Naïve Consumers and Financial Mistakes

Florian Exler* Alexander Hansak†
National Bank of Slovakia CERGE-EI

February 24, 2026

(Latest version here.)

Abstract

Financial contracts are complicated and consumers often do not grasp them in their entirety. This may lead to financial mistakes when borrowers do not fully internalize the cost of credit. We develop a quantitative framework that incorporates both sophisticated and naïve borrower behavior in a model of unsecured credit and equilibrium default. Borrowers select contracts that balance interest rates and penalty fees which make financial shocks - such as late payments or overdrawing - costly. While sophisticated borrowers correctly anticipate penalty fees, naïve borrowers face a higher risk without internalizing this fact. Thus, they make financial mistakes by choosing inefficiently high penalty fees. In equilibrium, penalty fees paid by naïves cross-subsidize interest rates for sophisticates. We use this framework to study consumer protection policies aimed at reducing financial mistakes such as borrowing limits, and also two unexplored features of the CARD Act: transparency requirements and penalty fee limits. More transparency makes financial contracts easier to understand, reducing the financial risk for naïve borrowers. Thus, naïves pay lower penalty fees. Fee limits directly ban high-fee contracts for everyone. Both policies reduce the expected revenue from naïve fee payments and consequently interest rates rise. In both cases, naïves make fewer financial mistakes and enjoy a welfare gain. Sophisticates, in contrast, suffer: Since naïves pay lower fees, sophisticates lose cross-subsidization and experience welfare losses.

JEL Codes: E21, E43, G18, G41, G51, K12

Keywords: Consumer Credit, Naïveté, Financial Mistakes, Bankruptcy, CARD Act, Cross-Subsidization

*National Bank of Slovakia, Imricha Karvaa 1, 813 25 Bratislava, Slovakia. ✉ florian.exler@nbs.sk, https://sites.google.com/site/floexler/

†CERGE-EI, Politických vězňů 7, 111 21 Prague 1, Czech Republic. ✉ Alexander.Hansak@cerge-ei.cz, https://www.sites.google.com/view/alexanderhansak

1 Introduction

Financial contracts can be quite complicated. According to the Wall Street Journal, credit card contracts in 2013 are up to 50 times longer than in 1980—20,000 words in 2013 versus 400 words in 1980.[1] Complicated contracts are harder to understand. Indeed, many cardholders fail to understand key aspects of their contracts, including when late payments trigger penalty fees (GAO, 2006).[2]

Despite being only partially internalized, penalty fees constitute a significant part of the cost of banking to consumers. In 2011, nearly 28% of US consumer checking accounts experienced non-sufficient funds or overdraft fees. Nearly one third of those accounts accrued more than 10 penalty items with an average fee of $225 (CFPB, 2014). Also credit card borrowers incur penalty fees for repaying financial obligations too late or for exceeding lines of credit. With late fees and over limit fees constituting the largest items, Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015) document that the average credit card holder paid $58 in fees per year.

While some of these penalty fees are certainly the consequence of a rational optimization of one's financial situation, some of these fees seem to occur by mistake. Missing important aspects of their credit contract,[3] consumers can lose track of financial commitments, misunderstand minimum payments, or forget to pay bills on time. These financial mistakes are quite common. For example, 52% of borrowers that were charged overdraft fees do not recall opting in to overdraft at all (PEW, 2014). Also, Stango and Zinman (2009) report that the median household could avoid more than half of the costs associated with checking and credit card accounts with minor changes in behavior. In addition to unexpected penalty fees, financial mistakes can cause financial distress. According to Warren, Sullivan, and Jacoby (2000), "credit card debt out of control" and "trouble in managing money" are the third and fourth most common reasons given for a consumer bankruptcy. Only job loss and medical reasons rank higher.

Due to their potentially high cost and severe consequences, financial mistakes are at the center of many pieces of regulation. The 2009 Credit Card Accountability Responsibility and Disclosure (CARD) Act in the U.S. is the most prominent recent example. Policymakers tackle financial mistakes along two dimensions: firstly, transparency requirements and reporting standards make contracts easier to understand, lowering the likelihood of incurring penalty fees by mistake. Secondly, the CARD Act caps penalty fees, reducing the cost of financial mistakes that trigger a penalty fee.

Despite the importance of penalty fees and financial mistakes for consumers and policymakers, state-of-the-art quantitative research cannot inform regulators about their

[1] See `https://www.wsj.com/articles/SB10001424127887324000704578386652879032748`.

[2] According to the United States Government Accountability Office, this lack of understanding is a consequence of lenders deliberately complicating their contracts. However, this paper will be agnostic about the motivation to design complicated contracts. We take their existence as given.

[3] See GAO (2006).

costs and consequences. Standard models of consumer debt simply abstract from penalty fees and have no role for financial mistakes. Without these important ingredients, transparency requirements or fee limits as mandated by the CARD Act seem futile. We address this gap by introducing a tractable credit-market framework in which borrowers choose contracts that trade off interest rates and penalty fees, and in which a subset of borrowers misperceive their exposure to shocks that trigger fees. The framework allows us to answer the following questions: What are the consequences of financial mistakes for credit market outcomes? Can regulation similar to the CARD Act achieve better outcomes? How do these policies influence the interaction between consumers that are more and less exposed to paying penalty fees?

Building on the theory of naïveté,[4] we set up a heterogeneous agents model of unsecured debt and default that is inhabited by two types of agents: sophisticates and naïves. While all borrowers are subject to financial shocks that trigger penalty fees, naïve borrowers are unaware of their increased exposure to these shocks and the ensuing penalty fees. Being naïve, these consumers behave just like sophisticated consumers and choose contracts with overly high penalty fees – what we term financial mistakes. Because borrowers behave identically, lenders cannot distinguish between the two types. This naturally leads to (partial) pooling of borrowers.[5] Lenders maximize profit and offer a menu of loan contracts. Conditional on a requested amount and borrower characteristics, lenders offer a continuum of interest rate and penalty fee combinations for both types of borrowers to choose from.

In equilibrium, lenders only offer contracts that yield equal expected revenue. Hence, debt contracts trade off lower interest rates for lower penalty fees. When choosing from the menu of loan contracts, naïve borrowers do not understand the true expected cost of penalty fees. They consequently choose contracts that carry too high penalty fees (that they are naïve about) because they prefer low interest rates. Naïve borrowers make financial mistakes: on average, they pay more for credit than they would if they knew their true exposure to financial shocks. Sophisticated consumers are less prone to financial shocks and incur fewer penalty fees. Being pooled with naïve agents, they benefit from the same set of contracts where low interest rates are cross-subsidized by high penalty fees. These high penalty fees are mainly borne by naïve consumers. Thus, sophisticated consumers face cheaper credit in the presence of naïve consumers than if they were by themselves.

As a practical application, we use this framework to analyze two important aspects of the 2009 CARD Act: (1) The CARD Act defines how late fees, interest rates, and minimum payments are to be reported and communicated to consumers. With stringent transparency requirements, standardized language, and clearer reporting standards

[4] Cf. Armstrong and Vickers (2012), Gabaix and Laibson (2006), and Heidhues and Kőszegi (2010).

[5] This approach has been established by Exler, Livshits, MacGee, and Tertilt (2024) in the context of over-optimistic consumers.

contracts are easier to understand. More understandable contracts reduce the risk of financial mistakes. (2) The CARD Act limits how lenders can reset interest rates in response to missed payments and restricts the amount of penalty fees to be charged. Limiting penalty fees for borrowers reduces the cost of financial mistakes. We find that both pieces of legislation have a similar impact on credit contracts offered in equilibrium. Revenues from penalty fees shrink, either because (1) consumers make fewer mistakes and thus pay lower fees or because (2) lenders are prevented from offering excessive penalty fees. Consequently, equilibrium interest rates rise and cross-subsidization from naïves to sophisticates falls under both policies.

Naïve consumers benefit from both reforms. Their lack of understanding becomes less consequential either (1) because they make fewer mistakes per se due to simpler contracts or (2) because they are forced to choose lower penalty fees and thereby avoid financial mistakes. Naïves pay too much for their credit prior to the reform and both policies reduce their cost of credit. Because sophisticated agents lose cross-subsidization, they stand to lose from these reforms. While inconsequential to their understanding of credit contracts, sophisticates lose out on cross-subsidization as a consequence of (1) transparency requirements. With naïves committing fewer financial mistakes and paying lower penalty fees, sophisticates face higher interest rates. They consequently suffer from transparency requirements. Similarly, (2) limiting penalty fees also reduces sophisticate welfare. Penalty fee limits force naïve borrowers to choose lower fees and reduce cross-subsidization. Thus, sophisticates experience welfare losses.

We also explore other policy tools such as borrowing limits. Because naïve consumers tend to carry higher debt burdens, restricting borrowing relative to income can help prevent over-indebtedness. However, the effects of implementing a debt service ratio are ambiguous: modest constraints may improve naïve borrowers' welfare by steering them away from exploitative contracts, but tighter limits can backfire. Faced with tight borrowing caps, naïve borrowers start to select contracts with even higher penalty fees to reduce nominal interest payments and comply with the constraints. These findings highlight the importance of carefully designing borrowing limits to avoid unintended consequences.

1.1 Related Literature

This paper contributes to several strands of literature. First, it relates to empirical studies evaluating the 2009 CARD Act. Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015) use panel data to analyze the effects of the 2009 CARD Act and estimate that regulations lead to a decrease in overall borrowing costs. Nelson (2020) also analyzes the CARD Act and finds that limits on markups and on lenders' ability to reprice interest rates based on new information reduced average transacted prices. However, he also documents price increases in some parts of the market, thereby unveiling possible cross-effects that might mitigate the legislator's intent. Similarly, Dou, Jagtiani, Maingi, and Ronen (2020) find that, following the introduction of the Act, the average ratio of total credit limits on consumer cards to total available credit declined more for non-prime than for prime borrowers – again suggesting that the regulation had heterogeneous effects across borrower types. However, such distributional implications can only be fully assessed within a structural framework. Consequently, we contribute to this literature by developing a structural model that captures two key components of the CARD Act reform – transparency regulations and fee payments. In addition to highlighting the role of financial mistakes, our quantitative approach allows us to uncover the underlying mechanisms, gauge effects through equilibrium pricing, and estimate the resulting welfare impacts.

Second, our paper extends the concept of naïve agents in credit markets to a setting with imperfect enforcement and equilibrium default. In doing so, we also provide a quantitative assessment of the relevance of naïveté for both credit market outcomes and regulatory policies. Naïveté has long been the focus of theoretical contributions. Armstrong and Vickers (2012) show that in markets with sophisticated and naïve consumers a pooling equilibrium may exist and that competition can work to subsidize the sophisticated at the expense of the naïve. This mechanism is also present in the work by Heidhues and Kőszegi (2015), who study naïveté-based discrimination and find that firms lend more than is socially optimal to increase unexpected payments from naïve consumers. Other related studies explore the behavioral foundations of financial mistakes. Heidhues and Kőszegi (2010) develop a model of loan-repayment in competitive credit markets with consumers who value immediate gratification. They show that non-sophisticated consumers take on credit which is cheap in the short term, but would then go on to overborrow and pay large penalties, thereby suffering considerable welfare losses. In Eliaz and Spiegler (2006), agents have dynamically inconsistent preferences, and the principal offers a menu of loan contracts in order to screen for sophistication and extract higher profits from naïve consumers through exploitative contracts. A related mechanism is explored by Gabaix and Laibson (2006), who propose a framework in which firms offer cheap baseline contracts in order to hook naïves and earn profits from shrouded prices for additional payments. However, all these models assume full commitment to repayment. In contrast, our paper incorporates the option to default, which significantly alters the interaction between sophisticated and naïve agents in credit markets. We quantitatively assess how default risk shapes credit contracts and influences market outcomes, offering new insights for both theory and policy.

Besides constituting an important theoretical tool to study credit markets and contract design, naïveté is also empirically relevant. DellaVigna and Malmendier (2004) study how firms respond to partially naïve time-inconsistent consumers. They show that contracts are designed to target consumers' misperception of future behavior. Agarwal, Chomsisengphet, Liu, and Souleles (2015) analyze an experiment from a large U.S bank that offered two contracts to customers. Borrowers could trade off fees and interest rates:

One contract offered a lower interest rate but came with an annual fee. The other eliminated the annual fee but charged a higher interest rate. The authors find that about 40% of consumers chose sub-optimally and ended up with higher total cost of credit. Stango and Zinman (2009) provide further evidence that many consumers misunderstand core features of borrowing and frequently choose dominated credit options with costs that could have easily been avoided. They find that the median household could save more than half of the costs associated with checking and credit card accounts by simply making minor changes in their long-run decision making. Lastly, Campbell, Grant, and Thorp (2022) analyze data from a controlled field trial studying costly credit card delinquency and show that simply sending a reminder to overdue credit card debtors significantly raises repayment rates as well as the amounts repaid by high credit score delinquents. These findings support the idea that observed borrowing behavior often reflects financial mistakes rather than rational responses to frictions or preferences.[6] Our model enables a quantitative assessment of the magnitude of financial mistakes made by naïve consumers, allowing us not only to analyze the welfare effects of credit market policies but also to evaluate their impact on consumer financial understanding.

Lastly, we expand the standard framework of unsecured credit and equilibrium default based on the seminal work by Chatterjee, Corbae, Nakajima, and Ríos-Rull (2007) and Livshits, MacGee, and Tertilt (2007).[7] First, we introduce penalty fees as a new dimension to credit contracts: borrowers can not only choose debt levels (and receive interest rate quotes as a function of their state and the amount borrowed), but they can also sign up for higher penalty fee payments to lower their interest rates. Secondly, we add a fraction of naïve agents to the economy that do not fully internalize their likelihood of paying these penalty fees. Consequently, they make mistakes when taking out loans. These novel features create a natural role for information requirements and penalty fee limits, which could not play a role in standard models of unsecured credit. In the context of the 2009 CARD Act, we leverage our new framework to investigate these regulatory features that have previously been un(der)studied.

There have been other approaches to incorporate behavioral consumers into a model of unsecured credit and equilibrium default. Nakajima (2017) introduces hyperbolic discounters but does not allow for any interactions between consumer types. Exler, Livshits, MacGee, and Tertilt (2024) model consumers that are over-optimistic about their future income and interact with rational consumers through the credit market. We follow the authors' assumptions on individual behavior to ensure that only pooling equilibria can exist. Chatterjee, Corbae, Dempsey, and Ríos-Rull (2023) develop a framework where

[6]There is also evidence that the interaction of more and less sophisticated consumers through pricing matters. Using transaction data of retail funds, Gao, Hu, Kelly, Peng, and Zhu (2020) document that naïve investors cross-subsidize sophisticated investors. This subsidization is especially pronounced when trading more complex structured funds.

[7]See Exler and Tertilt (2020) for a recent survey.

agents hold heterogeneous discount factors. Both, Chatterjee, Corbae, Dempsey, and Ríos-Rull (2023) and Exler, Livshits, MacGee, and Tertilt (2024) focus on how lenders learn about consumer types and do not have a role for information requirements or fee limits, as agents fully understand their credit contracts. Raveendranathan and Stefanidis (2025) analyze the ability to pay check mandated by the 2009 CARD Act, but also have no role for assessing transparency regulations. We see our work as complementing their findings.

The remainder of this paper is structured as follows: Section 2 presents the framework, Section 3 describes the calibration, and Section 4 explores the benchmark properties and mechanisms of the model. Section 5 then investigates two important components of the 2009 CARD Act and further policies, and Section 6 concludes.

2 The Model Framework

We propose a quantitative theory of unsecured debt and equilibrium default that includes naïve agents and offers a penalty fee vs. interest rate trade-off. We study an incomplete-market heterogeneous agent life-cycle model with idiosyncratic uncertainty about earnings, expenses and financial shocks. Interest rates are determined in equilibrium and reflect expected penalty fee revenues as well as write-offs from default.

2.1 Households

We set up a standard overlapping generations model that is populated by a continuum of households, with one model period representing one year. Individuals enter the model at age 21 and die with certainty at age 80. They derive utility from consumption in each period c_t, may experience a disutility from filing for bankruptcy χ, and maximize expected discounted lifetime utility,

$$\mathbb{E}_0 \left\{ \sum_{j=1}^{J} \beta^{j-1} \Big[u(c_j) - \delta_j \chi \Big] \right\},$$

where $\beta < 1$ is the time discount factor and δ_j is an indicator for filing for bankruptcy.

Each agent decides on consumption, whether to save or borrow, and can file for bankruptcy. The model abstracts from secured debt like mortgages and focuses on unsecured credit card debt. Borrowers can default in equilibrium by declaring Chapter 7 bankruptcy. Households face idiosyncratic income shocks, expense shocks, and financial shocks, while the risk-free interest rate is set exogenously.

Labor Productivity. Households are subject to idiosyncratic shocks to their labor productivity p, which represents the wage risk they face. Income y_j is given by $y_j =$

$p_j \cdot e(j)$, where $e(j)$ is an age-dependent efficiency premium. Individual labor productivity is comprised of a persistent AR(1) process z_j and a transitory white noise shock η_j. Hence, for a household of age j, productivity evolves according to

$$\begin{aligned} \log(p_j) &= z_j + \eta_j \\ z_j &= \rho z_{j-1} + \zeta_j, \end{aligned} \tag{1}$$

where $\rho \in [0,1]$, $\eta \sim \mathbb{N}(0, \sigma_\eta^2)$ and $\zeta \sim \mathbb{N}(0, \sigma_\zeta^2)$.

Expense Shocks. Households are subject to expense shocks. These shocks represent unforeseen expenses that arise from medical bills or family disruptions. Expense shocks are drawn from a discrete finite set $\kappa \in \Omega_\kappa = \{\kappa_0 = 0, \kappa_1, \ldots, \kappa_N\}$ and are modeled as independent and identically distributed with probabilities $p(\kappa_n)$. These expense shocks are unsecured claims and can be dispensed with through bankruptcy.

Financial Shocks. Furthermore, households that borrow on their credit cards are susceptible to financial shocks ε. Financial shocks represent households not meeting their minimal repayment requirements, paying their credit cards late, or over-borrowing on their accounts. If a household incurs such a shock, lenders will charge them a previously contracted upon additional fee $\phi \geq 1$, representing overdraft fees, late payment fees, etc. Thus, borrowers hit by a financial shock ε face additional charges of $\phi \cdot \varepsilon$. Here, ε can be interpreted as the fundamental cost of a lender due to borrowers mishandling their credit, and ϕ is a potential additional fee. These charges are not secured and can be defaulted upon. We model financial shocks as discrete iid shocks, drawn from $\varepsilon \in \Omega_\varepsilon = \{\varepsilon_0 = 0, \varepsilon_1, \ldots, \varepsilon_M\}$ with probabilities $p(\varepsilon_m)$.

Bankruptcy. Households can choose to default and not repay their debts, which includes any expense or financial shocks they incurred. Bankruptcy is modeled according to U.S. Chapter 7. It provides a "Fresh Start:" upon default, all outstanding unsecured debts and claims are forgiven.

However, bankruptcy is costly and requires a good faith effort to repay outstanding debts. This is captured in the model by garnishing a fraction γ of each bankrupt's current income and using it to (partially) repay creditors. Furthermore, bankrupts incur a utility cost χ, which represents other costs of defaulting, such as stigma.

2.2 Naïveté

There are two types of households in the economy: sophisticated consumers (S) and naïve consumers (N). Sophisticated and naïve consumers are identical along all dimensions, except with respect to the financial shock. Both types of agents incur financial shocks, i.e. both types sometimes pay late or borrow over limit. However, naïve borrowers are more prone to these financial shocks, while being naïve about their higher exposure.

In the model, the financial shocks of both types share a common support: $\varepsilon^S, \varepsilon^N \in \Omega_\varepsilon$. However, naïve borrowers face higher probabilities that a nonzero financial shock occurs:

$$p(\varepsilon^N = \varepsilon_m) > p(\varepsilon^S = \varepsilon_m) \quad \text{for all} \quad m > 0. \tag{2}$$

The relative difference in probabilities faced by naïve and sophisticated consumers is characterized by a probability spread, ψ, which is constant for all nonzero shocks:

$$\psi = \frac{p(\varepsilon^N = \varepsilon_1)}{p(\varepsilon^S = \varepsilon_1)} = \ldots = \frac{p(\varepsilon^N = \varepsilon_M)}{p(\varepsilon^S = \varepsilon_M)} \tag{3}$$

We define ψ as the extent of naïveté in the economy. If $\psi = 1$, there is no naïveté and all consumers face an identical shock process. If $\psi > 1$, naïve borrowers experience nonzero financial shocks more frequently.

Sophisticated borrowers know the distribution of their shock process and take into account the true expected shocks when taking out debt. In contrast, naïve consumers do not understand their higher exposure. They expect the same distribution that sophisticates face. Consequently, naïve borrowers underestimate their average financial shock:

$$\mathbb{E}^N(\varepsilon^N) = \mathbb{E}^S(\varepsilon^S) = \mathbb{E}(\varepsilon^S) < \mathbb{E}(\varepsilon^N), \tag{4}$$

where $\mathbb{E}^{N,S}(\cdot)$ denotes the type-specific expectation of consumers and $\mathbb{E}(\cdot)$ the expectation formed with the true underlying risks. In the case of a naïve agent, their expectation differs from a fully information rational expectation.

With identical expectations of financial shocks, naïve agents behave exactly like sophisticated consumers, conditional on all other states. Hence, lenders cannot distinguish between the two types nor separate them by offering distinct contracts. This assumption ensures a unique, tractable pooling equilibrium.[8] Furthermore, we abstract from learning for lenders and borrowers. Even if a consumer keeps experiencing financial shocks, none of the agents update their belief about the borrower's type but just deem this consumer very unlucky.[9]

To summarize, naïve borrowers are more exposed to financial shocks, biased in their expectation about future financial shocks, and pooled with sophisticated borrowers. In our benchmark and when conducting policy experiments, we decompose the effects of naïveté into the effects of higher financial shock risk and the effects of expectation bias.

[8]This setup is similar in spirit to Exler, Livshits, MacGee, and Tertilt (2024), where over-optimistic borrowers face worse income risk but are not aware of it.

[9]However, lenders can observe the debt level which contains information on the history of shocks. In Section 2.3 we discuss how lenders form expectations on the share of naïves for different debt levels in equilibrium.

When presenting our policy experiments, we also compare the results from our benchmark pooling equilibrium to the outcomes in a market where lenders perfectly observe borrower types and naïve and sophisticated borrowers are separated.

2.3 Credit Market

Loan Contracts. Households can borrow in one period loans and they have limited commitment to repay. A loan contract consists of the face value of the debt, $d' \in (0, \infty)$, the loan price, $q \in [0, 1]$ which translates into an interest rate $1/q - 1$, and penalty fees for financial shocks, $\phi \in [1, \infty)$. Borrowers can choose contracts with low penalty fees (e.g. $\phi = 1$) which reduce their exposure to a potential financial shock. Alternatively, they accept higher penalty fees ($\phi > 1$) in exchange for lower interest rates or higher debts.

When an individual takes out a loan, they choose from a menu of interest rates ($1/q - 1$) and associated penalty fees, ϕ, for each potential debt level d'. There is no asymmetric information about income, assets, or shocks, and lenders perfectly observe the current state of each borrower and the requested loan size. The only unobserved heterogeneity is borrower type: Whether a given borrower is naïve or sophisticated is unknown to all agents. However, lenders know that debt levels contain information about the fraction of naïves $\lambda(d')$ seeking that loan.

Lenders. Lenders operate in a perfectly competitive market with free entry. Conditional on observables, they earn zero profits in equilibrium. In other words, loan prices q perfectly reflect expected revenue for each borrower state, amount of debt d', and penalty fee ϕ. If a borrower repays their debts, revenue consists of two components: the repaid face value d' of the loan plus the expected revenue from penalty fee payments ϕ. If a borrower is hit by a financial shock ε and repays their debt, the lender earns additional revenue $\phi \cdot \varepsilon$. From that revenue, lenders pay the fundamental cost ε to meet the additional liquidity needs of a borrower that repays late or that over-borrows. Additionally, we assume a proportional transaction cost $(1 - \iota)$ for managing borrowers that do not abide by the loan terms. Consequently, the lender earns a net revenue of

$$\iota \cdot (\phi - 1) \cdot \varepsilon \tag{5}$$

from borrowers that are subject to the financial shock and repay.[10]

However, not all households repay their debts. If borrowers declare bankruptcy, a fraction γ of the household's current income y is garnished to repay parts of the outstanding debt. We assume that garnished income is split equally among all creditors. Thus, the

[10] One can also think of ι more broadly capturing the fact that additional profits derived from fee payments are not fully passed on as reductions in interest rates. However, in a perfectly competitive framework, we prefer the transaction cost interpretation.

amount recovered from defaulted loans can be written as

$$\rho(d,s) = \gamma y(s) \cdot \frac{d}{d+\varepsilon+\kappa}, \tag{6}$$

where $s = (j, z, \eta, \varepsilon, \kappa)$ summarizes the household's (exogenous) state.

Lenders' Expectations. To calculate the price of a loan, lenders form expectations about the borrower's future state $s' = (j', z', \eta', \epsilon', \kappa')$ using the probability measure $\mu(s' \mid d')$. For most state variables (j', z', η', κ'), the distribution is identical across all households. The distribution of financial shocks ε', however, depends on borrower type: naïve households face a higher probability of being hit by a positive shock than sophisticated households.

Since borrower type is unobservable, lenders cannot condition on it directly. However, they can form a belief $\hat{\lambda}(d')$ about the fraction of naïve borrowers who choose debt level d'. This induces a perceived distribution of financial shocks given debt d':

$$\mu_\epsilon(\cdot \mid d') = \Big(1 - \hat{\lambda}(d')\Big) F_S(\cdot) + \hat{\lambda}(d') F_N(\cdot), \tag{7}$$

where F_S and F_N denote the true distributions of ε for sophisticated and naïve households, respectively. Hence, the lender's overall expectation $\mu(s' \mid d')$ factorizes into exogenous distributions for (j', z', η', κ') and the belief-weighted mixture distribution $\mu_\epsilon(\cdot \mid d')$ for ε'.

In equilibrium, these beliefs will be correct: If beliefs and realizations differ, lenders observe systematic profit deviations and update their belief until a stationary fixed point $\hat{\lambda}(d') = \lambda(d')$ is reached. Hence, in equilibrium the conjectured distribution $\hat{\lambda}(d')$ coincides with the actual distribution of naïves across debt levels.

Loan Pricing. Let $\theta(d', \phi, s')$ denote a household's decision to default depending on the contracted repayment d' and fee ϕ. Also, let $\mu(s' \mid d')$ be the lender's expectation over next-period states, as discussed above.

The lenders' expected profit from extending a loan is then

$$\begin{aligned} \Pi(d', \phi, s) = -q(d', \phi, s)d' + \frac{1}{1+r+\tau} \int \theta(d', \phi, s')\rho(d', s') \\ + \Big(1 - \theta(d', \phi, s')\Big)\Big(d' + \iota(\phi - 1)\varepsilon'\Big)\, \mathrm{d}\mu(s' \mid d') \end{aligned} \tag{8}$$

where r denotes the exogenous refinance interest rate, and τ is a proportional transaction cost of loan creation.

Perfect competition and free entry lead to zero profits conditional on observables d'

and s. Consequently,

$$q(d', \phi, s) = \frac{1}{1 + r + \tau} \int \theta(d', \phi, s') \frac{\rho(d', s')}{d'} + \Big(1 - \theta(d', \phi, s')\Big) \left(1 + \frac{\iota(\phi - 1)\varepsilon'}{d'}\right) \mathrm{d}\mu(s' \mid d'). \tag{9}$$

Equation 9 is an implicit functional relationship $q = q(\phi, \cdot)$ between $q(\cdot)$ and ϕ for any debt level d' and exogenous household state s and has to hold in equilibrium. See Figure A.1 for an illustration of the relationship between $q, \phi,$ and d'.

Equation 9 can be thought of as a weighted average between what lenders recover in the case of bankruptcy $\rho(\cdot)$, and full repayment plus potential fees in the case of no default, weighted with the appropriate default risk $\theta(\cdot)$. If agents choose $\phi = 1$, they minimize their exposure to financial shocks and pay no additional fee. This comes at the cost of higher interest rates. A choice of $\phi > 1$ implies positive penalty fees that lenders charge in the case of financial shocks but might lead to lower interest rates. Lastly, note that there exists an endogenous limit on penalty fees, since too high fees always cause default: $\exists\ \bar{\phi} \in \mathbb{R}^{+}$: $\forall\ \phi > \bar{\phi} : \mathbb{E}[\theta(\cdot)] = 1$.

Total Cost of Credit Bias and Financial Mistakes. If a borrower seeks out a loan with debt level d', loan price q and punishment fees ϕ, the total cost of the credit, ToC, are all the payments resulting from interest and fees, that is

$$ToC = d' - qd' + (\phi - 1)\varepsilon \tag{10}$$

where $d' - qd'$ are interest payments and $(\phi - 1)\varepsilon$ are fee payments. Note that $\phi = 1$ means that no extra fees have to be paid. Hence, the cost of credit only exceeds interest payments if $\phi > 1$.

Due to underestimating their proneness to financial shocks, naïve borrowers also underestimate their fee payments, $\mathbb{E}^N(\phi \cdot \varepsilon^N) < \mathbb{E}(\phi \cdot \varepsilon^N)$. Consequently, for any given contract (d', q, ϕ) naïves misjudge the expected total cost associated with this contract. We define the *ToC Bias* as the percentage difference between the true expected total cost of credit and the naïve's biased expectation, expressed as:

$$ToC\ Bias = \frac{\mathbb{E}(ToC) - \mathbb{E}^N(ToC)}{\mathbb{E}^N(ToC)}, \tag{11}$$

where $\mathbb{E}^N(ToC)$ represents the naïve's biased expectation formed under the misperception of their true susceptibility to shocks, and $\mathbb{E}(ToC)$ represents the expectation formed using the true underlying process.

Because of this bias, naïve borrowers might fail to choose an optimal debt contract and opt for contracts with lower interest rates but higher penalty fees. As a result, their true expected total cost of credit will exceed the costs associated with an optimal contract that a social planner would choose for them. Let ToC^* denote the total cost of a credit contract that yields the same debt level d' as chosen by the naïve consumer but with a combination of loan price and fees (q^*, ϕ^*) that are optimal given the true underlying risks. The extent to which the actual payment of the naïve borrower exceeds the cost of this hypothetical optimal contract is called a *financial mistake*. Hence, in this framework a financial mistake is defined as

$$Mistake = \frac{\mathbb{E}(ToC) - \mathbb{E}(ToC^*)}{\mathbb{E}(ToC^*)}, \tag{12}$$

where the expectation $\mathbb{E}(\cdot)$ is formed using the true risks.

We illustrate these two measures in Figure 3 below when discussing the model mechanisms in the benchmark economy.

2.4 Dynamic Programming Problem

A household's decision variables are consumption c, debt level d' (where $d' < 0$ are positive asset holdings) and a financial contract when taking on a loan. A financial contract is a combination of loan prices q and fees for financial shocks ϕ. Lenders will offer different loan prices for different fees and the borrowers choose a contract by deciding on a fee level. Note, however, that this fee choice will only be important in the next period, when the borrower might be faced with a financial shock ε and then has to pay $\phi \cdot \varepsilon$. Also, since credit mishandling is only costly for those in debt, ε does not impact savers.

Let $V(d, \phi; s)$ denote the value function of a consumer in state $s = (j, z, \eta, \varepsilon, \kappa)$, who is holding debts d and chose a fee ϕ last period. If a consumer declares bankruptcy, all debt charges are dismissed ($d = 0$ and $\varepsilon = 0$), but a fraction γ of income is garnished and the household experiences a disutility χ. Also, no saving is allowed for next period, the household simply consumes all income that is not garnished.

With $q(d', \phi'; s)$ being the loan price for a loan of size d' with fees ϕ' for an individual in state s, the consumer's decision problem reads

$$\begin{aligned} V(d, \phi; s) = \max_{c, d', \phi'} & \left[u(c) + \beta \mathbb{E}^S \max \left\{ V(d', \phi'; s'), B(s') \right\} \right] \\ \text{s.t. } & c + d + \kappa + \phi \cdot \varepsilon \cdot \mathbb{1}_{d>0} \leqslant y(s) + q(d', \phi'; s) d' \end{aligned} \tag{13}$$

where $\mathbb{E}^S(\cdot)$ denotes the fact that all consumers think of themselves as sophisticated when forming expectations about future shocks and B is the value of filing for bankruptcy:

$$B(s) = u(c) - \chi + \beta \mathbb{E}^S \Big\{ V(0, \phi'; s') \Big\}$$
$$\text{s.t. } c = (1-\gamma) y(s) \tag{14}$$

Thus, even though we have two types of agents in the economy, the decision problem is the same for both sophisticated and naïve consumers since the latter do not realize their higher proneness to financial shocks. Consequently, the decisions of sophisticates and naïves will be the same based on their observable state.

Equilibrium Definition. An equilibrium consists of

- value functions $V(d, \phi; s)$ and $B(s)$,
- optimal decision rules for consumption $c(\cdot)$, next-period debt $d'(\cdot)$, fee choice $\phi'(\cdot)$, and default decisions $\theta(\cdot)$,
- a loan price schedule $q(d', \phi; s)$, and
- a stationary belief $\hat{\lambda}(d')$ about the share of naïve borrowers at each debt level,

such that:

1. Given $q(d', \phi; s)$, households solve the decision problem (13)–(14) and form optimal policies.
2. Given household policies and beliefs $\hat{\lambda}(d')$, the loan price schedule satisfies the zero-profit condition (9).
3. Beliefs are correct: the stationary distribution of debt choices implies $\hat{\lambda}(d') = \lambda(d')$ for all d'.

3 Calibration

This section presents the parameter values used to numerically solve the benchmark economy. Some parameters are set externally, while others are estimated using a Simulated Method of Moments approach to match important credit market data. We use data of the consumer credit card market from the Consumer Financial Protection Bureau (CFPB, 2014) and credit card account data from the Office of the Comptroller of the Currency in the US (OCC) as reported in Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015).

3.1 Externally Determined Parameters

Households live for 60 periods and one period represents one year. Individuals enter the model at age 21 and die with certainty at age 80. There is no formal retirement, but households' age-dependent experience premium drops after the age of 65. Per-period utility takes the form of a CRRA utility function

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma},$$

where the risk-aversion parameter, σ, is set to 2, which is standard in the literature. For the age-dependent efficiency premium, $e(j)$, we use the life-cycle component from Livshits, MacGee, and Tertilt (2007), using linear interpolation to create annual values from the original three-year periods. The risk-free rate is set exogenously to 1%.

Persistent labor productivity follows a simple AR(1) process where the autocorrelation parameter is set to $\rho = 0.99$ and the variance of the iid shock chosen to be $\sigma_\zeta^2 = 0.007$. The variance of the transitory shock is $\sigma_\eta^2 = 0.043$. All these values are within the range of standard literature. In particular, these are the same values employed by Livshits, MacGee, and Tertilt (2007).

For the parametrization of expense shocks to U.S data we follow the estimates of medical expenses, divorces, and unplanned parenthood from Livshits, MacGee, and Tertilt (2007) and Exler, Livshits, MacGee, and Tertilt (2024). Expense shocks can take on three values: $\kappa \in \{0, \kappa_1, \kappa_2\}$. The smaller shock corresponds to 26.4% and the larger one to 82.18% of average yearly income. Annualized probabilities $[\pi_1, \pi_2]$ of these shocks realizing are 2.37% and 0.15%, respectively.

Financial Shocks. Financial shocks are modeled to take on four different values: $\varepsilon \in \{0, \varepsilon_1, \varepsilon_2, \varepsilon_3\}$. These values are meant to represent no occurrence of a financial shock within a year, only one such occurrence, up to three occurrences, and more repeated occurrences. We use data on overdraft from the CFPB (2014) to estimate that a fraction of 12.5% of borrowers experiences only one item per year, while 9.4% of borrowers experience up to three items and 8.3% experience repeated items. Hence, we set $p(\varepsilon_1) = 0.125$, $p(\varepsilon_2) = 0.094$ and $p(\varepsilon_3) = 0.083$. These values represent the economy wide average and we construct different processes for naïve and sophisticated consumers below.

To pin down the size of these shocks, we express the average payments reported for one, up to three, and more occurrences in CFPB (2014) as a fraction of median income in the US. This yields shock sizes of $\varepsilon_1 = 0.0007$, $\varepsilon_2 = 0.0019$ and $\varepsilon_3 = 0.0090$ in terms of median income.

Naïveté. First, we need to proxy the fraction of naïves in the economy. Obviously, naïveté is not directly observable, so we proxy it by those consumers that pay exorbitantly

Table 1: Externally Determined Parameters

Economic Parameters		Value
CRRA consumption	σ	2
Persistent Wage Autocorrelation	ϱ	0.99
Persistent Wage Var	σ_ζ^2	0.007
Transitory Wage Var	σ_η^2	0.043
Risk Free Rate	r	1%
Expense Shocks		
Size of expense shocks	κ	$(0, 0.267, 0.8218)$
Probabilities	$[\pi_0, \pi_1, \pi_2]$	$(0.9748, 0.0237, 0.0015)$
Financial Shocks		
Size of shocks	ε	$(0, 0.0007, 0.0019, 0.0090)$
Unconditional probabilities	$Pr(\varepsilon)$	$(0.698, 0.125, 0.094, 0.083)$
Share of naïves	λ	0.3
Naïve probability spread	ψ	6

more penalty fees than the rest of the distribution. This is consistent within the model, as naïveté implies both a higher risk of financial shocks and a bias about this risk. In the OCC data, borrowers with FICO scores of 660 and less pay significantly more fees than borrowers above this threshold, see Figure A.3. Those borrowers comprise a share of 30% of the population. Consequently, we set the share of naïves, λ, equal to 30%.

Sub FICO 660 borrowers pay roughly six times more fees in the OCC data compared to higher FICO score borrowers. This lets us directly pin down the extent of naïveté, the probability spread between naïves and sophisticates (see Section 2.2). We set $\psi = 6$. The specific probability processes for naïves and sophisticated are then constructed in such a way that their weighted mean (with the share of naïves, λ, used as weight) results in the economy wide financial shock probabilities determined above. This yields the following probabilities: $p(\varepsilon^S) = (0.879, 0.05, 0.038, 0.033)$ and $p(\varepsilon^N) = (0.275, 0.3, 0.226, 0.199)$.

Externally determined parameters are summarized in Table 1.

3.2 Internally Determined Parameters

We are left with five parameters to determine internally: discount factor β, transaction cost of loan creation τ, efficiency of fee collection ι, the garnishment rate γ, and utility cost of default χ. We calibrate those five parameters to five data targets: Average interest payments to debt (AID), average fee payments to debt (AFD), fraction of borrowers, bankruptcy rate, and the average debt-to-income ratio. More technically, we choose these five parameters to minimize the sum of squared relative residuals between model moments

Table 2: Jointly-Targeted Moments

	Data (OCC)	Model
Avg. Interest / Debt (AID)	14.30%	12.97%
Avg. Fees / Debt (AFD)	6.70%	6.62%
Fraction Borrowers	25%	28.8%
Avg. Debt-to-Income Ratio	6%	5.2%
Bankruptcy Rate	0.45 %	0.45%
	Parameter	Value
Discount factor	β	0.9314
Transaction cost	τ	0.1344
Fee efficiency	ι	0.4025
Cost of default	χ	0.0193
Garnishment rate	γ	0.5479

and data targets.

While all parameters affect all moments jointly, we discuss parameters together with the moment(s) most affected by them. A discount factor $\beta = 0.93$ delivers a debt-to-income ratio of 5.2% vs. 6% in the data and an AID of 13% vs. 14.3% in the data. To further achieve this interest rate, a transaction cost for borrowing, $\tau = 0.13$ is chosen. This also allows us to match the share of borrowers: 29% in the model to 25% in the data.

A household who decides to file for bankruptcy experiences stigma costs in the form of a disutility of $\chi = 0.02$ in the period of default and a fraction $\gamma = 0.55$ of current income is garnished. These parameters deliver a fraction of 0.45% bankruptcies per year, exactly as in the data.

Finally, the fee payment efficiency ι, is chosen to match an average ratio of fee payments to debt of 6.6% versus the observed 6.7% in the OCC data.

Overall, the model fits the data very well. Especially achieving very high AFD numbers has proven challenging and lends credibility to the current approach. See Table 2 for a list of all internally determined parameters and the model and data moments.

4 Results

With the previously defined parameters, we solve the model quantitatively by backward iteration on the value function. After briefly discussing the benchmark outcomes in Section 4.1, Section 4.2 illustrates the effects of naïveté on the design of credit contracts and life cycle outcomes. Section 4.3 discusses cross-subsidization and its welfare effects and Section 4.4 highlights the importance of default for the mechanisms.

Table 3: Benchmark Outcomes

	Sophisticated	Naïve	Average
Avg. Interest / Debt (AID)	12.96%	12.99%	12.97%
Average Fees / Debt (AFD)	2.78%	15.59%	6.62%
Fraction Borrowers	28.25%	29.96%	28.76%
Bankruptcy Rate	0.33%	0.71%	0.45%
Debt-to-income ratio	4.8%	6.2%	5.2%
Cross-Subsidization	0.28%	−0.33%	
ToC Bias		64.6%	
Financial Mistake		64.5%	

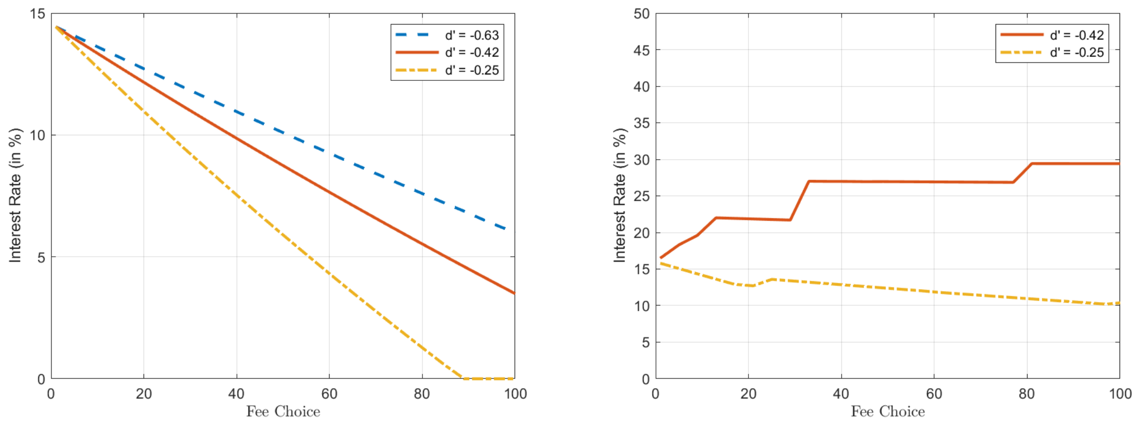
4.1 Benchmark Outcomes

Table 3 presents benchmark outcomes in our calibrated economy. The differences between sophisticated and naïve consumers are informative: Naïveté leads to higher debt (along the intensive and extensive margin) and higher default rates. Still, naïve borrowers pay virtually the same average interest rates. This is explained by substantially higher average penalty fees, with the average fee to debt ratio being nearly 16% compared to the 3% of sophisticated borrowers.

4.2 Illustration of Mechanism

Figure 1 shows an example of possible $(q(\phi, \cdot), \phi)$ schedules as defined in Equation 9 for different loan sizes d' with median income being normalized to 1. For the high income earner in Figure 1a we can see how a choice of higher fees can reduce the real interest rate. However, the $(q(\phi, \cdot), \phi)$ schedule does not have to be monotonic as can be seen in Figure 1b. There, the lender faces a low income earner who might default on future debts. Hence, at some point higher fees lead to an increase in default risk. Thus, the lender expects lower profits and consequently increases the interest rates. Note, however, that all the $(q(\phi, \cdot), \phi)$ schedules to the right of the first jump are strictly inferior to all the points to left. They combine higher fees with higher interest rates and would therefore never be chosen by the consumer. This finding ensures that – conditional on a loan size d' and consumer characteristics – the optimal contract is uniquely defined.

These $(q(\phi, \cdot), \phi)$ schedules become important when analyzing a market in which sophisticated and naïve borrowers are pooled together. Since lenders issue one contract for both types, balancing revenue from interest payments and penalty fees, the $(q(\phi, \cdot), \phi)$ schedules become steeper. This means that interest rates are falling more rapidly with the choice of higher fees, since the foregone interest payments are financed by the higher fee payments resulting from naïveté. Later sections will further elaborate on this cross-



(a) High income

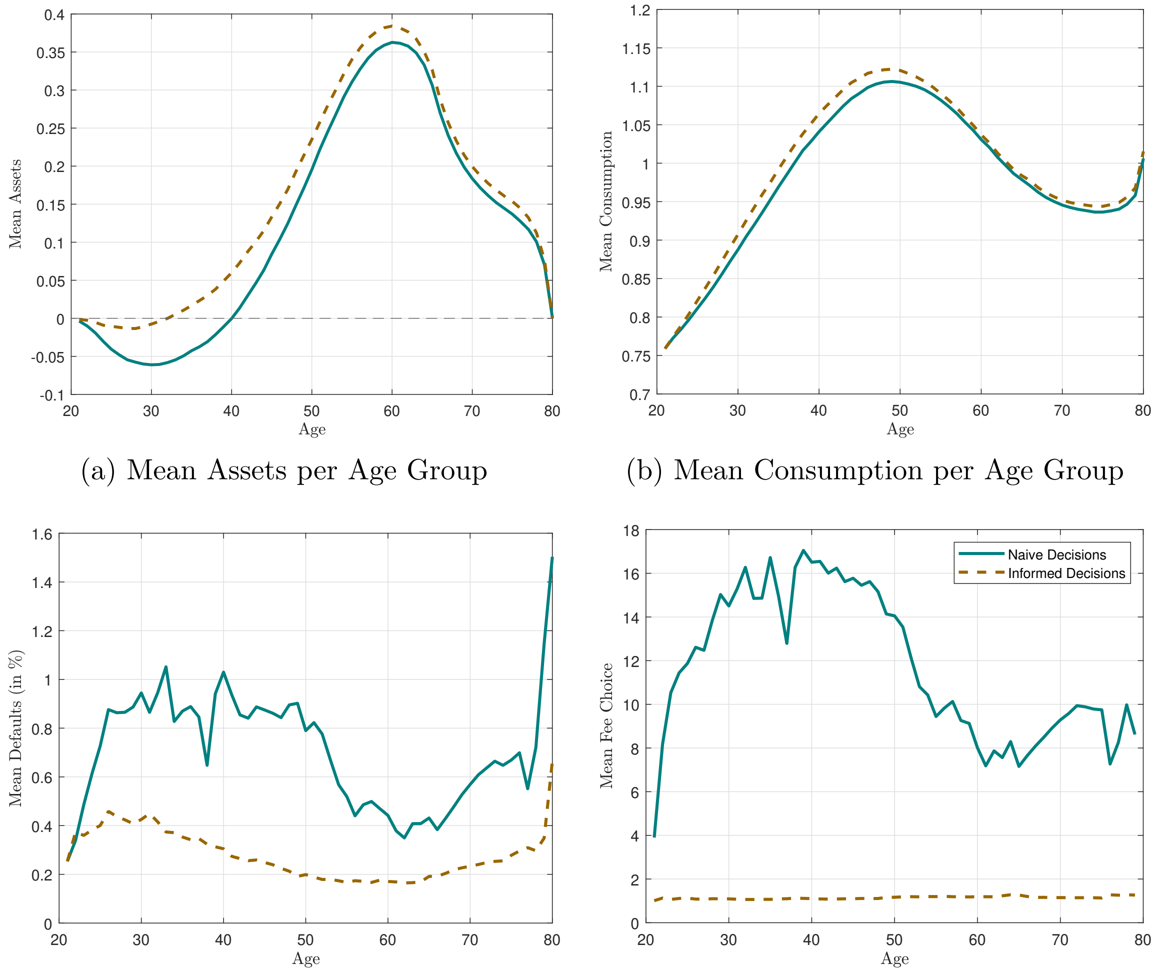
(b) Low income

Figure 1: Example of Equilibrium Credit Contracts (q,ϕ) for high (≈ 3.2) and low income (≈ 0.53), aged 50. Median income normalized to 1.

subsidization effect and show that sophisticates are indeed choosing different contracts in a pooling equilibrium when compared to a separating equilibrium.

To gauge the importance of misunderstanding the risk of financial shocks and fee payments in this framework, Figure 2 shows mean life-time decisions concerning assets, consumption, fees ϕ (conditional on taking out a loan) and bankruptcy for naïve consumers per age group. While the solid cyan line shows the actual decisions taken by naïve consumers, the dotted brown line shows what the optimal behavior would look like if the consumers were aware of the actual probabilities of experiencing financial shocks but faced the same prices. We call this hypothetical agent *informed naïve.* One can see how the naïve consumers choose significantly higher fees than the informed self, resulting in over-accumulation of debts. These two facts lead to higher default rates and lower consumption on average. Overall, the expected welfare of a naïve consumer is 79% lower in terms of CEV than that of an informed self – assessed with the same prices.

Financial Mistakes. We measure financial mistakes as the excess cost of credit that results from naïve borrowers not internalizing their true financial shock risk. In other words, the financial mistake is the difference between the total cost of credit associated with the contract that the naïve consumer chose and the optimal contract that an informed consumer would choose for the same debt level d. (see the discussion in Section 2.3 and Equation 12). This is illustrated in Figure 3. The graph shows the true expectation of the total cost of credit (solid brown line) and the naïve's expectation (dashed cyan line) for different contracts – varying by interest rates and fee choices – for a given loan size d. We can see that given their biased expectation of the resulting costs, the naïve consumer chooses a contract with very high fees (thinking that this choice minimizes their costs). However, as part 3a shows, the actual cost of credit exceeds the naïve's expectation by



(a) Mean Assets per Age Group (b) Mean Consumption per Age Group

(c) Mean Default per Age Group (d) Mean Fees per Age Group

Figure 2: Optimal (dotted brown) and actual decisions (solid cyan) of naïve consumers. Expected cost of Naïveté: 79% CEV.

50%. This is what we call *ToC Bias*.

Due to this bias, the choice of the naïve consumer might be different in a situation where they know their true risks. Figure 3b compares the naïve decision with the decision of a hypothetical informed naïve agent (who forms the correct expectation without affecting prices). We can see that for the given debt level the optimal contract would entail minimal fees ($\phi = 1$). The difference between the associated costs of this hypothetical optimal contract and the actual one chosen is 41%. This is the percentage that the naïve agent overpays for the given loan size due to underestimating their true risks. We call this a *financial mistake*.

4.3 Cross-subsidization and Welfare

Since lenders cannot separate naïve and sophisticated borrowers, they only issue one set of credit contract for both types. Lenders only offer contracts – i.e. interest rate and

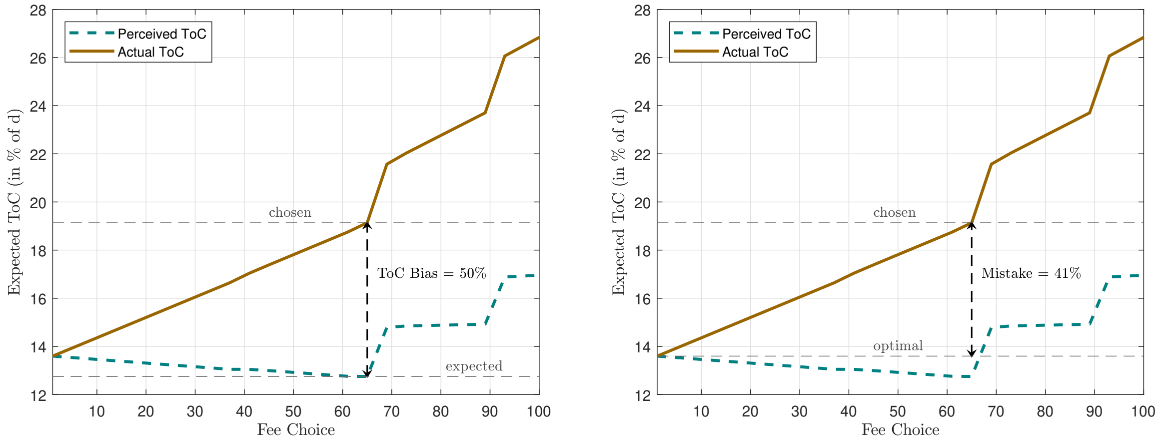


Figure 3: Illustration of ToC Bias and Contract Mistake of a naïve consumer at age 45 with high income seeking out a loan of two times the median income.

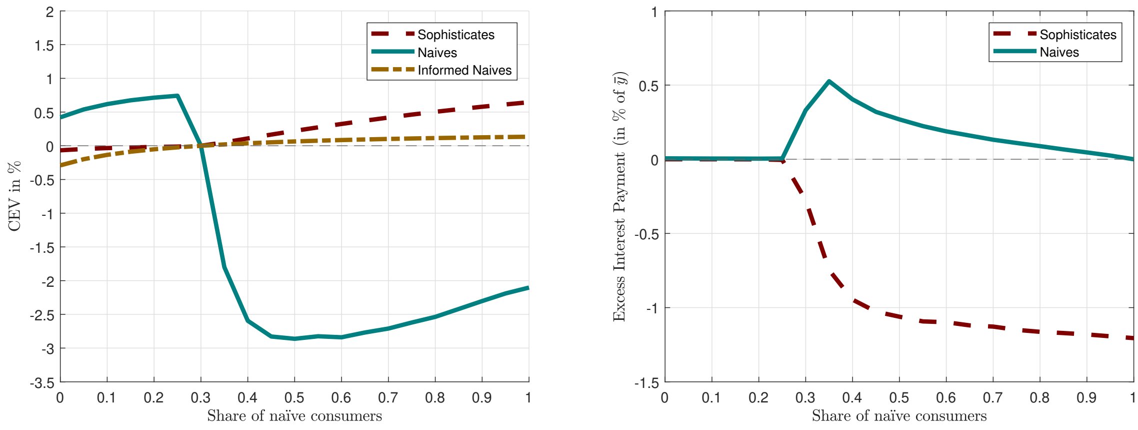
penalty fee combinations – that yield the same expected revenue. However, dealing with sophisticates only would lead to contracts with higher interest rates for each level of fees, since sophisticated consumer are less prone to mistakes. Furthermore, in equilibrium, sophisticates understand their exposure to financial shocks and want to insure against this risk, hence they will pick contracts with lower fees.

In contrast, dealing with naïve borrowers would lead to lower interest rates for each level of fees, as they are more likely to experience bad financial shocks and incur these fees. Furthermore, naïves underestimate their true proneness to these shocks and thus, in equilibrium they choose contracts with higher fees and lower interest rates than they should if they were aware of their risk.

In an economy where sophisticates and naïves are pooled, lenders have to price credit in such a way as to balance these two mechanisms according to the share of naïves in the market. The welfare effects of both types interacting through pooled credit prices are represented in Figure 4a, which plots welfare of sophisticated and naïve agents depending on the share of naïve consumers in the economy. Welfare changes are measured in consumption equivalence variation in percent and calculated relative to the benchmark economy with a share of naïves of 30%. The third line in Figure 4a depicts welfare for an atomistic *informed naïve* agent, i.e., an agent that faces the same prices and the same risk of financial shocks as every naïve borrower but understands their true exposure to these shocks and consequently picks optimal fee levels.

To gauge the importance of cross-subsidization for the observed welfare patterns, Figure 4b plots mean excess interest payments of both types of agents. Excess interest payments are defined as the difference in expected interest $(r(\cdot)d')$ paid by an agent in the pooling equilibrium relative to a hypothetical separating equilibrium in which lenders offer separate contracts for naïve and sophisticated borrowers.

First, note that welfare of sophisticates increases as the share of naïve consumers



(a) Welfare Change (b) Excess interest paid in pooling economy

Figure 4: Average welfare and the amount of cross-subsidization by consumer type and in dependence of the share of naïve consumers in the economy.

increases. Sophisticates profit from the presence of naïve agents in the economy. They profit from cross-subsidization, since pooled interest rates will fall more rapidly as fees increase because more naïves are around. Figure 4b shows that this effect manifests in decreasing average interest rates (or, equivalently, increasingly negative excess interest payments) compared to the absence of naïves. Average interest paid falls as the share of naïves in the pooled market rises. This happens because lenders shift larger parts of their revenue to fees for the financial mistakes made by naïves, thereby increasing fees and lowering interest rates. Since the decisions made by sophisticates are optimal regarding their risk of financial shocks, they benefit relatively more from lower interest rates than they are hurt by higher fees.

In contrast, welfare of naïves is highest in an economy where there are mostly sophisticates and falls sharply as the share of naïves rises, even becoming highly negative before slightly rising back up again. This might seem counter-intuitive because excess interest payments drop as the fraction of naïves increases and there are fewer sophisticates to enjoy cross-subsidization (cf. Figure 4b). However, there is a strong opposing effect reducing naïve welfare. Banks shift their focus from interest payments to penalty fees as the probability of dealing with a naïve agent increases. Abstracting from naïveté and suboptimal fee choices, we first focus on an informed naïve. Her welfare is highest in an economy with only naïve consumers because she can self-insure against her higher risk of mistakes by not choosing very low fees while also enjoying contracts with lower interest rates (similar to sophisticates). As the share of naïves in the economy decreases, the set of equilibrium contracts shifts towards higher interest rates. Thus, contracts for the informed naïve get worse and welfare drops.

In contrast, naïve borrowers fail to recognize their true proneness to financial shocks and therefore do not insure themselves adequately by selecting contracts with lower penalty fees. As a result, they choose inferior contracts that leave them vulnerable to financial shocks by exhibiting large fee payments. Consequently, as lenders shift more revenue collection toward fee payments, naïve borrowers are disproportionally affected and their welfare declines, because their total cost of credit exceeds their expectations. However, with higher shares of naïves in the economy, their welfare begins to increase again, suggesting that naïve borrowers are slightly better off in a fully naïve economy than in one where sophisticates and naïves each share about 50% of the population. This is due to a change in the contract space. As naïve borrowers become more prominent in the economy, the overall default risk rises. Hence, at some point, lenders are not able to increase fee charges any further, since naïve borrowers would simply default on their debt instead of repaying. Consequently, as the share of naïve consumers in the economy approaches 1, average fee payments do not rise any further (in fact, they fall marginally), while interest rates are still kept very low. As a result, naïve welfare improves slightly compared to the worst point where they are pooled with sophisticates. See Figure A.4 for an illustration of these mechanisms.

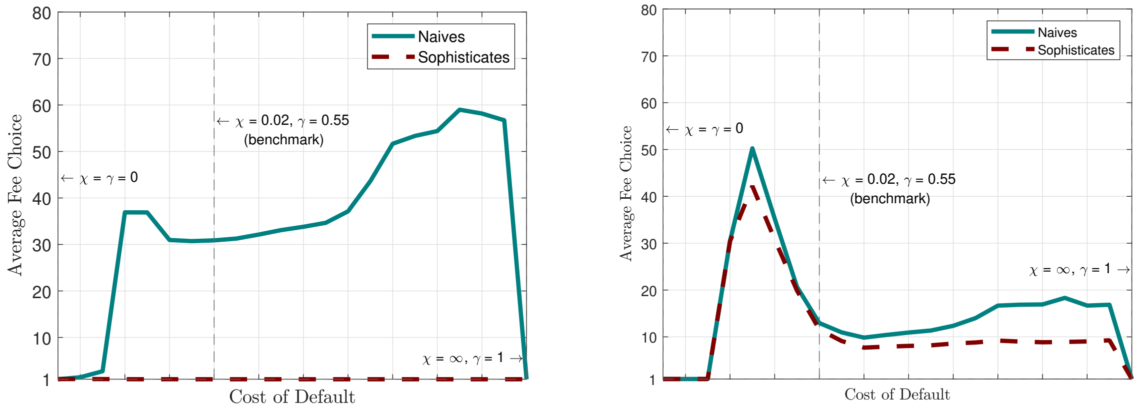
Lastly, the positive welfare gain of almost 1% CEV for naïves in an economy dominated by sophisticates can be attributed to a better set of available contracts and forced insurance via the absence of high-fee options. Since lenders focus more on sophisticates, mostly low-fee contracts are offered and naïve borrowers are effectively protected from their own mistakes. Although they might prefer lower interest rates and higher fee charges ex ante, the limited contract space forces them into safer choices, leaving them ex post better off. As the share of naïves drops below 20%, however, welfare begins to moderately decline again. This is due to the same mechanism driving the welfare of an informed naïve. With the share of naïves being so low, the contracts only focus on interest rates and contracts start becoming worse relative to the optimal point. This change also leads to borrowers shying away from fees altogether as can be seen in Figure 4b, where excess interest payments drop to zero for lower shares of naïves.

Overall, this discussion clearly shows that if the fraction of naïve consumers is high enough, sophisticated borrowers are being cross-subsidized by the mistakes made by naïve individuals. At the same time, naïves can also benefit from a higher share of sophisticates in the market, because this leads to a contract space that forces naïves to choose more optimally as their preferred contracts become unavailable.

4.4 Default Matters

When analyzing credit contracts and the interaction between sophisticated and naïve agents, default is important.[11] Default arises because agents have less than perfect com-

[11] We extend Heidhues and Kőszegi (2010) to incorporate default and better gauge the magnitude of cross-subsidization and its welfare effects.



(a) No Pooling

(b) Pooling

Figure 5: Effects of default costs on average fee choice.

mitment to repay. Agents only repay their debts if it is *ex-post* rational to do so. This limits the amount of late fees a lender can charge and consequently limits cross-subsidization. This has important welfare consequences.

With lower default costs, households face less commitment to repay their debts which is taken into account by creditors when pricing the loans. If filing for bankruptcy is completely free, only very low levels of debt can be sustained in equilibrium and lenders will charge high interest rates. In contrast, if bankruptcy is costly enough to ensure that nobody ever wants to default, households will take on loans only if they are sure that they will be able to repay. This leads to an equilibrium where default is theoretically possible, yet never occurs. Hence, borrowing interest rates are risk-free.

Sophisticates can only profit from lower interest rates due to the presence of naïves if a positive amount of debt can be sustained in equilibrium. If default is costless, lenders are only willing to provide loans at prohibitively high costs and nobody can take out a loan. Consequently, costless default leads to a situation without debt and without cross-subsidization. Similarly, if default costs are very high households will try to minimize default risks and thus choose debt contracts without fees ($\phi = 1$). Again, this leads to a situation without cross-subsidization since sophisticates can only profit from the presence of naïves if the lender expects additional revenue from fee payments.

Figures 5 and 6 show the effects of different levels of default on average fee choice and excess interest payments. To that end, we interpolate between zero default cost and prohibitively high default cost. Default costs have two components: the fraction of garnished income γ and the utility cost χ. Costless default corresponds to $\gamma = \chi = 0$, while in the benchmark we set $\gamma = 0.55$ and $\chi = 0.02$ which leads to an average default rate of 0.45%. Prohibitively high default costs are $\gamma = 1$ and $\chi = \infty$.

Figure 5a shows the effects of higher defaults costs on the average fee choice in an economy without pooling where there is no interaction between sophisticated and naïve

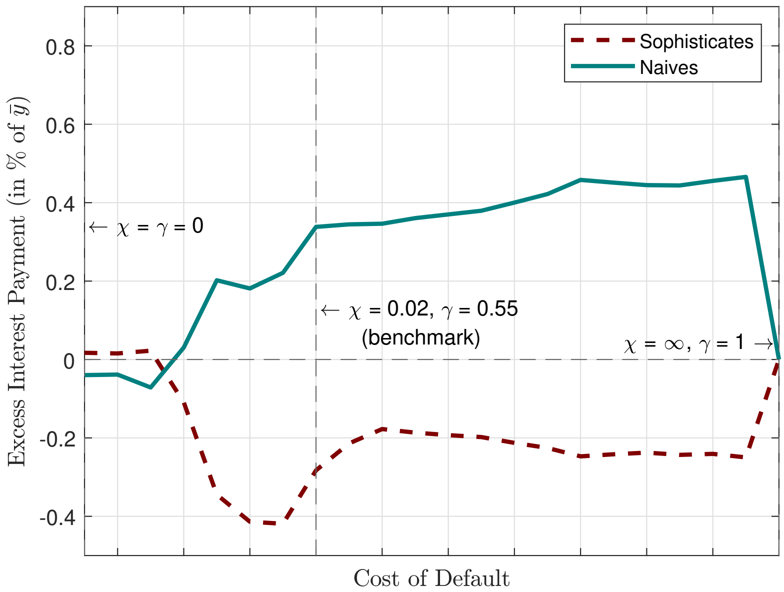


Figure 6: Excess Interest Payments and Default Cost

households. We can see that in such a separating environment choices of naïve and sophisticated borrowers differ drastically. While sophisticates never choose any fees, the average fee choice of naïves rises with the cost of default. Because of limited commitment to repay debts, lenders' expectations of profits from naïve decisions rises as default becomes more costly and thus less likely. Hence, they offer contracts with low interest rates and high fees. Since naïves underestimate their tendency to incur these financial costs, they choose high fees to benefit from the low interest rates. As the costs of filing for bankruptcy increase further, however, the average fee choice falls again. Now, borrowers never want to file for bankruptcy and thus, even naïve consumers start to insure themselves against possible default risks. The choices of sophisticates, in contrast, are unaffected by the cost of default. Since there is no cross-subsidization without pooling, sophisticates don not profit from choosing higher fees and remain at $\phi = 1$.

Figure 5b again shows the effects of changing default costs, but here sophisticates and naïves are pooled and face the same contracts. One striking difference fromFigure 5b is that now sophisticates always choose contracts with higher fees than in the separating economy, exposing themselves to higher costs associated with financial shocks. However, because of pooling, sophisticates can profit from these contracts. At the same time, naïve borrowers choose lower fees (a peak of ca. 50 when pooled vs. nearly 60 when separated) when in the presence of sophisticates. This comes from the fact that – through cross subsidization – naïve interest rates drop slower with increasing fees. Pooled naïves thus find higher fees less attractive and choose lower average levels. Also, note that higher default costs again lead to a decrease in average fee choice. If default becomes so costly that nobody wants to file for bankruptcy anymore, the average chosen fee in the economy even drops to 1, meaning that consumers want to insure against financial shocks. Agents ensure that unexpected financial costs cannot lead to a situation in which they cannot repay and have to default.

The effect of default costs on cross-subsidization is depicted in Figure 6. It plots the costs of default against the excess interest payments compared to an economy without pooling, taking as given the penalty fee choices discussed before. We can see that as long as agents hold debt and there is default in the economy, naïves cross-subsidize sophisticates.[12] This supports the idea that sophisticates choose higher fees when pooled together with naïves only to profit from the lower interest rates. If default becomes too expensive, average penalty fees drop sharply and cross-subsidization all but vanishes. Thus, cross-subsidization crucially hinges on the occurrence of default in equilibrium. In a model with a continuum of credit contracts that trade off interest rates for penalty fees, naïve agents would otherwise self-insure against their perceived risk of mistakes and break the cross-subsidization between them and sophisticates.[13]

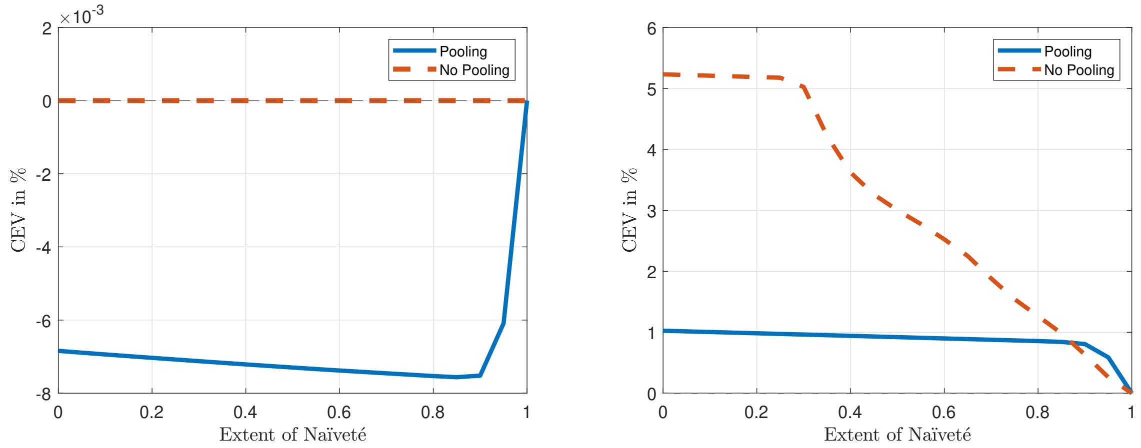
5 Policy Experiments

Policymakers frequently pass legislation to reduce the likelihood of financial mistakes (e.g., standardized language, information requirements, or transparency rules for contracts). Such policies are difficult to study in traditional quantitative models because they have no effect on perfectly informed rational agents. In our framework, however, they map naturally into changes in the informational friction that differentiates sophisticated and naïve borrowers. We can easily simulate different information policies by scaling the extent of naïveté in our model, i.e., reducing the difference in shock risk between sophisticated and naïve consumers. In addition, the model allows us to study interventions that directly restrict contractual terms or borrowing, such as limits on penalty fees and debt-service constraints.

Throughout this section, we analyze these interventions in terms of their effects on welfare (calculated as consumption-equivalent variation relative to the benchmark) and on market outcomes such as default rates and interest rates. To make the mechanisms more transparent, we report results both for the benchmark pooling economy and for a counterfactual separating economy in which lenders can condition contracts on borrower type. Comparing the two enables us to disentangle direct behavioral responses from effects operating through equilibrium re-pricing and cross-subsidization.

[12] In the case of free default, debt is not sustainable in equilibrium. In the case of infinitely costly default, agents will never default. In both of these extreme cases, cross-subsidization breaks down.

[13] We do not see a theoretical argument why prohibitively high bankruptcy cost must always lead to no fees and no cross-subsidization. Rather, in our calibration, financial shocks are large enough that some borrowers sometimes cannot pay them. However, in the opposite case of free bankruptcy and no sustainable debt in equilibrium, there can never be cross-subsidization.



(a) sophisticates (b) naïves

Figure 7: Welfare Effects of Transparency Requirements.

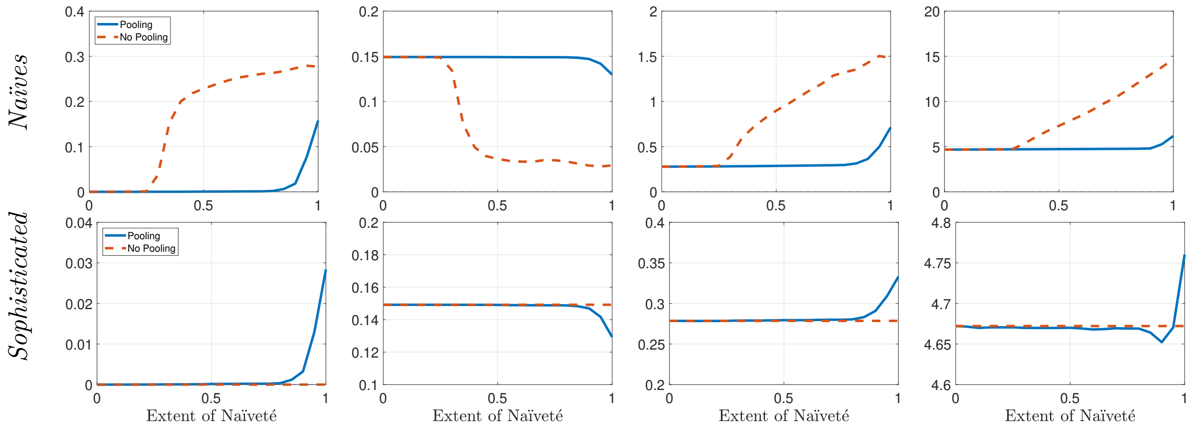
Note: Extent of naïveté is normalized: 1 is benchmark ($\psi = 6$), and 0 is no naïveté ($\psi = 1$).

Finally, beyond welfare and market aggregates, we track how different policies affect both the ToC bias and the incidence of financial mistakes among naïve borrowers. This allows us to evaluate interventions not only by their welfare consequences, but also by their ability to guide naïve borrowers toward more informed financial decisions.

5.1 Transparency Requirements

Credit market regulations and legislations which are aimed at reducing the likelihood of financial shocks by making contracts easier to understand should have an effect on naïve consumers, since they will become better at understanding the true underlying risks they are facing. Sophisticated consumers, in contrast, will not be affected since they are already aware of the true risks. Hence, we capture the effect of an improvement in the transparency of financial contracts as a reduction in the spread between the probabilities of financial shocks of sophisticates and naïves. With ψ denoting this spread as defined in Equation 3, we vary the extent of naïveté from the benchmark ($\psi = 6$) to no naïveté ($\psi = 1$) where naïves face exactly the same shocks as sophisticates. The welfare effects of this experiment are shown in Figure 7 in terms of consumption equivalence (CEV) relative to the benchmark in percent, and Figure 8 reports important credit market outcomes.

In terms of welfare, the effect of improving the transparency of financial contracts is unambiguously positive for naïves as seen in Figure 7b. With full transparency – and hence no more naïveté– welfare of naïve borrowers can be improved by up to 1% in the pooling equilibrium and by up to 5% in the hypothetical separating equilibrium. This is not surprising, given that the policy directly lowers the likelihood of adverse shocks for naïve borrowers. However, note that these welfare gains rise sharply at the beginning, while becoming flatter as transparency is already high and naïveté is rather weak. Improving transparency by just 20%, for example, already leads to welfare gains of approximately



(a) Avg. Fees to Debt (b) Avg. Interest Rate (c) Defaults (in %) (d) Avg. DTI (in %)

Figure 8: Credit Market Outcomes of Transparency Requirements.

0.9% and 1.2% in CEV, with and without pooling, respectively. Interestingly, these welfare gains are much stronger in the hypothetical separated market, where lenders can tailor contacts to each type. In pooling markets, naïve borrowers still benefit, but to a lesser extent, since the presence of sophisticated consumers means that the available contracts already reflect some degree of forced fee discipline. Thus, in the pooling equilibrium, average fee payments by naïves are lower and average interest rates are higher than in the separating equilibrium (see Figure 8). Consequently, being pooled with sophisticates already forces naïves towards better contracts and the additional gain from transparency regulation is smaller than in the separating economy where lenders are able fully exploit the naïve behavior by chargin maximum fees.

For sophisticated borrowers, the welfare effects of higher transparency requirements differ greatly depending on whether they are in an economy with or without naïve consumers, as shown in Figure 7a. The policy has no effect in the economy without pooling, since lenders can perfectly distinguish between sophisticated and naïve borrowers. Sophisticated borrowers have already considered their true risks when deciding on a credit contract and thus, their decisions are unaffected by the extent of naïveté, because banks do not change their offers to sophisticates. Looking at the pooling equilibrium, in contrast, reveals that sophisticates are hurt by the new policies when they are in an economy with naïves. Since improved transparency of financial contracts reduces the revenue banks can expect from fee payments from naïve borrowers, sophisticates lose some cross-subsidization on their interest payments (cf. Figure 8b). Hence, a legislation aimed at reducing the likelihood of financial mistakes will help naïves while hurting sophisticates, meaning that overall welfare effects are ambiguous and depend on the fraction of naïves in the economy.

To further illustrate the mechanisms, Figure 8 displays some important credit market outcomes for sophisticated and naïve borrowers. Figure 8a documents a significant reduction in fees paid by naïves when transparency requirements are introduced. Interestingly, in an economy where naïve and sophisticated borrowers are pooled, this reduction is less pronounced as equilibrium fees are lower to begin with. Naïves choose lower fees because – being pooled with sophisticates that benefit from cross-subsidization – they face interest rates that are less elastic to higher fees. Sophisticates only ever choose penalty fees in the pooling economy, which is when they experience a reduction in average fees paid per debt.

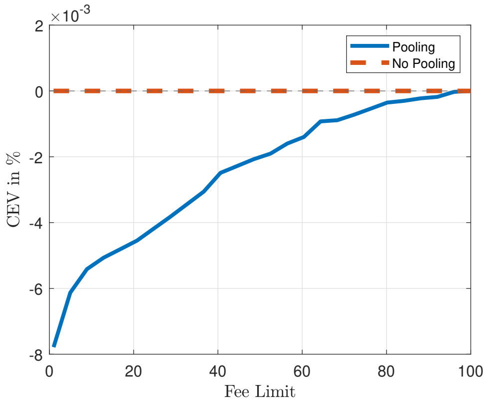
As all equilibrium credit contracts offer the same expected return, lower average fees come with higher average interest rates, cf. Figure 8b. Nevertheless, the shift in contract structure has beneficial effects on risk exposure. Transparency regulation, by limiting exploitative contract features, reduces borrowers' exposure to financial shocks and lowers the likelihood of default. Figure 8c shows that default rates for naïve borrowers in the pooling economy fall by more than 50%, while the impact on sophisticates remains modest.

Lastly, Figure 8d explores the impact of transparency regulations on borrower indebtedness, as measured by the debt-to-income (DTI) ratio. In the pooling economy, the policy has only negligible effects on the DTI levels. This is because the contract space is already constrained by the need to account for both sophisticated and naïve borrowers, limiting the scope for exploitative designs. In contrast, the separated economy exhibits a sharp decline in DTI among naïve borrowers as transparency requirements increase. In the separated benchmark setting, contracts targeted at naïve borrowers feature low interest rates and high fee payments to exploit their limited understanding of the true financial risks. With stronger transparency, naïve borrowers become more aware of these hidden costs and are less willing to accept high-fee contracts. Lenders respond by shifting revenue toward higher interest rates, which in turn dampens the incentive to over-borrow. As a result, borrowing falls significantly and the average DTI drops.

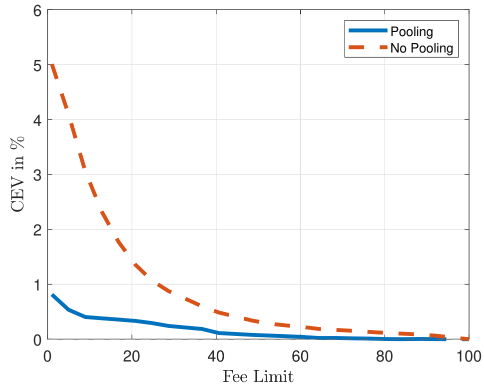
5.2 Limiting Fees

Another possible intervention by policymakers is to impose restrictions on the maximum amount of fees which can be charged for mishandling one's credit (i.e., when incurring a financial shock). Again, such policies are easily simulated within our model by limiting the feasible range of fees, ϕ, in contract offers via imposing an upper bound $\bar{\phi}$. Figure 9 summarizes the welfare effects of various fee limits measured in consumption equivalence relative to the benchmark economy in percent, while Figure 10 reports important credit market outcomes. Again, effects are calculated for an economy with and without pooling.

As with transparency requirements, Figure 9a shows that a fee limit has no effects on sophisticated borrowers in a separated economy. Since sophisticates already insure themselves against the risk of financial shocks by only accepting low-fee contracts, the cap never binds they are unaffected by restrictions. However, when sophisticates are pooled with naïve borrowers, the effects of fee limits become unambiguously negative for them. In pooling markets, sophisticates benefit from contracts with low interest rates and high penalty fees, as lenders anticipate disproportionately higher fee revenue from



(a) sophisticates



(b) naïves

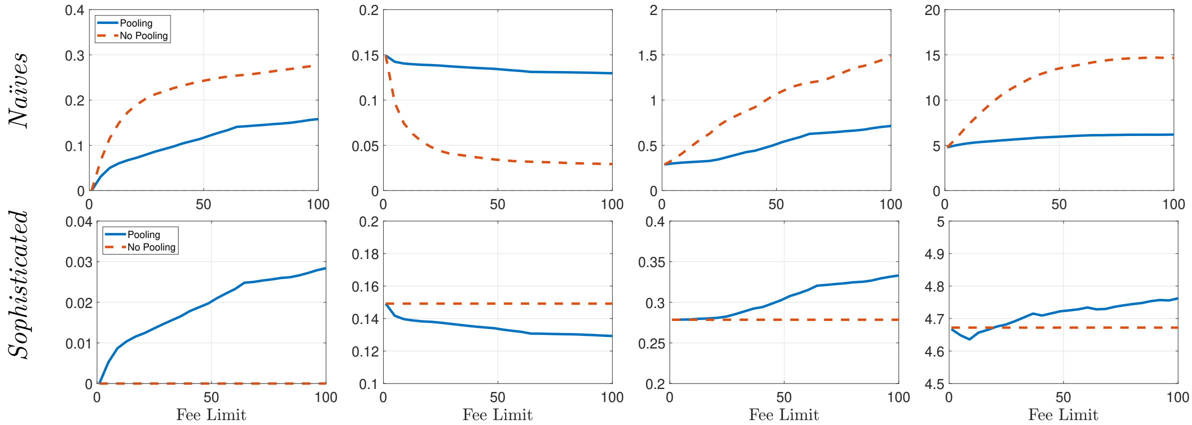
Figure 9: Welfare effects of fee limits. (Benchmark: $\overline{\phi} = 101$)

naïve borrowers. With the introduction of a fee limit, sophisticates are forced to choose different contracts that yield lower fees and higher interest rates (cf. Figures 10a and 10b), meaning that they lose cross-subsidization. Consequently, sophisticates pay more for their credit than in the benchmark economy and they experience welfare losses.

Turning to Figure 9b we see that naïve borrowers benefit from fee caps in both pooling and separating economies, but the gains are larger when contracts are separated. This is because in the separated economy contracts are geared more towards fee payments, since lenders do not have to account for the presence of sophisticates (as can be seen in Figure 10a). Hence, limiting possible fee payments has a stronger effect on contracts in the separating equilibrium than in the pooling equilibrium. The mechanism, however, is the same in both cases: naïves are profiting from the fact that upper bounds on fees reduce the size of their potential mistakes. Stricter fee limits force naïve borrowers to avoid mistakes and choose credit contracts which are closer to the optimal contracts they would choose if they knew their true risks, i.e. contracts with higher nominal interest rates but lower fees in case of a financial shock (cf. Figures 10a and 10b).

Lastly, comparing Figure 9a and Figure 9b shows that welfare effects are evolving at different rates for sophisticated and naïve consumers. While the loss of welfare for sophisticates exhibits a steady fall, welfare gains for naïves rise slowly at first but very sharply at tighter limits. The overall welfare effects of such a policy therefore crucially depend on the specific choice of the fee limit. Note, however, that welfare gains by naïves always seem to outweigh the welfare losses of sophisticates, even when accounting for their different shares in the economy.

To illustrate the underlying mechanisms in more detail, Figure 10 again summarizes important credit market outcomes. Figures 10a and 10b mirror the mechanisms observed under transparency regulation: as fee limits tighten, the average fee burden relative to debt falls, while average interest rates rise. This trade-off reflects the lenders' effort to maintain profitability as they lose fee-based revenue: they compensate by increasing the



(a) Avg. Fees to Debt (b) Avg. Interest Rate (c) Defaults (in %) (d) Avg. DTI (in %)

Figure 10: Credit Market Outcomes of Fee Limits. (Benchmark: $\overline{\phi} = 101$)

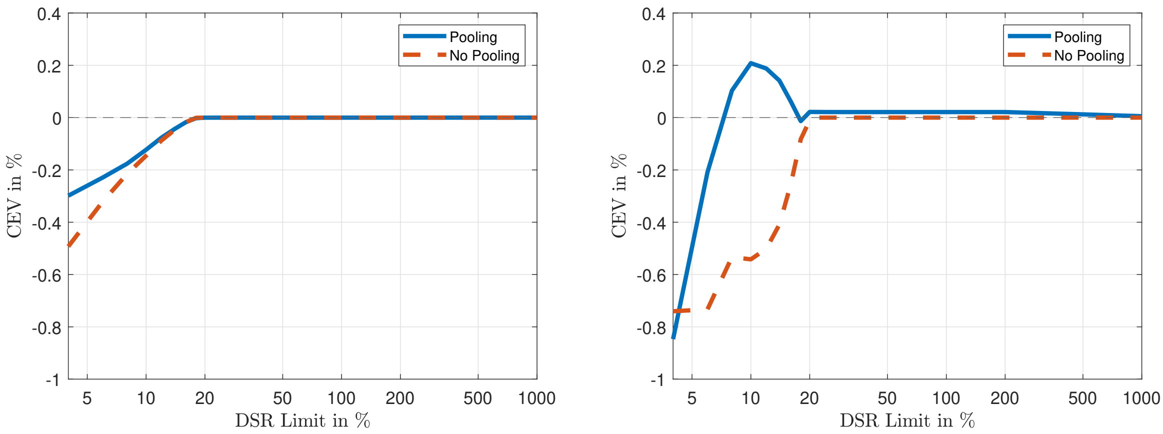
interest rate charged across contracts. However, the degree of this adjustment is much more muted than under transparency reforms where most of the effect is already realized with a modest reduction in naïveté. In contrast, fee limits must be tightened almost fully ($\overline{\phi} = 1$) to achieve a comparable shift in contract pricing. This highlights the bluntness of fee caps as a policy tool: they need to be highly restrictive to meaningfully alter contract structures.

Figure 10c shows that tighter fee limits also lead to a substantial reduction in default rates, particularly for naïve borrowers. As contracts become less backloaded with fees, borrowers face smaller payment spikes when shocks occur. This makes debt burdens more predictable and manageable, improving overall repayment behavior. At the same time, 10d shows that average debt-to-income (DTI) ratios decline slightly, suggesting that borrowers adjust by taking on marginally less debt — likely a response to the higher interest rates resulting from tighter fee constraints.

Taken together, the results in Figure 10 demonstrate that limiting penalty fees reshapes the structure of credit contracts in important ways: it reduces reliance on fees, raises interest charges, and reduces the risk of default, especially for naïve consumers. However, these benefits are only realized under fairly stringent limits. Unlike transparency policies, which work by correcting borrower beliefs and thus improve contract selection directly, fee caps operate through external constraints that restrict the menu of available contracts. As such, they are less efficient in achieving similar outcomes —- particularly in economies with pooling — unless pushed to their most extreme form.

5.3 Borrowing Limits

Naïve borrowers' misperception of their susceptibility to financial shocks not only draws them into unfavorable contracts with high potential fee payments but also results in excessive debt accumulation over the life cycle (see Figure 2). This observation motivates the



(a) Sophisticates

(b) Naïves

Figure 11: Welfare effects of implementing a limit on the debt service ratio.

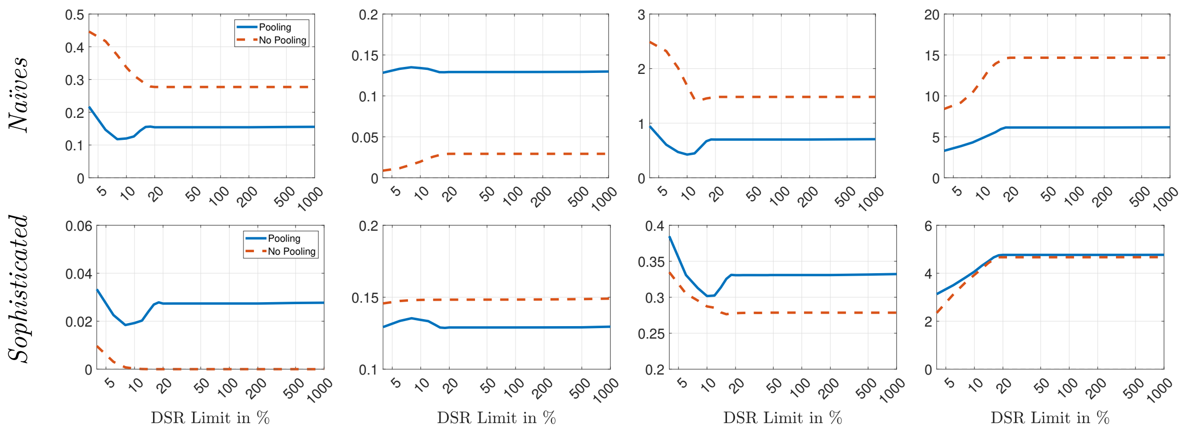
evaluation of policies that focus not on altering contract terms but on limiting borrowing based on consumer characteristics.

In this section, we examine a limit on the *debt service ratio* (DSR), defined as the ratio of interest payments on new debt to income. The interest paid on debt d' with chosen contract q is $(1-q)d'$. The debt service ratio (DSR) measures these interest payments relative to income. Thus, a DSR limit is given by

$$q(d', \phi, s) = \begin{cases} q(d', \phi, s), & \text{if } (1-q(\cdot))d'/y(s) \leq \bar{D} \\ 0, & \text{else.} \end{cases} \tag{15}$$

Figure 11 illustrates the welfare implications of introducing such a DSR limit with varying degrees of tightness. Focusing first on sophisticated consumers, Figure 11a reveals a consistent decline in overall welfare as the DSR constraint tightens. Sophisticated borrowers are already making optimal decisions based on correct beliefs, so any external restriction on their choice set reduces their flexibility and thus hurts them. The overall welfare loss reaches nearly 0.4% under strict limits. Interestingly, this effect is slightly stronger in the hypothetical separating equilibrium than when being pooled with naïves. This suggests that even with tight DSR limits sophisticated borrowers are still profiting from some cross-subsidization that softens the negative effects of the policy.

For naïve borrowers (Figure 11b), the effects are more nuanced and not only depend on the market structure – pooling vs separating equilibrium — but also on the specific choice of the DSR limit. In a pooling economy, where lenders cannot distinguish between borrower types, modest DSR limits (around 10% to 20%) can slightly improve naïve welfare. These constraints reduce exposure to high-fee contracts by flattening the pricing schedule, thereby mitigating some of the financial mistakes made by naïve borrowers. The modest constraint reduces excessive borrowing and limits exposure to worst-case contracts



(a) Avg. Fees to Debt (b) Avg. Interest Rate (c) Defaults (in %) (d) Avg. DTI (in %)

Figure 12: Credit Market Outcomes of Borrowing Limits.

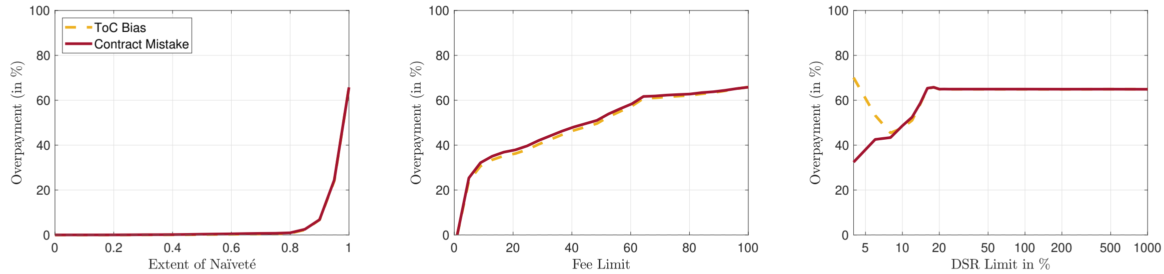
without fully cutting off credit access. However, tighter limits backfire: as naïves seek to remain within the constraint, they accept contracts with higher upfront fees in exchange for lower interest rates. This increases their total borrowing cost and reduces welfare, with a loss of up to -0.8% for very tight limits.

In a separating economy, where lenders tailor contracts specifically to naïves, implementing a DSR limit exclusively has negative implications for borrowers' welfare. Because lenders already offer high-fee, low-interest contracts that exploit naïve misperceptions, imposing a DSR limit forces naïve borrowers to accept even worse fee structures or reduce their debt levels significantly. Either response raises their cost of credit and leads to a welfare decline across the board.

The underlying mechanisms behind these welfare changes become clearer when examining the credit market outcomes shown in Figure 12, which disaggregates the impact of DSR limits on fees, interest rates, default rates, and debt levels, across both pooling and separated economies.

Implementing a DSR limit monotonically drives up average fee payments in the separated economy as borrowers try to lower their interest payments (Figure 12a). This fee hike corresponds to only marginal reductions in interest rates (Figure 12b), suggesting that the trade-off is costly. However, in an economy with pooling, limiting the DSR *might* decrease fee payments. Interest payments only rise ever so slightly, as defaults are also reduced (cf. Figure 12c). The average outstanding debt, see Figure 12d decreases. Thus, limiting the debt service ratio in the presence of pooling might lead to a more efficient equilibrium with lower defaults and, thus, lower deadweight losses. These effects overcompensate naïve borrowers for being constrained in their choices and can actually generate a welfare gain. This effect is absent in the economy without pooling.

Together, these findings suggest that borrowing limits based on debt service ratios can improve overall market efficiency and reduce defaults in economies with mixed borrower types. In these settings, moderate limits lead to less overborrowing and fewer default



(a) Transparency requirements (b) Fee limits (c) DSR limits

Figure 13: Policy effects on average size of over-payment by naïves.

risks, helping Naïves indirectly even if they do not fully understand contract risk by forcing them to avoid choices they would otherwise make. Importantly, these benefits can outweigh the loss of flexibility for naïve borrowers — but only under conditions where pooling softens the adverse trade-offs between fees and interest.

However, in fully separating markets where naïves are already segmented and exploited through pricing, DSR constraints offer little protection and can even exacerbate financial mistakes. Therefore, while DSR-based borrowing limits may have potential in mixed-type markets, they are not a comprehensive solution for improving decision-making quality or fully mitigating behavioral biases.

5.4 Policy Interventions and Financial Mistakes

We have shown that transparency regulations and the imposing of fee limits are effective tools to improve naïve welfare, while the effects of a DSR limit are ambiguous and depend not only on the market structure (pooling vs separated) but also on the level of the limit itself. However, policymakers might not only care about improving overall welfare, but also about helping naïve borrowers avoid mistakes. We therefore analyze the effects of different policies on the average size of mistakes of naïve borrowers. Mistakes are measured as the amount by which the expected total cost of a chosen credit contract exceeds the expected cost of the optimal contract (cf. Equation 12 and Figure 3 in Section 4.2). We also report the Total Cost of Credit (ToC) Bias of the naïve consumer which is defined as percentage difference between the true expected cost of credit and the naïve 's biased expectation (cf. Equation 11 in Section 2.3).

Figure 13 shows the change in the average financial mistakes (i.e. the average over-payment percentage) and the average ToC Bias of naïve borrowers in a pooling economy for all three policies presented above: reducing the likelihood of financial shocks, limiting fees, and imposing DSR constraints. First, note how both the contract mistake and the ToC bias are nearly 70% in the benchmark economy, suggesting that naïve borrowers on average pay nearly 70% more for their credit than when choosing the optimal contract. Also, in the extreme case of completely avoiding mistakes – either by directly abolishing naïveté or by prohibiting lenders to charge any fees – the contract mistake is zero.

Going more into detail, Figure 13a shows the effect of transparency regulations, which reduce the extent of naïveté in the population. As the level of bias falls, both the contract mistake and the ToC bias decline rapidly (and at the exact same rate). A modest 20% reduction in naïveté already nearly eliminates both the contract mistake and the ToC bias. The figure highlights how even partial improvements in borrower understanding lead to substantial reductions in financial errors.

In contrast, Figure 13b examines the effect of fee limits. Here, the reduction in mistakes is much less pronounced. When fees are only moderately capped, the average over-payment remains largely unchanged. Only when fees are nearly eliminated do we observe a significant drop in the contract mistake. Again, the ToC bias follows a similar pattern. This suggests that fee caps are a relatively blunt instrument: unless applied very aggressively, they do little to curb the fundamental bias that leads naïves to select costly contracts.

Finally, Figure 13c presents the impact of borrowing constraints, specifically limits on the debt service ratio (DSR). The effects here are more complex. Moderate limits slightly reduce contract mistakes and ToC bias, but these improvements are limited in scope. As the constraint tightens, the contract mistake initially declines, but then the ToC bias rises again. This reversal occurs because tight DSR limits push naïves toward contracts with lower interest rates but higher fees - contracts that appear cheaper based on their biased expectations, but are in fact more expensive in the long run. The bias worsens even as total cost mistakes decline modestly. This illustrates the policy trade-off: constraints may limit certain mistakes, but they can also distort the perceived attractiveness of remaining contract options. Thus, while DSR limits can reduce the cost of financial mistakes made by naïve borrowers, they do so simply by restricting their access to inferior contracts but without improving their underlying understanding.

Overall, these results suggest that the most effective way to reduce financial mistakes is to target their root cause: the misperception of risk. Transparency reforms that reduce the degree of naïveté are more efficient and less distortionary than price regulation or borrowing constraints. Fee caps and DSR limits may still have value, but their effects on mistake reduction are limited unless implemented very stringently —- and even then, they risk introducing new distortions in borrower behavior.

6 Conclusion

We propose a novel quantitative theory of naïveté in the credit market. We incorporate naïveté into a standard framework of unsecured credit and equilibrium default. The model gives rise to policy interventions as naïve borrowers misunderstand their contracts and make financial mistakes. In equilibrium, they pay high penalty fees that benefit sophisticated borrowers through cross-subsidized interest rates. At the same time, naïve borrowers can also benefit from the presence of sophisticates, as pooled markets force lenders to design contracts that are acceptable to sophisticates, thus indirectly protecting naïves from making inferior choices.

Our framework offers a novel tool for evaluating policies aimed at improving credit card contract design. For example, we find that the 2009 CARD Act tackles two important dimensions of financial mistakes. Firstly, it makes contracts easier to understand and thereby reduces financial mistakes. Through the lens of our model, the welfare effects of such a policy are ambiguous as sophisticated consumers are hurt by higher interest rates, while naïves make less mistakes. The overall welfare effect therefore crucially hinges on the fraction of naïve consumers in the economy.

Secondly, the CARD Act limits the amount of penalty fees that lenders can charge. This limit has a clear and considerably positive impact on naïve borrowers. They are forced to accept contracts with lower fees, which also reduces the size of potential mistakes from wrong contract choices. Hence, by banning high-fee contracts that naïves could otherwise have wrongly chosen, the limit protects them against their future risks, while they themselves would not have done so. Sophisticates, in contrast, are again hurt by such a policy. By reducing the possible size of mistakes made by naïves, cross-subsidization of lower interest rates is also reduced and sophisticates end up paying more for their credit than in the benchmark economy. Also, the welfare gains of naïves and the losses of sophisticates evolve at different rates. Hence, the overall welfare effects of such a policy can be highly ambiguous and depend not only on the fraction of naïves in the economy, but also on the tightness of the limit.

Comparing the effects of these two policies on the size of the mistakes made by naïve consumers, we find that transparency requirements are more effective as sizable results are more easily achieved. Also, in light of the welfare effects, transparency requirements seem to have a less strong negative impact on sophisticates. Hence, while limiting fees might be the easier policy, improving the language used in financial contracts and raising standards for transparency and information requirements seems to be the better solution.

Lastly, as a third avenue, we analyze a debt service ratio (DSR) limit as a form of restricting over-borrowing. Although this policy can reduce the financial burden of naïves by steering them away from risky contracts, it does so without improving their understanding of the underlying risks. Furthermore, positive effects are only found in a specific range, while a tight limit even exacerbates the problem by pushing naïves toward fee-heavy contracts that appear deceptively cheap in order to comply with the borrowing constraint. Moreover, in separating economies where lenders already target naïves with backloaded fees, DSR limits unambiguously lower welfare. Thus, while DSR limits may reduce some forms of overpayment, they do not correct the underlying misperceptions and are unlikely to be effective unless carefully calibrated.

References

Agarwal, Sumit, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S. Souleles (Apr. 2015). "Do Consumers Choose the Right Credit Contracts?" *The Review of Corporate Finance Studies* 4.2, pp. 239–257.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel (2015). "Regulating Consumer Financial Products: Evidence from Credit Cards." *The Quarterly Journal of Economics* 130.1, pp. 111–164.

Armstrong, Mark and John Vickers (2012). "Consumer Protection and Contingent Charges." *Journal of Economic Literature* 50.2, pp. 477–493.

Campbell, Daniel, Andrew Grant, and Susan Thorp (Sept. 2022). "Reducing credit card delinquency using repayment reminders." *Journal of Banking & Finance* 142, p. 106549.

CFPB (2014). *Data Point: Checking account overdraft.* CFPB Data Point. Consumer Financial Protection Bureau.

Chatterjee, Satyajit, Dean Corbae, Kyle Dempsey, and José-Víctor Ríos-Rull (2023). "A Quantitative Theory of the Credit Score." *Econometrica* 91.5, pp. 1803–1840.

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull (2007). "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default." *Econometrica* 75.6, pp. 1525–1589.

DellaVigna, Stefano and Ulrike Malmendier (May 2004). "Contract Design and Self-Control: Theory and Evidence." *The Quarterly Journal of Economics* 119.2, pp. 353–402.

Dou, Yiwei, Julapa Jagtiani, Quinn Maingi, and Joshua Ronen (Aug. 2020). *The Credit Card Act and Consumer Debt Structure.* Working Paper. Federal Reserve Bank of Philadelphia.

Eliaz, Kfir and Ran Spiegler (2006). "Contracting with Diversely Naive Agents." *Review of Economic Studies* 73.3, pp. 689–714.

Exler, Florian, Igor Livshits, James MacGee, and Michèle Tertilt (Dec. 2024). "Consumer Credit with Over-Optimistic Borrowers." *Journal of the European Economic Association.*

Exler, Florian and Michèle Tertilt (2020). "Consumer Debt and Default: A Macro Perspective." In: *Oxford Research Encyclopedia of Economics and Finance.*

Gabaix, Xavier and David Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *The Quarterly Journal of Economics* 121.2, pp. 505–540.

GAO (2006). *CREDIT CARDS. Increased Complexity in Rates and Fees Heightens Need for More Effective Disclosures to Consumers.* Report to the Ranking Minority Member, Permanent Subcommittee on Investigations, Committee on Homeland Security and Governmental Affairs, U.S. Senate. United States Government Accountability Office.

Gao, Pengjie, Allen Hu, Peter Kelly, Cameron Peng, and Ning Zhu (Nov. 2020). *Exploited by Complexity*. Working Paper. SSRN.

Heidhues, Paul and Botond Kőszegi (2010). "Exploiting Naivete about Self-Control in the Credit Market." *American Economic Review* 100.5, pp. 2279–2303.

Heidhues, Paul and Botond Kőszegi (2015). "On the Welfare Costs of Naiveté in the US Credit-Card Market." *Review of Industrial Organization* 47.3, pp. 341–354.

Livshits, Igor, James MacGee, and Michèle Tertilt (2007). "Consumer Bankruptcy: A Fresh Start." *American Economic Review* 97.1, pp. 402–418.

Nakajima, Makoto (2017). "Assessing bankruptcy reform in a model with temptation and equilibrium default." *Journal of Public Economics* 145, pp. 42–64.

Nelson, Scott (Jan. 2020). "Private Information and Price Regulation in the US Credit Card Market." Unpublished Manuscript, Chicago Booth.

PEW (2014). *Overdrawn: Persistent Confusion and Concern About Bank Overdraft Practices*. PEW Report. The PEW Charitable Trust.

Raveendranathan, Gajendran and Georgios Stefanidis (2025). "The Unprecedented Fall in U.S. Revolving Credit." *International Economic Review* 66.1, pp. 393–451.

Stango, Victor and Jonathan Zinman (May 2009). "What Do Consumers Really Pay on Their Checking and Credit Card Accounts? Explicit, Implicit, and Avoidable Costs." *American Economic Review* 99.2, pp. 424–29.

Warren, Elizabeth, Teresa Sullivan, and Melissa Jacoby (Apr. 2000). "Medical Problems and Bankruptcy Filings." *Norton Bankruptcy Law Adviser* 5.

A Additional Figures

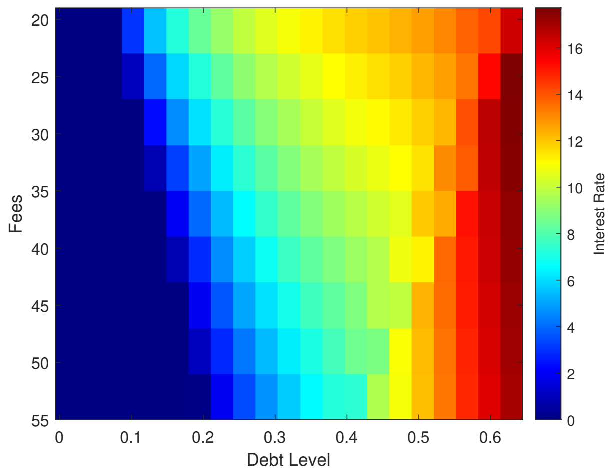


Figure A.1: Example contract space

Note: Choosing higher fees does not necessarily lower interest rates. In this example, for debts above 0.5, borrowers cannot credibly commit to repay in the event of a financial shock. Thus, higher fees can lead to *higher* interest rates.

Figure A.1 shows an example of the contract space and the relationship between interest rates and fees at different debt levels for a middle-aged borrower with median persistent idiosyncratic productivity.

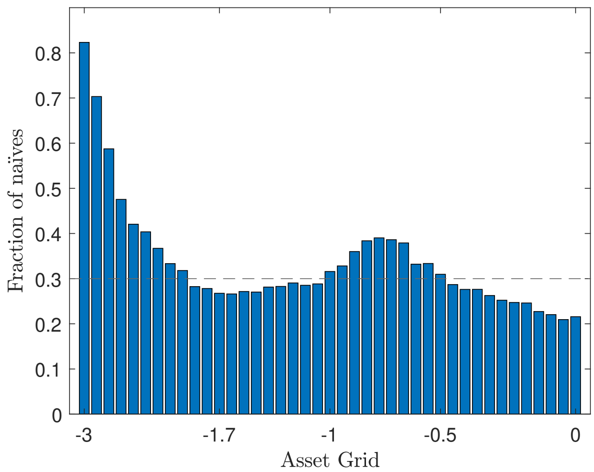


Figure A.2: Example of distribution of naïve consumers over different debt levels (share of naïves in whole economy $\lambda = 0.3$).

Figure A.2 depicts the distribution of naïve borrowers over debt levels. The fraction of naïves typically rises with higher debt levels, while the fraction of naïves at low debt levels is smaller than the economy wide average.

Figure A.3 shows the data on fee payments relative to FICO scores. Above a score of 660, borrowers pay constantly low fees relative to the amount borrowed. Below 660

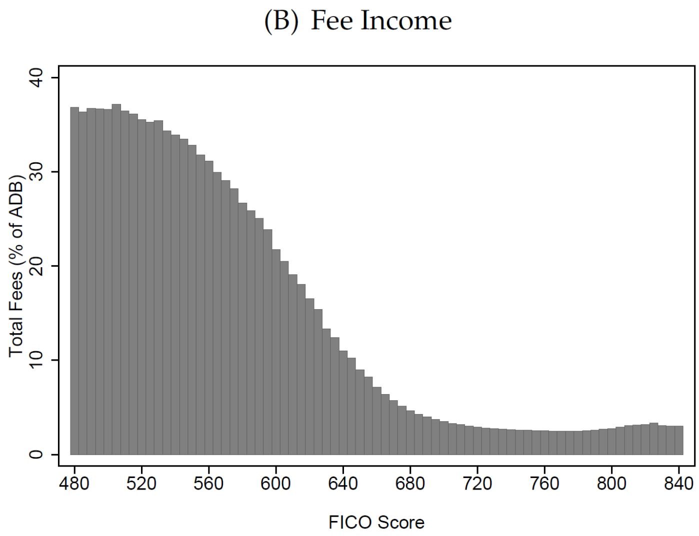
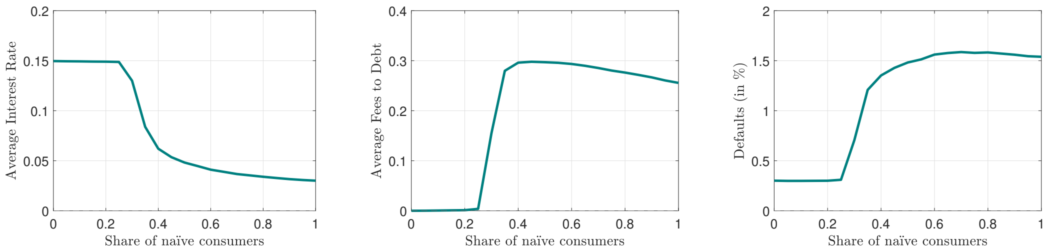


Figure A.3: Fee Payments by FICO Score
Source: Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015, Figure 2 (B))



(a) Average interest rate (b) Average fees to debt (c) Average default rate

Figure A.4: Average credit market outcomes for naïve borrowers depending on the share of naïves in the economy.

however, fee payments rise sharply. The share of borrowers with scores below 660 is 30% in the data.

Figure A.4 shows important credit market outcomes – average interest rates, average fee payments relative to debt, and average default rates – for naïve borrowers depending on the share of naïves in the economy.