A Simple Dynamic Theory of Credit Scores Under Adverse Selection

Dean Corbae
University of Wisconsin - Madison

Andrew Glover
University of Texas - Austin

February 15, 2015

Abstract

We study a dynamic model of unsecured credit markets with adverse selection and an endogenous signal of a borrower's riskiness (modeled as a credit score). Credit contracts are statically constrained efficient in our model, which is achieved by limiting the borrowing of low-risk borrowers while subsidizing the interest rate for the high-risk borrowers. A higher credit score (i.e. higher prior that the borrower is low risk) relaxes the constraint on low-risk borrowers and increases the subsidization for high-risk, which means that utility for both types increases with their credit scores. We calibrate the model to salient features of the unsecured credit market and consider the welfare consequences of different information regimes.

1 Introduction

Credit scores are an important factor for determining eligibility and terms for loans of all types. Despite their widespread use, there is little theoretical research that explains their existence, use in credit markets, dynamics, or social value. Specifically, why does a household's credit score affect the credit terms that he receives once lenders control for his debt levels? Why do lenders appear to charge mark-ups on consumer loans, even though the industry is competitive? Why do credit contracts specify a constant interest rate and a hard limit, rather than allowing households to borrow up to a natural limit at varying rates? We add to this small literature by providing a simple environment in which an individual's propensity to default is private information, but his repayment history provides an informative signal of this propensity.

Our simple environment is best thought of as describing revolving debt. That is, there is a mismatch between the time within a period when a household wishes to consume and when he receives income. Unsecured credit is therefore used to smooth consumption on a rolling basis. Households are heterogeneous in their utility cost of default (i.e., the stigma from default) which makes a given loan more risky for some type households than for others. Households with a high stigma will be less likely to default on a given amount of debt in any state of the world than a low stigma household, and so will be a lower risk borrower. [1] If lenders could perfectly observe a household's risk type then the equilibrium would involve an actuarially fair interest rate function for each type household and a level of debt consistent with household optimization. [2]

In our model, however, the household's risk type is not public information. Optimal credit market contracts are non-linear: they specify an amount that can be borrowed and an amount that must be paid conditional

[1] We will refer to a low-risk household as a low-risk household and a high-risk household as a high-risk household.

[2] Actuarially fair in this context means that the amount of resources delivered to the household is equal to the expected amount paid back. The expectation under full information conditions on both the amount of debt taken *and* risk type.

on repayment. This implies an interest rate, but does not allow households to choose debt freely at that interest rate nor does that interest rate have to be actuarially fair for all levels of debt. Low-risk households may face endogenous *credit limits*: a maximal amount of credit which is strictly below the amount he would choose at the interest rate charged on his debt. Furthermore, the implied interest rate may not be actuarially fair, even conditional on the information available: low-risk households may end up cross-subsidizing high-risk households because this allows lenders to relax credit limits. This in turn gives both household types an endogenous incentive to improve their credit scores: low-risk households get smoother consumption while high-risk households get lower interest rates through cross-subsidization. [3]

Our theory is able to answer the above open questions for models of unsecured credit. Average interest rates fall and average credit limits rise as score rise. Default (repayment) has a negative (positive) effect on a households score, with magnitudes that depend on the initial score (i.e., somebody with a high score who defaults would suffer a larger fall in their score than somebody who had a low score in the first place). Our theory can also generate apparent mark-ups on some credit card contracts that are do purely to informational rents, which previous studies have explained via exogenous intermediation costs. Furthermore, our theory is relatively simple and tractable compared to the literature. This is because our credit market is intra-temporally constrained efficient, which allows us to solve sequential planners problems for the credit market allocation.

In addition to being computationally attractive, this feature allows us to give insight on policies that limit the amount of information used in unsecured credit markets. That is, since the market can fully respond to the new policy regime we can be certain that our welfare numbers are not tainted by ad-hoc restrictions on credit markets.

[3] A credit score here is a prior on the probability that the household is a high-risk type, so one may wish to think of it as a "type score". However, with two risk types and constant repayment rates there is a bijection between this score and the prior on the probability of repayment, which is more in line with a credit score. Given this equivalence for our purposes, we use the more common "credit score" language.

We consider the welfare consequences of two alternative policies. The first is simply a way of bounding the value of additional information, in that we consider what happens if stigmas are suddenly fully public. This obviously increases the ex-ante utility of a new born household (before he realizes his risk type), but redistributes amongst those alive when stigmas become public. low-risk households who have been unlucky and have low scores get a large welfare gain while high-risk households with high scores suffer a large welfare loss.

The second experiment is more directly related to potential policies which could restrict information available in credit markets. We consider the extreme example of completely eliminating credit scores, thereby restricting all priors to the population shares of each risk type. For those alive when this policy is introduced there is straightforward redistribution: anybody with a low score benefits and those with a high score lose. What may be surprising is that a new born household may enjoy an ex-ante welfare *gain* from such a policy. Whether there is a gain or a loss boils down to magnitudes: high-risk households have higher ex-ante welfare without credit scores while low-risk households gain. These low-risk gains are muted though, because low-risk households have volatile consumption in the benchmark economy even though they have higher credit scores on average.

1.1 Relation to Literature

The novel feature of this paper is that we consider a richer contract space than the current literature on unsecured credit markets under private information. This makes our approach technically easier along some dimensions but harder along others. The two papers most closely related to ours are Athreya, et al [1] and Chatterjee, et al [3] so we briefly describe how our approach differs from these. Two important differences are that we consider a much simpler environment (among other things quasi-linear preferences) which allows for a straightforward characterization of the default decision as well as a different equilibrium concept which simplifies the inference problem facing creditors.

Athreya, et al restrict the credit history to be in $\{0, 1\}$ by recording just whether the household repaid in the previous period and they restrict lenders to offer linear prices, which in that context are price functions faced by borrowers who can otherwise choose debt freely. The difficulty in that framework is computing the default rate for a given level of debt, since lenders use the level of debt requested to infer a borrower's private type and the equilibrium concept requires zero profits for each level of debt. While this works well under full information, it opens the possibility of sub-optimality under private information since a lender would be able to make positive profits by offering a non-linear contract such as those we study due to the usual cream-skimming argument as in Rothschild and Stiglitz [**?**]. They therefore search for equilibria in the spirit of Wilson [**?**], who pointed out that if firms anticipate that contracts rendered unprofitable by cream-skimming will be withdrawn from the market, they will also anticipate that the cream skimming contract will end up attracting both high- and low-risk types and therefore turn out to be unprofitable. Given these beliefs, Wilson showed that firms cannot offer separating contracts on which they anticipate earning non-negative profits and so equilibrium with pooling contracts can exist.

The paper by Chatterjee, et al retains the assumption of Wilson contracts, but is closer to ours, since their contracts are inherently non-linear and condition on a dynamic assessment of the borrowers creditworthiness. In their paper, a menu of one-period contracts is offered to households conditional on how much they want to borrow from a finite set of debt levels, which depend on the credit score. The credit score is based on a Bayesian assessment of the borrower's type conditional on asset market behavior like past debt balances and repayments. This is a challenging inference problem, but must be done in order to calculate zero-profit prices. We take a different approach; since our equilibrium concept finds the set of debt levels endogenously, it avoids the inference problem by jointly determining and pricing the debt grid in order to separate households by default risk. More importantly, our debt grid is chosen so that the lender makes zero profits on average over all risk-types with a given score, but not necessarily for each loan.

We use the equilibrium concept described in Netzer and Scheuer [6]. They study the robust sub-game perfect equilibrium of a sequential game of private information between firms competing for one period loans. The salient assumption which we share with Netzer and Scheuer that lenders can offer a menu of contracts (multiple contracts per lender). These menus determine a finite number of debt levels from which households can choose, but in principle any positive amount of debt can be listed as part of the menu. This game reduces to solving a type of planner's problem with incentive compatibility constraints and generates separating equilibria with cross-subsidization. We think the assumption of multiple contracts is a better description of the unsecured credit market, but also demonstrate theoretical and computational advantages to this assumption. We find this environment ideal both because it has a strong micro foundation and is extremely tractable.

Our contracts are represented by menus which cater to households of different types: they have an (implicit) interest rate, but also a borrowing level (which will turn out to be a constraint for some households). These menus may depend on a household's observable credit score (indeed, they will in equilibrium). This additional richness is important because, in equilibrium, low-risk households repay in more second sub-period states than high-risk households. This outcome, coupled with expected utility preferences, means that a single-crossing property holds for the indifference curves over credit and interest rates. Compared to a high-risk household, a low-risk household requires a larger increment in the amount lent to him for a given increment in his debt due. This is used to separate households by type, period by period, using a combination of credit constraints (on the low-risk households) and interest rate subsidies (for the high-risk households). Importantly, the credit limit for the low-risk household is truly a credit *constraint* in this model, which differs from Chatterjee, et al [2] where rising interest rates induce a natural credit limit above which households would never *choose* to borrow.

In addition to these important assumptions about the equilibrium concept and contracting environment, we have also simplified the model via a judicious choice of intra-period timing and household preferences. Specifically, we do not allow for inter temporal savings and all debt is intra-period, which makes the household's credit score the only inter temporal state variable. We also use quasi-linear preferences, which ensures that the repayment choice has a simple cut-off property over expenditure shocks. These modeling choices share many features with Lagos and Wright [?], but the application is quite different.

The paper proceeds by first describing the economic environment, then characterizing the equilibrium and finally demonstrating the results of our numerical experiments.

2 Environment

We study a dynamic economy in which households must borrow in credit markets subject to adverse selection.

2.1 Population and Preferences

Time is discrete with index $t = 0, 1, ...T$, where T is potentially infinite. Each period consists of two sub-periods. There are two classes of agents with a unit measure of each in existence at any point of time. We will refer to the first class as households and they will be our main point of interest. The second class we will refer to as lenders.

A household could be one of two types indexed by $i \in \{h, \ell\}$. There are μ_i of each type of household at each point in time, with $\mu_h + \mu_\ell = 1$. An individual household dies probabilistically each period with probability $1 - \delta$ and δ new agents are born in each period (μ_i of whom are type i). Households discount at rate γ and we will write the effective discount factor as $\beta = \gamma\delta$.

Households value consumption in both sub-periods ($c_{1,t}$ and $c_{2,t}$) and dislike work (n_t), which they can only provide in the second sub-period. Households can take an action $d_t \in \{0, 1\}$ in the second sub-period which bear some disutility. In any given period t, preferences are given by

$$u_i(c_{1,t}, c_{2,t}, n_t, d_t) = \log c_{1,t} + \log c_{2,t} - n_t - \lambda_i d_t. \tag{1}$$

We will call λ_i the household's "stigma" parameter and let $\lambda_h > \lambda_\ell$.

Lenders can provide labor only in the first sub-period and value consumption only in the second sub-period. Survival and discount factors for lenders are the same as households, but lenders are risk neutral over consumption:

$$U^{\text{lender}} = \mathbb{E} \sum_{t=0}^{\infty} \beta^t (c_{2,t} - n_{1,t}) \tag{2}$$

We will assume that the amount of consumption good produced per unit of labor is always unitary, so that the assumption of linear disutility of labor for lenders means that they can create enough consumption good to fulfill loan contracts (i.e., we do not put an upper bound on $n_{1,t}$).

2.2 Technology and Shocks

There is a technology available in the second sub-period of each period t for converting labor into consumption goods at a one-to-one rate. This technology is available to all households. The output from this technology is not storable across periods.

There is one source of uncertainty in the model, which is an exogenous expenditure shock (e.g. medical expenditure) in the second sub-period. The shock, $\tau_{t,j} \in \{\tau_1, ...\tau_J\}$ with probabilities $\{\phi_1, ...\phi_J\}$ is iid over time and households.

2.3 Markets, Legal Framework, and Information

Since a household's earnings or output comes at the end of the second sub-period and it wants to consume the non-storable good in the first subperiod, a household would like to borrow against its future resources. In exchange for lenders delivering Q_t units of the consumption good in the first subperiod, households promise to repay b_t in the second sub-period. However, the legal system allows a household to renege on that promise, freeing them from repaying both b_t and τ_t. We will call this action "default" (i.e. set $d_t = 1$). While default frees the household from payments $(b_t + \tau_t)(1 - d_t)$, it also bears type dependent stigma costs $\lambda_i d_t$.

We are interested in credit extension when a household's type is private information. We will assume that defaults are observable. That is, at the beginning of a period t after which a household has set $d_k = 1$ for some set of $k \in \{0, ..., t-1\}$, the household is associated with a public history:

$$d^t = (d_0, ..., d_{t-1}) \tag{3}$$

Lenders observe a household's history of default, d^t, which allows them to infer the probability that a household has high stigma costs of default (i.e. parameter λ_h). We we will refer to this probability as $s_t(d^t)$. Lenders can also infer the probability that a household of type i repays b_t via rational expectations, which is denoted by $\pi_{i,t}(b_t, d^t)$. For a household with history d^t, the lender's expected profits from delivering Q_t units of consumption to a household of type i in the first sub-period and receiving promised repayment of b_t units in the second sub-period is then:

$$\Pi_{i,t}(Q_t, b_t, d^t) = -Q_t + \pi_{i,t}(b_t, d^t) b_t \tag{4}$$

Note that the history of repayment choices affects the expected profit of lenders only through the probability of a household being of the high type and on the household's solvency probability.

2.4 Timing

As previously mentioned, a period is divided into two sub-periods. The timeline can be seen as follows:

- Sub-Period One: Households borrow Q_t and consume $c_{1,t}$. The amount borrowed is chosen along with repayment promise, b_t, from a set of options posted by lenders. Lenders work sufficient hours to meet their lending responsibilities.

- Sub-Period Two:
  - Households receive expenditure shock τ_t
  - Households choose labor supply n_t and whether to pay debt and expenditure shock $(b_t + \tau_t)(1 - d_t)$. Notice that default is on both of these jointly.
  - Households consume $c_{2,t}$.

3 Equilibrium

3.1 Household Decision Problems

Households enter the first sub-period without savings (by assumption). Suppose that a household has credit contract given by (Q_t, b_t). Then the household's budget set is given by

$$c_{1,t} \leq Q_t \tag{5}$$
$$c_{2,t} \leq n_t - (b_t + \tau_t)(1 - d_t) \tag{6}$$

Since the household's utility is strictly increasing in consumption, we can substitute out budget constraints and write indirect utility over credit market contracts. For $t = T$ this gives:

$$\mathcal{U}_{i,T}(Q, b, d^T) = \log Q + \sum_{j=1}^{J} \phi_j \left\{ \max_{n, d \in \{0,1\}} \log\left(n - (1-d)(b + \tau_{T,j})\right) - n - \lambda_i d \right\} \tag{7}$$

For a household with history d^T, it chooses its credit contract (Q, b) from a set of those offered by intermediaries (described below) denoted $\mathbb{C}_T(d^T)$. Then we can define:

$$V_{i,T}(d^T) = \max_{(Q,b) \in \mathbb{C}_T(d^T)} \mathcal{U}_{i,T}(Q, b, d^T) \tag{8}$$

This in turn allows us to recursively define indirect utility at $t < T$ as:

$$\mathcal{U}_{i,t}(Q,b,d^t) = \log Q + \tag{9}$$
$$\sum_{j=1}^{J} \left\{ \phi_j \max_{n,d\in\{0,1\}} \log\left(n-(1-d)(b+\tau_j)\right) - n - \lambda_i d + \beta V_{i,t+1}\left((d^t,d)\right) \right\}.$$

Similar to above, this gives the value function over d^t:

$$V_{i,t}(d^t) = \max_{(Q,b)\in\mathbb{C}_t(d^t)} \mathcal{U}_{i,t}(Q,b,d^t) \tag{10}$$

3.2 Credit Market

As stated previously, we use the Netzer and Scheuer concept to generate our credit market contracts. There is a sequential game in the background, but for our purposes we need only characterize the equilibrium contracts arising form such a game. We assume that the set of contracts available, $\mathbb{C}_t(d^t)$,is determined from a sequence of programming problems which are analogous to the static model of insurance under adverse selection in Netzer and Scheuer.

$$\mathbb{C}_t(d^t) = \underset{Q_h,b_h,Q_\ell,b_\ell}{\operatorname{argmax}} \mathcal{U}_{h,t}\left(Q_h,b_h,d^t\right) \tag{11}$$
$$\text{s.t.}$$
$$\mathcal{U}_{\ell,t}\left(Q_\ell,b_\ell,d^t\right) \geq \mathcal{U}_{\ell,t}\left(Q_h,b_h,d^t\right) \tag{12}$$
$$\mathcal{U}_{h,t}\left(Q_h,b_h,d^t\right) \geq \mathcal{U}_{h,t}\left(Q_\ell,b_\ell,d^t\right) \tag{13}$$
$$s_t(d^t)\left[-Q_h+\pi_{h,t}(b_h,d^t)b_h\right] + \left(1-s_t(d^t)\right)\left[-Q_\ell+\pi_{\ell,t}(b_\ell,d^t)b_\ell\right] \geq 0 \tag{14}$$
$$\mathcal{U}_{\ell,t}\left(Q_\ell,b_\ell,d^t\right) \geq \max_b \mathcal{U}_{\ell,t}\left(\pi_{\ell,t}(b,d^t)b,b,d^t\right) \tag{15}$$

This maximization problem is equivalent to that of a social planner who is assigned to each pool of households with history d^t subject to incentive compatibility constraints, score-specific resource feasibility, and a participation constraint for the high-risk households. The first two constraints above are just the incentive compatibility, which specifies that a low-risk household prefers a contract tailored to him and likewise for a high-risk household.

The third contract is the resource feasibility for contracts conditioned on the observable signal $s_t(d^t)$. This allows for cross-subsidization on the interest rate within a score, but not across different scores. The final constraint forces the planner to deliver to the high-risk household at least as much utility as what he would get with actuarially fair interest rates.

3.3 Definition of Equilibrium

A Bayesian Competitive Equilibrium for this economy consists of:

1. A Bayesian probability function of histories: $\forall t, d^t : s_t(d^t)$ which gives the probability that a household is low-risk given his history d^t along with the policy functions of other households and credit contract sets.

2. Sets of available credit contracts: $\forall t, d^t : \mathbb{C}_t(d^t)$ which contain contracts that solve the planner's problems defined above.

3. Household credit market choices: $\forall t, d^t, i \in \{h, \ell\} : (Q_{i,t}(d^t), b_{i,t}(d^t) \in \mathbb{C}_t(d^t)$ and labor and default choices $\forall t, d^t, i \in \{h, \ell\} : n_{i,t}(\cdot, d^t) : \{0, ..., \tau_N\} \rightarrow \mathbb{R}_+, d_{i,t}(\cdot, d^t) : \{0, ..., \tau_N\} \rightarrow \{0, 1\}$ which solve the household problems given the sets of contracts available and the updating function.

4 Equilibrium Existence and Characterization

Throughout the remainder of the paper we will look at the case of $J = 3$. This is the minimum number of shocks required to have signals affect allocations in equilibrium and for signals to affect default decisions. We will build the equilibrium by beginning with the case of $T = 1$ and then adding additional periods.

4.1 Equilibrium For $T = 1$

We start with the full information version of the economy as a useful benchmark, then move to the version with adverse selection.

4.2 Full Information Benchmark

Under full information, the only difference will be in the credit market. If a household's type is observable, then a competitive market for loans means that each type must get an actuarially fair price so that:

$$Q_{i,T} = \pi_{i,T}(b)b_i \tag{16}$$

where we have suppressed time subscripts for this static model. Once we know the amount of debt taken by the household we can find his hours and default choice by solving:

$$\max_{d,h} \log(h - (1-d)(b+\tau)) - \lambda_i d - h \tag{17}$$

This gives hours $h = 1 + (1-d)(b+\tau)$ and default decision $d = 1 \leftrightarrow b + \tau > \lambda_i$. Furthermore, it allows us to rewrite Equation 8 as:

$$\mathcal{U}_{i,T}(Q,b) = \log Q - 1 - \sum_{j=1}^{3} \phi_j (b+\tau_j) \mathbb{I}_{\{b+\tau_j \leq \lambda_i\}} - \lambda_i \sum_{j=1}^{3} \phi_j \mathbb{I}_{\{b+\tau_j > \lambda_i\}} \tag{18}$$

Which means that the solvency probability is known and is equal to:

$$\pi_{i,T}(b) = \begin{cases} 1 & : b \leq \lambda_i - \tau_3 \\ \phi_1 + \phi_2 & : \lambda_i - \tau_3 < b \leq \lambda_i - \tau_2 \\ \phi_1 & : \lambda_i - \tau_2 < b \leq \lambda_i - \tau_1 \\ 0 & : \lambda_i - \tau_1 < b \end{cases} \tag{19}$$

which gives a further simplification of Equation 8:

$$\mathcal{U}_{i,T}(Q,b) = \log Q - 1 + \pi_{i,T}(b)b - \begin{cases} \sum_{j=1}^{3} \phi_j \tau_j & : b \leq \lambda_i - \tau_3 \\ \phi_1\tau_1 + \phi_2\tau_2 + \phi_3\lambda_i & : \lambda_i - \tau_3 < b \leq \lambda_i - \tau_2 \\ \phi_1\tau_1 + (\phi_2 + \phi_3)\lambda_i & : \lambda_i - \tau_2 < b \leq \lambda_i - \tau_1 \\ \lambda_i & \lambda_i - \tau_1 < b \end{cases} \tag{20}$$

Using Equation 16 we can solve for the full information credit contract by solving:

$$\max_b \mathcal{U}_{i,T}(\pi_{i,T}(b)b, b) \tag{21}$$

which is equivalent to solving for the optimal $b^k_{i,T}$ in each sub-region $\mathbb{B}^k_{i,T}$ define as:

$$\mathbb{B}^k_{i,T} = \begin{cases} (-\infty, \lambda_i - \tau_3] & : k = 1 \\ (\lambda_i - \tau_3, \lambda_i - \tau_2] & : k = 2 \\ (\lambda_i - \tau_2, \lambda_i - \tau_1] & : k = 3 \\ (\lambda_i - \tau_1, \infty) & : k = 4 \end{cases} \tag{22}$$

Doing so gives the debt choice in each region:

$$b^k_{i,T} = \begin{cases} \min\{1, \lambda_i - \tau_3\} & : k = 1 \\ \min\{(\phi_1 + \phi_2)^{-1}, \lambda_i - \tau_2\} & : k = 2 \\ \min\{\phi_1^{-1}, \lambda_i - \tau_1\} & : k = 3 \\ 0 & : k = 4 \end{cases} \tag{23}$$

The optimal level of debt can then be found from:

$$b_{i,T} = \text{argmax}\left\{\mathcal{U}_{i,T}(\pi_{i,T}(b^k_{i,T})b^k_{i,T}, b^k_{i,T})\right\} \tag{24}$$

This problem has two features. First, the debt levels are priced at the solvency rate consistent with each value. Second, the optimal level of debt must be chosen from this finite set of levels (at most Jvalues). This brings us to our first set of assumptions and result:

Assumption 1 *The high stigma household does not repay τ_3 units of debt but does repay repay $\phi_1^{-1}+\tau_2$ units and hence $(\phi_1+\phi_2)^{-1}+\tau_2$ as well. That is:*

$$(\phi_1+\phi_2)^{-1}+\tau_2<\phi_1^{-1}+\tau_2\leq\lambda_h<\tau_3 \tag{25}$$

This assumption guarantees that the high-type household takes debt $(\phi_1+\phi_2)^{-1}$ and repays when he receives shocks τ_1 and τ_2. The stronger assumption guarantees that the high type would also repay the level of debt taken by the low-type household.

Assumption 2 *The high-risk household repays $\phi_1^{-1}+\tau_1$ units of debt but does not repay τ_2 units. That is:*

$$\phi_1^{-1}+\tau_1\leq\lambda_\ell<\tau_2 \tag{26}$$

This assumption guarantees that the high-risk household takes debt ϕ_1^{-1} and repays only when he gets shock τ_1.

Proposition 1 *Under Assumptions (1) and (2), the full-information equilibrium with $T=1$ is given by:*

$$Q_{h,T}=1 \tag{27}$$

$$Q_{\ell,T}=1 \tag{28}$$

$$b_{h,T}=\frac{1}{\phi_0+\phi_1} \tag{29}$$

$$b_{\ell,T}=\frac{1}{\phi_0} \tag{30}$$

The assumptions on λ_ℓ guarantee that $b_{\ell,T}^k<0$ for $k=1,2$ and $b_{\ell,T}^3=\phi_1^{-1}$. Since $b<0$ is not possible, the only possible solution is our claim. The assumptions on λ_h guarantee that $b_{h,T}^1<0$ and $b_{h,T}^2=(\phi_1+\phi_2)^{-1}$. Furthermore, $b_{h,T}^3=\phi_1^{-1}$ and chooses $b_{h,T}^2$. This is true since there are only two cases. The first is that $b_{h,T}^3+\tau_2\leq\lambda_h$ (i.e. the low-risk household would repay $b^3_{h,T}$ when $\tau = \tau_2$). Even if $b^3_{h,T} + \tau_2 > \lambda_h$, the low-risk household prefers $b^2_{h,T}$ since:

$$\begin{aligned} \mathcal{U}^2_{h,T} - \mathcal{U}^3_{h,T} = \\ -1 - \phi_1\tau_1 - \phi_2\tau_2 - \phi_3\lambda_h - [-1 - \phi_1\tau_1 - (\phi_2 + \phi_3)\lambda_h] \\ = \phi_2 (\lambda_h - \tau_2) > 0 \end{aligned} \tag{31}$$

4.3 Private Information Credit Market

We now assume that lenders cannot directly observe a household's type, but instead know the probability that he is the low-risk type, which in the $T = 1$ model is given by the population fraction of low-risk households, μ_h.

The first thing to note before we start solving the private information credit market problem is that the low type's repayment probability is ϕ_1 for debt up to $\lambda_\ell - \tau_1$ and zero for any debt above that level. The low-risk household's repayment probability is $\phi_1 + \phi_2$ for debt up to $\lambda_h - \tau_2$ and ϕ_1 for debt above that level but below $\lambda_h - \tau_1$ and zero beyond that. Given that the full-information level of debt for the low-risk household is in the region $(0, \lambda_h - \tau_2]$, we will assume that to be the case here as well and then verify that solution ex-post. Assuming that $(b_h, b_\ell) \in (0, \lambda_h - \tau_2] \times (0, \lambda_\ell - \tau_1]$ allows us to simplify the programming problem as follows:

$$\max \log Q_h - (\phi_1 + \phi_2) b_h - \Gamma_h \tag{32}$$

$$\text{s.t.}$$

$$\log Q_\ell - \phi_1 b_\ell - \Gamma_\ell \geq \log Q_h - \phi_1 b_h - \Gamma_\ell \tag{33}$$

$$\log Q_h - (\phi_1 + \phi_2) b_h - \Gamma_h \geq \log Q_\ell - (\phi_1 + \phi_2) b_\ell - \Gamma_h \tag{34}$$

$$\mu_h [-Q_h + (\phi_1 + \phi_2) b_h] + (1 - \mu_h) [-Q_\ell + \phi_1 b_\ell] \geq 0 \tag{35}$$

$$\log Q_\ell - \phi_1 b_\ell - \Gamma_\ell \geq \max_b \log(\phi_1 b) - \phi_1 b - \Gamma_\ell \tag{36}$$

where the Γ_i are given by:

$$\Gamma_h = -1 - \phi_1\tau_1 - \phi_2\tau_2 - \phi_3\lambda_h \tag{37}$$

$$\Gamma_\ell = -1 - \phi_1\tau_1 - (\phi_2 + \phi_3)\lambda_\ell \tag{38}$$

Now, this can be transformed into a concave programming problem by making the change of variables $u_i = \log(Q_i)$ and $\beta_i = -b_i$. The only constraint that haS to be checked is the feasibility constraint, but that describes a convex set since the inverse utility function is exponential. That is, with the change of variables the constraint becomes:

$$g(u_h, b_h, u_\ell, b_\ell) = \mu_h \left[-e^{u_h} + (\phi_1 + \phi_2) b_h\right] + (1 - \mu_h) \left[-e^{u_\ell} + \phi_1 b_\ell\right] \geq 0 \quad (39)$$

Taking $x^1 = (u_h^1, b_h^1, u_\ell^1, b_\ell^1), x^2 = (u_h^2, b_h^2, u_\ell^2, b_\ell^2), \theta \in (0,1), x^\theta = \theta x^1 + (1 - \theta) x^2$ we can see that:

$$g(x^\theta) > \theta g(x^1) + (1 - \theta) g(x^2) \quad (40)$$

This means we can solve the problem for different values of μ_h and perform comparative statics. Importantly, the values and allocations are continuous (but not necessarily smooth) in μ_h. This is the first step in solving for $T > 1$ since the household's score will act as μ_h in period T.

4.4 Characterization

Since we know that the original problem is equivalent to a concave one and is therefore well behaved, we will work with it instead. The first thing to note is that the low-type household will always get at least $\bar{U}_{\ell,T} = -1$ utils, and may get more than that. Therefore, we can introduce the additional utility that he receives as Δ and let that be a choice variable. That allows us to drop the final (participation) constraint for the low type. In addition, we will drop the incentive compatibility constraint for the high-type and check that is slack ex-post. Finally, we will solve the problem for a general fraction of low-risk households, s, since the solution will be valid at $s = \mu_h$

as well. Then we have the problem:

$$\max_{Q_h, b_h, \Delta, Q_\ell} \log Q_h - (\phi_1 + \phi_2) b_h \tag{41}$$

s.t.

$$\bar{U}_\ell + \Delta \geq \log Q_h - \phi_1 b_h \tag{42}$$

$$s\left[-Q_h + (\phi_1 + \phi_2) b_h\right] + (1-s)\left[\log Q_\ell - Q_\ell - \bar{U}_\ell - \Delta\right] \geq 0 \tag{43}$$

We can see that $Q_h < 1$, which is the amount of first sub-period consumption delivered to the low-risk household under full information. In this sense, the low-risk household is borrowing constrained but pays an interest rate which is higher than actuarially fair given his repayment rate [4] At its highest the interest rate implies that the low-risk household has to pay back debt $\frac{1}{\phi_1}$. Notice that Q_ℓ only enters on the right hand side of a positive inequality, thus can be eliminated by solving $\max_Q \log(Q) - Q \rightarrow Q_\ell = 1$ and $b_\ell = \frac{1}{\phi_1}[1 - \Delta]$. It is easy to see that the two inequalities must bind at an optimum, so we can eliminate Δ and b_h by using $(\phi_1 + \phi_2) b_h = Q_h + \frac{1-s}{s}\Delta$ and $\phi_1 b_h = \log Q_h - \Delta + 1$ to get:

$$\Delta = \frac{s\phi_1}{s(\phi_1 + \phi_2) + (1-s)\phi_1}\left[\frac{\phi_1 + \phi_2}{\phi_1}\log Q_h - Q_h + \frac{\phi_1 + \phi_2}{\phi_1}\right] \tag{44}$$

This can then be substituted in for $\pi_h b_h$ which transforms the maximization problem into simply:

$$\max_Q \left[s(\phi_1 + \phi_2) - (1-s)\phi_2\right] \log Q - Qs(\phi_1 + \phi_2) \tag{45}$$

This gives the solution:

$$Q_h = \frac{s(\phi_1 + \phi_2) - (1-s)\phi_2}{s(\phi_1 + \phi_2)} \tag{46}$$

Notice that this equation has intuitive comparative statics. If s rises then the household is more likely to be the high type, and therefore the lender

[4]Note that he may even pay a higher rate than the high-risk household, although this is not the case in our example.

is better able to cross-subsidize the (relatively few) low types, which allows for better consumption smoothing for the high type. This reaches the full info limit when $s \to 1$ and we can see that $\lim_{s\to 1} Q_h = 1$. Away from $s = 1$ we would expect Q_h to rise with s, which is exactly what we get:

$$\frac{\partial Q_h}{\partial s} = \frac{1}{s^2(\phi_1 + \phi_2)^2} > 0 \tag{47}$$

This equation holds so long as $\Delta > 0$, which cannot be true in every parameterization and for every s. We need the following to hold:

$$\frac{\phi_1 + \phi_2}{\phi_1} \log \frac{s(\phi_1 + \phi_2) - (1-s)\phi_2}{s(\phi_1 + \phi_2)} > \frac{s(\phi_1 + \phi_2) - (1-s)\phi_2}{s(\phi_1 + \phi_2)} - \frac{\phi_1 + \phi_2}{\phi_1} \tag{48}$$

In order to evaluate these expressions, we need $s(\phi_1 + \phi_2) > (1-s)\phi_1 \to s > \underline{s} = \frac{\phi_1}{2\phi_1 + \phi_2}$. Approaching this point from above, we know that the left hand side is smaller (approaches $-\infty$) than the right hand side (approaches $-\frac{\phi_1+\phi_2}{\phi_1}$). On the other hand, if $s = 1$ then the left hand side is equal to zero while the right-hand side is equal to $\frac{-\phi_2}{\phi_1+\phi_2} < 0$. Therefore, we know that there are at least some values of s near one where the inequality holds. In fact, we can show that there is a unique s^* for which the inequality holds if $s > s^*$. Define:

$$h(s) = \frac{\phi_1 + \phi_2}{\phi_1} \log Q_h(s) - Q_h(s) + \frac{\phi_1 + \phi_2}{\phi_1} \tag{49}$$

where I have just resubstituted $Q_h(s)$. Then we can calculate:

$$h'(s) = \left[\frac{\phi_1 + \phi_2}{\phi_1} Q_h(s)^{-1} - 1\right] \frac{\partial Q_h}{\partial s} \tag{50}$$

which is positive for $s > \underline{s}$, since $Q_h(s) \leq 1$ and $\frac{\phi_1+\phi_2}{\phi_1} > 1$.

For values of $s < s^*$ we cannot use $\Delta > 0$ but instead have that $\log(Q_\ell) - \phi_1 b_\ell = -1$ which means that $Q_\ell = \phi_1 b_\ell$ and $Q_h = (\phi_1 + \phi_2) b_h$. The low-type incentive compatibility then gives the high-type allocation by solving:

$$-1 = \log(Q_h) - \frac{\phi_1}{\phi_1 + \phi_2} Q_h \tag{51}$$

Note that using $Q_h(s)$ and the above equation coincide at $s*$, which means that the allocation is continuous in s but kinks at s^*

5 Two-Period Model

We now consider the $T = 2$ case. We will take the above as the model in the second (and last) period and add an initial period which is identical except that the household realizes that his actions will affect his future utility. We therefore begin by writing the household's choice problem in the second sub-period of period $t = 1$:

$$\max_{d,h} \log\left(h - (1-d)(b+\tau)\right) - h - d\left[\lambda_i + V_{i,2}(0) - V_{i,2}(1)\right] + V_{i,2}(0) \quad (52)$$

Importantly, notice that the cost of default now depends on the difference between entering the second period with a default flag, in which case the credit market allocation will be the one associated with $s_T(1)$, versus entering without the flag, in which case the credit market allocation will be the one with $s_T(0)$. Therefore, in a dynamic setting, the utility cost of default depends on both the exogenous stigma but also on the endogenous change in future credit market terms. In order to evaluate these functions we need to know the values of the posteriors, which are consistent with Bayes rule. There are three possibilities, only one of which is possible in equilibrium:

1. $s_2(0) = s_2(1)$. *This is not possible.* This would mean that $V_{i,2}(0) = V_{i,2}(1)$, which means that the cost of default is just λ_i as in the static model. Since the above assumptions ensure that the high-risk household defaults when $\tau \geq \tau_2$ but the low-risk household remains solvent at $\tau = \tau_2$ in the $T = 1$ model, we get:

$$s_2(0) = \frac{\mu_h(\phi_1 + \phi_2)}{\mu_h(\phi_1 + \phi_2) + \mu_\ell \phi_1} \quad (53)$$

$$s_2(1) = \frac{\mu_h \pi_\infty}{\mu_h \pi_\infty + \mu_\ell(\phi_2 + \pi_\infty)} \quad (54)$$

   From which it is clear that $s_2(0) > s_2(1)$.

2. $s_2(0) < s_2(1)$. *This is not possible.* The period T utility of both types of households is increasing in the score s, therefore this would imply that $V_{i,2}(0) < V_{i,2}(1)$. Under our assumptions, the the high-risk household defaults on $\tau = 1$ in the static model, so he must also default on $\tau = 1$ in this case. If the high stigma household now defaults when $\tau = 1$, then both are defaulting in the same states and $s_2(0) = s_2(1)$, a contradiction. If he continues to repay then the posterior must update positively according to equations 53 and 54, which is also a contradiction.

3. $s_2(0) > s_2(1)$. *This is the only possible equilibrium outcome.* We will provide an assumption below so that the low-risk household still defaults on τ_3 and the high-risk household defaults on τ_2 and τ_3, just as in the static model. Then the posteriors are identical to those in equations 53 and 54.

The assumptions which guarantee an equilibrium as described in possibility three are very similar to the static model, except they must reflect the new dynamic incentives to repay debt. Those incentives are endogenous and not easily found analytically, so we use extreme values to find sufficient conditions. The fundamental insight is that the highest T utility for both households is achieved if $s_2 = 1$. At that score a household gets contract $(Q, b) = (1, \frac{1}{\phi_1+\phi_2})$, which is just the full-info allocation for the low-risk household. The worst contract is associated with $s_2 = 0$ and is equal to $(1, \frac{1}{\phi_1})$, which is the high-risk household's full-info contract. With these points in mind, we make two additional assumptions:

Assumption 3 *In the second sub-period of $t = 1$ in the economy with $T = 2$, the low-risk household does not repay τ_3 units of debt, but does repay both $(\phi_1 + \phi_2)^{-1} + \tau_2$ and $\phi_1^{-1} + \tau_2$ units. That is:*

$$\phi_1^{-1} + \tau_2 \leq \lambda_h \tag{55}$$

and

$$\lambda_h + \beta(\phi_1 + \phi_2)\left[\frac{1}{\phi_1} - \frac{1}{\phi_1 + \phi_2}\right] < \tau_3 \tag{56}$$

Assumption 4 *The high-risk household repays* $\phi_1^{-1} + \tau_1$ *units of debt but does not repay* τ_2 *units. That is:*

$$\phi_1^{-1} + \tau_1 \le \lambda_\ell \tag{57}$$

and

$$\lambda_\ell + \beta\phi_1 \left[\frac{1}{\phi_1} - \frac{1}{\phi_1 + \phi_2}\right] < \tau_2 \tag{58}$$

These assumptions are built to give the same solvency rates in $t = 1$ in the $T = 2$ economy as in the $T = 1$ economy. The resulting programming problem is:

$$\max \log Q_h - (\phi_1 + \phi_2) b_h - \Gamma_{h,1} \tag{59}$$

$$\text{s.t.}$$

$$\log Q_\ell - \phi_1 b_\ell - \Gamma_{\ell,1} \ge \log Q_h - \phi_1 b_h - \Gamma_{\ell,1} \tag{60}$$

$$\log Q_h - (\phi_1 + \phi_2) b_h - \Gamma_{h,1} \ge \log Q_\ell - (\phi_1 + \phi_2) b_\ell - \Gamma_{h,1} \tag{61}$$

$$\mu_h \left[-Q_h + (\phi_1 + \phi_2) b_h\right] + (1 - \mu_h)\left[-Q_\ell + \phi_1 b_\ell\right] \ge 0 \tag{62}$$

$$\log Q_\ell - \phi_1 b_\ell - \Gamma_{\ell,1} \ge \max_b \log(\phi_1 b) - \phi_1 b - \Gamma_{\ell,1} \tag{63}$$

where the terms $\Gamma_{i,1}$ constants are given by:

$$\Gamma_{h,1} = -1 - \phi_1\tau_1 - \phi_2\tau_2 + \beta V_{h,2}(0) - \phi_3\left[\lambda_h + \beta(V_{h,2}(0) - V_{h,2}(1))\right] \tag{64}$$

$$\Gamma_{\ell,1} = -1 - \phi_1\tau_1 + \beta V_{\ell,2}(0) - (\phi_2 + \phi_3)\left[\lambda_\ell + \beta(V_{\ell,2}(0) - V_{\ell,2}(1))\right] \tag{65}$$

Proposition 2 *Suppose that Assumptions 3 and 4 hold in the economy with* $T = 2$*. Then an equilibrium exists and has first period allocations* $\{(Q_{\ell,1}, b_{\ell,1}), (Q_{h,1}, b_{h,1})\}$ *which solve the programming problem described by (59) through (63), second period scores consistent with Equations (53) and (54), and second period contracts which solve the programming problem described by (32) through (36).*

The proof is simple, we just need to make sure that the household default decisions are consistent with the second period utilities under these posteriors. That means that we need:

$$\tau_2+b^*_{\ell,1} > \lambda_\ell+\beta V_{\ell,2}\left(\frac{\mu_h(\phi_1+\phi_2)}{\mu_h(\phi_1+\phi_2)+\mu_\ell\phi_1}\right)-\beta V_{\ell,2}\left(\frac{\mu_h\pi_\infty}{\mu_h\pi_\infty+\mu_\ell(\phi_2+\pi_\infty)}\right) \tag{66}$$

The lowest possible value of $b^*_{\ell,1}$ is $\frac{1}{\phi_1+\phi_2}$, which is the low-risk household's full information debt. This is larger than zero. As described above, the largest that the difference in future utilities could be is $\beta\phi_1\left[\frac{1}{\phi_1}-\frac{1}{\phi_1+\phi_2}\right]$, thus the second inequality in Assumption (4) guarantees that the high-risk household would default when $\tau \geq \tau_2$. Since $b^*_{\ell,1} \leq \frac{1}{\phi_1}$ thanks to the cross-subsidization of the high-risk households credit contract, the first inequality of Assumption (**??**) guarantees that he will remain solvent for $\tau = \tau_1$.

Similar reasoning applies for the low-risk household. The first inequality in Assumption (3) guarantees that he will repay any level of debt he may receive in the proposed equilibrium since $b^*_{h,1} \leq \frac{1}{\phi_1}$, which is the high-risk's debt under full information. The second inequality in the assumption guarantees that he would default on $\tau = \tau_3$ (again, using the bound on discounted second period utilities).

6 T-Period Model

We now consider the general version of $T > 2$. We again provide assumptions so that the low-risk household repays for $\tau \in \{\tau_1, \tau_2\}$ and the high-risk household repays for only $\tau = \tau_1$.

Assumption 5 *In the second sub-period of $t = 1$ in the economy lasting T periods, the low-risk household does not repay τ_3 units of debt, but does repay both $(\phi_1+\phi_2)^{-1}+\tau_2$ and $\phi_1^{-1}+\tau_2$ units. That is:*

$$\phi_1^{-1}+\tau_2 \leq \lambda_h \tag{67}$$

and

$$\lambda_h+\beta(\phi_1+\phi_2)\left[\frac{1-\beta^{T-1}}{1-\beta}\right]\left[\frac{1}{\phi_1}-\frac{1}{\phi_1+\phi_2}\right] < \tau_3 \tag{68}$$

This also guarantees that, in the second sub-period of $t = T-1$ in the *economy lasting T, the low-risk household does not repay τ_3, but does repay both $(\phi_1+\phi_2)^{-1}+\tau_2$ and $\phi_1^{-1}+\tau_2$ units, i.e. Assumption (3) is implied by this one.*

Assumption 6 *In the second sub-period of $t=1$ in the economy lasting T periods, the high-risk household repays $\phi_1^{-1}+\tau_1$ but does not repay τ_2. That is:*

$$\phi_1^{-1}+\tau_1 \leq \lambda_\ell \tag{69}$$

and

$$\lambda_\ell + \beta\phi_1\left[\frac{1-\beta^{T-1}}{1-\beta}\right]\left[\frac{1}{\phi_1}-\frac{1}{\phi_1+\phi_2}\right] < \tau_2 \tag{70}$$

This also guarantees that, in the second sub-period of $t=T-1$ in the economy lasting T, the high-risk household does not repay τ_2, but does repay $\tau_1^{-1}+\tau_1$, i.e. Assumption (4) is implied by this one.

With these assumptions, we have a proposition similar to the $T=1$ and $T=2$ cases. That is, the solvency probabilities are $\phi_1+\phi_2$ for the low-risk household and ϕ_1 for the high-risk household, so that the score updates are:

$$s_{t+1}\left((d^t,0)\right) = \frac{(\phi_1+\phi_2)s_t(d^t)}{(\phi_1+\phi_2)s_t(d^t)+\phi_1\left(1-s_t(d^t)\right)} \tag{71}$$

$$s_{t+1}\left((d^t,1)\right) = \frac{\phi_3 s_t(d^t)}{\phi_3 s_t(d^t)+(\phi_2+\phi_3)\left(1-s_t(d^t)\right)} \tag{72}$$

And credit market contracts solve the following for all t and d^t:

$$\max \log Q_h - (\phi_1+\phi_2)b_h - \Gamma_{h,t} \tag{73}$$

$$\text{s.t.}$$

$$\log Q_\ell - \phi_1 b_\ell - \Gamma_{\ell,1} \geq \log Q_h - \phi_1 b_h - \Gamma_{\ell,t} \tag{74}$$

$$\log Q_h - (\phi_1+\phi_2)b_h - \Gamma_{h,t} \geq \log Q_\ell - (\phi_1+\phi_2)b_\ell - \Gamma_{h,t} \tag{75}$$

$$s_t(d^t)\left[-Q_h+(\phi_1+\phi_2)b_h\right] + \left(1-s_t(d^t)\right)\left[-Q_\ell+\phi_1 b_\ell\right] \geq 0 \tag{76}$$

$$\log Q_\ell - \phi_1 b_\ell - \Gamma_{\ell,t} \geq \max_b \log\left(\phi_1 b\right) - \phi_1 b - \Gamma_{\ell,t} \tag{77}$$

where the terms $\Gamma_{i,t}$ constants are given by:

$$\Gamma_{h,t} = -1 - \phi_1\tau_1 - \phi_2\tau_2 + \beta V_{h,t+1}\left((d^t,0)\right) - \phi_3\left[\lambda_h + \beta V_{h,t+1}\left((d^t,0)\right) - \beta V_{h,t+1}\left((d^t,1)\right)\right] \quad (78)$$
$$\Gamma_{\ell,t} = -1 - \phi_1\tau_1 + \beta V_{\ell,t+1}\left((d^t,0)\right) - (\phi_2+\phi_3)\left[\lambda_\ell + \beta V_{\ell,t+1}\left((d^t,0)\right) - \beta V_{\ell,t+1}\left((d^t,1)\right)\right] \quad (79)$$

where we assume that $d^1 = 0$ and $s_1(d^1) = \mu_h$. Further, $\Gamma_{i,T}$ differs from $t < T$ by:

$$\Gamma_{h,t} = -1 - \phi_1\tau_1 - \phi_2\tau_2 - \phi_3\lambda_h \quad (80)$$
$$\Gamma_{\ell,t} = -1 - \phi_1\tau_1 - (\phi_2+\phi_3)\,\lambda_\ell \quad (81)$$

This leads us to our final proposition.

Proposition 3 *Suppose that Assumptions (5) and (6) hold in the economy lasting T periods. Then an equilibrium exists and has $t-$period allocations which are functions of a household's score, $\left\{(Q_{\ell,t}(s), b_{\ell,t}(s))_{t=1}^T, (Q_{h,t}(s), b_{h,t}(s))_{t=1}^T\right\}$ that solve the programming problems described by (73) through (77) where score updating is done via Equations (71) and (72).*

This holds as $T \to \infty$ as well, which is the case we use for quantitative experiments, but first we find the stationary distribution in that case.

6.1 Law of Motion of Distribution

We augment the model now to have death so that we can get a stationary distribution of scores in the case of $T \to \infty$. This doesn't change anything from our analysis above except that we must now interpret part of the discount factor as a survival probability. Let such a survival probability be δ, and assume that $1-\delta$ new agents enter the economy each period, each of whom has a high stigma with probability μ.

We first take a given score, s, and ask what values of s_{-1} would have given that score. There are at least two possible values, which come from the updating functions under solvency and default. We therefore define two correspondences, s^0_{-1} and s^1_{-1}:

$$s = \frac{(\phi_1+\phi_2)s^0_{-1}(s)}{(\phi_1+\phi_2)s^0_{-1}(s)+\phi_1\left(1-s^0_{-1}(s)\right)} \tag{82}$$

$$s = \frac{\phi_3 s^1_{-1}(s)}{\phi_3 s^1_{-1}(s)+(\phi_2+\phi_3)\left(1-s^1_{-1}(s)\right)} \tag{83}$$

As can be seen, these are functions since we can solve for them as:

$$s^0_{-1}(s) = \frac{\phi_1 s}{\phi_1+\phi_2+s(\phi_1-\phi_1-\phi_2)} \tag{84}$$

$$s^1_{-1}(s) = \frac{(\phi_2+\phi_3)s}{\phi_3+s(\phi_2+\phi_3-\phi_3)} \tag{85}$$

This means that, given a CDF for the $i-$type distribution over credit scores, $F_{i,n}$, we can define the operator:

$$T_h[F](s) = \delta\left\{\int_0^{s^0_{-1}(s)}(\phi_1+\phi_2)dF(x)+\int_0^{s^1_{-1}(s)}\phi_3 dF(x)\right\}+(1-\delta)\mathbb{I}_{\{s\geq\mu_h\}} \tag{86}$$

$$T_\ell[F](s) = \delta\left\{\int_0^{s^0_{-1}(s)}\phi_1 dF(x)+\int_0^{s^1_{-1}(s)}(\phi_2+\phi_3)dF(x)\right\}+(1-\delta)\mathbb{I}_{\{s\geq\mu_h\}} \tag{87}$$

But of course these can be simplified to:

$$T_h[F](s) = \delta\left\{(\phi_1+\phi_2)\,F\left(s^0_{-1}(s)\right)+\phi_3 F\left(s^1_{-1}(s)\right)\right\}+(1-\delta)\mathbb{I}_{\{s\geq\mu_h\}} \tag{88}$$

$$T_\ell[F](s) = \delta\left\{\phi_1 F\left(s^0_{-1}(s)\right)+(\phi_2+\phi_3)\,F\left(s^1_{-1}(s)\right)\right\}+(1-\delta)\mathbb{I}_{\{s\geq\mu_h\}} \tag{89}$$

Notice that these operators map continuous CDFs on $[0,1]$ into continuous CDFs on $[0,1]$ since $s^0_{-1}(0)=s^1_{-1}(0)=0$ and $s^0_{-1}(1)=s^1_{-1}(1)=1$. Furthermore, the operators satisfy Blackwell's sufficient conditions for a contraction mapping. That is, letting $p_h=\phi_1+\phi_2$ and $p_\ell=\phi_1$, we can check both monotonicity and discounting. First, let $F(s)\geq G(s)$ for some continuous

CDFs F and G. Then we have:

$$T_i[F](s) - T_i[G](s) = \tag{90}$$
$$\delta \left(p_i[F(s^0_{-1}(s)) - G(s^0_{-1}(s))] + (1-p_i)[F(s^1_{-1}(s)) - G(s^1_{-1}(s))] \right) \geq 0$$

And for any CDF F and $a > 0$:

$$T_i[F+a](s) = T_i[F](s) + \delta a \tag{91}$$

We therefore know that there is a unique $F_i(s)$ for $i \in \{h, \ell\}$.

7 Quantitative Experiments

We now provide some quantitative results for the model with $J = 3$ and $T = \infty$. Given the parsimonious nature of our model, there are in effect only three free parameters, the two repayment rates of each risk-type and the measure of households who are born as low risk, which we choose to match an overall repayment rate of 70.77%, a subprime repayment rate of 24.26% and a population fraction of subprime households of 27%. There are also two parameters fixed from outside of the model: $\beta = 0.96$ since we think of a period as a year and $\delta = 0.1$ since we think of credit scores resetting every ten years. [5]

7.1 Household Allocations

The household credit market allocations vary over time as a function of the household's credit history, which itself maps to their score $s(d^t)$. We therefore plot these allocations as a function of the score, which is the payoff relevant state variable in the economy. There are two distinct regions of the score domain: (low) values for which there is zero cross-subsidization in

[5] Note that there are actually six parameters that govern the solvency rates, the two stigma parameters, the two expenditure shock levels, and the probability of each expenditure shock occurring. However, once we set these to satisfy the conditions that ensure low-risk households repay for $\tau < \tau_3$ while high-risk only repay for $\tau = 0$, we are effectively only choosing the probabilities that those shocks occur.

the credit market and (high) values for which there is cross-subsidization.

This can be seen in the interest rate for the high-risk household by looking at his interest rate in Figure (1). As noted in the characterization section, the high-risk household always gets the full-information level of credit, which is equal to one with our preferences. For low credit scores, he gets the actuarially fair interest rate (here equal to 400% because his solvency rate is 20%). As his score rises he begins to get lower interest rates, reaching the limiting value associated with the full-information low-risk contract rate of 11.11% (we assume that the low-risk solvency rate is 90%).

The low-risk household's credit allocation has somewhat more interesting dynamics as can be seen in Figure (2). The first thing to note is that he is credit constrained so long as his score is less than one, which is a direct implication from the incentive compatibility constraint on the high-risk household. For low scores, this constraint is strongest and the low-risk household's first sub-period consumption is 60% below the full-information level (which is one due to our utility function). Once his score crosses a threshold the credit allocation begins to subsidize the high-risk contract, which allows for more credit to be extended to the low-risk household, eventually reaching the full-information level as his score reaches one. This loosening of credit constraints is the efficient way of delivering higher utility to the low-risk household, but actually leads to a region of increasing interests rates. This must occur since the interest rate for low scores is actuarially fair (11.11%) and then reaches that rate again as his score reaches one. While this may seem counterintuitive, it is some solace to note that the region of scores over which it occurs is quite small and that the unconditional average interest rate falls over the entire range of credit scores.

7.2 The Distribution of Scores

While the household allocations are interesting to know what the model predicts *can* occur, it is more interesting to know what actually *does* occur. That is, what is the prediction for the actual distribution of scores, and therefore credit contracts.

We first show the distribution over scores for each type. As can be seen in Figure (3), the low-risk households tend to have higher scores than the unconditional share of low-risk households in the population (which is 72% in this example). This CDF shows that roughly 15% of low-risk households have scores below 0.72 and only 20% have scores below 0.9. The high-risk households tend to have scores below 0.72, which can be seen in Figure (4). Over two-thirds have scores below 0.1 and almost all of these households have scores below 0.72. This is because the solvency rates are so different for the two risk types, so the Bayesian updating underlying the credit scores implies quite rapid sorting along this dimension.

We can use these to think about what credit contracts actually occur in equilibrium. We do so by overlaying the empirical PDF of scores onto the credit allocation plots from before. The most interesting thing to notice is in Figure (5), where we can see that most low-risk households have scores above the level where interest rates begin to decline with score (roughly 93%). This means that the interest rate would be decreasing with credit scores over most of the low-risk population, even if we conditioned on the household's type (it is decreasing with credit scores over the entire range if do not condition on risk type, which we think is the more empirically relevant measure).

We can also see in Figure (6) that very few low-risk households have the tightest credit limit (only about 1% in fact). For the high-risk households, we already know that the credit extended is constant, but we show the distribution over interest rates in Figure (7). Most high-risk households have interest rates substantially higher than the full-information low-risk rate of 11.11%.

7.3 Apparent Mark-Ups From Aggregating Contracts

The heterogeneity within scores leads to interesting aggregate dynamics as well. The existing literature on unsecured debt often looks only at interest rates and default rates and use an aggregate zero-profit pricing condition to infer mark-ups. In this model lenders earn zero profits on each contract, but not necessarily on each loan. As the composition of loans changes across credit scores the model predicts that a mark-up can arise due to aggregation error.

As can be seen in Figure (8), the model is broadly consistent with data from credit scoring agencies as the interest rate is declining as scores rise.[6] It is also interesting to note how aggregating over household types can lead to apparent markups. That is, suppose that we computed the solvency rate for households with a given score, s_0. This is quite simply $0.9s_0 + 0.2(1 - s_0)$ in our calibration. If we then calculate the average "price" for households with that score, which is $s_0 \frac{Q_h(s_0)}{b_h(s_0)} + (1 - s_0) \frac{Q_\ell(s_0)}{b_\ell(s_0)}$ then we can define the mark-up as the difference between the average solvency rate and the average price.

The resulting calculations can be seen in Figure (9). For scores below the threshold of where high-risk households are subsidized by low-risk there is no markup, but for scores above that value there is an apparent mark-up due to this cross subsidization. For this calibration the markup reaches as high as 10% for some scores, though it is quite a bit smaller on average at 1.12%.

7.4 Welfare Consequences of Alternative Information Structures

The above experiments show that our model is capable of matching salient features of the data which makes our model a good laboratory to understand the welfare consequences of policies which affect how much information is available in the credit market. We consider two possibilities: a policy that increases the amount of information available and one that restricts it.

[6] We are being careful not to call our score a "credit" score, since it is literally the probability of being a low-risk household. However, we show below that it would be equivalent to use a "credit" score defined as the probability of default, in the special case of just two types and a constant solvency rate for each type.

7.4.1 More Information Extreme: Full Information

The first thing that should be clear is that more information does not help everybody at the time it is made available. In order to see this, suppose that a policy could be enacted which moved the economy to full-information. This would immediately change the credit market allocations to the full-information contracts, which would deliver higher utility to all of the low-risk households, but would reduce the utility of any high-risk household with a high enough credit score to receive a subsidized interest rate. Intuitively, these high-risk households are currently receiving surplus utility from their scores and therefore would lose if that surplus was taken from them. We plot the welfare consequences of switching to full-information across all credit scores in Figure (10).

On the other hand, a utilitarian policy maker might care about the ex-ante welfare from such a policy rather than the distributional effect on households at the time it was implemented. When we perform this calculation, we find a positive welfare effect equal to approximately 1.17% of first sub-period consumption each period. [7]

7.4.2 Less Information Extreme: No Signals

The opposite case is a policy that limits the amount of information available to lenders. This could be through forcing credit agencies to keep shorter records or limiting the information collected. We consider the extreme case of getting rid of credit scores altogether. This again has distributional effects at the time of implementation, as households with high-scores lose that information and therefore receive worse terms. However, a low-risk household tends to lose much more, and in fact even lose if his score is below μ_h since he is very likely to have a high score in the future. A high-risk household loses less and often gain since he is likely to have a low score in the future. These observations can all be made by looking at the welfare gains/losses in Figure (11).

[7] This is calculation is performed under the assumption that a new-born household's likelihood of becoming a low-risk type is equal to μ_h.

Once again, however, we are more interested in the welfare gains for an ex-ante household being born into each economy with fair odds of becoming either type. Perhaps surprisingly, we find a welfare *gain* of 0.25% of first sub-period consumption. This is because all of the high-risk households gain from getting rid of signals ex-ante and the losses of low-risk households are offset somewhat by the smoothing of first sub-period consumption. That is, in the economy with scores a low-risk household experiences volatility in his credit score over time which generates uninsurable movements in his first sub-period consumption. Without scores he receives a lower level of first sub-period consumption forever, but it is constant. [8]

8 Robustness and Alternative Models

8.1 Occasionally Binding Credit Limits

We say that the low-risk household has a binding credit limit in this paper because, given the interest rate on his debt, he would choose to take more debt than entailed by his contract if he had that option. This is an interesting novelty of our model, but is stronger than what we observe in reality since most people do not consistently exhaust their credit limit. Here we show a version in which the low-risk household does not exhaust his limit in every period, but is still credit constrained.

The new assumption is that there is a possibility of expenditure shock, η, in the first sub-period, which affects the household's marginal utility of funds at that point. The probability of this shock occurring is constant across households, given by γ. However, contracts are still signed before this shock occurs, so that the household can only borrow up to Q units independent of the shock. Denoting q_1 the amount borrowed if the shock occurs and q_0 the amount if it doesn't, the household borrows $q_i \leq Q$ and pays a specified rate R on the debt (if he remains solvent). [9] We maintain

[8] The difference in average levels is not very large in this parameterization, but this volatility mechanism is robust to that feature.

[9] There are numerous ways to write the contracts, this one is simple and provides sufficiently rich instruments.

the assumptions which guarantee that the low-risk household is solvent in the second sub-period for shocks τ_1 and τ_2, while the high-risk household is solvent only for shock τ_1.

In the model with $T = 1$, we can write indirect utilities over contracts as:

$$\begin{aligned} V_h(Q,R) &= \max_{q_0 \leq Q, q_1 \leq Q} \gamma \log(q_0) + (1-\gamma)\log(q_1 - \eta) \quad (92) \\ &- (\phi_1 + \phi_2)\, R\,(\gamma q_0 + (1-\gamma) q_1) - \phi_1 \tau_1 - \phi_2 \tau_2 - \phi_3 \lambda_h \\ V_\ell(Q,R) &= \max_{q_0 \leq Q, q_1 \leq Q} \gamma \log(q_0) + (1-\gamma)\log(q_1 - \eta) \quad (93) \\ &- \phi_1 R\,(\gamma q_0 + (1-\gamma) q_1) - \phi_1 \tau_1 - (\phi_2 + \phi_3)\lambda_\ell \end{aligned}$$

The first-order conditions on q_0 and q_1 guarantee that $q_0 = q_1 - \eta$, and since neither q_0 or q_1 can exceed Q, therefore $q_1 = Q$ and $q_0 = q_1 - \eta = Q - \eta$. The indirect utilities are therefore simplified to:

$$\begin{aligned} V_h(Q,R) &= \log(Q-\eta) - (\phi_1 + \phi_2)\, RQ - \phi_1\tau_1 - \phi_2\tau_2 - \phi_3\lambda_h \quad (94) \\ V_\ell(Q,R) &= \log(Q-\eta) - \phi_1 RQ - \phi_1\tau_1 - (\phi_2 + \phi_3)\lambda_\ell \quad (95) \end{aligned}$$

Under full information, the interest rate is actuarially fair (i.e., $R_h = \frac{1}{\phi_1+\phi_2}$ and $R_\ell = \frac{1}{\phi_1}$). Then the credit limit solves the following equation for both household types:

$$\frac{1}{Q-\eta} = 1 \quad (96)$$

Thus, the full-information credit limit is $Q_i = 1 + \eta$, which means that neither household is ever constrained, independent of the first sub-period shock. With adverse selection there is under provision of lending to the low-risk household (i.e., $Q_h < 1+\eta$). Therefore, the low-risk household borrows less than his limit when he avoids the expenditure shock but maxes out his limit (which is constrained) when he does receive the shock. The high-risk household is always unconstrained, though he does borrow less when he avoids the expenditure shock.

This example shows that the economy with first sub-period expenditure shocks can be mapped into the economy without such shocks, but with utility function $\log(c-\eta)$. Similar arguments can then be made for versions of this economy with $T > 1$.

8.2 Inter-temporal Savings

We do not consider and economy with inter-temporal savings directly, but we can consider what a household would choose to do if he was allowed to begin saving at some interest rate r. That is, he knows that his actions will not affect the options available to him in equilibrium. Such a household can choose to save a units in the second sub-period of period t and will have an additional $a(1+r)$ units of consumption in the first sub-period of period $t+1$. The Euler Equations for a household who has a high stigma would be:

$$1 \geq \beta(1+r)\frac{1}{Q_h\left(d^t,0\right)+(1+r)a} \text{ and } = \text{ if } a > 0 \tag{97}$$

$$1 \geq \beta(1+r)\frac{1}{Q_h\left(d^t,1\right)+(1+r)a} \text{ and } = \text{ if } a > 0 \tag{98}$$

Similar equations hold for the high-risk household, except that we already know his value of Q is always equal to one:

$$1 \geq \beta(1+r)\frac{1}{1+(1+r)a} \text{ and } = \text{ if } a > 0$$

We can therefore obtain a sufficient condition for all households to set $a = 0$.

Proposition 4 *If* $r \leq \min_{t\in\{1,\dots,T\}} \min_{d^t,d_{t+1}\in\{0,1\}} \frac{Q_h\left((d^t,d_t)\right)}{\beta} - 1$, *then any household who is given the choice to save between the second sub-period of any period* t *and first sub-period of period* $t+1$ *will choose* $a = 0$.

This follows from the Euler Equations above.

8.3 Equivalence Between Type Score and Credit Score

In the equilibrium with two types and constant solvency rates, it is possible to use the probability that a household repays his debt as a state variable because it is directly related to the natural state variable, which is the probability that the household is the high stigma type. In order to see this, note that the probability that a household defaults given his score is:

$$p(s) = s\,(\phi_1 + \phi_2) + (1 - s)\phi_1 \tag{99}$$

Which can be inverted to give:

$$s(p) = \frac{p - \phi_1}{\phi_2} \tag{100}$$

The law of motion for p is then given by:

$$p'(d, p) = p\left(s'(d, s(p))\right)$$

which gives:

$$p'(0, p) = \frac{(\phi_1 + \phi_2)^2(p - \phi_1) + \phi_1^2(\phi_1 + \phi_2 - p)}{(\phi_1 + \phi_2)(p - \phi_1) + \phi_1(\phi_1 + \phi_2 - p)} \tag{101}$$

$$p'(1, p) = \frac{(\phi_1 + \phi_2)\pi_\infty(p - \phi_1) + \phi_1(\phi_1 + \phi_2)(\phi_1 + \phi_2 - p)}{\pi_\infty(p - \phi_1) + (\phi_1 + \phi_2)(\phi_1 + \phi_2 - p)} \tag{102}$$

With this equivalence we can plot the distribution of credit scores similarly to type scores, which can be seen in Figure (**??**).

8.4 Continuous Expenditure Shocks

The specification we study allows us to prove that the credit-market programming problem is well behaved, but it doesn't allow for the default rate to vary smoothly with the information structure or model parameters. We therefore study a version with continuous shocks by assuming that the first-order approach is valid. We first redefine that indirect utilities and value functions:

$$\mathcal{U}_{i,T}(Q,b,d^T) = \log Q + ... \tag{103}$$

$$\int_0^\infty \max_{n,d} \{\log(n-(1-d)(b+\tau)) - h - \lambda_i d\} \, dF(\tau)$$

$$V_{i,T}(d^T) = \max_{(Q,b)\in\mathbb{C}_T(d^T)} \mathcal{U}_{i,T}(Q,b,d^T) \tag{104}$$

And then for $t < T$:

$$\mathcal{U}_{i,T}(Q,b,d^t) = \log Q + ... \tag{105}$$

$$\int_0^\infty \max_{n,d} \left\{\log(n-(1-d)(b+\tau)) - h - \lambda_i d + \beta V_{i,t+1}\left((d^t,d)\right)\right\} dF(\tau)$$

$$V_{i,t}(d^t) = \max_{(Q,b)\in\mathbb{C}_t(d^t)} \mathcal{U}_{i,t}(Q,b,d^t) \tag{106}$$

Denote $\tilde{\lambda}_{i,T}(d^t) = \lambda_i$ and for $t < T$:

$$\tilde{\lambda}_{i,t}(d^t) = \lambda_i + \beta\left[V_{i,t+1}\left((d^t,0)\right) - V_{i,t+1}\left((d^t,1)\right)\right] \tag{107}$$

Then the household sets $d = 1$ if and only if:

$$b + \tau > \lambda_{i,t}(d^t) \tag{108}$$

and we can simplify the indirect utilities to:

$$\mathcal{U}_{i,t}(Q,b,d^T) = \log Q + \int_0^{\tilde{\lambda}_{i,t}(d^t)-b} F(\tau)d\tau - \lambda_i + \beta V_{i,t+1}\left((d^t,1)\right)$$

In principle we could use these functions in the credit market problem as before and solve the problem numerically. While we have not been able to solve this version in full generality, we can do so under the assumption that high-risk households will default on any positive level of debt. [10] In this case we know that $b_\ell = 0$ and the participation constraint for the high-risk household cannot bind since he receives zero credit in his full information contract, so we drop that constraint. Furthermore, the incentive compati-

[10] This can be guaranteed by setting λ_ℓ sufficiently low (potentially negative).

bility constraint immediately implies that $Q_\ell = Q_h$ so the only remaining condition is the zero-profit constraint. This gives:

$$Q_h = Q_\ell = sF\left(\tilde{\lambda}_{h,t}(d^t) - b_h\right) b_h \tag{109}$$

Therefore the entire credit-market programming problem simplifies to simply:

$$\max_{b_h} \log\left(sF\left(\tilde{\lambda}_{h,t}(d^t) - b_h\right) b_h\right) + \int_0^{\tilde{\lambda}_{h,t}(d^t)-b_h} F(\tau)d\tau \tag{110}$$

Results from this specification are forthcoming.

9 Conclusion

This paper provides a parsimonious and tractable model of adverse selection in credit markets and shows how credit scores can be used to reduce the distortions arising from such an information friction. The model rests on a solid game-theoretic micro foundation and is able to generate qualitative features of the unsecured credit market. Initial policy experiments using a calibrated version of the model suggest that large improvements in the informational content of credit scores may be beneficial on average, while having initial distributional effects. Surprisingly, restricting the usage of credit scores may also be beneficial due to other forms of market incompleteness.

References

[1] Athreya, K. X.Tam, and E. Young (2012) "A Quantitative Theory of Information and Unsecured Credit ", *American Economic Journal: Macroeconomics*, 4,p.153-183.

[2] Chatterjee, S. D. Corbae, M. Nakajima, and V. Rios-Rull (2007), "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default "*Econometrica* 75(6), 1525-1589

[3] Chatterjee, S. D. Corbae, and V. Rios-Rull (2008) "A Finite-Life Private-Information Theory of Unsecured Consumer Debt "*Journal of Economic Theory*, 142, p.149-177.

[4] Chatterjee, S., D. Corbae, and V. Rios-Rull (2014) "A Theory of Credit Scoring and Competitive Pricing of Default Risk", mimeo.

[5] Livshits, I., J. MacGee, and M. Tertilt (2014) "The Democratization of Credit and the Rise in Consumer Bankruptcies", mimeo.

[6] Netzer, N. and F. Scheuer (2014) "A Game Theoretic Foundation of Competitive Equilibria with Adverse Selection", *International Economic Review*, 55, p. 399-422.

[7] Wilson, Charles (1977) "A Model of Insurance Markets With Incomplete Information "*Journal of Economic Theory*, 16, p.167-207

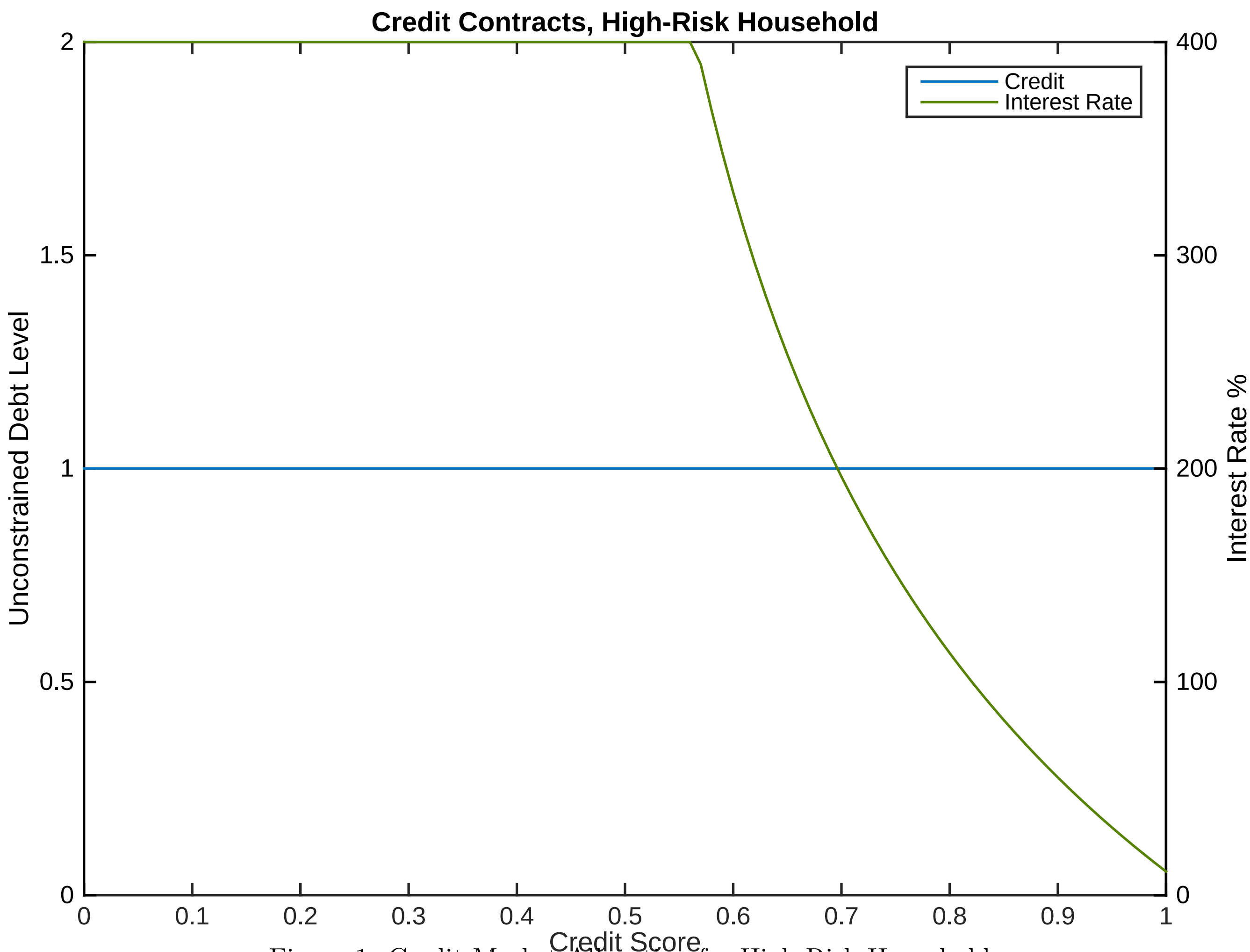

Figure 1: Credit Market Allocation for High-Risk Households

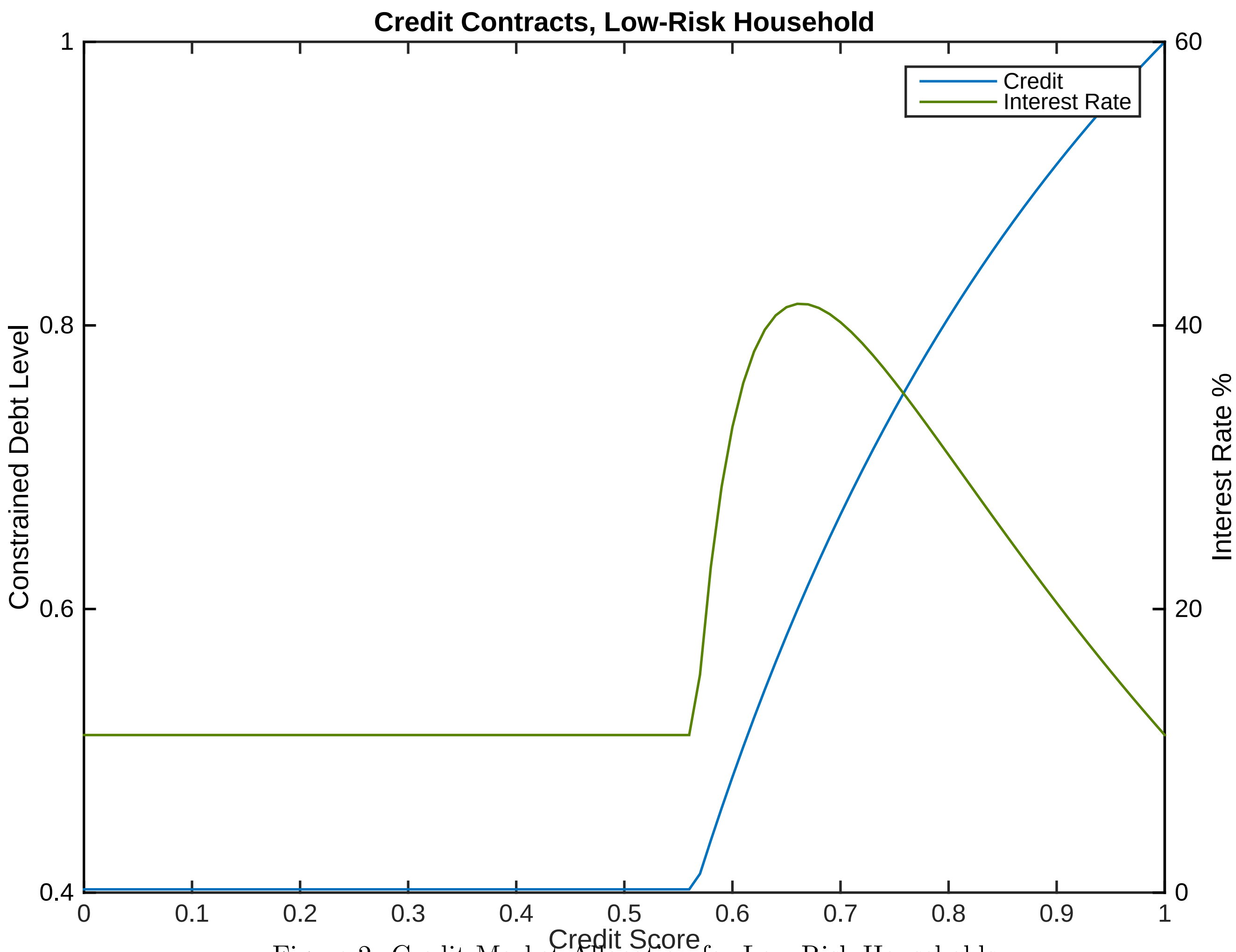

Figure 2: Credit Market Allocation for Low-Risk Households

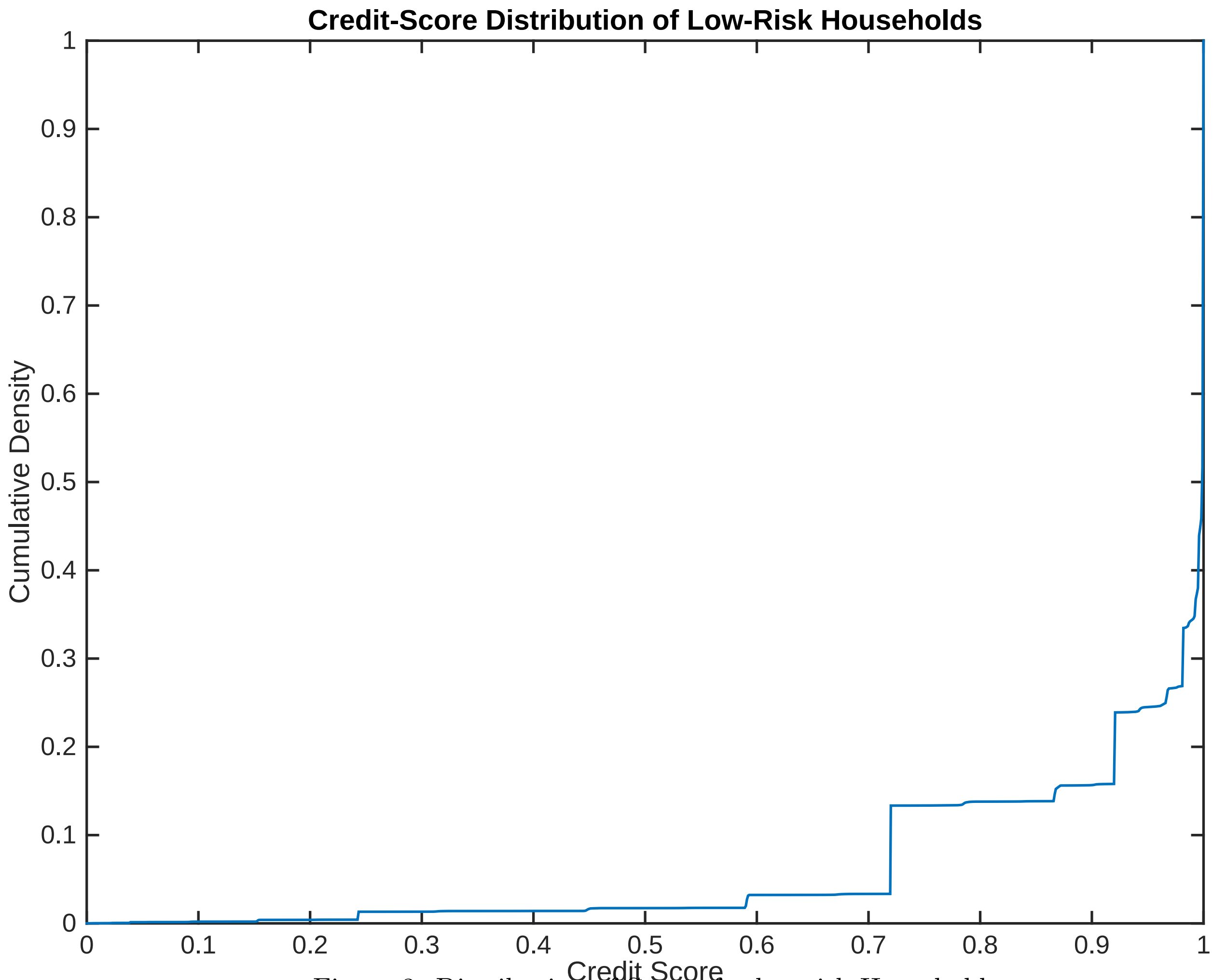

Figure 3: Distribution of Scores for low-risk Households

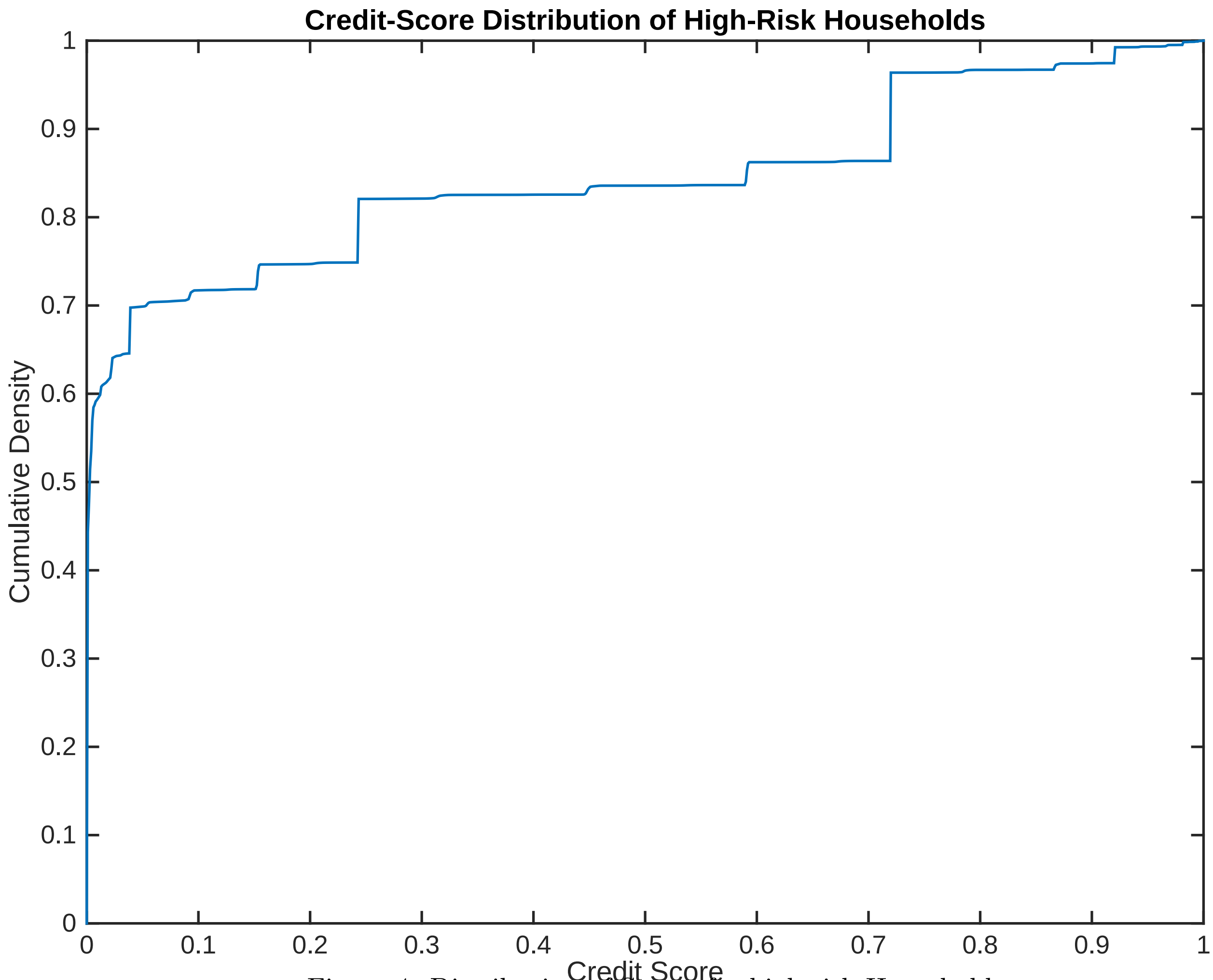

Figure 4: Distribution of Scores for high-risk Households

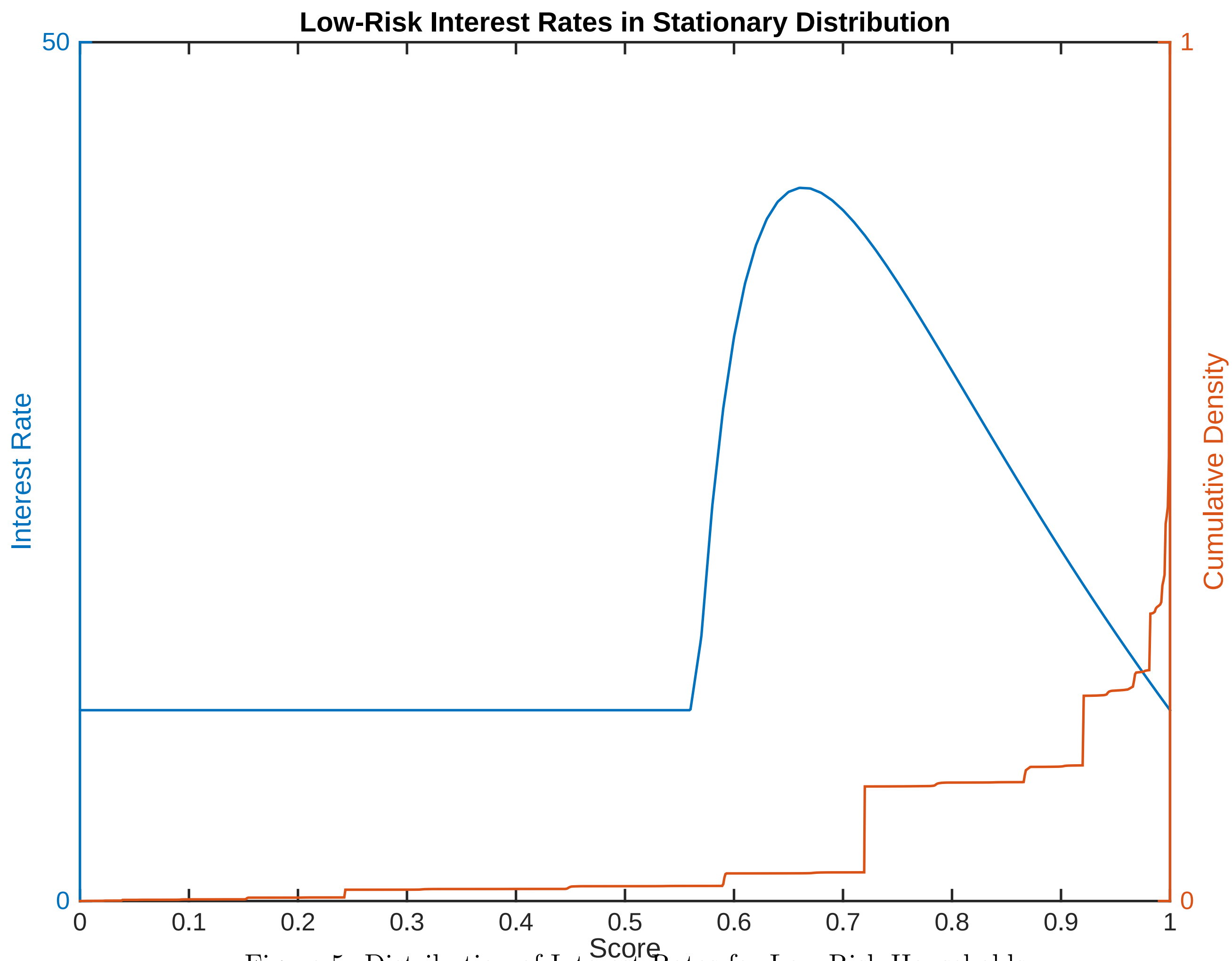

Figure 5: Distribution of Interest Rates for Low-Risk Households

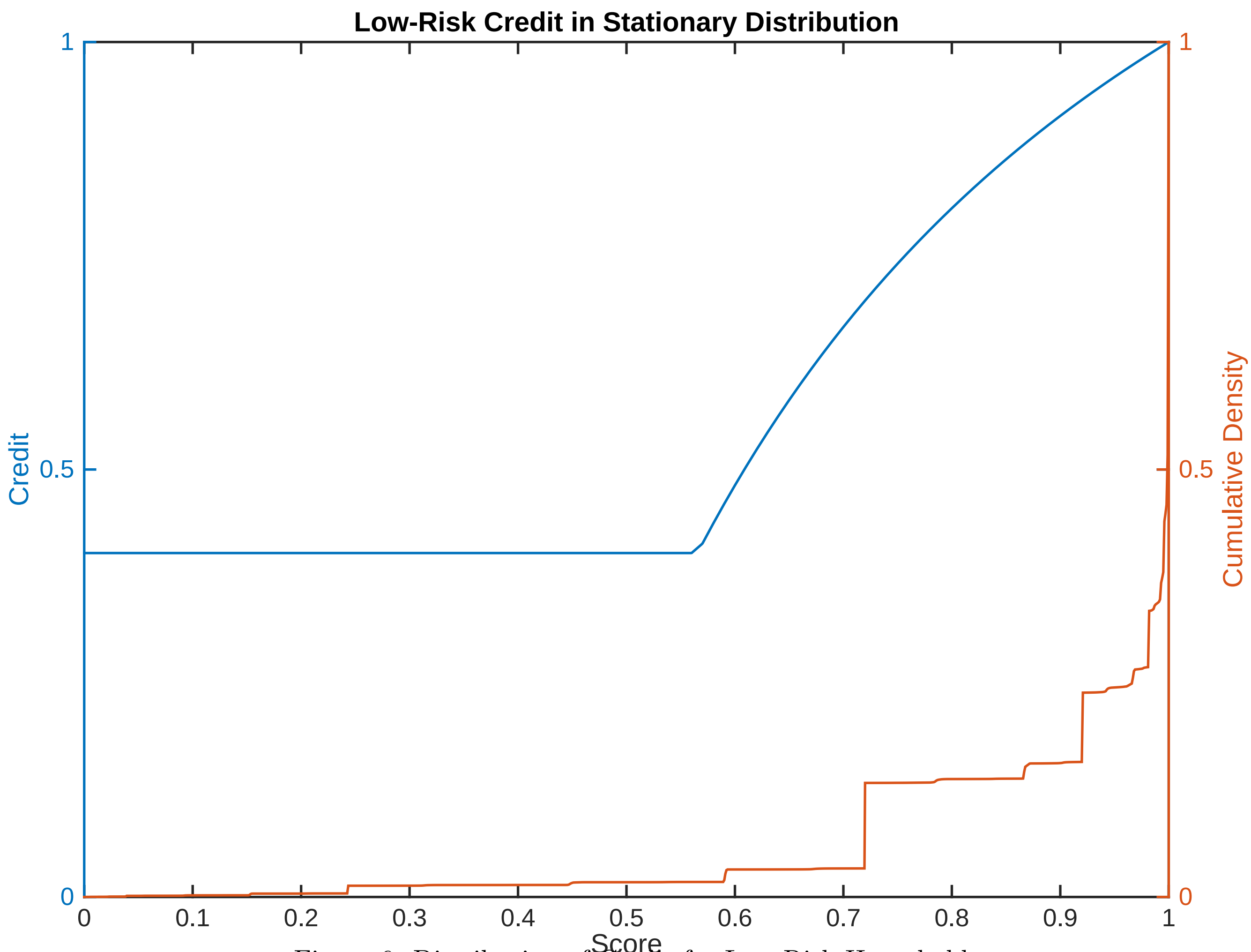

Figure 6: Distribution of Credit for Low-Risk Households

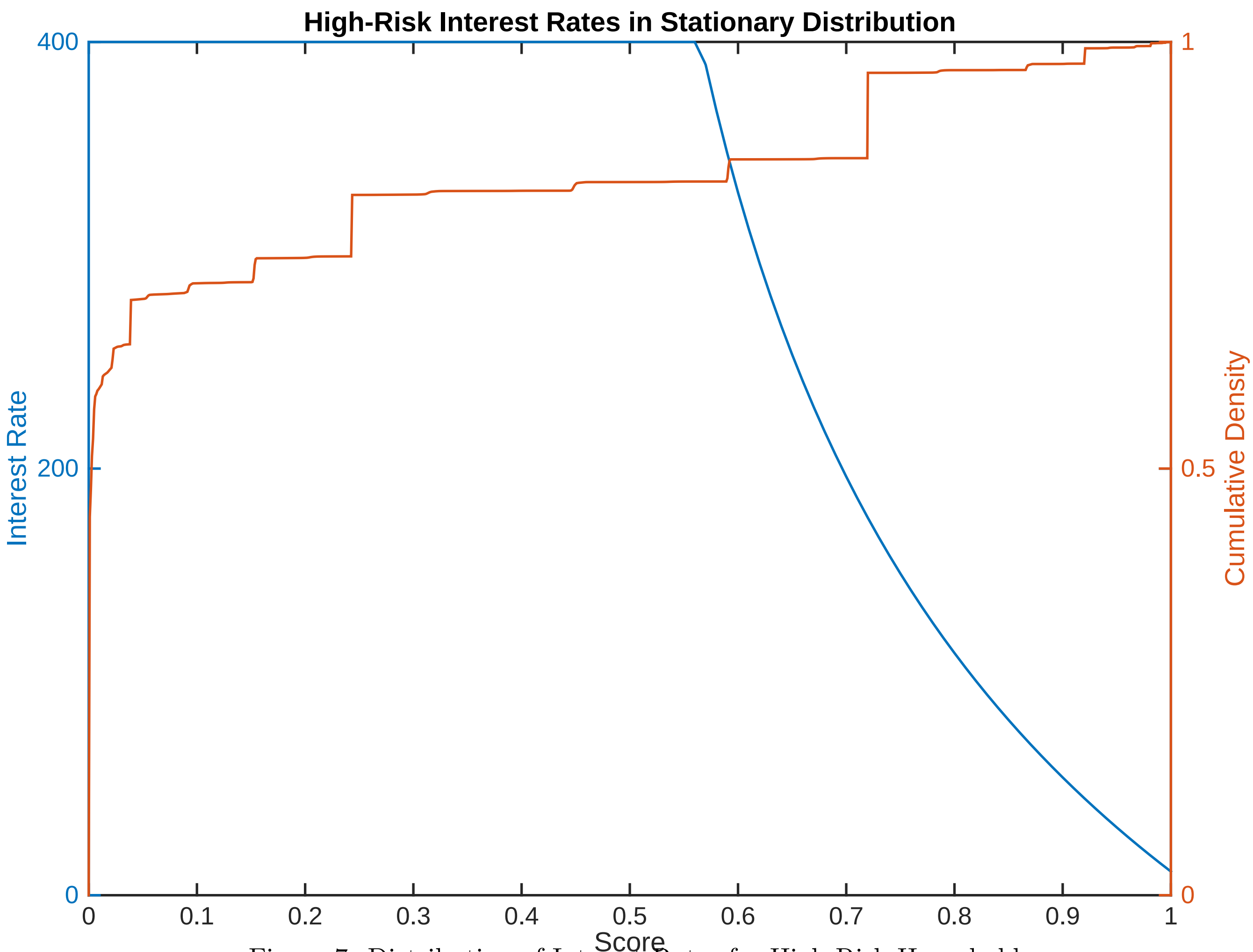

Figure 7: Distribution of Interest Rates for High-Risk Households

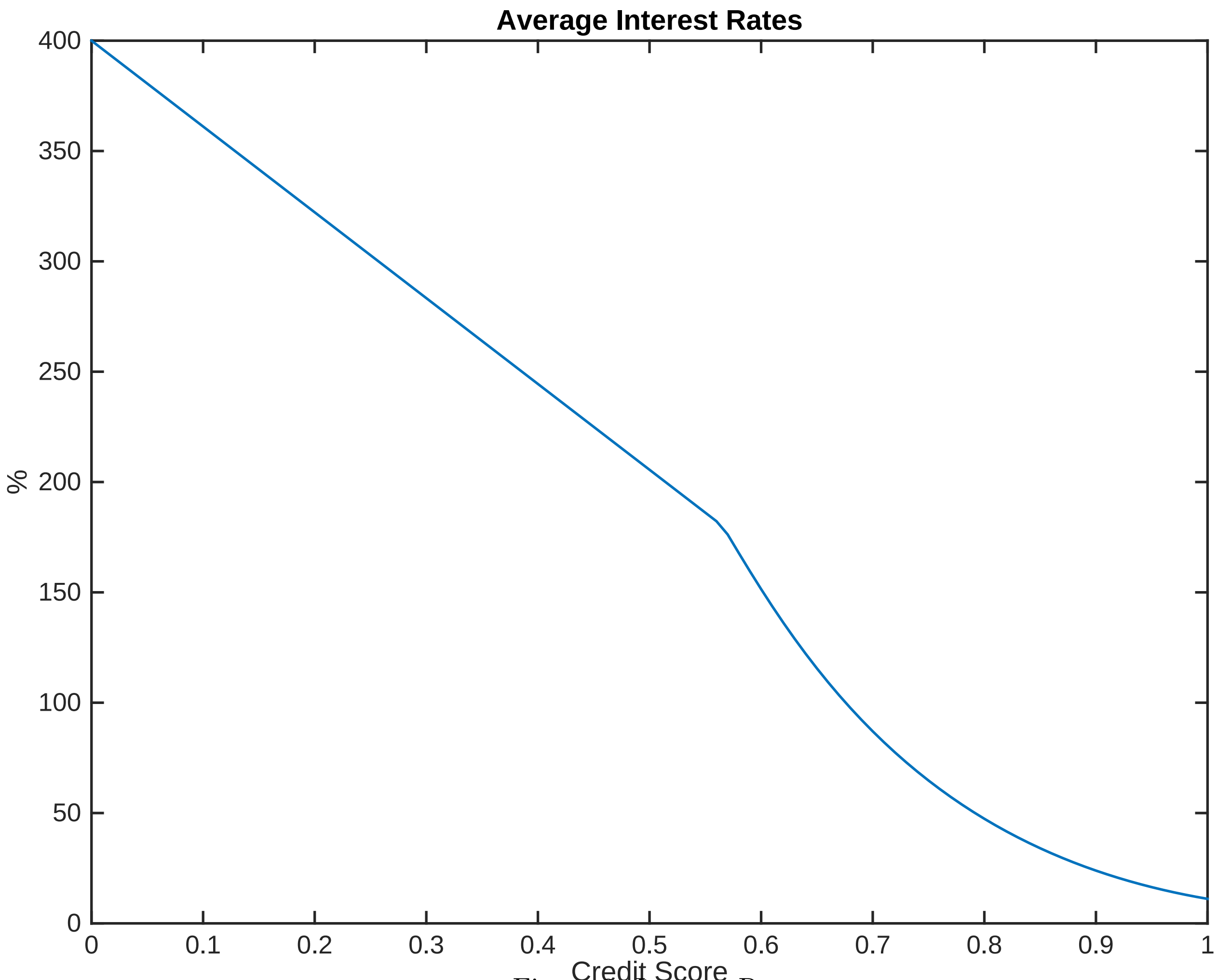

Figure 8: Interest Rates

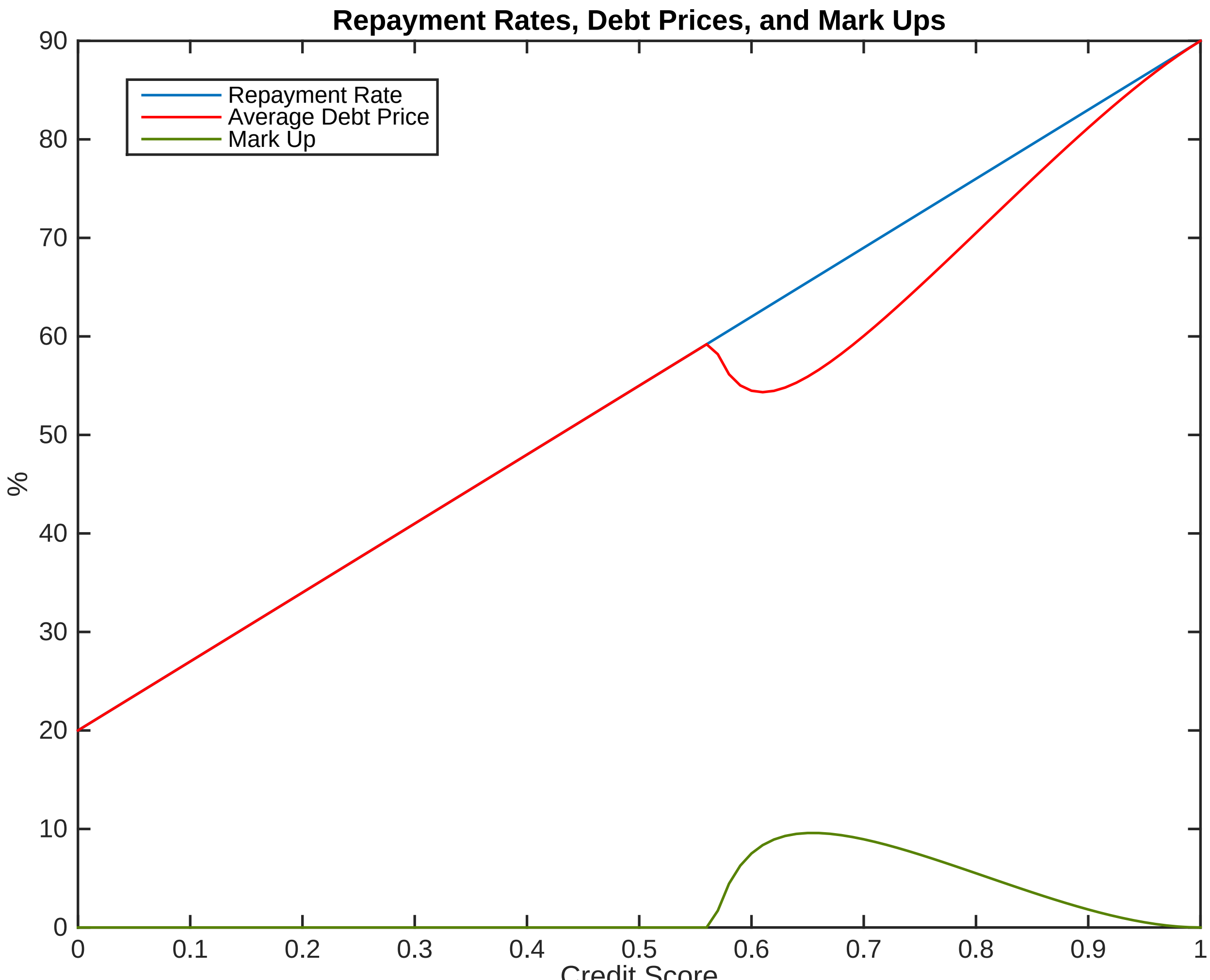

Figure 9: Interest Rates, Default Rates, and Apparent Mark-Ups

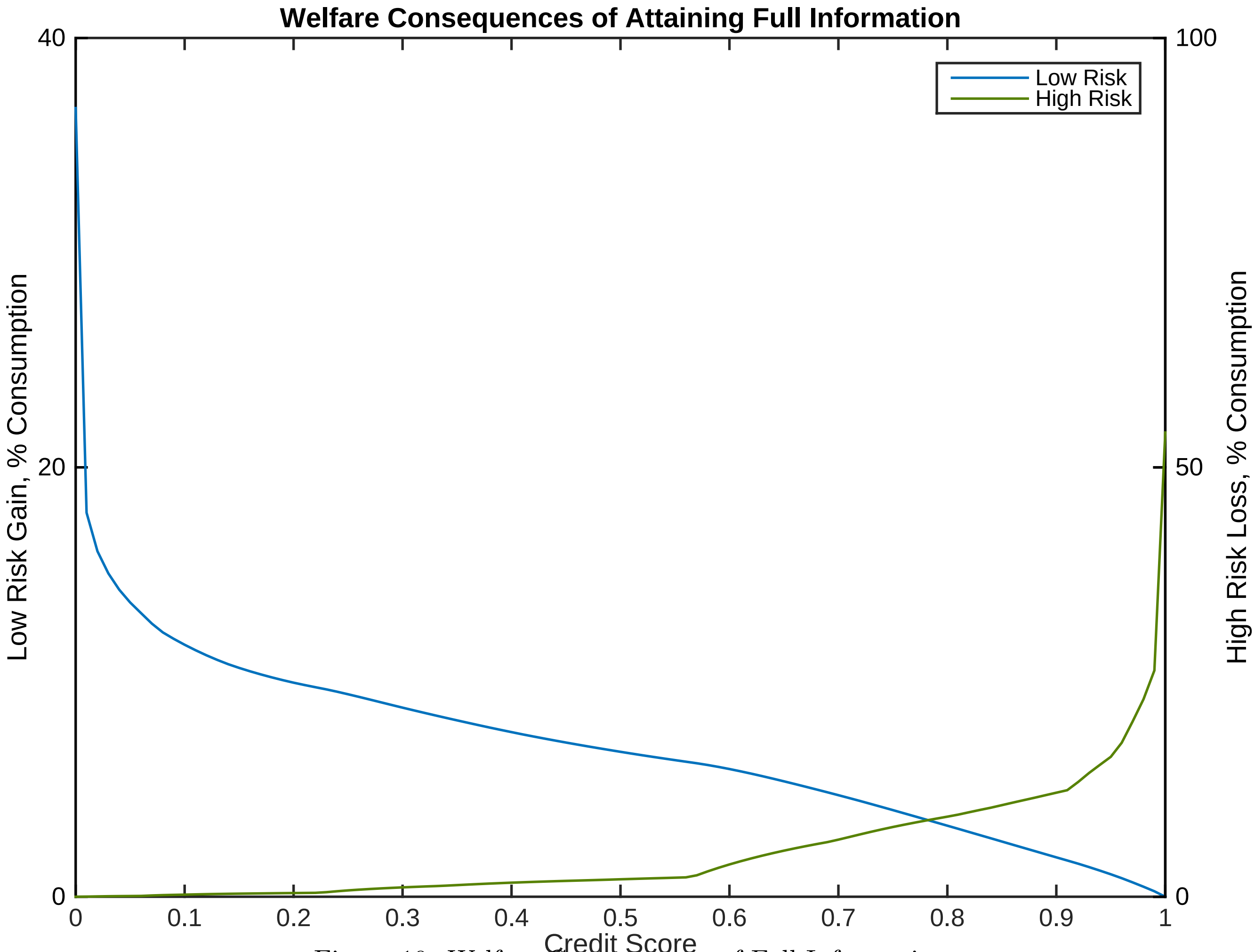

Figure 10: Welfare Consequences of Full Information

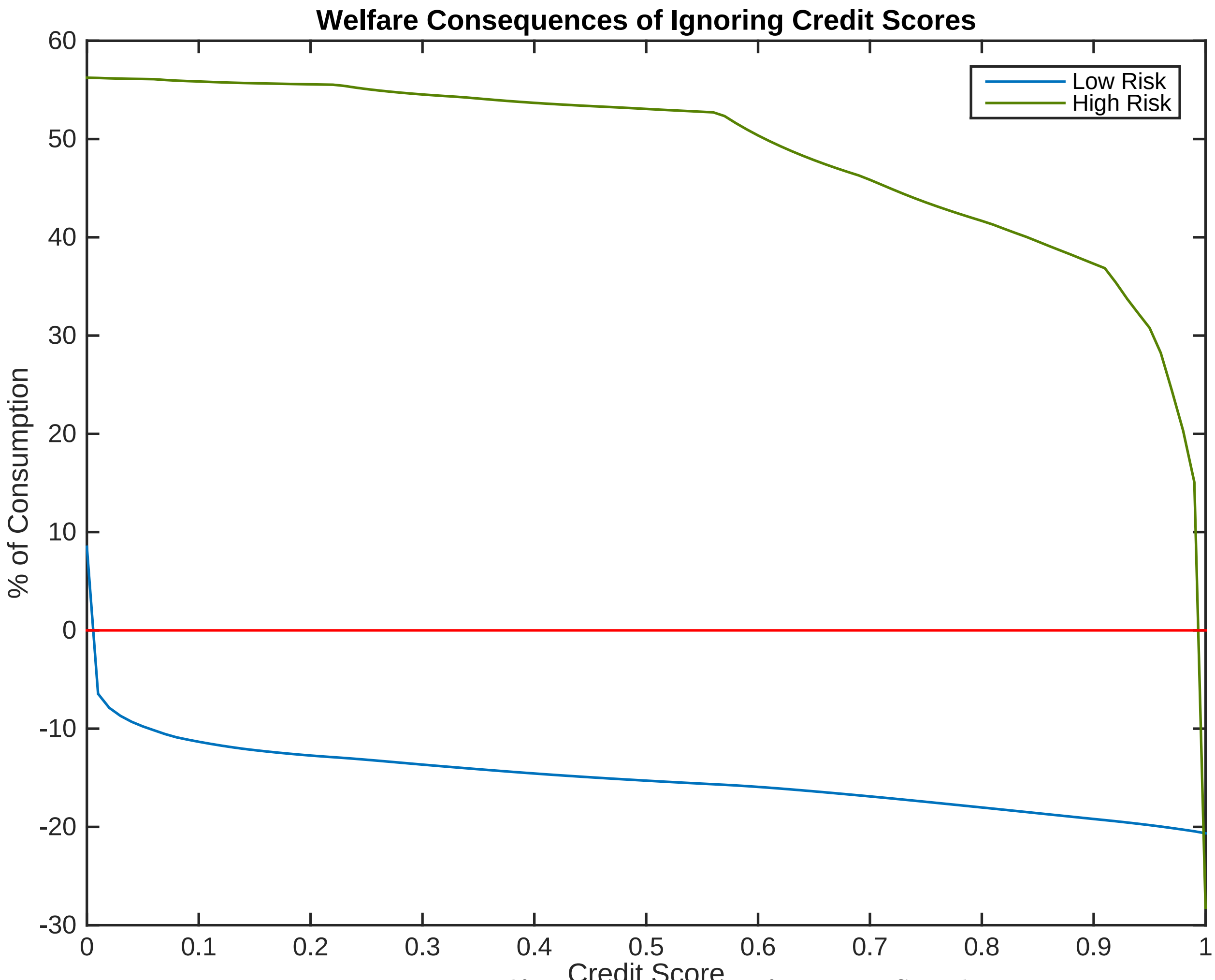

Figure 11: Welfare Consequences of Ignoring Signals

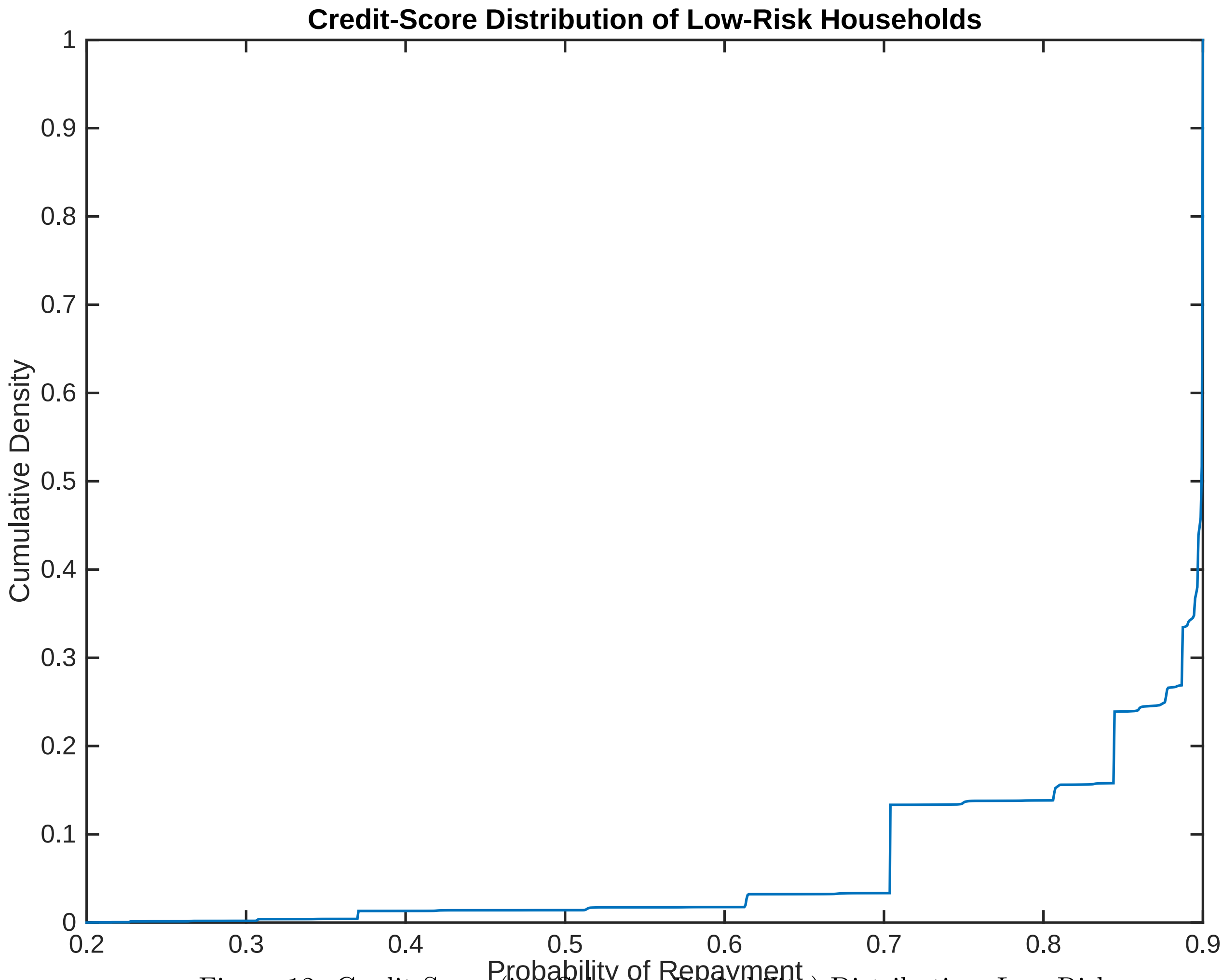

Figure 12: Credit Score (i.e. Solvency Probability) Distribution, Low-Risk

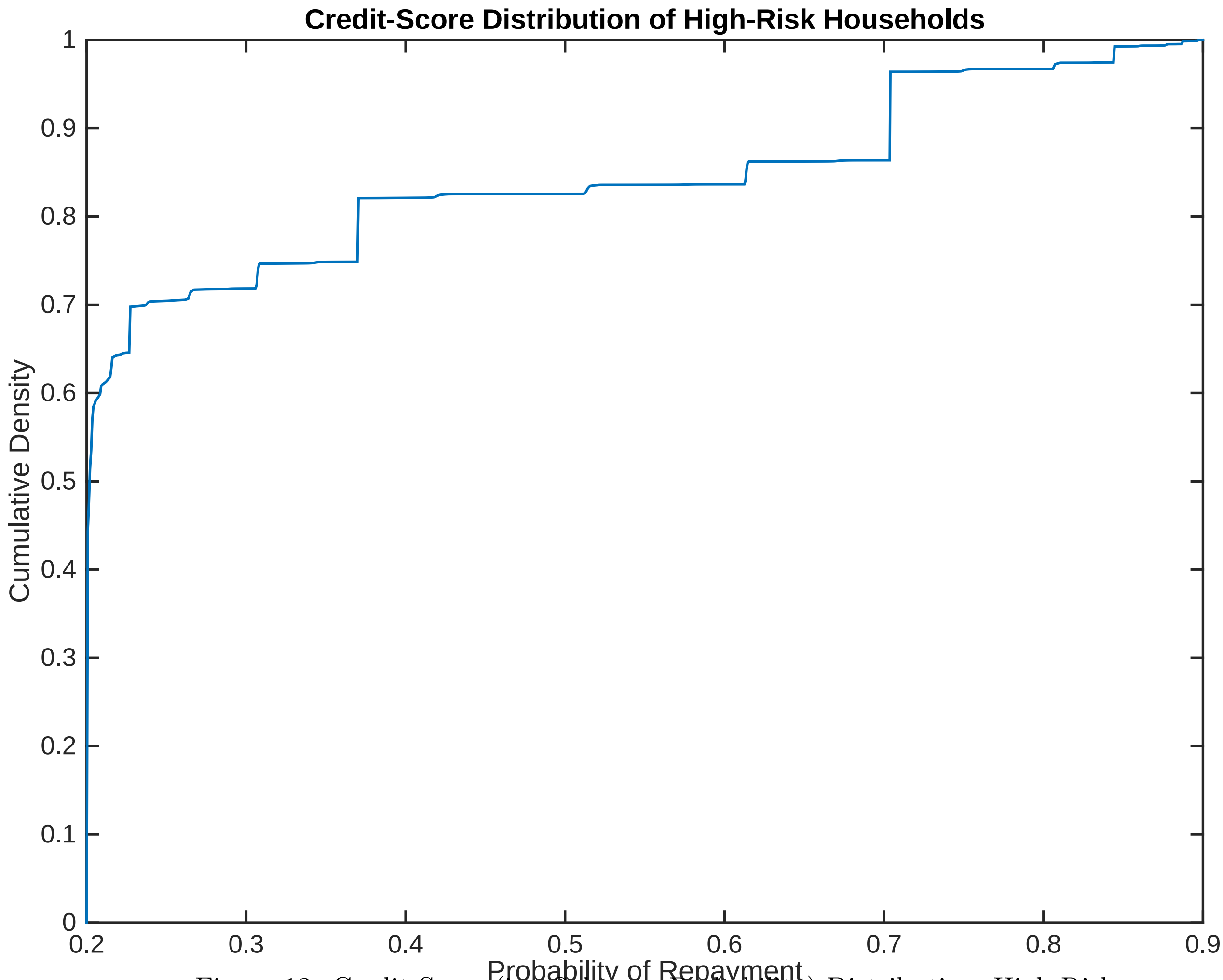

Figure 13: Credit Score (i.e. Solvency Probability) Distribution, High-Risk