# Credit Scores and Inequality Across the Life Cycle *

Satyajit Chatterjee † Dean Corbae ‡ Kyle Dempsey §
José-Víctor Ríos-Rull ¶

This draft: August 1, 2024

**Abstract**

Credit scores are a primary screening device for the allocation of credit, housing, and sometimes even employment. In the data, credit scores grow and fan out with age; at the same time, income and consumption inequality also increase with a cohort's age. We postulate a simple model with hidden information to explore the joint determination of credit scores, income, and consumption over an individual's lifetime which can replicate these empirical facts. We use the model to understand the role of technologies like big data or legal restrictions limiting information on certain adverse events intended to increase credit market access.

**PROPOSAL FOR NBER MACRO ANNUAL**

**PRELIMINARY AND INCOMPLETE**

*Keywords*: Adverse Selection, Moral Hazard, Bayesian Learning.

*JEL Classification Numbers:* D82, E21, G51.

*The views expressed in this paper are those of the authors and do not necessarily reflect views of the Federal Reserve Bank of Philadelphia or of the Federal Reserve System.

†Federal Reserve Bank of Philadelphia

‡University of Wisconsin-Madison and NBER

§The Ohio State University

¶University of Pennsylvania, CAERP, CEPR, NBER, and UCL.

# 1 Introduction

## 1.1 Motivation, goals, and anticipated contributions

Credit scores are a fundamental ingredient of a borrower's access to credit and housing. They are also widely used elsewhere: in pre-employment credit screens, the determination of insurance rates, and even in the choice of partners. Broadly, they are often seen as proof of "character," even though adverse events outside a person's direct control (like hospitalizations that result in financial distress) may enter an individual's credit report and weaken one's score. Despite their widespread use, credit scores are conspicuously absent from standard quantitative models used to evaluate consumer credit policy recommendations.

In this paper, we analyze the link between credit scores and inequality. Specifically, we measure how credit scores, and the reputational incentives which give rise to them, contribute to consumption and income inequality both in the cross-section and over the life cycle. This exercise is potentially confounded by feedback from an individual's unobservable type (which is assessed in credit scores) to income; this link has been missing in much of the literature.

To conduct this exercise, we develop a rich quantitative model of unsecured credit and human capital acquisition with both adverse selection and moral hazard. Both ingredients are necessary. Adverse selection (i.e. an unobservable, payoff-relevant type) implies that dynamically updated reputations – "type scores," which map directly into credit scores – are an essential part of the equilibrium in the credit market. Moral hazard (i.e. a costly unobservable choice) implies that there is endogenous feedback between these credit scores and income (and therefore consumption) both across individuals and over the life cycle for a given individual. Capturing the interaction between these forces in a structural model allows us to disentangle their respective contributions, which is crucial for policy analysis.

While our framework can encapsulate many potential implementations of one's unobservable "character," for ease of exposition we model it as patience in the present paper.[1] All else equal, a patient type is more likely to exert costly effort to build human capital and increase future income, as well as to repay debt given persistent default costs. Both forces generally lead those with high scores to have access to credit at relatively favorable terms and have lower leverage. An additional long run goal of this project is to use cross-sectional and life cycle data on the joint evolution of income, consumption, credit, and credit scores to discern the relevant personal characteristics which credit scores seek to ascertain.

[1] One may interpret patience as conscientiousness within OCEAN (openness, conscientiousness, extroversion, agreeableness, and neuroticism) psychological measures. That is, conscientiousness within OCEAN is defined to be the desire to be careful, diligent and to regulate immediate gratification with self-discipline. Traits include ambition, discipline, consistency and reliability.

In addition to measuring how credit scores contribute to inequality and welfare differences across individuals, we will also use the model for policy analysis. Our framework is uniquely well-suited to assess the implications of policies which expand or restrict the information set upon which credit scores can be conditioned. For instance, since April 2023 the three nationwide credit reporting companies - Equifax, Experian, and TransUnion - have taken steps to remove all medical collections under $500. With this change, it's estimated that roughly half of those with medical debt on their reports will have it removed from their credit history.[2] Here we can assess how the pooling of individuals with adverse events stricken from their records affects the cost of credit and the unintended cross-subsidization which may arise.

Along this dimension, our paper will illustrate how policies which expand information access affect credit access and income and consumption inequality. In this proposal, we develop and solve a simple, two-period version of the full model we will build. We use this simple model, as well as a "full information" or "big data" counterfactual in which credit scoring is irrelevant, to show that adverse selection and moral hazard can actually exacerbate income inequality while mitigating consumption inequality. These effects arise because the motivation to signal one's type and improve one's reputation strengthens the incentive to accumulate human capital and raise lifetime income. This finding provides an example of the scope for interesting and nuanced policy implications which arise from properly accounting for the joint relationship between reputation and income.

We are able to accomplish this by making a a simple but powerful methodological contribution which can be used in a variety of quantitative applications. The model combines screening and dynamic signaling and naturally gives rise to issues associated with "off-equilibrium-path" beliefs. To avoid these issues, we incorporate extreme value shocks on decisions which cloud inference about unobservable types and generically imply a semi-separating or partial pooling equilibrium. This ensures that lenders' beliefs following any feasible action are determined in equilibrium (reminiscent of Selten (1975) and Myerson (1978)): thus, there is no need to supply off-the-equilibrium-path beliefs. Further, the shocks eliminate multiplicity of equilibria that can arise in signalling games from variation in off-equilibrium-path beliefs. Finally, the assumption that the shocks are drawn from a type one extreme value distribution delivers computational tractability as in Rust (1987).

## 1.2 Related literature

There is a rich literature on quantitative dynamic models of unsecured credit with bankruptcy. Early examples include Livshits et al. (2007) and Chatterjee et al. (2007). In these papers,

[2] www.consumerfinance.gov/about-us/blog/medical-debt-anything-already-paid-or-under-500-should-no-longer-be-on-your-credit-report/.

while a credit score (the probability of repayment on a given size loan) can be constructed in a full information model like those above, the history of past asset market actions plays no role in that construction. In this paper and other work Chatterjee et al. (2023), we extend these models to include hidden information which requires us to index credit prices to the market assessment of individual types. Here, by contrast, past asset market actions are informative about an individual's repayment probability, and these actions are summarized dynamically in assessments which resemble credit scores. A related paper is Athreya et al. (2012), which assumes anonymous markets where *current* asset choices are observed, but no information about past behavior can be used to infer one's unobservable type.

A number of papers have examined the role of improvements in information technology on credit access. In these papers, the "technology" is a noisy signal of a borrower's true characteristics and the "improvement" is an increase in signal precision. These include Narajabad (2012) and Livshits et al. (2016), and Sanchez (2018). For instance, Narajabad (2012) compares polar cases: one in which the credit market lacks information, resulting in a pooling equilibrium, and one in which there is sufficient information to separate borrowers. Livshits et al. (2016) consider a simple asymmetric information model with costly contracting where borrowers know their types but uninformed lenders receive a noisy signal of a borrower's type. As signal precision improves, the level of pooling in a given contract falls. Kovrijnykh et al. (2023) build a simple but powerful model to rationalize the empirical fact that when there is little past information about the borrower's creditworthiness, non-exclusive lenders with heterogeneous signals about a borrower's type can learn from others' approval decisions. By contrast, the credit history of an established borrower may provide sufficient information to supercede the need for others' signals.

While previous quantitative theory models imposed exogenous punishment, we incorporate dynamic reputation as a means of disciplining borrowers along the lines of Diamond (1989). Our reputational environment, where everyone optimizes but people have hidden knowledge about their preferences, is closely linked to repeated games with incomplete information. Reputation in debt markets in which one player is a commitment type have been recently studied by Amador and Phelan (2021). The fact that reputation in one market may discipline behavior in another market has been considered in Cole and Kehoe (1998), Chatterjee et al. (2008), and Corbae and Glover (2024).

## 1.3 Outline of the proposal

The rest of this proposal is structured as follows. Section 2 presents two motivating facts for our analysis: (i) credit rankings increase and become more dispersed with age; and (ii)

credit scores and incomes both play an important role in determining terms of credit. We conjecture, and seek to explore in this project that these relationships contribute to observed patterns of consumption and income inequality. Our goal for an NBER Macroeconomics Annual paper is to explore these links in the data extensively and provide a new set of facts for macroeconomic researchers.

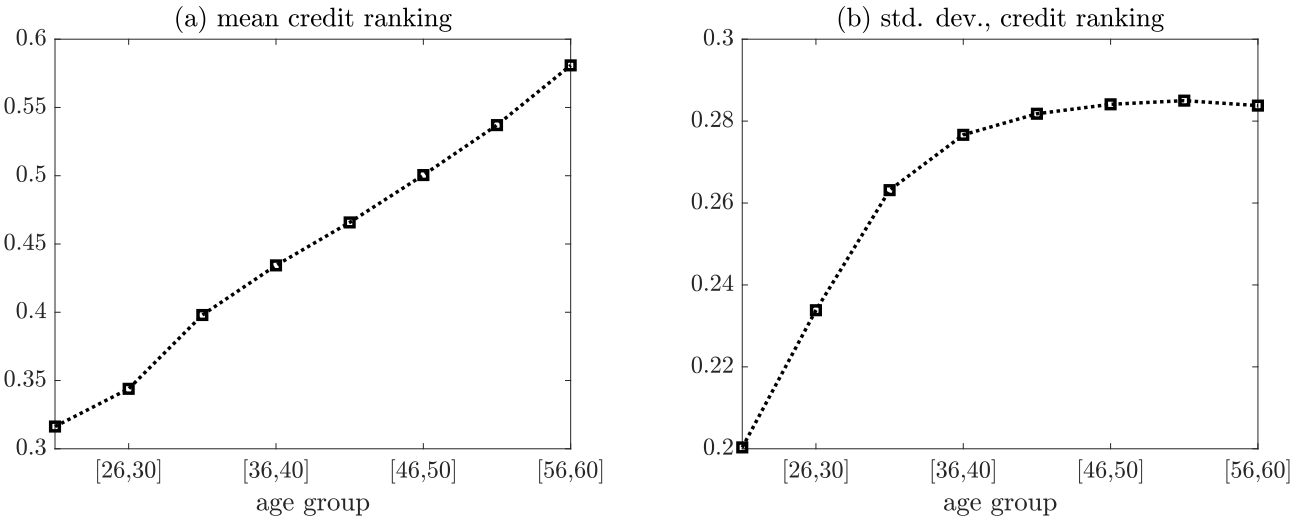
Section 3 provides, as a proof of concept, the simplest possible life cycle model with adverse selection and moral hazard capable of confronting these facts. We also use this simple model to illustrate our methodological contribution to solving dynamic quantitative Bayesian equilibrium models with hidden information. We construct a numerical illustration of this model to show how the intensity of information and moral hazard frictions impact credit assessments and income and consumption inequality. Lastly, we consider a "full information" version of our proposed economy to provide an intuitive frame for assessing policies which change the credit scorer's information set. Our goal for an NBER Macroeconomics Annual paper is to provide the simple model to illustrate the methodological contribution and then to extend the model quantitatively to confront the data in Section 2.

Finally, Section 4 provides a preliminary set of policy issues surrounding the use of "big data" or the unintended consequences of well-intentioned restrictions on information made available to credit reporting agencies we hope to address in an NBER Macroeconomics Annual paper.

# 2 Stylized Facts

**Overview** By way of motivation, we present two sets of stylized facts. First, we document that individuals' credit rankings – an ordinal, more easily interpretable transformation of credit scores – tend to increase as individuals age, while simultaneously becoming more disperse. That is, individuals' assessed creditworthiness tends to increase with age, and these assessments become more informative. We comment on the potential relationship between these patterns and analogous ones for income and consumption documented in Heathcote et al. (2010). Second, we use administrative data on credit card accounts to document that, while credit rankings and income have a strong correlation, they both appear to matter separately for credit pricing. We conjecture that similar relationships hold across other outcomes of interest at the individual level.

**Fact 1: Credit rankings grow and become more dispersed with age.** Figure 1 shows the age profile of individuals' average credit ranking (panel (a)) and of the dispersion in these rankings (panel (b)) in the data. We define an individual's credit *ranking* – as



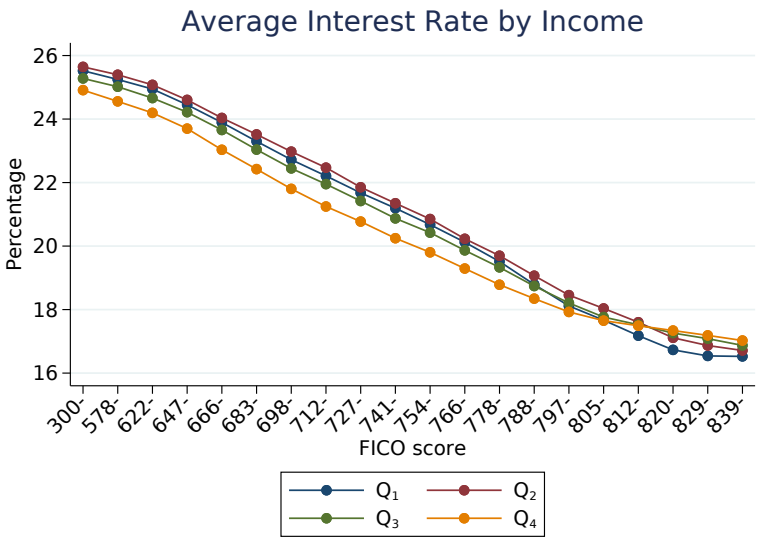

**Notes:** The credit ranking data is based on author calculations using FRBNY CCP/Equifax data, whose construction is detailed in Online Appendix C of Chatterjee et al. (2023).

Figure 1: **Credit Ranking Age Profile**

opposed to her credit *score* – as the percentile in the distribution of credit scores. Since credit scores are ordinal, credit rankings are as well, but the units are easier to interpret.[3] Importantly, both the age-specific mean and standard deviation of credit rankings rise with age. The latter result is consistent with learning: individuals with short credit histories tend to be pooled, while longer histories lead to more "separation" of good from bad risks by the rating agencies.

This paper will quantify how much these credit ranking dynamics contribute to income and consumption inequality. Heathcote et al. (2010) address income and consumption inequality over the life cycle: concretely, Figure 14 of their paper documents rising cross-sectional variance across the age profile in both income and consumption. How much does this explain, or to what extent can this be explained by, the credit ranking patterns we document in Figure 1? In a related paper, Krueger and Perri (2006) separate "across-group" variation, which owes to observable differences like education, from "within-group" variation, the residual which includes idiosyncratic income, and explore grouping people on observables like age and income along with unobservables like type. Importantly, building a structural model with both credit scores and unobservables which affect human capital will allow us to

[3]We avoid "cardinal" measures like a FICO score 760 and instead create "ordinal" rankings. For this we use a 2% sample (approximately 150,000 observations) of the FRBNY CCP/Equifax panel (described in detail in Appendix C of the Online Supplemental Material in Chatterjee et al. (2023)). With this data, we create credit rankings, defined as the percentile ranking of an individual's Risk Score relative to the overall sample distribution of Risk Scores, and group individuals between the ages of 21 and 60 in 5 year bins. We then compute the means and standard deviations of credit rankings within each age bin.




**Notes:** This figure is constructed using the Federal Reserve Board's Y-14M data as described in Appendix A of Dempsey and Ionescu (2023).

Figure 2: **Credit Card APRs by FICO Score and Income Quartile**

decompose how much of the correlation between credit scores and income is due to direct effect of income lowering the likelihood of default versus the indirect effect that human capital choices depend on underlying type differences. Such a decomposition might be critical for policy analysis, since the former has little to do with inference and information restrictions, while the latter does.

**Fact 2: Credit scores and income are correlated and both contribute to terms.** Figure 2, taken from Dempsey and Ionescu (2023), uses administrative data from the Federal Reserve Board's Y-14M data set to plot the relationship between credit scores, income, and the prices paid on consumer credit.[4] Two key facts emerge from this figure. First, conditional on FICO score, there is significant residual dispersion in APR based on income. For example, even conditional on a high FICO score of 790, the gap between the highest and lowest income quartiles is 1.5 percentage points (pp). Second, conditional on income quartile, there is also significant dispersion based on FICO. For example, conditional on being in the highest income quartile, moving from the 75th percentile FICO score to the 25th is associated with a 5 pp increase in APR.

Since credit score and income are themselves highly correlated (not shown pending approval), a natural question is: what causes these pricing patterns? Our project will aim to

[4]The figure splits the data into 80 bins: 20 FICO bins $\times$ 4 income quartiles, and each data point corresponds to the within-bin average of the APR on the associated credit card account.

systematically evaluate how these pricing patterns might change in respond to policy interventions in the consumer credit market, which are likely to effect the underlying relationship between these two key variables of interest, shedding light on this crucial question.

# 3 A Simple Model

Here we build quite possibly the simplest life cycle model of adverse selection and moral hazard consistent with the facts in Figures 1 and 2.

## 3.1 Environment

An individual lives two periods $t \in \{0, 1\}$. There is a unit measure of two types $\tau \in \{H, L\}$ of individuals who differ in their patience where $\beta_H > \beta_L$. The fraction of type $\tau = H$ in the population is $\rho$. Type is unobservable.

Individuals are endowed with income $y_t$. At $t = 0$, $y_0 = y_\ell$ for all individuals, and each agent chooses whether or not to exert effort $e_0 \in \{0, 1\}$. This unobservable, costly effort choice influences the probability that the individual receives date 1 income $y_1 \in \{y_h, y_\ell\}$, where $y_h > y_\ell$. Specifically, we assume that exerting effort ($e_0 = 1$) raises the probability of receiving high income at $t = 1$; formally, defining $\pi_{y_1|e_0} \equiv \mathbb{P}(y_1 \mid e_0)$, we assume that $\pi_{y_h|e_0=1} > \pi_{y_h|e_0=0}$.[5] The $t = 0$ utility cost for exerting effort is $-\psi \cdot e_0$, while the gain associated with higher future expected income is discounted by $\beta_\tau$.

Given that agents discount the future ($\beta_\tau < 1$) and start with low income, we assume they can borrow in an unsecured consumer credit market at $t = 0$ to smooth consumption. Specifically, at $t = 0$, individuals can choose to borrow ($a_1 = -a$) or not ($a_1 = 0$). We assume that $y_\ell > a$ so that it is always budget feasible to pay back an individual's debt, so a decision to default is strategic and not due to an empty budget set.

Lenders observe individuals' credit market actions ($t = 0$ borrowing choice and $t = 1$ default decision $d_1 \in \{0, 1\}$), as well as their income $y_t$. They do not, however, observe an individual's type $\tau$. Thus, lenders use all relevant observable information to form assessments (Bayesian posteriors) about an individual's type which they use to correctly price unsecured credit. These assessments begin with the $t = 0$ prior that the individual is of type $H$, $s_0 \equiv \mathbb{P}(\tau = H) = \rho$. In this environment, the posterior that an individual is of type $H$ at the beginning of date 1 after observing her credit market action $a_1$ and income realization

[5] Throughout this paper, we denote probability with $\mathbb{P}(\cdot)$ and expectations with $\mathbb{E}(\cdot)$.

$y_1$ is

$$\begin{aligned} s_1(a_1, y_1) &= \mathbb{P}(\tau = H \mid a_1 \in \{-a, 0\}, y_1) \\ &= \frac{\mathbb{P}((a_1, y_1) \mid \tau = H) s_0}{\mathbb{P}((a_1, y_1) \mid \tau = H) s_0 + \mathbb{P}((a_1, y_1) \mid \tau = L)(1 - s_0)} \end{aligned} \tag{1}$$

Similarly, the posterior after observing $d_1$ is $s_2(d_1, a_1, y_1, s_1) \equiv \mathbb{P}(\tau = H \mid d_1, a_1, y_1, s_1)$.[6] The fact that the likelihood of observing a given $y_1$ depends on the unobservable effort choice $e_0$ implies moral hazard. This, as well as adverse selection, complicates the creditworthiness inference problem.

Competitive lenders condition the price of debt on these assessments as well as any other observables in their information set which might affect the likelihood of default. We denote the price of debt $q(a_1)$ at $t = 0$, which is pinned down by a zero profit condition to be described in Section 3.2.

Finally, we assume individuals have risk averse preferences over consumption $c_t$ given by $U(c_t)$ and have preferences over their reputation $\phi(s_2)$ at the end of period $t = 1$. We assume simply that $\phi$ is an increasing function of $s_2$. One can think of this reduced form preference over reputation as capturing how a parent's reputation may affect their access to other markets like insurance as in Chatterjee et al. (2008), labor as in Corbae and Glover (2024), or even children's access to future credit much like an intergenerational gift. The perpetual youth model in Chatterjee et al. (2023) provides a micro-foundation for $\phi(\cdot)$ via an individual's value function which is increasing in their posterior assessment due to enhanced credit access.

To summarize timing, we have:

1. At $t = 0$, an individual chooses $a_1$ and $e_0$ taking as given $q(a_1)$, understanding it will affect the lender's assessment of their type at stage (2a)
2. At $t = 1$,
   (a) $y_1$ realized and $s_1(a_1, y_1)$ evaluated
   (b) $d_1$ choice made and $s_2(d_1, a_1, y_1, s_1)$ evaluated

[6] For notational simplicity, we suppress dependence of $s_1$ on the prior $s_0 = \rho$ since it is simply taken parametrically.

## 3.2 Bayesian Equilibrium

The individual problem in period $t = 0$ is

$$\begin{aligned}(a_1^*(\tau), e_0^*(\tau)) &= \arg \max_{a_1 \in \{-a,0\}, e_0 \in \{0,1\}} U(c_0) - \psi e_0 + \beta_\tau \mathbb{E}_{y'|e_0} \left[ V_1(a_1, y_1, s_1; \tau) \right] \\ \text{subject to:} \quad & c_0 + q(a_1) a_1 = y_\ell \end{aligned} \tag{2}$$

At date 1, conditional on being in debt ($a_1 = -a$), the individual solves

$$\begin{aligned} V_1(a_1, y_1, s_1; \tau) &= \max_{d_1 \in \{0,1\}} U(c_1) + \beta_\tau \phi \left( s_2(d_1, a_1, y_1, s_1) \right) \\ \text{subject to:} \quad & c_1 = y_1 + (1 - d_1) a_1 \end{aligned} \tag{3}$$

The solution to (3) is simple: $d_1^*(-a, y_i, s_1; \tau) = 0$ if $U(y_i - a) + \beta_\tau \phi \left( s_2(d_1 = 0, -a, y_i, s_1) \right) \geq U(y_i) + \beta_\tau \phi \left( s_2(d_1 = 1, -a, y_i, s_1) \right)$. In terms of signalling theory, we can write this as:

$$\phi \left( s_2(d_1 = 0, -a, y_i, s_1) \right) - \phi \left( s_2(d_1 = 1, -a, y_i, s_1) \right) \geq \frac{U(y_i) - U(y_i - a)}{\beta_\tau} \tag{4}$$

where the left hand side of (4) is the reputational benefit of paying back while the right hand side is the cost of paying back. Notice that with $\beta_H > \beta_L$, the signaling cost is higher for the $L$ type. Notice further that given a concave utility function then $U(y_h) - U(y_h - a) < U(y_\ell) - U(y_\ell - a)$ so that the cost of paying back in higher in low income states, making it more likely that default happens in low income states.

At $t = 0$, a competitive lender must assess an individual's probability of repayment $1 - d_1^*$ on a loan of size $-a$ in all possible future states; the only current observable is the choice of $a_1 = -a$ itself. Given this probability, the lender must only discount the expected payoff back to the present date. Therefore, the equilibrium loan price must satisfy

$$q(-a) = \frac{1}{1 + r} \mathbb{P}(d_1 = 0 \mid a_1 = -a) \tag{5}$$

Notice that given the timing and informational assumptions, lenders have a difficult inference problem in (5) when posting zero profit prices at $t = 0$. Specifically, the equilibrium default decision depends not only on observable $a_1$ but also on not yet observed $y_1$ whose likelihood itself depends on the unobservable effort choice $e_0$ at $t = 0$. Hence $\mathbb{P}(d_1 = 0 \mid a_1 = -a)$ poses a complicated inference problem for the credit scorer due to the interaction of moral hazard and adverse selection.[7] Crucially, this interaction influences the age profile of lifetime earnings. If there is separation between types (e.g. $e_0^*(H) = 1$ while $e_0^*(L) = 0$ since type $H$

[7] This interaction was absent in our earlier work Chatterjee et al. (2023).

is willing to bear the current cost because they value the *future* gain in income more than type $L$), then upon observing a realization $y_h$, a Bayesian lender will update their posterior $s_1$ upward. On the other hand, if there is pooling of decisions (i.e. $e_0^*(H) = e_0^*(L)$), then the lender must form at least one off-the-equilibrium path beliefs about the individual's type.

## 3.3 Bayesian Issues and a Fix

**The Issue** We now explain a methodological contribution of our framework which has relevance for the empirical contribution of our paper. We illustrate it by working backwards on one portion of the individual's problem: the decision to default in (3) on $a_1 = -a$. Consistent with the determinants of the default decision in (4), consider individuals of both types who start with the same score $s_1$ and receive the same $y_1$. Suppose type $H$ never defaults, while type $L$ defaults when $y_1 = y_\ell$ but not when $y_1 = y_h$:

| $d_1^*(-a, y_i, s_1; \tau)$ | $\tau = H$ | $\tau = L$ |
|---|---|---|
| $y_h$ | 0 | 0 |
| $y_\ell$ | 0 | 1 |

Under this conjecture, in the state $y_1 = y_h$ where there is pooling, we have:

$$s_2(d_1 = 1, -a, y_h, s_1) = \frac{0 \cdot s_1}{0 \cdot s_1 + 0 \cdot (1 - s_1)} = \frac{0}{0}, \tag{6}$$

$$s_2(d_1 = 0, -a, y_h, s_1) = \frac{1 \cdot s_1}{1 \cdot s_1 + 1 \cdot (1 - s_1)} = s_1. \tag{7}$$

It is evident from (6) that we have an off-equilibrium-path (o-e-p) beliefs issue when $d_1 = 1$. In that case, many researchers simply assume since there is no new information,

$$s_2^{o-e-p}(d_1 = 1, -a, y_h, s_1) = s_1. \tag{8}$$

Since there is separation when $y_1 = y_\ell$, type is perfectly revealed:

$$s_2(d_1 = 1, -a, y_\ell, s_1) = \frac{0 \cdot s_1}{0 \cdot s_1 + 1 \cdot (1 - s_1)} = 0 \tag{9}$$

$$s_2(d_1 = 0, -a, y_\ell, s_1) = \frac{1 \cdot s_1}{1 \cdot s' + 0 \cdot (1 - s_1)} = 1 \tag{10}$$

Is this conjectured behavior that is similar to what we see in the data (i.e. no default in high income states with infrequent - in this case fraction $1 - \rho$ - default in low income states) consistent with a Bayesian equilibrium given what many researchers consider reasonable off-

equilibrium-path beliefs? In the $y_1 = y_\ell$ case where beliefs are well defined because types are separated, for the high type to repay and low type to default, we need, respectively:

$$[\phi(s_2(d_1 = 0, -a, y_\ell, s_1)) - \phi(s_2(d_1 = 1, -a, y_\ell, s_1))] = \phi(1) - \phi(0) \geq \frac{U(y_\ell) - U(y_\ell - a)}{\beta_H}$$
$$[\phi(s_2(d_1 = 0, -a, y_\ell, s_1)) - \phi(s_2(d_1 = 1, -a, y_\ell, s_1))] = \phi(1) - \phi(0) < \frac{U(y_\ell) - U(y_\ell - a)}{\beta_L}.$$

Both conditions are possible under certain parameterizations, especially with sufficiently low $\beta_L$ .

For the troublesome $y_1 = y_h$ case where agents are pooled and not all feasible actions are taken in equilibrium, in order for both types to default we need:

$$[\phi(s_2(d_1 = 0, -a, y_h, s_1)) - \phi(s_2(d_1 = 1, -a, y_h, s_1))] = \phi(s_1) - \phi(s_2^{o-e-p}) \geq \frac{U(y_h) - U(y_h - a)}{\beta_L} \tag{11}$$

So what conditions on o-e-p beliefs in $\phi(s_2^{o-e-p})$ make this an equilibrium? Notice that if we take the off-the-equilibrium path belief in (8) we have $s_2^{o-e-p} = s_1$ in which case (11) is given by:

$$\phi(s_1) - \phi(s_1) = 0 < \frac{U(y_h) - U(y_h - a)}{\beta_\tau}, \forall \tau. \tag{12}$$

In that case, a standard assumption for o-e-p beliefs results in non-existence of Bayesian equilibrium.

**The Fix:** There are three discrete choice problems in our model: (i) whether to borrow or not ($a_1 \in \{-a, 0\}$), (ii) whether to expend unobservable effort to raise future mean income ($e_0 \in \{0, 1\}$), and (iii) whether to default or not ($d_1 \in \{0, 1\}$). We assume, as in the standard discrete choice literature,that an individual's discrete choice is subject to unobservable shocks drawn from a type one ext distribution.

For example, in the discrete default choice described above, there are numerous shocks (besides income realizations) like unobserved (to the econometrician working at the credit scoring agency) medical expenditures or marital disruptions. Specifically, we now assume that in $y_1 = y_h$ states where an agent has debt $-a$, there are unobservable preference shocks over the default choice ($\varepsilon^{d_1}$) drawn from an extreme value function with scale parameter set to $\alpha$. In that case, we modify (3) to

$$V_1(a_1, y_1, s_1; \tau) = \max_{d_1 \in \{0,1\}} \mathbb{E}\left[U(c_1) + \varepsilon^{d_1} + \beta_\tau \phi(s_2(d_1, a_1, y_1, s_1))\right] \tag{13}$$

Letting $V_1^{d_1}(a_1, y_1, s_1; \tau)$ denote the value of choosing a given action $d_1$, the solution to

(13) yields choice probabilities given by, for the no-default case:

$$\sigma^{d_1=0}(-a, y_i, s_1; \tau) = \frac{\exp\left(\frac{V_1^{d_1=0}(-a,y_i,s_1;\tau)}{\alpha}\right)}{\sum_{d_1}\exp\left(\frac{V_1^{d_1}(-a,y_i,s_1;\tau)}{\alpha}\right)} \tag{14}$$

$$= \left[1+\exp\left[\frac{U(y_i)+\beta_\tau\phi\left(s_2(d_1=1,-a,y_i,s_1)\right)-U(y_i-a)-\beta_\tau\phi\left(s_2(d_1=0,-a,y_i,s_1)\right)}{\alpha}\right]\right]^{-1}.$$

Then, it is simple to see from the last line of (14) that as long as $\phi'(s_2) > 0$, type $H$ individuals are less likely to default than type $L$ when default lowers the Bayesian posterior of a person's type: that is,

$$\begin{aligned} &U(y_i)+\beta_H\phi\left(s_2(d_1=1,-a,y_i,s_1)\right)-U(y_i-a)-\beta_H\phi\left(s_2(d_1=0,-a,y_i,s_1)\right) \\ &< U(y_i)-\beta_L\phi\left(s_2(d_1=1,-a,y_i,s_1)\right)-U(y_i-a)-\beta_L\phi\left(s_2(d_1=0,-a,y_i,s_1)\right) \\ \iff &\phi\left(s_2(d_1=1,-a,y_i,s_1)\right) < \phi\left(s_2(d_1=0,-a,y_i,s_1)\right) \end{aligned} \tag{15}$$

Finally, notice that the choice probabilities in (14) are strictly bounded in $(0,1)$ so there is no longer an issue with supplying off-equilibrium-path beliefs associated with Bayesian posteriors for the perfectly separating case in (6). Specifically, posteriors are now given by

$$s_2(d_1,-a,y_i,s_1) = \frac{\sigma^{d_1}(-a,y_i,s_1;H)s_1}{\sigma^{d_1}(-a,y_i,s_1;H)s_1+\sigma^{d_1}(-a,y_i,s_1;L)(1-s_1)}.$$

The fact that $V_1(a_1,y_1,s_1;\tau)$ is increasing in $\beta_\tau$ implies $\sigma^{d_1=0}(-a,y_i,s_1;H) > \sigma^{d_1=0}(-a,y_i,s_1;L)$ which yields $s_2(d_1=0,-a,y_i,s_1) > s_2(d_1=1,-a,y_i,s_1)$ as in the last line of (15).

## 3.4 Model Properties

Here we illustrate some of properties of this simple model and consider the role of information in income and consumption inequality.

### 3.4.1 The role of adverse selection

Underlying our framework is an information friction: different types have different preferences, and therefore different propensities to exert effort and default. We begin, therefore, by assessing how credit market and effort decisions depend on the severity of this information friction. For illustrative purposes, we vary the intensity of the information friction by varying the "distance" between the two types: that is, we fix $\beta_H$ and study equilibria over a range of $\beta_L \in [\underline{\beta}_L, \overline{\beta}_L \equiv \beta_H]$. These results are simply a numerical illustration for an arbitrary set of

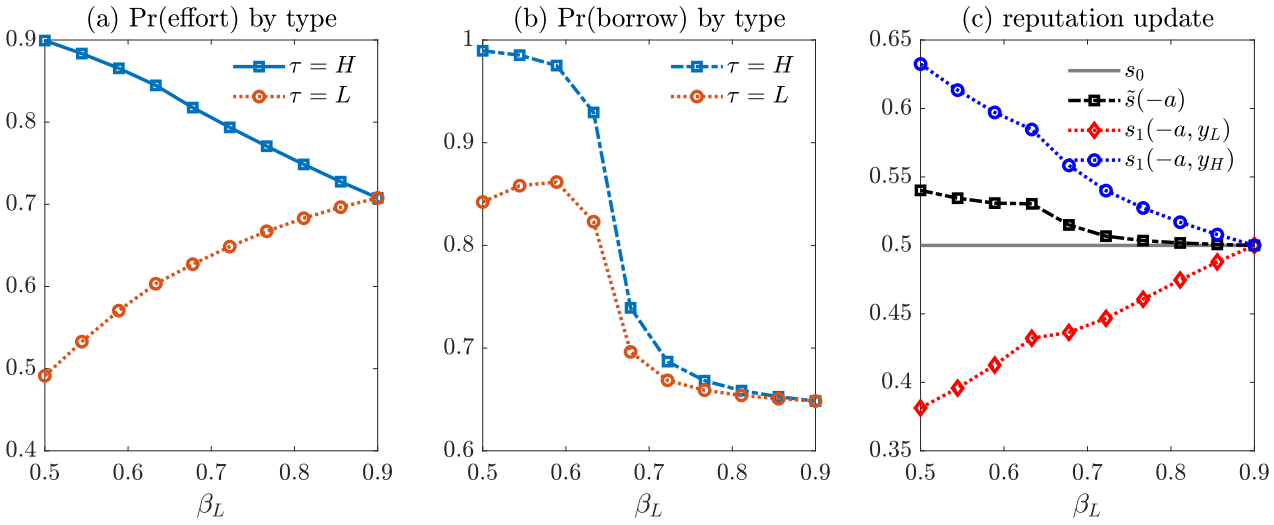



Figure 3: **Illustration of Role of Adverse Selection in Simple Model**

model parameters.

Figure 3 summarizes two key decisions early in life - observable borrowing and unobservable effort choices - as well as the dynamics of reputation in response to various outcomes. Beginning with the decision rules, a clear and intuitive pattern emerges. Panel (a) shows that patient types ($\tau = H$) are more likely to exert costly effort than impatient types. This is because they value the associated increase in future income (and therefore consumption) more. Since their date 1 income is higher relative to their date 0 income, they borrow and consume more at date 0, as reflected in panel (b). This result would be surprising in the absence of moral hazard: we'd expect the less patient agents to be more willing to borrow and increase current consumption at a given interest rate. Crucially, however, the interaction with moral hazard – that is, the endogenous correlation between effort, income, and the borrowing decision – overturns this simplistic dynamic.

These observations all hold for a given fundamental type gap $\beta_H - \beta_L$. Figure 3 shows further that these differences become even more pronounced as this gap grows. Crucially, the mechanism here is not merely that the increasingly impatient types exert less effort and therefore borrow less: at the same time, the patient types – whose fundamental degree of patience is unchanged – exert more effort and borrow more. Why is this? Panel (c) demonstrates that this change in behavior is directly attributable to reputation: as the gap between types grows, actions become more revelatory, and therefore the Bayesian posteriors diverge. Since agents value their reputations, then, high types are even more likely on the margin to take the actions which differentiate them from their counterparts.[8]

[8]The opposite is also true: as it becomes even more costly for the low types to mimic the high types, they do so less.

How are reputations updated? In this example, the population is split evenly between high and low types; that is, the prior $s_0 = 0.5$. In the absence of new information, the assessment of one's type remains at $s_0$ (flat grey line). It is useful to examine how lenders' assessments change as new information is revealed sequentially. To this end, we consider an "interim" type score update *after* the date 0 asset choice, but *before* the realization of income at $y_1$. Formally, we define $\tilde{s}(a_1) \equiv \mathbb{P}(\tau = H \mid a_1)$, plotted in the black line in panel (c). As discussed above, as the type gap grows, patient types become more likely to borrow both in absolute terms and *relative* to the impatient types. Therefore, the act of borrowing becomes more revelatory and leads to a larger revision of assessed type.

Agents unobservable effort choice at date 0 governs the likelihood of each income realization at date 1. Hence, upon observing $y_1$, lenders further update the type assessment from $\tilde{s}(a_1)$ to $s_1(a_1, y_1)$. Panel (c) plots these revisions for each possible $y_1$, conditional on borrowing at date 0. Since effort increases the likelihood of high income and patient types are more likely to exert effort, lenders revise assessed type upward upon observing $y_1 = y_h$ (blue line). Conversely, observing low income leads to a downward revision (red line). Notably, this downward revision outweighs the initial upward revision from borrowing. This is because the gap between choice probabilities between the types is larger for effort than for borrowing, and so this decision is more informative. This is sensible, since the borrowing decision is confounded by the fundamental impatience of the low types, which makes them more inclined to borrow. Importantly, this pattern from our numerical illustration is consistent with what is observed in Figure 1: as time goes on, further informative realizations lead to more diffuse assessments of type.

### 3.4.2 The role of moral hazard

Relative to Chatterjee et al. (2023), one of the innovations in this framework is the presence of moral hazard. In this section, we examine how moral hazard affects the performance of the model by considering the analyses of the previous subsection for two levels of the effort cost $\psi$: the baseline level, and a lower level. When $\psi$ is lower, effort is less costly, and so the moral hazard friction is weaker. Moreover, rather than considering the same set of decisions as above, we examine how this friction affects empirically observable objects of interest.

Panel (a) of Figure 4 plots the loan rate $r(-a) \equiv \frac{1}{q(-a)} - 1$ across levels of $\beta_L$ for the baseline and low moral hazard cases. Two clear patterns emerge. First and foremost, loan rates are more sensitive to the severity of moral hazard than to the type gap. Specifically, the low moral hazard model features markedly cheaper credit. This is because when effort is cheaper, it is undertaken more, leading to higher income *on average* in the low moral hazard case. Given a standard leverage effect, then, agents are less likely to default conditional on

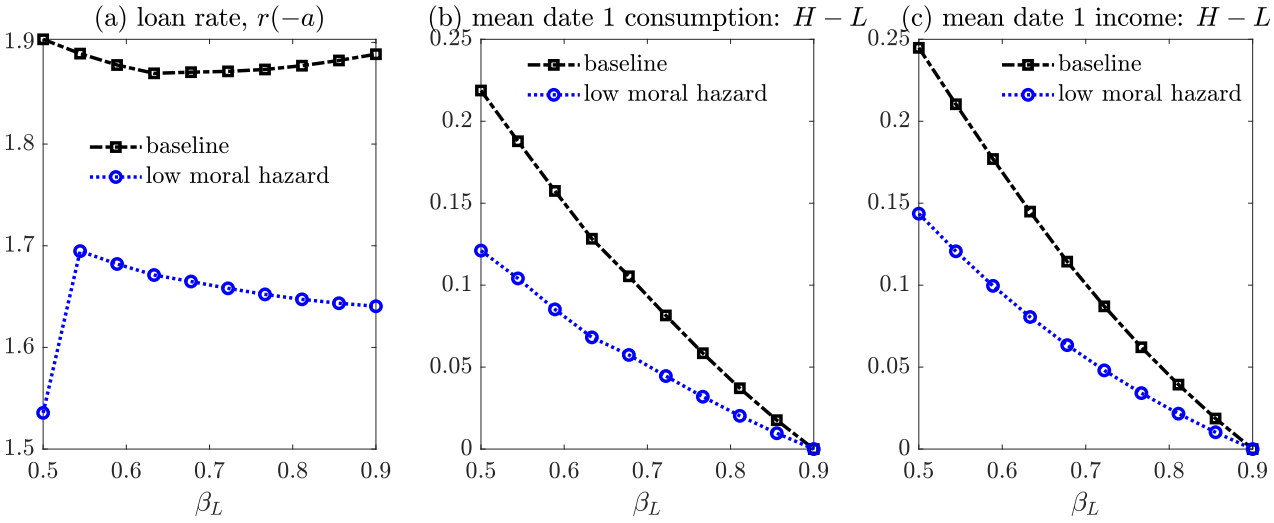



Figure 4: **Illustration of Role of Moral Hazard in Simple Model**

borrowing, which drives down the cost of borrowing. Second, loan rates are non-monotone in the type gap for a given level of moral hazard.

Panels (b) and (c) provide two measures of inequality: the date 1 gap between high and low type average consumption and income, respectively. Again, we get a clear insight: the more severe is the moral hazard friction, the greater is the degree of inequality. It should be noted that, while not shown in Figure 3, moral hazard increases not only consumption and income inequality in the cross-section, but also affects a given individual's intertemporal consumption smoothing motives.

### 3.4.3 The role of information

Towards consideration of our intended policy counterfactuals in which we will analyze the equilibrium effects of changes in credit market information restrictions, in this section we compare our baseline model to a "full information" version which has typically been studied in the quantitative literature like Livshits et al. (2007) and Chatterjee et al. (2007). Notably, while intermediate cases are possible, the version we present here is the simplest: both type and effort are observable. This means that both the adverse selection and moral hazard frictions analyzed above are absent in this version of the model. Concretely, this version has no role for type scores, and the date 0 loan pricing function can condition on both $e$ and $\tau$ in addition to $a_1$. Figure 5 presents a selection of results comparing this model to the baseline.

We begin by considering the effort decision in panel (a). Two findings emerge immediately. First, for each type, there is more effort exerted in the baseline model than in the full information case. This follows from the fact that the marginal benefit of exerting effort is higher in the version of the model with moral hazard: individuals gain not only from the

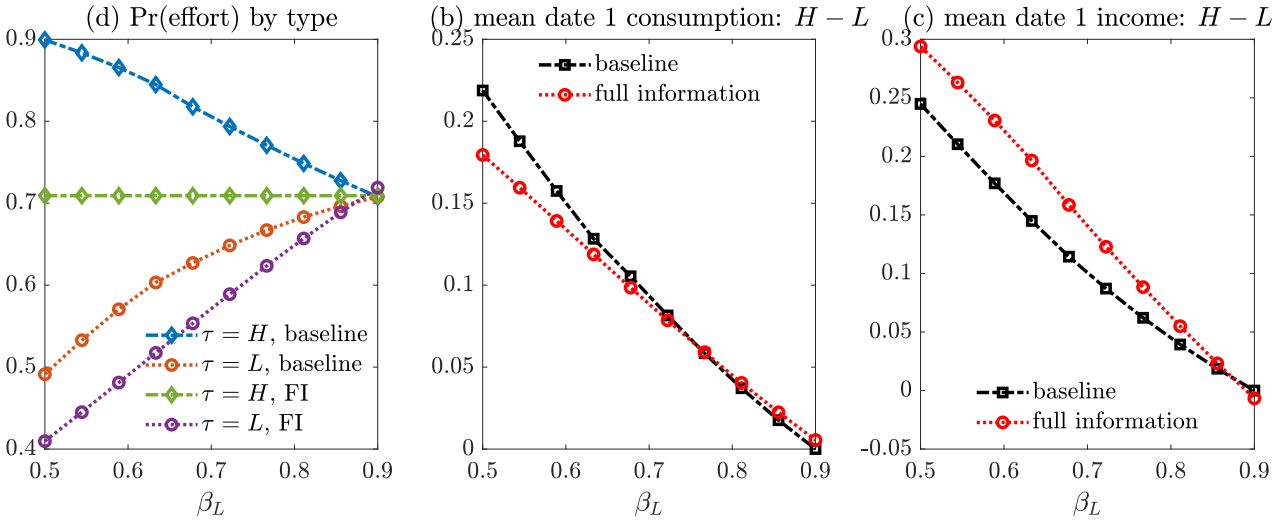



Figure 5: **Illustration of Role of Information in Simple Model**

expected increase in future income, but from the associated increase in reputation stemming from signalling that they are the patient $H$ type. This latter effect is of course absent in the full information model. Second, in the full information model the behavior of the patient type is invariant to the level of patience for the impatient type. This again stems from the lack of reputational considerations. Panel (a) also makes clear the incentive effects of the link between credit and labor markets: an impatient $L$ type has no incentive to take effort in order to mimic a patient type to enhance their reputation in the credit market.

How do these forces affect our key outcomes of interest, consumption and income inequality? Panels (b) and (c) explore these outcomes, respectively. The key finding from these figures is that the tight link between income and consumption is broken: that is, even though the full information model features more income inequality than the baseline, the baseline still features more consumption inequality. This force emerges due to the different behavior of credit prices across the two models. This finding exemplifies the scope for interesting and nuanced policy implications once the joint relationship between reputation and income is properly accounted for.

## 3.5 Mapping to Credit Scores

In the economy described thus far, an individual's reputation is encapsulated in her type score $s$. In U.S. credit markets, it is encapsulated in a *credit score*. A credit score is an index that is positively related to the likelihood of repayment. Here we show that under certain conditions, an equilibrium with discrete choice shocks as described in the previous subsection can be implemented via an arrangement in which lenders use a model equivalent of a credit

score to assess the probability of repayment on a loan conditional on current observables in their information set.

Specifically, we formalize the notion that a credit score depicts a consumer's creditworthiness by defining it to be the probability of repayment $\mathbb{P}(d_1 = 0 \mid a_1 = -a)$ defined in equation (5). Note that since the default decision comes in state $(-a, y_i, s_1; \tau)$ which includes the posterior $s_1$, the conditional expectation $\mathbb{P}(d_1 = 0 \mid a_1 = -a)$ must account for all those future states.[9] In the simple model above, all individuals at $t = 0$ start in an identical observable state $\omega$ with zero assets, the lowest earnings level, and common priors. In Chatterjee et al. (2023) we show an equivalence between a related type-scoring environment and one which uses credit scores if there is a one-to-one and onto mapping between $s(\omega) \in [0, 1]$ and $\mathbb{P}(d_1 = 0 \mid \omega) \in [0, 1]$, holding fixed the other observables in $\omega$. In that case, wherever $s(\omega)$ appears in the theoretical model, it can be replaced by $\mathbb{P}(d_1 = 0 \mid \omega)$. As discussed in Section 2, since $\mathbb{P}(d_1 = 0 \mid a_1 = -1)$ lies in $[0, 1]$ while credit scores in the data lie in $[300, 850]$, we simply use ordinal "credit rankings" in both the model and the data.

# 4 Policy Implications

What are the implications of

- technological innovation (i.e. "big data")
- laws restricting adverse events from being on an individual's credit record
- consolidation of credit relevant information as business models evolve (e.g. one's credit card company is now also one's bank)
- the rise of non-regulated credit scorers (i.e. fin-tech companies)

for credit access and income and consumption inequality?

The answers to these questions obviously depend on the degree to which credit scores accurately separate types and incentivize reputation acquisition. In the early stages of the pandemic, income support and forbearance programs led to major shifts in the credit score distribution documented by Driscoll et al. (2024). The authors show that very large increases in subprime delinquency rates are in large part driven by pandemic-era shifts in the credit score distribution, or what they call "upward credit score migration" as the share of borrowers with subprime scores fell from about 23% to 18% as the those individuals migrated up and

[9] Notably, we can always use $\mathbb{P}(d_1 = 0 \mid a_1 = -a)$ to run a counterfactual to calculate the probability of repayment even if the person chose not to borrow (i.e. chose $a_1 = 0$) since they will have a different $s_1(y_1, a_1 = 0)$.

were pooled with near and prime borrowers. The authors run a counterfactual which shows that the large increase in subprime delinquency rates would not have arisen without the migration.

Innovation in consumer finance can also degrade the information content of credit scores by creating "hidden debts." When Buy-Now-Pay-Later (BNPL) lending became more common, there was a question of whether these new types of debts were properly reported since the big credit card banks were not doing it. These "hidden debts" were showing up when people defaulted on them but not otherwise.

These real world examples illustrate how technological innovations or legal restrictions on information can affect the credit score distribution. Whether these changes lead to more separation or more pooling have important general equilibrium effects and consequences for inequality and welfare. Through the lens of our model, we can decompose the implications of changes in information sets and evaluate policy impacts in the credit scoring industry which have hitherto been left unexplored.

# References


Amador, M. and C. Phelan (2021): "Reputation and Sovereign Default," *Econometrica*, 89, 1979–2010.

Athreya, K., X. S. Tam, and E. R. Young (2012): "A Quantitative Theory of Information and Unsecured Credit," *American Economic Journal: Macroeconomics*, 4, 153–183.

Chatterjee, S., D. Corbae, K. Dempsey, and J.-V. Rios-Rull (2023): "A Quantitative Theory of the Credit Score," *Econometrica*, 91, 1803–1840.

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

Chatterjee, S., D. Corbae, and J.-V. Ríos-Rull (2008): "A finite-life private-information theory of unsecured consumer debt," *Journal of Economic Theory*, 142, 149–177.

Cole, H. and P. Kehoe (1998): "Models of Sovereign Debt: Partial Versus General Reputations," *International Economic Review*, 39, 55–70.

Corbae, D. and A. Glover (2024): "Employer Credit Checks: Poverty Traps versus Matching Efficiency," *Review of Economic Studies*, forthcoming.

Dempsey, K. and F. Ionescu (2023): "Borrowing Premia in Unsecured Credit Markets," Unpublished.

Diamond, D. W. (1989): "Reputation Acquisition in Debt Markets," *Journal of Political Economy*, 97, 828–862.

Driscoll, J., J. Flagg, B. Katcher, and K. Sommer (2024): "The Effects of Credit Score Migration on Subprime Auto Loan and Credit Card Delinquencies," Unpublished.

Heathcote, J., F. Perri, and G. Violante (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13, 15–51.

Kovrijnykh, N., I. Livshits, and A. Zetlin-Jones (2023): "Building Credit Histories," Unpublished.

Krueger, D. and F. Perri (2006): "Does Income Inequality Lead to Consumption Inequality ? Evidence and Theory," *Review of Economic Studies*, 73, 163–193.

Livshits, I., J. MacGee, and M. Tertilt (2007): "Consumer Bankruptcy: A Fresh Start," *The American Economic Review*, 97, 402–418.

Livshits, I., J. C. MacGee, and M. Tertilt (2016): "The Democratization of Credit and the Rise in Consumer Bankruptcies," *Review of Economic Studies*, 83, 1673–1710.

Myerson, R. B. (1978): "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 73–80.

Narajabad, B. N. (2012): "Information Technology and the Rise of Household Bankruptcy," *Review of Economic Dynamics*, 15, 526–550.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

Sanchez, J. M. (2018): "The Information Technology Revolution and the Unsecured Credit Market," *Economic Inquiry*, 56, 914–930.

Selten, R. (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 25–55.