American Economic Review 100 (December 2010): 2279–2303
http://www.aeaweb.org/articles.php?doi=10.1257/aer.100.5.2279

Exploiting Naïvete about Self-Control in the Credit Market

By Paul Heidhues and Botond Kőszegi*

We analyze contract choices, loan-repayment behavior, and welfare in a model of a competitive credit market when borrowers have a taste for immediate gratification. Consistent with many credit cards and subprime mortgages, for most types of nonsophisticated borrowers the baseline repayment terms are cheap, but they are also inefficiently front loaded and delays require paying large penalties. Although credit is for future consumption, nonsophisticated consumers overborrow, pay the penalties, and back load repayment, suffering large welfare losses. Prohibiting large penalties for deferring small amounts of repayment—akin to recent regulations in the US credit-card and mortgage markets—can raise welfare. (*JEL* D14, D18, D49, D86)

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.[1] 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 naïvely underestimate the extent of this taste.[2] Yet the formal relationship between a taste for immediate gratification and consumer behavior and welfare in the credit market remains largely unexplored and unclear. Existing work on contracting with time inconsistency (DellaVigna and Ulrike Malmendier 2004; Botond Kőszegi 2005;

* Heidhues: ESMT European School of Management and Technology GmbH, Schlossplatz 1, 10178 Berlin, Germany (e-mail: paul.heidhues@esmt.org); Kőszegi: University of California, Berkeley, Department of Economics, 508-1 Evans Hall #3880, Berkeley, CA 94720 (e-mail: botond@econ.berkeley.edu). First version: November 2007. We thank Stefano DellaVigna, Ted O'Donoghue, Arthur Fishman, Marina Halac, Dwight Jaffee, Ulrich Kamecke, Sebastian Kranz, Tymofiy Mylovanov, Georg Nöldeke, Matthew Rabin, and Tymon Tatur for very helpful discussions, and two anonymous referees and audiences at the AEA Meetings in San Francisco, Behavioral Models of Market Competition Conference in Bad Homburg, Berkeley, Bielefeld, Bocconi, Central Bank of Hungary, Chicago Booth School of Business, Cornell, Düsseldorf, ECARES, the ENABLE Conference in Amsterdam, Groningen, Heidelberg, Helsinki School of Economics, the HKUST Industrial Organization Conference, Hungarian Society for Economics Annual Conference, ITAM, UCL, LSE, Maastricht, Mannheim, Michigan, the Network of Industrial Economists Conference at Oxford, NYU Stern School of Business, the SFB/TR meeting in Gummersbach, Vienna/IHS, Yale, and Zürich for comments. Heidhues gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft through SFB/TR-15. Kőszegi thanks the National Science Foundation for financial support under Award #0648659.

[1] See, for instance, Lawrance M. Ausubel (1997), Thomas A. Durkin (2000), Kathleen C. Engel and Patricia A. McCoy (2002), Oren Bar-Gill (2004), Elizabeth Warren (2007), and Bar-Gill (2008).

[2] David I. Laibson, Andrea Repetto, and Jeremy 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, Stephan Meier and Charles Sprenger (2009) 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, Repetto, and Tobacman (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 naïvete 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 (Haiyan Shui and Ausubel 2004) and the post-introductory interest rate (Ausubel 1999), suggesting that they end up borrowing more than they intended or expected. Paige Marta Skiba and Tobacman (2007) 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 naïve time inconsistency. For further discussions as well as evidence for a taste for immediate gratication in other domains, see Stefano DellaVigna (2009).

Kfir Eliaz and Ran 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 nonsophisticated 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 nonsophisticated 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 I 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 $-k(q) - \beta k(r)$ 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 nonsophisticated borrower's contract choice. To allow for self 0 to be overoptimistic regarding her future self-control, we follow Ted O'Donoghue and Matthew 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 nonlinear 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 II 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 nonsophisticated 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 nonsophisticated 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 nonsophisticated 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, nonsophisticated but not-too-naïve 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 naïve, 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 nonsophisticated 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 III, 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 nonsophisticated 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 nonsophisticated borrower chooses. As before, even if a nonsophisticated 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 nonsophisticated borrowers, if nonsophisticated borrowers are not too naïve, then the restricted market has higher total welfare.

In Section IV, we generalize our basic model—in which a nonsophisticated borrower believes with certainty that her taste for immediate gratification is above β—as well as other existing models of partial naïvete 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 nonsophisticated 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 V, 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 of refinancing is inefficient. If borrowers do not know at the time of contracting whether they can repay fast, a model of sequential screening (Pascal Courty and Hao Li 2000) or postcontractual 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 VI, 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 nonsophisticated borrowers will accept, and whether and how borrowers might learn about their time inconsistency. Proofs are in the Web Appendix.

I. A Model of the Credit Market

A. *Set-up*

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.[3] 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,[4] 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.

[3] 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 III demonstrates that when β is unknown and $k'(0)$ is sufficiently low, the bounds are not binding.

[4] 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 (Celia Ray Hayhoe et al. 2000, Susan Reda, "2003 Consumer Credit Survey." *Stores Magazine*, November.)

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 Ted O'Donoghue and Matthew Rabin's (2001) formulation of partial naïvete, we assume that self 0 believes with certainty that self 1 will maximize $-k(q) - \hat{\beta}k(r)$, 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 naïvete about the time inconsistency, and more generally $\hat{\beta}$ is a measure of sophistication. Because the O'Donoghue-Rabin specification of partial naïvete using degenerate beliefs is special, in Section IV we allow borrower beliefs to be any distribution, and show that so long as a nonsophisticated borrower attaches nontrivial 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 IV 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.[5] 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 nonlinear 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.[6] An *unrestricted credit contract* is a consumption level c along with a 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.

[5] By standard indicators of competitiveness, the subprime loan origination market seems quite competitive: no participant has more than 13 percent 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 IIIB's and Section IV's results on the sorting of consumers according to their beliefs in period 0 do take advantage of our competitiveness assumption.

[6] While the effects of relaxing exclusivity warrant further research, in general it would not eliminate our main points regarding nonsophisticated 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 (2001) and is consistent with evidence in Haiyan Shui and Ausubel (2004), in a more realistic, long-horizon setting nonsophisticated borrowers may procrastinate for a long time before finding or taking advantage of favorable refinancing opportunities. And even if a nonsophisticated 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.

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.[7] Since a borrower's behavior in period 0 can depend only on $\hat{\beta}$, the competitive equilibrium will be a set of contracts $\{(c_i, \mathcal{C}_i)\}_{i\in\{2,\dots,I\}}$ for the possible $\hat{\beta}$ types β_2 through β_I.[8] 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 1: *The maps $q_i, r_i : \{\beta_1, \dots, \beta_I\} \to \mathbb{R}_+$ are jointly* incentive compatible *for $\mathcal{C}_i$ if $(q_i(\beta), r_i(\beta)) \in \mathcal{C}_i$ for each $\beta \in \{\beta_1, \dots, \beta_I\}$, 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: *A* competitive equilibrium *is a set of contracts $\{(c_i, \mathcal{C}_i)\}_{i\in\{2,\dots,I\}}$ and incentive-compatible maps $(q_i(\cdot), r_i(\cdot))$ for each $\mathcal{C}_i$ with the following properties:*

1. *[Borrower optimization] For any $\hat{\beta} = \beta_i \in \{\beta_2, \dots, \beta_I\}$ and $j \in \{2, \dots, I\}$, 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.[9]

The last, nonredundancy, 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

[7] This approach is similar in spirit to Michael Rothschild and Joseph E. Stiglitz's (1976) definition of competitive equilibrium with insurance contracts. By thinking of borrowers as sellers of repayment schedules C, lenders as buyers of these schedules, and c as the price of a schedule C, we can modify Pradeep Dubey and John Geanakoplos's (2002) competitive-equilibrium framework for our setting in a way that yields the same contracts as Definition 2.

[8] 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.

[9] 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.

not affect any of our predictions regarding outcomes and welfare.[10] Due to the nonredundancy condition, the competitive-equilibrium contracts we derive exclude most options by assumption; in particular, nonsophisticated 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 nonlinear 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.

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; Jonathan Gruber and Kőszegi 2002; O'Donoghue and Rabin 2006, for example) and identify welfare with long-run, period-0 preferences.[11] 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.[12] 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 nonsophisticated 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 V) the large penalties are unlikely to be serving a neoclassical purpose, and we are also unaware of unmodeled "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 $\{(q, r) \mid q + r/R = L \text{ and } q, r \leq M\}$, where M is an 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,

[10] For general distributions of β and $\hat{\beta}$, our definition of nonredundancy would have to be more inclusive. Specifically, it would have to allow for a repayment schedule in C_i to be the expected choice from C_i of a consumer type not choosing (c_i, C_i)—because such an option could play a role in *preventing* the consumer from choosing (c_i, 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 IIIB already fully sorts consumers according to $\hat{\beta}$.

[11] 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 B. Douglas Bernheim and Antonio Rangel (2004a, 2004b) can be interpreted as saying that a taste for immediate gratification is often a mistake not reflecting true welfare.

[12] 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 Manuel Amador, George-Marios Angeletos, and Iván Werning (2006) show, however, this intervention is suboptimal if consumers are subject to ex post shocks in their financial circumstances.

and for which we require $k'(M) > 1/\beta$.[13] As we note below, many other ways of eliminating disproportionately large penalties have the same or similar welfare effect.

B. *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 nonsophisticated $(I = 2, p_2 < 1)$. 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 nonredundant 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 nonsophisticated 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 nonsophisticated 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:

LEMMA 1: *Suppose $\hat{\beta}$ is known $(I = 2)$, the possible βs 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.*

II. Nonlinear Contracting with Known β and $\hat{\beta}$

We begin our analysis of nonlinear contracting with the case when both β and $\hat{\beta}$ are known. We show that nonsophisticated 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.

[13] 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.

A. *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 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 IIC), 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 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 nonsophisticated borrower, for whom $\hat{\beta} > \beta$. Applying Lemma 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

$$\max_{c, q, r, \hat{q}, \hat{r}} q + r - c \tag{1}$$

$$\text{such that } c - k(\hat{q}) - k(\hat{r}) \geq \underline{u}, \tag{PC}$$

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

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

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

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

$$\text{such that } c - k(\hat{q}) - k(\hat{r}) = \underline{u}, \tag{PC}$$

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

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

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

and gives the following proposition:

PROPOSITION 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 nonsophisticated 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 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 nonsophisticated 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.[14] Unlike 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 trade-off 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 nonsophisticated 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

[14] 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 naïve decisionmakers borrow more or less than sophisticated ones depends on the per-period utility function. In our setting, nonsophisticated consumers borrow more for any $k(\cdot)$.

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 nonsophisticated 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.[15] 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.[16]

The properties of the nonsophisticated 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.[17] 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

[15] 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 nonredundancy condition in Definition 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.

[16] 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 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.

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

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.[18] As emphasized by Ronald Paul Hill and John C. Kozup (2007) and especially Elizabeth 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 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 are obtained for refinancing into a larger new loan for the purposes of extracting cash.[19]

B. *A Welfare-Increasing Intervention*

Given nonsophisticated 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 3: *A sophisticated borrower* $(\hat{\beta} = \beta)$ *is equally well off in the restricted and unrestricted markets. If a nonsophisticated 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 nonsophisticated but not-too-naïve borrowers into discretely mispredicting their behavior, and hence raises these borrowers' welfare. For any interest rate R, a slightly naïve 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.

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

[18] Yuliya S. Demyanyk and Otto Van Hemert (2008) report that 54.5 percent of US subprime mortgages postulated a prepayment penalty.

[19] 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.

(Colin Camerer et al. 2003): it greatly benefits nonsophisticated 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 3 still holds if we allow contracts with a "focal" installment plan $\overline{q}, \overline{r}$ and a higher interest rate when repaying less than $\overline{q}$ 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 United States are interpretable in terms of Proposition 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 apply for only 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 nonsophisticated but not-too-naïve consumers' overborrowing, so if there is a nontrivial 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.[20] Our model predicts that these opponents may well be right in predicting a decreased amount of credit, but also says that inasmuch as this happens, it will benefit rather than hurt consumers—because consumers were borrowing too much to start with.[21]

Proposition 3 holds in general only for sufficiently sophisticated borrowers because both restricted and unrestricted contracts can lead a very naïve borrower to severely overestimate how much she will be willing to pay back in period 1. If many consumers are very naïve 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

[20] 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.

[21] If we relax the simplifying assumption that $k'(0) < \beta$, the exclusion from the market mentioned above occurs in our model for a nonsophisticated but not-too-naïve 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.

commonly used utility functions, in fact, nonsophisticated borrowers are better off in a restricted market with an interest-rate cap of even zero than in an unrestricted market:[22]

PROPOSITION 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, nonsophisticated 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 nonsophisticated borrowers, because it 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 nonsophisticated borrowers in the population.

C. *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 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 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 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

[22] These utility functions guarantee that with linear contracts, nonsophisticated 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 4.

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.[23]

Although a nonsophisticated 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 III, 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.

III. Nonlinear 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 II survive. First, because sophisticated and nonsophisticated 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.

A. *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 nonsophisticated borrowers have the same beliefs in period 0, they accept the same contract. The following proposition identifies key features of this contract.

PROPOSITION 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*

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

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

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

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), 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 (3), the nonsophisticated 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 trade-off 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 trade-off 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 trade-off takes the period-0 perspective. But since a nonsophisticated 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 trade-off takes the period-1 perspective.

The difference between the sophisticated and nonsophisticated 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 nonsophisticated 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 nonsophisticated 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 nonsophisticated borrowers are not too naïve, eliminating disproportionately large penalties for deferring small amounts of repayment is still welfare improving:

PROPOSITION 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 nonsophisticated 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 nonsophisticated borrowers are not too naïve, 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 nonsophisticated 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 Xavier Gabaix and Laibson (2006): because nonsophisticated borrowers are more profitable, in a competitive equilibrium it must be that firms make money on nonsophisticated 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 nonsophisticated types in the population.

B. *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 IIIA 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 nonsophisticated 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 IIIA, 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 1: $\underline{u}_i$ is increasing in β_i.

Condition 1 states that if $\hat{\beta}$ were 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 naïve about it. Since firms compete more fiercely for such profitable borrowers, they drive up the utility of sophisticated borrowers.[24]

We argue that under Condition 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 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 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 naïve, this TV will be too nice, and she will fail to pay it off.

[24] Condition 1 is clearly nonempty. 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 nontrivial probability of being nonsophisticated, Condition 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 nonsophisticated with sufficiently higher probability.

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

PROPOSITION 8: (Period-0 Screening). *Suppose Condition* 1 *holds. Then, in the unique competitive equilibrium with* $\hat{\beta}$ *unobserved, each consumer accepts the same contract as when* $\hat{\beta}$ *is observed.*

IV. 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 naïvete 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 straightforward 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 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*

$$k'(q) = 1; \qquad k'(r) = \frac{1}{F(\beta) + (1 - F(\beta))\beta}. \tag{4}$$

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 9 generalizes many of the central points regarding outcomes and welfare we have made in this paper. In particular, nonsophisticated 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.[25]

Equation (4) in the proposition also clarifies that the extent to which a nonsophisticated 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

[25] 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.

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, 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 $F_n \rightarrow F^*$ in distribution (or, equivalently, $F_n \rightarrow F^*$ in probability), and this statement does *not* imply that $F_n(\beta) \rightarrow F^*(\beta) = 1$. In fact, this implication seems extremely special, especially for sequences that approach F^* from the direction of overoptimistic beliefs.

Intuitively, a nonsophisticated borrower has much lower utility than a sophisticated borrower if she assigns a nontrivial 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 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 Geir B. 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 naïvete, 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 naïvete, 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 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.[26]

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 II and III, such an intervention benefits near-sophisticated borrowers with $F(\beta)$ nontrivially 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.[27]

V. Related Literature

A. *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), Daniel Gottlieb (2008), and Elif Incekara Hafalir (2008) on specific contract forms and Kőszegi (2005) and Eliaz and Spiegler (2006) on general nonlinear 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 nonsophisticated 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 nonsophisticated 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

[26] 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.

[27] As we have discussed in Section II, if many consumers are very naïve 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.

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 naïvete 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 naïve 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 naïve consumers, they may not have an incentive to debias the latter ones. While both forms of naïvete 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 nonsophisticated borrowers receive—an aspect that seems very realistic for credit markets.

Michael D. Grubb (2007) 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 (2007), 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.

B. *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 by 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 postcontractual hidden knowledge (Jean-Jacques Laffont and David 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 to 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.[28]

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 Andrew Weiss (1981) and Helmut 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 appear only 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; Sujit Chakravorti and William R. Emmons 2003). In fact, consumers who regularly pay off their balances are sometimes referred to in the industry as "deadbeats" or "freeloaders" (Chakravorti and Alpa 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 were 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.[29] In contrast to these penalties that serve only a preventive role and that agents rarely or never pay in equilibrium, in our model nonsophisticated borrowers do pay the

[28] 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 nonrefundable ticket to the leisure traveler. Analogously, a lender should offer an expensive flexible mortgage to 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.

[29] In the classical case of moral hazard, see James A. Mirrless (1999) and Patrick Bolton and Mathias Dewatripont (2005, 140).

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.

VI. 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 naïvete justifies intervention. Although the intervention we propose is welfare improving in the sense typically used in economics (social welfare), in the spirit of libertarian paternalism's (Richard Thaler and Cass Sunstein 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 naïvete with which we are familiar) starts from exogenously given beliefs, it cannot easily accommodate such learning and meta-sophistication.[30] 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.

REFERENCES

Ali, S. Nageeb. 2009. "Learning Self-Control." Unpublished.
Amador, Manuel, Iván Werning, and George-Marios Angeletos. 2006. "Commitment Vs. Flexibility." *Econometrica*, 74(2): 365–96.
Asheim, Geir B. 2008. "Procrastination, Partial Naïvete, and Behavioral Welfare Analysis." http://www.esop.uio.no/research/working-papers/procrastinationAsheim.pdf.
Ausubel, Lawrence M. 1991. "The Failure of Competition in the Credit Card Market." *American Economic Review*, 81(1): 50–81.
Ausubel, Lawrence M. 1997. "Credit Card Defaults, Credit Card Profits, and Bankruptcy." *American Bankruptcy Law Journal*, 71(2): 249–70.
Ausubel, Lawrence M. 1999. "Adverse Selection in the Credit Card Market." http://www.ausubel.com/larry/creditcard-papers/adverse.pdf.
Bar-Gill, Oren. 2004. "Seduction by Plastic." *Northwestern University Law Review,* 98(4): 1373–1434.
Bar-Gill, Oren. 2008. "The Law, Economics and Psychology of Subprime Mortgage Contracts." New York University Law and Economics Research Paper 08–59.
Bernheim, B. Douglas, and Antonio Rangel. 2004a. "Addiction and Cue-Triggered Decision Processes." *American Economic Review*, 94(5): 1558–90.

[30] The only paper we know that systematically studies whether individuals will learn their taste for immediate gratication is S. Nageeb Ali (2009). 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 (2009) theory as deepening the puzzle of how learning affects the behavior of time-inconsistent individuals.

Bernheim, B. Douglas, and Antonio Rangel. 2004b. "Saving and Cue-Triggered Decision Processes." Unpublished.
Bester, Helmut. 1985. "Screening vs. Rationing in Credit Markets with Imperfect Information." *American Economic Review*, 75(4): 850–55.
Bester, Helmut. 1987. "The Role of Collateral in Credit Markets with Imperfect Information." *European Economic Review*, 31(4): 887–99.
Bolton, Patrick, and Mathias Dewatripont. 2005. *Contract Theory*. Cambridge, MA: MIT Press.
Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism.'" *University of Pennsylvania Law Review,* 151(3): 1211–54.
Chakravorti, Sujit, and William R. Emmons. 2003. "Who Pays for Credit Cards?" *Journal of Consumer Affairs,* 37(2): 208–30.
Chakravorti, Sujit, and Alpa Shah. 2001. "A Study of the Interrelated Bilateral Transactions in Credit Card Networks." *Federal Reserve Bank of Chicago Policy Studies,* EPS-2001-2.
Courty, Pascal, and Hao Li. 2000. "Sequential Screening." *Review of Economic Studies*, 67(4): 697–717.
DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–72.
DellaVigna, Stefano, and Ulrike Malmendier. 2004. "Contract Design and Self-Control: Theory and Evidence." *Quarterly Journal of Economics*, 119(2): 353–402.
Demyanyk, Yuliya S., and Otto Van Hemert. 2008. "Understanding the Subprime Mortgage Crisis." Unpublished.
Dubey, Pradeep, and John Geanakoplos. 2002. "Competitive Pooling: Rothschild-Stiglitz Reconsidered." *Quarterly Journal of Economics*, 117(4): 1529–70.
Durkin, Thomas A. 2000. "Credit Cards: Use and Consumer Attitudes, 1970–2000." *Federal Reserve Bulletin*, 86(9): 623–34.
Eliaz, Kfir, and Ran Spiegler. 2006. "Contracting with Diversely Naïve Agents." *Review of Economic Studies*, 73(3): 689–714.
Engel, Kathleen C., and Patricia A. McCoy. 2002. "A Tale of Three Markets: The Law and Economics of Predatory Lending." *Texas Law Review,* 80(6): 1255–1381.
Gabaix, Xavier, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 121(2): 505–40.
Gottlieb, Daniel. 2008. "Competition over Time-Inconsistent Consumers." *Journal of Public Economic Theory*, 10(4): 673–84.
Grubb, Michael D. 2007. "Selling to Overconfident Consumers." Social Science Research Network Working Paper 721701.
Gruber, Jonathan, and Botond Kőszegi. 2004. "Tax Incidence when Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes." *Journal of Public Economics*, 88(9–10): 1959–87.
Hafalir, Elif Incekara. 2008. "Credit Card Competition and Naïve Hyperbolic Consumers." http://www.andrew.cmu.edu/user/incekara/Paper2009.pdf.
Hayhoe, Celia Ray, Lauren J. Leach, Pamela R. Turner, Marilyn J. Bruin, and Francis C. Lawrence. 2000. "Differences in Spending Habits and Credit Use of College Students." *Journal of Consumer Affairs,* 34(1): 113–33.
Hill, Ronald Paul, and John C. Kozup. 2007. "Consumer Experiences with Predatory Lending Practices." *Journal of Consumer Affairs,* 41(1): 29–46.
Kőszegi, Botond. 2005. "On the Feasibility of Market Solutions to Self-Control Problems." *Swedish Economic Policy Review,* 12(2): 71–94.
Laffont, Jean-Jacques, and David Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model.* Princeton, NJ: Princeton University Press.
Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443–77.
Laibson, David, and Leeat Yariv. 2007. "Safety in Markets: An Impossibility Theorem for Dutch Books." http://www.hss.caltech.edu/~lyariv/Papers/DutchBooks.pdf.
Laibson, David, Andrea Repetto, and Jeremy Tobacman. 2007. "Estimating Discount Functions with Consumption Choices over the Lifecycle." National Bureau of Economic Research Working Paper 13314.
Meier, Stephan, and Charles Sprenger. 2010. "Present-Biased Preferences and Credit Card Borrowing." *American Economic Journal: Applied Economics*, 2(1): 193–210.
Mirrlees, James A. 1999. "The Theory of Moral Hazard and Unobservable Behaviour: Part I." *Review of Economic Studies*, 66(1): 3–21.

O'Donoghue, Ted, and Matthew Rabin. 2001. "Choice and Procrastination." *Quarterly Journal of Economics*, 116(1): 121–60.

O'Donoghue, Ted, and Matthew Rabin. 2006. "Optimal Sin Taxes." *Journal of Public Economics*, 90(10–11): 1825–49.

Reda, Susan. 2003. "2003 Consumer Credit Survey." *Stores Magazine,* November.

Renuart, Elizabeth. 2004. "An Overview of the Predatory Mortgage Lending Process." *Housing Policy Debate,* 15(3): 467–502.

Rothschild, Michael, and Joseph E. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, 90(4): 630–49.

Shui, Haiyan, and Lawrence M. Ausubel. 2004. "Time Inconsistency in the Credit Card Market." In *14th Annual Utah Winter Finance Conference,* ed. Avner Kalay, and Jim S. Schallheim.

Skiba, Paige Marta, and Jeremy Tobacman. 2007. "Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default." http://bpp.wharton.upenn.edu/speaker%20files/Tobacman.Jeremy.3309.jobmarketpaper.pdf.

Stiglitz, Joseph E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review*, 71(3): 393–410.

Thaler, Richard H., and Cass R. Sunstein. 2003. "Libertarian Paternalism." *American Economic Review*, 93(2): 175–79.

Warren, Elizabeth. 2007. "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,* US Senate, January 25, 2007.

This article has been cited by:

1. Frank A. Sloan, Lindsey M. Eldred, Yanzhi Xu. 2014. The behavioral economics of drunk driving. *Journal of Health Economics* **35**, 64-81. [CrossRef]
2. Alexei Alexandrov, Özlem Bedre-Defolie. 2014. The Equivalence of Bundling and Advance Sales. *Marketing Science* **33**:2, 259-272. [CrossRef]
3. Randy Hodson, Rachel E. Dwyer, Lisa A. Neilson. 2014. Credit Card Blues: The Middle Class and the Hidden Costs of Easy Credit. *The Sociological Quarterly* n/a-n/a. [CrossRef]
4. Hunt Allcott, Sendhil Mullainathan, Dmitry Taubinsky. 2014. Energy Policy with Externalities and Internalities. *Journal of Public Economics* . [CrossRef]
5. Karna Basu. 2013. Commitment Savings in Informal Banking Markets. *Journal of Development Economics* . [CrossRef]
6. Brishti Guha, Prabal Roy Chowdhury. 2013. Micro-finance competition: Motivated micro-lenders, double-dipping and default. *Journal of Development Economics* **105**, 86-102. [CrossRef]
7. M. J. Garmaise. 2013. The Attractions and Perils of Flexible Mortgage Lending. *Review of Financial Studies* **26**:10, 2548-2582. [CrossRef]
8. Myong-Il Kang, Shinsuke Ikeda. 2013. TIME DISCOUNTING AND SMOKING BEHAVIOR: EVIDENCE FROM A PANEL SURVEY *. *Health Economics* n/a-n/a. [CrossRef]
9. Ryan Bubb, Alex Kaufman. 2013. Consumer biases and mutual ownership. *Journal of Public Economics* **105**, 39-57. [CrossRef]
10. Laura Ricaldi, Michael S Finke, Sandra J Huston. 2013. Financial literacy and shrouded credit card rewards. *Journal of Financial Services Marketing* **18**:3, 177-187. [CrossRef]
11. Fabian Herweg, Konrad Mierendorff. 2013. UNCERTAIN DEMAND, CONSUMER LOSS AVERSION, AND FLAT-RATE TARIFFS. *Journal of the European Economic Association* **11**:2, 399-432. [CrossRef]
12. Abhijit Vinayak Banerjee. 2013. Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?. *Annual Review of Economics* **5**:1, 487-519. [CrossRef]
13. Sendhil Mullainathan, Joshua Schwartzstein, William J. Congdon. 2012. A Reduced-Form Approach to Behavioral Public Finance. *Annual Review of Economics* **4**:1, 511-540. [CrossRef]
14. M. K. Chen, A. Schwartz. 2012. Intertemporal Choice and Legal Constraints. *American Law and Economics Review* **14**:1, 1-43. [CrossRef]
15. J.K. DokkoMortgage Choice 326-329. [CrossRef]
16. TOMOMI TANAKA, TAKESHI MUROOKA. 2012. SELF-CONTROL PROBLEMS AND CONSUMPTION-SAVING DECISIONS: THEORY AND EMPIRICAL EVIDENCE*. *Japanese Economic Review* no-no. [CrossRef]
17. John Gathergood. 2011. Self-control, financial literacy and consumer over-indebtedness. *Journal of Economic Psychology* . [CrossRef]
18. Alexia Gaudeul, Robert Sugden. 2011. Spurious Complexity and Common Standards in Markets for Consumer Goods. *Economica* n/a-n/a. [CrossRef]