forms of competition, showing that our results on random sales rely on some amount of market power. First, we consider perfect ex-ante competition for consumers, as for example when consumers decide which supermarket or restaurant to frequent. Two retailers simultaneously commit to their price distributions, and after observing the distributions, the consumer decides which retailer to visit and forms expectations about her consumption outcomes. We assume that if indifferent, the consumer visits each firm with positive probability. Finally, a price is drawn from each retailer's price distribution, and the consumer decides whether to buy at her previously chosen retailer's price. We assume that the two retailers have identical costs $c < v$, and that they use pure strategies (i.e., they do not mix between distributions). Then:

Proposition 4.3. *For any $\eta > 0$, $\lambda > 1$, the unique equilibrium with ex-ante competition is for each firm to choose the deterministic price c.*

Proposition 4.3 says that if there is perfect competition, firms do not use a manipulative price distribution, but instead choose the deterministic price equal to cost. The reason is simple: a manipulative price distribution would lead the consumer to visit the other retailer.

Second, 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. 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.

4.5.2 Price Stickiness

As has been intuited by researchers for a long time and shown for instance by Sibly (2002) and Heidhues and Kőszegi (2008), consumer loss aversion often creates "price stickiness"—an unresponsiveness of prices to changes in cost or demand circumstances. While the main point of this paper is that loss aversion can create the opposite incentive—to introduce uncertainty into a deterministic environment despite facing a consumer who dislikes this uncertainty—we conjecture that the price variation we identify in this paper is consistent with stickiness in the regular price, and in price stickiness in a competitive environment. Intuitively, not only does a monopolist not need variation in the regular price (as we explained above), it has an incentive to keep the regular price sticky to induce the consumer to buy at the regular price in addition to the sale prices. If the regular price was uncertain, the consumer would experience a gain if it turned out relatively low and a loss if it turned out relatively high. Due to loss aversion, she would feel the loss more heavily, making her less willing to buy at an uncertain regular price. Similarly, because a consumer dislikes uncertainty in the price, to attract her from a competitor a firm has an incentive to eliminate variation in the price, leading to sticky prices under ex-ante competition. These intuitions suggest that our model is consistent with the puzzling combination of stickiness and flexibility in prices.

To demonstrate these forces toward price stickiness formally in our model, it is necessary to introduce features that in a classical setting would lead to price variation. A natural way to do so is to assume that demand is downward sloping and the firm's cost is uncertain. We have, however, been unable to analyze models with these features in general, and even special cases raise considerable technical issues. We describe here two restrictive cases we have analyzed in detail in our working paper (Heidhues and Kőszegi 2010b). In the monopoly case, 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 $s \in [0, 1)$, p_L, p_H, α_L, and α_H satisfying $p_H > p_L + 2\bar{\alpha}$ and $0 \leq \alpha_L, \alpha_H \leq \bar{\alpha}$. In this setting, we show that if $\bar{\alpha}$ is sufficiently small, the optimal price distribution has a sales-and-regular-prices structure ($s > 0$) and a single regular price ($\alpha_H = 0$), and if in addition the firm's marginal cost is sufficiently narrowly distributed, sales prices are flexible ($\alpha_L = \bar{\alpha}$). These findings contrast with those in the corresponding classical model, where for sufficiently narrowly distributed costs sale prices would not be used ($s = 0$), but the regular price would adjust to cost shocks ($\alpha_H > 0$).

In the competition case, we consider a variant of our model in Section 4.5.1

in which there is a mass of consumers whose consumption value is distributed continuously on the interval $[0, \overline{v}]$, with positive density everywhere, and the firms have identical cost distributions uniformly distributed on the interval $[c_L, c_H]$ with density d. We show that if d is sufficiently large, then for any $\Delta > 0$ the unique symmetric equilibrium is for each firm to choose the deterministic price $(c_L + c_H)/2$.[12]

4.5.3 Further Extensions and Modifications

An implicit assumption of our model above is that it is costless for the consumer to observe the price in period 1. In contrast, consumers often have to go out of their way to learn a particular product's price. We formally analyze a variant of our model with such price-discovery costs in our working paper (Heidhues and Kőszegi 2010b) and demonstrate that for low price-discovery costs, the limit-optimal price distribution is very similar to the one we find in Proposition 4.1, with one important difference: the monopolist charges a price of zero with small probability. Intuitively, the possibility of a "free sample" makes non-buying non-credible despite price-discovery costs because the consumer—even if she had been expecting not to do so—would want to pay the small price-discovery cost in period 1 to see whether she can get the free sample.[13] In contrast, when price-discovery costs are high, it becomes too costly or impossible to manipulate the consumer into buying against her will through a sales-and-regular-price strategy, so that the firm switches to deterministic pricing. This is easiest to see when price-discovery costs are greater than $\underline{p}$: in this case, a strategy of never buying is always credible, so that it is impossible to manipulate a consumer into buying against her will. Our framework therefore has the novel prediction that sales are more likely when price discovery costs are low. This is arguably the case in supermarkets for the marginal consumer of any given product—so long as these consumers are visiting the supermarket to buy other products anyhow—but is arguably not

[12] Note that sticky pricing is not an equilibrium in this model when consumers have classical 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.

[13] While we provide an explanation for free samples, this prediction is not robust to realistic variations of our model in which a free sample would generate extra money-losing demand, for instance by attracting low-valuation consumers or by inducing consumers to store. When these considerations are important, the firm will use a positive (but low) price atom instead or switch to deterministic pricing.

the case for many other retailers.

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 not, 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 from the perspective of our model each v can be thought of as a different pricing situation, in each of which 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 (different) sticky price for each v. This prediction is consistent with matinees in movie theaters and cyclical sales of many products for which the sale price is also sticky. At the same time, our model does not explain why prices do not seem to change in response to some other predictable changes in demand.

4.6 Related Theories of Sales

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, and each consumer decides 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.[14] 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.

There is also a set of models in industrial organization in which the oligopolistic environment leads firms to play mixed strategies.[15] In all of these papers, each firm is left with a "captive" group of consumers who will not buy from a

[14] 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.

[15] See, for example, Shilony (1977), Varian (1980) Gal-Or (1982), Davidson and Deneckere (1986), Baye, Kovenock and de Vries (1992).

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.[16]

A distinct testable implication of our model relative to most models of sales above is that the monopolist makes higher profits at high prices than at low prices. Intertemporal-price-discrimination models predict the opposite: in those models, 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 a low price only if compensated by higher current profits. And by the nature of a mixed-strategy equilibrium, mixed-strategy models predict equal profits for low and high prices. While we are unaware of systematic empirical evidence on profits in sale periods relative to regular-price periods, it seems plausible that for at least some types of sales profits are higher at regular prices. For instance, sales that are subject to strict "limited availability" presumably generate most profits once the item on sale runs out.[17] Furthermore, unlike our theory the above models of sales do not seem to predict stickiness in the regular price when demand is downward sloping and there are cost shocks.

Although in its current form our model is not intended as a basis for a macroeconomic model studying monetary policy, in proposing a possible mechanism for sales and price stickiness it is related to such models. 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 implies a price distribution similar to ours, it leaves unanswered why there would be

[16] In a model that is similar to the mixed-strategy oligopoly pricing models but in which consumers' purchase decisions are based on a naive sampling procedure, Spiegler (2012) predicts a price pattern similar to ours.

[17] One paper suggesting that sales can increase profits at higher prices is Slade (1999). Slade investigates the prices set for saltine crackers by grocery stores in a small US town. 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. The positive effect is meant to capture a stock of goodwill, which she argues could arise through "consumer habit formation, product awareness, or brand loyalty," but might also be due to loss aversion as in this paper. In her empirical implementation, she finds evidence that low past prices increase current sales. At the same time, Anderson and Simester (2010) find that mail-order customers who have recently bought an item at a high price and later see the item with its price deeply reduced are less likely to order in the future, making sales with deep discounts unprofitable in this industry.

two different kinds of prices with different menu costs. Our theory provides a kind of micro-foundation for these reduced-form assumptions, and makes a variety of additional predictions on the effects of competition, price-discovery costs, and other forces.[18]

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 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. Furthermore, unless pass-through is very high, their model (unlike ours) predicts frequent sales of considerable magnitude only when there are frequent and considerable changes in marginal costs. And 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.

4.7 Conclusion

While our model provides a potential explanation for a number of pricing patterns, there are some patterns it cannot 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. In addition, given that volume is for some

[18] It is important to note, however, that—because gain-loss utility is based on consumption utility and hence is defined over real variables—in our model the price would scale with nominal variables, and our result on the stickiness of the regular price is stickiness in the real rather than nominal regular price. This implies that in contrast to the above and many other macroeconomic models, our model does not generate a force that in itself leads to monetary non-neutrality. Whether and how variants of our model—e.g., one in which consumers are loss-averse over nominal prices, and hence the nominal regular price is sticky—lead to monetary non-neutrality, and whether and how loss aversion interacts with other forces that can generate monetary non-neutrality, is beyond the scope of this paper.

items much higher during sales than when the regular price is charged, it is likely that storage 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.

Appendix: Proofs

Some Preliminaries

First, we introduce some notation we will use throughout our proofs. For any market price distribution Π, 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 Π and buying at all prices less or equal to p_l is:

$$\begin{aligned} EU(p_l;\Pi) &= Q_l v - Q_l P_l - \eta(\lambda-1)Q_l(1-Q_l)v \\ &- \eta(\lambda-1)Q_l(1-Q_l)P_l - \eta(\lambda-1)\sum_{l'=1}^{l}\sum_{l''=1}^{l'} q_{l''}q_{l'}(p_{l'}-p_{l''}). \end{aligned} \tag{4.6}$$

Finally, buying for all prices less or equal to p_l is a personal equilibrium if, 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

$$p_l \leq \frac{1+\eta(1-Q_l)+\eta\lambda Q_l}{1+\eta\lambda}\, v + \frac{\eta(\lambda-1)}{1+\eta\lambda}\, Q_l P_l < p_{l+1}. \tag{4.7}$$

Discrete Version of Proposition 4.1

To establish Proposition 4.1, 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:[19]

Proposition 4.4. *For any $\eta > 0$, $\lambda > 1$, and Δ satisfying $0 < \Delta < v - \underline{p}$, 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 $p_{z^*} > p_z$, where $\underline{p} - 2\Delta < p_1 \leq \underline{p}$, 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 $1 - \sum_{l=1}^{z} q_l$.*

Proof. We begin by introducing the formal versions of what in the text we call sale prices and regular prices. Let $Q_z \geq 0$ be the highest probability such that in any PE, the consumer buys the product with probability of at least Q_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. We think of the prices up to and including p_z as the sale prices, and the higher prices as the regular prices. Hence, Q_z is the probability of buying on sale, and $Q_{z^*} - Q_z$ is the probability of buying at the regular prices.

It is useful to first outline the broad steps in our proof. There are two major steps, and several substeps. The main parts of the proof correspond to Steps 1 through 5 that provide the intuition for Proposition 4.1 in the text. For technical reasons, however, the order of the steps is not exactly the same, and there are also other steps.

Part A. 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. This is the key part of the proof. We will use the following substeps:

A(i). We show that there must be a single atom on the interval $(p_z, p_{z^*}]$—that is, there must be a single regular price. This corresponds to Step 5 in the main text.

A(ii). We establish the (intuitively obvious) result that the consumer buys with probability 1 ($Q_{z^*} = 1$).

[19] Proposition 4.4 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 $p_1 = \underline{p}$, $p_{l+1} - p_l = \Delta$ for any $l < z$, and $q_z \leq A\Delta/(v + p_z)$.

From here, the proof corresponds to Steps 1 through 4 of the intuition in the text.

A(iii). We show that with a deterministic price, the firm cannot earn revenue greater than v. (Step 1.)

A(iv). We show that if Δ is sufficiently small, the firm charges at least two prices with positive probability, establishing that $Q_z > 0$. (Step 2.)

A(v). We show that the sale prices are at most 2Δ apart, and have the property that if the consumer expected to buy up to some sale price, she would just be willing to buy at the next sale price. (Step 3.)

A(vi). We establish that it is not optimal to set $Q_z = 1$. (Step 4.)

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

Part A.

A(i). 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 $\overline{p} = (1+\eta\lambda)/(1+\eta)$. 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 4.7, 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 4.6 that $EU(p_{z^*}; F) < EU(p_a; F')$, and by construction, $EU(p_l; F) = EU(p_l; F')$ for any $l < z^*$. 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.

A(ii). 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 4.7 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 4.7 that buying at all prices is a PE after the price change. Now using Equation 4.6 and the fact that with price distribution F the consumer prefers the PE in which she buys up to p_{z^*}, one has

$$\begin{aligned}
EU(p_z; F) &= Q_z v - Q_z P_z - \eta(\lambda - 1) Q_z (1 - Q_z) v \\
&- \eta(\lambda - 1) Q_z (1 - Q_z) P_z - \eta(\lambda - 1) \sum_{l'=1}^{z} \sum_{l''=1}^{l'} q_{l''} q_{l'} (p_{l'} - p_{l''}) \\
&\leq Q_z v - Q_z P_z + q_{z^*}(v - p_{z^*}) - \eta(\lambda - 1)(Q_z + q_{z^*})(1 - Q_z - q_{z^*}) v \\
&- \eta(\lambda - 1)(1 - Q_z - q_{z^*})(Q_z P_z + q_{z^*} p_{z^*}) \\
&- \eta(\lambda - 1) \left(\sum_{l'=1}^{z} \sum_{l''=1}^{l'} q_{l''} q_{l'} (p_{l'} - p_{l''}) + q_{z^*} \sum_{l=1}^{z^*} q_l (p_{z^*} - p_l) \right) \\
&= EU(p_{z^*}; F).
\end{aligned}$$

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

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

Dividing by q_{z^*}, one has

$$0 \leq v - p_{z^*} - \eta(\lambda - 1) \left((1 - 2Q_z - q_{z^*}) v + (1 - q_{z^*}) p_{z^*} - 2 Q_z P_z \right). \tag{4.8}$$

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 $z^* = z$.

A(iii). First, consider $z = 0$. In that case, the monopolist charges a single deterministic price. Note that in any PE, the consumer gets what she expected, so that her total utility is equal to her consumption utility.[20] This means that for any $p > v$, the ex-ante optimal strategy is not to buy. We show that for such prices, not buying is a PE, so that it must be the PPE. 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

$$p > \frac{1+\eta}{1+\eta\lambda} \cdot v \equiv \underline{p}.$$

Since this inequality is satisfied for any $p > v$, at these prices it is a PE for the consumer not to buy.

A(iv). We now establish that if $\Delta < v - \underline{p}$, 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 $\underline{p}$ and weight $1 - \epsilon$ on

$$p_{z^*} = v + \frac{2\eta(\lambda - 1)\epsilon\underline{p}}{1 + \eta(\lambda - 1)\epsilon}.$$

It is easy to check that (either directly or using Equation 4.10 below) that for a sufficiently small ϵ buying at both prices is the PPE. Hence, with this

[20] 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 with probability 1.

pricing distribution the firm's revenue is:

$$(1-\epsilon)v + (1-\epsilon)\frac{2\eta(\lambda-1)}{1+\eta(\lambda-1)\epsilon}\epsilon\underline{p} + \epsilon\underline{p}. \tag{4.9}$$

For $\epsilon = 0$ the revenue is equal to v. Taking the derivative with respect to ϵ and evaluating it at $\epsilon = 0$ yields

$$-v + \underline{p}(2\eta(\lambda-1)+1) = \frac{\eta(\lambda-1)+2\eta^2(\lambda-1)}{1+\eta\lambda}\cdot v > 0.$$

A(v). 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 4.7 and 4.8, the highest p_{z^*} at which this holds is

$$p_{z^*} = \min\left\{\frac{v + \frac{\eta(\lambda-1)}{1+\eta\lambda}Q_zP_z}{1-\frac{\eta(\lambda-1)}{1+\eta\lambda}q_{z^*}}, v + \frac{2\eta(\lambda-1)Q_zP_z}{1+\eta(\lambda-1)Q_z}\right\} \tag{4.10}$$

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 4.7 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

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

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 $\Pr_F(p_z)<2A\Delta/(v+p_z)$. Suppose by contradiction that $\Pr_F(p_z)\geq 2A\Delta/(v+p_z)$. 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 $z'=z+1$, $z^{*\prime}=z^*+1$, with the distribution F' created from F by shifting up the weight $\Pr_F(p_z)-A\Delta/(v+p_z)$ 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 $Q'_{z'}P'_{z'}>Q_zP_z$, 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 2Δ. 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 $z^{*\prime}=z^*+1$, 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 $Q'_{z'}P'_{z'}>Q_zP_z$, 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 $\underline{p}-2\Delta<p_1\leq\underline{p}$. Clearly, if $p_1>\underline{p}$, there is a PE in which the consumer does not buy. We are left to show that $p_1>\underline{p}-2\Delta$. Suppose otherwise. Then, since $p_2-p_1<2\Delta$, we must have $p_2<\underline{p}$. 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 $Q'_{z'}P'_{z'}>Q_zP_z$, a contradiction.

A(vi). We are thus left to rule out that $Q_z=1$. Suppose, toward a contradiction, that for an optimal price distribution F, $Q_z=1$. By A(iv), $\sum_{l=1}^{z}q_lp_l>v$, and hence $p_z>v$. By A(v), p_z is such that if the consumer

expected to buy at prices up to p_{z-1}, she would just be indifferent between buying and not buying at p_z. Using Inequality 4.7, this means

$$p_z = \frac{1+\eta(1-Q_{z-1})+\eta\lambda Q_{z-1}}{1+\eta\lambda}v + \frac{\eta(\lambda-1)}{1+\eta\lambda}Q_{z-1}P_{z-1}. \tag{4.11}$$

Notice that the above implies

$$p_z < \frac{1+\eta(1-Q_z)+\eta\lambda Q_z}{1+\eta\lambda}v + \frac{\eta(\lambda-1)}{1+\eta\lambda}Q_zP_z = v + \frac{\eta(\lambda-1)}{1+\eta\lambda}P_z.$$

Now consider a price distribution F' obtained from F by moving the atom at p_z to some $p'_z > p_z$, while leaving all other atoms and all weights unchanged. By continuity of both sides in p'_z, the above inequality implies that for p'_z sufficiently close to p_z, it is a PE for the consumer to buy at all prices when facing F'. We now show that for p'_z sufficiently close to p_z, when facing F' the consumer prefers the PE of buying at all prices to the PE of buying only up to price p_{z-1}, so that always buying is the PPE. This completes the proof because F' yields higher revenue than F, contradicting the optimality of F.

Using Expression 4.6, we want to show that

$$\begin{aligned} v - P'_z - \eta(\lambda-1)\sum_{l'=1}^{z}\sum_{l''=1}^{l'} q_{l''}q_{l'}(p'_{l'}-p'_{l''}) &> Q_{z-1}v - Q_{z-1}P_{z-1} \\ - \eta(\lambda-1)Q_{z-1}(1-Q_{z-1})(v+P_{z-1}) &- \eta(\lambda-1)\sum_{l'=1}^{z-1}\sum_{l''=1}^{l'} q_{l''}q_{l'}(p'_{l'}-p'_{l''}). \end{aligned}$$

Rearranging, this is equivalent to

$$q_zv - q_zp'_z + \eta(\lambda-1)q_z(1-q_z)(v+P_{z-1}) - \eta(\lambda-1)q_z\left(\sum_{l''=1}^{z} q_{l''}(p'_z - p'_{l''})\right) > 0,$$

which, using that $p'_{l''} = p_{l''}$ for $l'' < z$ and therefore $\sum_{l''=1}^{z-1} q_{l''}p'_{l''} = Q_{z-1}P_{z-1}$, simplifies to

$$p'_z < v + \frac{2\eta(\lambda-1)}{1+\eta(\lambda-1)Q_{z-1}}Q_{z-1}P_{z-1}.$$

Because the right-hand side of this inequality is strictly greater than the right-hand side of Equation 4.11, this inequality must hold for p'_z sufficiently close to p_z.

Part B.

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 π^*. 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 4.10. 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 Δ apart, and the consumer does not buy for any price about $\overline{p}$, 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_z. In addition, by Equation 4.10, which is continuous in p_l and q_l, in PPE the consumer is willing to buy also at p_{z^*}. Hence, when facing F, the PPE is for the consumer to buy at all prices, so that the firm achieves profit level π^*—a contradiction. □

Proofs of Propositions in Text

Proof of Proposition 4.1.

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_l^n, q_l^n, and $Q_{z^n}^n$ analogously to Proposition 4.4. Assume first that $Q_{z^n}^n$ converges to some s; we will establish this below.

Trivially, as Δ decreases the optimal profits must weakly increase since the firm could always choose the same distribution as it did for a higher value of Δ. 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 4.4, 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 4.4, consider the distribution on $[\underline{p}, p_{max}]$ 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 $p_{max}(s)$

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

Since $\underline{p} - 2\Delta^n < p_1^n \leq \underline{p}$, we have $p_1^n \to \underline{p}$. We prove that $p_{z^n}^n \to p_{max}(s)$. We have

$$\begin{aligned} Q_{z^n}^n &= \sum_{l=1}^{z^n} q_l^n = q_{z^n}^n + A \sum_{l=1}^{z^n-1} \frac{p_{l+1}^n - p_l^n}{v + p_l^n} \\ &= q_{z^n}^n + A \sum_{l=1}^{z^n-1} \left[\int_{p_l^n}^{p_{l+1}^n} \frac{1}{v+p} dp + \underbrace{\int_{p_l^n}^{p_{l+1}^n} \left(\frac{1}{v+p_l^n} - \frac{1}{v+p} \right) dp} \right] \end{aligned} \tag{4.12}$$

We work on the sum of the underbraced term:

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

Notice that this is positive and (since $p_{l+1}^n - p_l^n < 2\Delta^n$) it is less than

$$\sum_{l=1}^{z^n-1} \int_{p_l^n}^{p_{l+1}^n} \frac{2\Delta^n}{(v+p_l^n)(v+p)} dp < \sum_{l=1}^{z^n-1} \frac{2(p_{l+1}^n - p_l^n)\Delta^n}{v^2} = \frac{2(p_{z^n}^n - p_1^n)\Delta^n}{v^2},$$

which approaches zero as $n \to \infty$. Taking the limit of Equation 4.12, plugging in that the sum of the underbraced terms approaches zero, and using that $q_{z^n}^n \to 0$ as $n \to \infty$, we get

$$s = \lim_{n\to\infty} A \int_{p_1^n}^{p_{z^n}^n} \frac{1}{v+p} dp = \lim_{n\to\infty} A \int_{\underline{p}}^{p_{z^n}^n} \frac{1}{v+p} dp.$$

This implies that $p_{z^n}^n \to p_{max}(s)$ as $n \to \infty$.

Next, we show that for a sufficiently large n, we have $p_{z^n+1}^n > p_{max}(s)$. We know that $p_{z_n}^n$ satisfies the condition that if the consumer expected to buy up to price $p_{z^n-1}^n$, she would just be indifferent to buying at $p_{z^n}^n$. This is equivalent to

$$\begin{aligned} p_{z^n}^n &= \frac{(1+\eta+\eta(\lambda-1)Q_{z^n-1}^n)v + \eta(\lambda-1)Q_{z^n-1}^n P_{z^n-1}^n}{1+\eta\lambda} \\ &\leq \frac{(1+\eta+\eta(\lambda-1)Q_{z^n}^n)v + \eta(\lambda-1)Q_{z^n}^n P_{z^n}^n}{1+\eta\lambda} \end{aligned}$$

Given that $p^n_{z^n} \to p_{max}(s)$ and $Q^n_{z^n} \to s < 1$, this and Equation 4.10 imply that for a sufficiently large n, we have $p^n_{z^n+1} > p_{max}(s)$.

Clearly, for any $x \leq \underline{p}$, $H(x) = \lim_{n\to\infty} F^n(x) = 0$. Now take any x satisfying $\underline{p} < x < p_{max}(s)$. So long as $p^n_{z^n} > x$, which holds for n sufficiently large, we have

$$\begin{aligned} F^n(x) &= \sum_{l, p^n_l \leq x} q^n_l = A \sum_{l, p^n_l \leq x} \frac{p^n_{l+1} - p^n_l}{v + p^n_l} \\ &= A \sum_{l, p^n_l \leq x} \left[\int_{p^n_l}^{p^n_{l+1}} \frac{1}{v+p} dp + \underbrace{\int_{p^n_l}^{p^n_{l+1}} \left(\frac{1}{v+p^n_l} - \frac{1}{v+p} \right) dp} \right]. \end{aligned} \tag{4.13}$$

By the same argument as above, the sum of the underbraced term approaches zero as $n \to \infty$, and we must have $\max_l \{p^n_l | p^n_l \leq x\} \to x$ as $n \to \infty$. Hence, taking the limit of Equation 4.13, we have

$$\lim_{n\to\infty} F^n(x) = A \int_{\underline{p}}^{x} \frac{1}{v+p} dp = H(x).$$

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

Next, notice that in order for F^n to converge in distribution, the sequence p_{z^n+1} must converge. Let the limit be p. Applying Equation 4.10, $p > p_{max}(s)$. We have shown that the limiting distribution has the properties in the proposition.

To conclude the proof, it remains to show that $Q^n_{z^n}$ 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 $Q^{n_1}_{z^{n_1}}$ and $Q^{n_2}_{z^{n_2}}$ 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 $[\underline{p}, p_{max}(s_1)]$ and has an isolated atom, while the limit of F^{n_2} is distributed continuously on $[\underline{p}, p_{max}(s_2)]$ and has an isolated atom. But this means that the sequence F^n does not converge in distribution, a contradiction. □

Proof of Proposition 4.2. From the proof of Proposition 4.4, for $\Delta < v - \underline{p}$ 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 4.3. We first show that firms earn zero expected profits in equilibrium. Suppose otherwise. Then, we can assume without loss of generality that firm 1 earns weakly lower profits than firm 2, so that firm 2 earns positive expected profit. To make positive expected profit, firm 2 must attract the consumer with positive probability. Let F_2 be firm 2's price distribution, with price atoms $p_1^2, \ldots, p_S^2$, and weight q_s^2 on p_s^2. Similarly to above, let p_z^2 be the highest price at which the consumer buys in any PE, and $p_{z^*}^2$ the highest price at which she buys with positive probability. We distinguish two cases.

First, suppose $p_{z^*}^2 = p_z^2$. Then, consider the price distribution F_1 obtained from F_2 by shifting all prices up to p_z^2 down by $\epsilon > 0$, and putting the rest of the weight on any price strictly greater than $(1+\eta\lambda)v/(1+\eta)$. Then, the consumer's unique PE when facing price distribution F_1 is to buy up to the price $p_z^2 - \epsilon$. In addition, the consumer's ex-ante expected utility is strictly higher under F_1 than under F_2, so that firm 1 attracts the consumer with probability 1. For a sufficiently small $\epsilon > 0$, this deviation therefore increases firm 1's expected profits, contradicting equilibrium.

Second, suppose that $p_{z^*}^2 > p_z^2$. Consider the price distribution F_1 obtained from F_2 by (i) keeping the prices and weights up to p_z^2 the same; (ii) replacing the prices $p_{z+1}^2, \ldots, p_{z^*}^2$ with a single atom at the average $p_{z+1}^1 = (q_{z+1}^2 p_{z+1}^2 + \cdots + q_{z^*}^2 p_{z^*}^2)/(q_{z+1}^2 + \cdots + q_{z^*}^2) - \epsilon$; (iii) putting the rest of the weight on any price strictly greater than $(1+\eta\lambda)v/(1+\eta)$. The consumer's unique PPE when facing the price distribution F_1 is to buy up to the price p_{z+1}^1, and in this PPE the consumer obtains ex-ante expected utility strictly greater than that when facing F_2. Hence, firm 1 attracts the consumer with probability 1. For a sufficiently small $\epsilon > 0$, this deviation therefore increases firm 1's expected profits, contradicting equilibrium. This completes the proof that firms earn zero expected profits in equilibrium.

To complete the proof, we show that firms charge a deterministic price of c in equilibrium. Observe that if the consumer goes to a firm that makes zero expected profits, the consumer's ex-ante expected utility is weakly lower than $v - c$: the consumer's expected consumption utility is at most $v - c$, and by Expression 4.6 her gain-loss utility is non-positive. Furthermore, if the consumer faces a stochastic price, by the same logic (and using that in that case she faces uncertainty either in the product or in the price dimension) her ex-ante expected utility is strictly lower than $v - c$.

Now suppose, toward a contradiction, that firm 2 charges a (non-degenerate) stochastic price. Then, the consumer cannot strictly prefer firm 1 to firm 2: if this was the case, then by the above observation the ex-ante expected utility the consumer would obtain from firm 2 would be strictly lower than $v - c$; but

in that case, firm 1 could make positive profits by offering a deterministic price of $c + \epsilon$ for a sufficiently small $\epsilon > 0$. Because the consumer does not strictly prefer firm 1, the consumer visits firm 2 with positive probability. Again by the above observation, then, firm 2 gives the consumer an ex-ante expected utility strictly below $v - c$. Hence, once again firm 1 has a profitable deviation, a contradiction. □

Bibliography

Abeler, Johannes, Armin Falk, Lorenz Götte, and David Huffman, "Reference Points and Effort Provision," *American Economic Review*, 2011, *101* (2), 470–492.

Aghion, Philippe and Mark Schankerman, "On the Welfare Effects and Political Economy of Competition-Enhancing Policies," *Economic Journal*, 2004, *114* (498), 800–824.

Ainslie, George, *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*, Cambridge, UK: Cambridge University Press, 1992.

Ali, S. Nageeb, "Learning Self-Control," *Quarterly Journal of Economics*, 2011, *126* (2), 857–893.

Amador, Manuel, George-Marios Angeletos, and Iván Werning, "Commitment vs. Flexibility," *Econometrica*, 2006, *74* (2), 365–396.

Anderson, Eric T. and Duncan I. Simester, "Price Stickiness and Customer Antagonism," *Quarterly Journal of Economics*, 2010, *125* (2), 729–765.

Asheim, Geir B., "Procrastination, Partial Naivete, and Behavioral Welfare Analysis," 2008. Working Paper, University of Oslo.

Athey, Susan, Kyle Bagwell, and Chris Sanchirico, "Collusion and Price Rigidity," *Review of Economic Studies*, April 2004, *71* (2), 317–349.

Ausubel, Lawrence M., "The Failure of Competition in the Credit Card Market," *American Economic Review*, 1991, *81* (1), 50–81.

——, "Credit Card Defaults, Credit Card Profits, and Bankruptcy," *American Bankruptcy Law Journal*, 1997, *71* (2), 249–270.

——, "Adverse Selection in the Credit Card Market," 1999. Working Paper, University of Maryland.

Bar-Gill, Oren, "Seduction by Plastic," *Northwestern University Law Review*, 2004, *98* (4), 1373–1434.

——, "The Law, Economics and Psychology of Subprime Mortgage Contracts," 2008. NYU Law and Economics Research Paper No. 08-59.

Barberis, Nicholas, Ming Huang, and Tano Santos, "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 2001, *116* (1), 1–53.

Barberis, Nick, Ming Huang, and Richard H. Thaler, "Individual Preferences, Monetary Gambles and Stock Market Participation: A Case of Narrow Framing," *American Economic Review*, 2006, *96* (4), 1069–1090.

Baye, Michael R., Dan Kovenock, and Casper G. de Vries, "It Takes Two to Tango: Equilibria in a Model of Sales," *Games and Economic Behavior*, 1992, *4* (4), 493–510.

—— **, John Morgan, and Patrick Scholten**, "Price Dispersion in the Large and in the Small: Evidence from an Internet Price Comparison Site," *Journal of Industrial Economics*, 2004, *52* (4), 463–496.

Beck, Jonathan, "Fixed, Focal, Fair? Book Prices with Optional Resale Price Maintenance," 2004. WZB Discussion Paper SP II 2004-15.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 1964, *9*, 226–232.

Benartzi, Shlomo and Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 1995, *110* (1), 73–92.

Bernheim, B. Douglas and Antonio Rangel, "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 2004, *94* (5), 1558–1590.

—— **and** ——, "Saving and Cue-Triggered Decision Processes," 2004. Mimeo, Stanford University.

Bester, Helmut, "Screening vs. Rationing in Credit Markets with Imperfect Information," *American Economic Review*, 1985, *75* (4), 850–855.

——, "The Role of Collateral in Credit Markets with Imperfect Information," *European Economic Review*, 1987, *31* (4), 887–899.

Bils, Mark and Peter J. Klenow, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 2004, *112* (5), 947–985.

Blinder, Alan S., Elie R. D. Canetti, David E. Lebow, and Jeremy B. Rudd, *Asking About Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation, 1998.

Bolton, Patrick and Mathias Dewatripont, *Contract Theory*, MIT Press, 2005.

Camerer, Colin F., Linda Babcock, George Loewenstein, and Richard H. Thaler, "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics*, 1997, *112* (2), 407–441.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism'," *University of Pennsylvania Law Review*, 2003, *151* (3), 1211–1254.

Caplin, Andrew and John Leahy, "Psychological Expected Utility and Anticipatory Feelings," *Quarterly Journal of Economics*, 2001, *116* (1), 55–79.

Carlton, Dennis W., "The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?," in Richard Schmalensee and R.D. Willig, eds., *Handbook of Industrial Organization*, Vol. 1 Elsevier Science Publishers B.V. 1989, pp. 909–946.

Cecchetti, Stephen G., "The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines," *Journal of Econometrics*, 1986, *31* (3), 255–274.

Chakravorti, Sujit and Alpa Shah, "A Study of the Interrelated Bilateral Transactions in Credit Card Networks," 2001. FRB Chicago Policy Studies No. EPS-2001-2.

—— **and William R. Emmons**, "Who Pays for Credit Cards?," *Journal of Consumer Affairs*, 2003, *37* (2), 208–230.

Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi, "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 2003, *93* (1), 15–37.

Competition Commission of the United Kingdom, "The Supply of Recorded Music: A Report on the Supply in the UK of Prerecorded Compact Discs, Vinyl Discs and Tapes Containing Music," URL: http://www.competition-commission.org.uk/rep_pub/reports/1994/356recordedmusic.htm#full 1994.

Conlisk, John, Eitan Gerstner, and Joel Sobel, "Cyclic Pricing by a Durable Goods Monopolist," *Quarterly Journal of Economics*, 1984, *99* (3), 489–505.

Courty, Pascal and Hao Li, "Sequential Screening," *Review of Economic Studies*, 2000, *67* (4), 697–717.

Crawford, Vincent P. and Juanjuan Meng, "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *American Economic Review*, 2011, *101* (5), 1912–1932.

Davidson, Carl and Raymond Deneckere, "Long-Run Competition in Capacity, Short-Run Competition in Price, and the Cournot Model," *RAND Journal of Economics*, 1986, *17* (3), 404–415.

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 2009, *47* (2), 315–372.

—— **and Ulrike Malmendier**, "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Demyanyk, Yuliya and Otto Van Hemert, "Understanding the Subprime Mortgage Crisis," 2008. Working Paper, `http://ssrn.com/abstract=1020396`.

Dubey, Pradeep and John Geanakoplos, "Competitive Pooling: Rothschild-Stiglitz Reconsidered," *Quarterly Journal of Economics*, 2002, *117* (4), 1529–1570.

Durkin, Thomas A., "Credit Cards: Use and Consumer Attitudes, 1970-2000," Federal Reserve Bulletin 2000.

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo, "Reference Prices, Costs and Nominal Rigidities," 2009. Working Paper.

Einav, Liran and Barak Y. Orbach, "Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry," *International Review of Law and Economics*, 2007, *27* (2), 129–153.

Eliaz, Kfir and Ran Spiegler, "Contracting with Diversely Naive Agents," *Review of Economic Studies*, 2006, *73* (3), 689–714.

Ellison, Glenn, "A Model of Add-On Pricing," *Quarterly Journal of Economics*, 2005, *120* (2), 585–637.

——, "Bounded Rationality in Industrial Organization," *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, 2006, *2*, 142–174.

Engel, Kathleen C. and Patricia A. McCoy, "A Tale of Three Markets: The Law and Economics of Predatory Lending," *Texas Law Review*, 2002, *80* (6), 1255–1381.

Erickson, Gary M. and Johny K. Johansson, "The Role of Price in Multi-attribute Product Evaluations," *Journal of Consumer Research*, 1985, *12* (2), 195–199.

Ericson, Keith M. Marzilli and Andreas Fuster, "Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior," 2009. Working Paper, Harvard University.

Estelami, Hooman and Donald R. Lehmann, "The Impact of Research Design on Consumer Price Recall Accuracy: An Integrative Review," *Journal of the Academy of Marketing Science*, 2001, *29* (1), 36–49.

Farber, Henry S., "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," *Journal of Political Economy*, 2005, *113* (1), 46–82.

——, "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers," *American Economic Review*, 2008, *98* (3), 1069–1082.

Farrell, Joseph and Paul Klemperer, "Coordination and Lock-In: Competition with Switching Costs and Network Effects," in Mark Armstrong

and Rob Porter, eds., *Handbook of Industrial Organization*, Vol. 3, North-Holland, 2007.

Frederick, Shane and George Loewenstein, "Hedonic Adaptation," in Daniel Kahneman, Ed Diener, and Norbert Schwarz, eds., *Well-Being: The Foundations of Hedonic Psychology*, New York, NY, US: Russell Sage Foundation, 1999, pp. 302–329.

——, ——, and Ted O'Donoghue, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 2002, *40* (2), 351–401.

Fudenberg, Drew and David K. Levine, "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 1989, *57* (4), 759–778.

Gabaix, Xavier and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

Gal-Or, Esther, "Hotelling's Spatial Competition as a Model of Sales," *Economics Letters*, 1982, *9* (1), 1–6.

Genesove, David, "The Nominal Rigidity of Apartment Rents," *Review of Economics and Statistics*, 2003, *85* (4), 844–853.

Gottlieb, Daniel, "Competition over Time-Inconsistent Consumers," *Journal of Public Economic Theory*, 2008, *10* (4), 673–684.

Grubb, Michael D., "Selling to Overconfident Consumers," *American Economic Review*, 2009, *99* (5), 1770–1805.

Gruber, Jonathan and Botond Kőszegi, "Tax Incidence when Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes," *Journal of Public Economics*, 2004, *88* (9-10), 1959–1987.

Hafalir, Elif Incekara, "Credit Card Competition and Naive Hyperbolic Consumers," 2008. Working Paper, Carnegie Mellon University.

Hall, R. L. and C . J. Hitch, "Price Theory and Business Behavior," *Oxford Economic Papers*, 1939, *2*, 12–45.

Hardie, Bruce G. S., Eric J. Johnson, and Peter S. Fader, "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science*, 1993, *12* (4), 378–394.

Hastings, Justine and Ebonya Washington, "The First of the Month Effect: Consumer Behavior and Store Responses," *American Economic Journal: Economic Policy*, 2010, *2* (2), 142–62.

Hayhoe, Celia Ray, Lauren J. Leach, Pamela R. Turner, Marilyn J. Bruin, and Frances C. Lawrence, "Differences in Spending Habits and Credit Use of College Students," *Journal of Consumer Affairs*, 2000, *34* (1), 113–133.

Heidhues, Paul and Botond Kőszegi, "New Evidence on Focal Pricing," 2007. In Preparation.

___ **and** ___ , "Competition and Price Variation When Consumers Are Loss Averse," *American Economic Review*, 2008, *98* (4), 1245–1268.

___ **and** ___ , "Exploiting Naivete about Self-Control in the Credit Market," *American Economic Review*, 2010, *100* (5), 2279–2303.

___ **and** ___ , "Regular Prices and Sales," 2010. ESMT Working Paper 10-008.

Herweg, Fabian and Konrad Mierendorff, "Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs," *Journal of the European Economic Association*, forthcoming.

___ **, Daniel Müller, and Philipp Weinschenk**, "Binary Payment Schemes: Moral Hazard and Loss Aversion," *American Economic Review*, 2010, *100* (5), 2451–2477.

Hill, Ronald Paul and John C. Kozup, "Consumer Experiences with Predatory Lending Practices," *Journal of Consumer Affairs*, 2007, *41* (1), 29–46.

Huffman, David and Matias Barenstein, "A Monthly Struggle for Self-Control? Hyperbolic Discounting, Mental Accounting, and the Fall in Consumption Between Paydays," 2005. Working Paper, IZA.

Jacobson, Robert and Carl Obermiller, "The Formation of Expected Future Price: A Reference Price for Forward-Looking Consumers," *Journal of Consumer Research*, 1990, *16* (4), 420–432.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

____ **, Jack L. Knetsch, and Richard H. Thaler**, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 1990, *98* (6), 1325–1348.

____ **,** ____ **, and** ____ , "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *Journal of Economic Perspectives*, 1991, *5* (1), 193–206.

Kalwani, Manohar U. and Chi Kin Yim, "Consumer Price and Promotion Expectations: An Experimental Study," *Journal of Marketing Research*, 1992, *29*, 90–100.

Kashyap, Anil K., "Sticky Prices: New Evidence from Retail Catalogs," *Quarterly Journal of Economics*, 1995, *110* (1), 245–274.

Kehoe, Patrick J. and Virgiliu Midrigan, "Temporary Price Changes and the Real Effects of Monetary Policy," 2008. Federal Reserve Bank of Minneapolis Working Paper #661.

Kőszegi, Botond, "On the Feasibility of Market Solutions to Self-Control Problems," *Swedish Economic Policy Review*, 2005, *12* (2), 71–94.

____ , "Utility from Anticipation and Personal Equilibrium," *Economic Theory*, 2010, *44* (3), 415–444.

____ **and Matthew Rabin**, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1166.

____ **and** ____ , "Reference-Dependent Risk Attitudes," *American Economic Review*, 2007, *97* (4), 1047–1073.

____ **and** ____ , "Reference-Dependent Consumption Plans," *American Economic Review*, 2009, *99* (3), 909–936.

Klenow, Peter J. and Oleksiy Kryvtsov, "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?," *Quarterly Journal of Economics*, 2008, *123* (3), 863–904.

Knittel, Christopher R. and Victor Stango, "Price Ceilings as Focal Points for Tacit Collusion: Evidence from Credit Cards," *American Economic Review*, 2003, *93* (5), 1703–1729.

Laffont, Jean-Jacques and David Martimort, *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press, 2001.

Laibson, David I., "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

—— **and Leeat Yariv**, "Safety in Markets: An Impossibility Theorem for Dutch Books," 2007. Working Paper, Caltech.

—— **, Andrea Repetto, and Jeremy Tobacman**, "Estimating Discount Functions with Consumption Choices over the Lifecycle," 2007. NBER Working Paper #13314.

MacDonald, James M. and Daniel Aaronson, "How Firms Construct Price Changes: Evidence from Restaurant Responses to Increased Minimum Wages," *American Journal of Agricultural Economics*, 2006, *88* (2), 292–307.

Macera, Rosario, "Intertemporal Incentives Under Loss Aversion," 2012. Working Paper.

Maskin, Eric and Jean Tirole, "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica*, 1988, *56* (3), 571–599.

McMillan, Robert S., "Estimating Demand for Differentiated Products when Choice is Continuous and Complementarities Exist: An Application to the Puzzle of Uniform Pricing." PhD dissertation, Stanford University August 2004.

Meier, Stephan and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

Mirrlees, James A., "The Theory of Moral Hazard and Unobservable Behaviour: Part I," *Review of Economic Studies*, 1999, *66* (1), 3–21. Originally published as a working paper in 1975.

Nakamura, Emi and Jón Steinsson, "Price Setting in Forward-Looking Customer Markets," 2009. Working Paper.

Novemsky, Nathan and Daniel Kahneman, "The Boundaries of Loss Aversion," *Journal of Marketing Research*, 2005, *42* (2), 119–128.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

___ **and** ___ , "Choice and Procrastination," *Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

___ **and** ___ , "Optimal Sin Taxes," *Journal of Public Economics*, 2006, *90* (10-11), 1825–1849.

Pesendorfer, Martin, "Retail Sales: A Study of Pricing Behavior in Supermarkets," *Journal of Business*, 2002, *75* (1), 33–66.

Rabin, Matthew, "Psychology and Economics," *Journal of Economic Literature*, 1998, *36* (1), 11–46.

___ , "Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion," in Daniel Kahneman and Amos Tversky, eds., *Choices, Values, and Frames*, New York: Russell Sage Foundation, 2000, chapter 11, pp. 202–208.

___ , "Risk Aversion and Expected-utility Theory: A Calibration Theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

___ **and Richard Thaler**, "Anomalies: Risk Aversion," *Journal of Economic Perspectives*, 2001, *15* (1), 219–232.

Reda, Susan, "2003 Consumer Credit Survey," *Stores Magazine*, November 2003.

Renuart, Elizabeth, "An Overview of the Predatory Mortgage Lending Process," *Housing Policy Debate*, 2004, *15* (3), 467–502.

Rotemberg, Julio J., "Customer Anger at Price Increases, Changes in the Frequency of Price Adjustment and Monetary Policy," *Journal of Monetary Economics*, 2005, *52* (4), 829–852.

___ , "Altruistic Dynamic Pricing with Customer Regret," *Scandinavian Journal of Economics*, 2010, *112* (4), 646–672.

___ , "Fair Pricing," *Journal of the European Economic Association*, 2011, *9* (5), 952–981.

—— **and Michael Woodford**, "The Cyclical Behavior of Prices and Costs," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1B, Amsterdam; New York and Oxford: Elsevier Science, North-Holland, 1999, pp. 1051–1135.

Rothschild, Michael and Joseph E. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 1976, *90* (4), 630–649.

Salop, Steven C., "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 1979, *10* (1), 141–156.

Shapiro, Jesse M., "Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle," *Journal of Public Economics*, 2005, *89* (2-3), 303–325.

Shilony, Yuval, "Mixed Pricing in Oligopoly," *Journal of Economic Theory*, 1977, *14*, 373–388.

Shui, Haiyan and Lawrence M. Ausubel, "Time Inconsistency in the Credit Card Market," 2004. Available at http://ssrn.com/abstract=586622.

Sibly, Hugh, "Loss averse customers and price inflexibility," *Journal of Economic Psychology*, August 2002, *23* (4), 521–538.

Skiba, Paige Marta and Jeremy Tobacman, "Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default," 2008. Vanderbilt Law and Economics Research Paper No. 08-33.

Slade, Margaret, "Sticky Prices in a Dynamic Oligopoly: An Investigation of (s, S) Thresholds," *International Journal of Industrial Organization*, 1999, *17* (4), 477–511.

Smith, Alec, "Lagged Beliefs and Reference-Dependent Utility," 2008. Working Paper, University of Arizona.

Sobel, Joel, "The Timing of Sales," *Review of Economic Studies*, July 1984, *51* (3), 353–368.

—— , "Durable Goods Monopoly with Entry of New Consumers," *Econometrica*, 1991, *59* (5), 1455–1485.

Spiegler, Ran, "Consumer Bounded Rationality and Rigidity/Flexibility Retail Price Patterns," 2012. Working Paper.

Stiglitz, Joseph E., "Competition and the Number of Firms in a Market: Are Duopolies More Competitive than Atomistic Markets?," *Journal of Political Economy*, 1987, *95* (5), 1041–1061.

—— **and Andrew Weiss**, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 1981, *71* (3), 393–410.

Sunstein, Cass and Richard Thaler, "Libertarian Paternalism," *American Economic Review*, 2003, *93* (2), 175–179.

Sweezy, P., "Demand Conditions under Oligopoly," *Journal of Political Economy*, 1939, *47*, 568–573.

Sydnor, Justin, "(Over)insuring Modest Risks," *American Economic Journal: Applied Economics*, 2010, *2* (4), 177–99.

Thaler, Richard, "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior & Organization*, 1980, *1* (1), 39–60.

—— , "Some Empirical Evidence on Dynamic Inconsistency," *Economics Letters*, 1981, *8*, 201–207.

Tversky, Amos and Daniel Kahneman, "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 1991, *106* (4), 1039–1061.

—— **and** —— , "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–324.

Varian, Hal R., "A Model of Sales," *American Economic Review*, 1980, *70* (4), 651–659.

Warren, Elizabeth, "Examining the Billing, Marketing, and Disclosure Practices of the Credit Card Industry, and Their Impact on Consumers," Testimony before the Committee on Banking, Housing and Urban Affairs of the United States Senate; January 2007.

Winer, Russell S., "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 1986, *13* (2), 250–256.