*Econometrica*, Vol. 91, No. 5 (September, 2023), 1803–1840

# A QUANTITATIVE THEORY OF THE CREDIT SCORE

SATYAJIT CHATTERJEE
Research Department, Federal Reserve Bank of Philadelphia

DEAN CORBAE
Department of Economics, University of Wisconsin-Madison and NBER

KYLE DEMPSEY
Department of Economics, The Ohio State University

JOSÉ-VÍCTOR RÍOS-RULL
Department of Economics, University of Pennsylvania, CAERP, CEPR, NBER, and UCL

What is the role of credit scores in credit markets? We argue that it is, in part, the market's assessment of a person's unobservable type, which here we take to be patience. We postulate a model of persistent hidden types where observable actions shape the public assessment of a person's type via Bayesian updating. We show how dynamic reputation can incentivize repayment. Importantly, we show how an economy with credit scores implements the same equilibrium allocation. We estimate the model using both credit market data and the evolution of individuals' credit scores. We conduct counterfactuals to assess how more or less information used in scoring individuals affects outcomes and welfare. If tracking of individual credit actions is outlawed, poor young adults of low type benefit from subsidization by high types despite facing higher interest rates arising from lower dynamic incentives to repay.

KEYWORDS: Credit scores, unsecured consumer credit, bankruptcy, persistent hidden information.

## 1. INTRODUCTION

CREDIT SCORES are a fundamental ingredient of a borrower's access to credit. In the United States, credit bureaus and credit rating agencies serve this function for individual borrowers. Similar agencies exist in many other countries. Credit scores affect borrowing terms and change with credit use and repayments. Despite their widespread use in actual credit markets, credit scores are conspicuously absent from standard quantitative models of consumer default, which are typically more concerned with allocations than the contractual arrangements that generate them.

Satyajit Chatterjee: satyajit.chatterjee@phil.frb.org
Dean Corbae: dean.corbae@wisc.edu
Kyle Dempsey: dempsey.164@osu.edu
José-Víctor Ríos-Rull: vr0j@econ.upenn.edu

This project has been a long time in the making. An earlier version circulated under the title "Credit Scoring and the Competitive Pricing of Default Risk." We thank three referees for very helpful comments on an earlier version of this paper and Isaiah Andrews as well as Tsung-Hsien Li and Jan Sun for this version. We thank all the economists who have had an important input into the final product: Murat Tasci, Pablo D'Erasmo, Daphne Chen, Jake Zhao, and Kuan Liu. We thank Hongchao Zhang for kindly sharing with us his DFBOLS Fortran code. We thank Cole Drier and Michael Slonkosky for research assistance. We also thank the many seminar and conference participants who commented on earlier versions of the paper. Finally, Corbae and Ríos-Rull wish to thank the National Science Foundation for support under Grants SES-0751380 and SES-0351451. 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.

In this paper, we provide a theory of the joint behavior of unsecured credit and credit scores which accounts for both allocations and arrangements. Reputations are formed in the presence of hidden information about a persistent, credit-relevant individual characteristic, which we take to be patience. The incentive to maintain a good reputation plays a central role, shaping borrowing and saving behavior over individuals' lifetimes.

Our theory is founded on the premise that an individual's true propensity to repay—that is, the individual's true *type*—is hidden from her creditors, and it is the presence of this *persistent hidden information* that makes an individual's history of actions relevant for lenders. Our theory is dynamic: at any point in time, lenders use a person's observable history of actions to perform a Bayesian update of her type; individuals understand this and choose actions mindful of the consequence any action has on the future beliefs of lenders. A loss of reputation, rather than stigma or exogenous exclusion from future borrowing, is the only dynamic punishment from default. Specifically, an individual's credit score falls upon default and she subsequently faces worse borrowing terms. Our theory is competitive: information available to any lender is available to all lenders and there is free entry into the business of lending. Finally, our theory respects a key feature of the institutional arrangement under which unsecured consumer credit is extended in the United States: at some monetary cost, individuals can choose to have their debts discharged via Chapter 7 bankruptcy.

We make several contributions. First, we extend the theory of unsecured credit to accommodate persistent hidden information about individual types. Our model environment is rich enough to cover four of the five characteristics lenders use to assess creditworthiness: character (reflected in credit history), capacity (reflected in debt-to-income ratio), capital (wealth), and conditions (amount of the loan).[1] Competition drives lending contracts to be indexed by all observable borrower and loan characteristics relevant for predicting the probability of default on a loan. When there is hidden information, a new individual characteristic becomes relevant: a borrower's type probability vector—in the terminology of this paper, the borrower's *type score*—indicating the probability that a person is of each of the different types existing in the economy. The Bayesian update of an individual's type score conditions on all relevant observables: the individual's current type score, her current net wealth, all the relevant information to forecast future earnings, and, of course, her current action (save, borrow, or default). One way to interpret the large number of conditioning variables is that the lender is using "big data." Our framework easily encompasses "small data" cases in which lenders observe only some strict subset of actions (an instance of a"small data" world is explored in Section 6.2).

Second, after proving an equilibrium with type scores exists, we show that a market arrangement which uses credit scores replicates the same equilibrium allocation without any reference to type scores. Specifically, we use the type score to define a *credit score*—an object that yields a ranking of individuals with regard to their probability of default on a particular contract. Such an ordinal ranking is widely used by credit bureaus. We provide an easily verifiable sufficient condition such that the equilibrium under the arrangement that uses credit scores to index contracts has the same allocation as the equilibrium of our baseline economy with type scores. Just as agents take prices as given in standard competitive equilibrium models, in our equilibrium with credit scores individuals and lenders take credit-score-dependent prices and the distribution of future credit scores conditional on their current state and actions as given; they do not need to know what is behind such

[1]The fifth, collateral, is not relevant for unsecured credit. See https://www.investopedia.com/terms/f/five-c-credit.asp.

functions, just that they exist. In doing so, we provide a theory of the credit score itself and of how it evolves over time in response to fundamentals. In this context, we take to heart that the actual market arrangement is a form of data and our equivalence result allows for the use of such data for empirical purposes.

Third, we take our model to the data, estimating preference parameters (specifically a stochastic process for unobservable discount factors) from the joint behavior of credit scores over an individual's lifetime and aggregate credit market moments.[2] It is here that our decision to model age variation in the evolution of earnings and hidden characteristics pays off. For these estimates, we verify that the sufficient condition which guarantees equivalence between the type-score economy and credit-score economy holds. We find what we believe are important properties of the U.S. population with regard to (hidden) patience as revealed by the properties of the credit market: (i) the difference in discount factors between patient and impatient people is 13% annually; (ii) slightly less than one third of people are born patient, but the share of patient people rises with age; and (iii) patience is persistent but not permanent (the transitions between types occur with an average duration of between 4 and 5 years). Random changes in unobserved type (i.e., in patience) along with transitory variation in unobserved shocks to earnings and extreme value shocks to utility, prevent fast learning about an individual's type. As we quantify later in Section 5.3 on selection and reputation effects, the 13% difference in unobservable discount factors makes it costly for low discount factor (i.e., high-risk) types to mimic the asset market behavior of high discount factor (i.e., low-risk) types. Thus, our estimates suggest there is sufficient scope for signaling (and separation). The force for separation induces our estimates of the variance of the transitory shocks to preferences, especially for the default decision, to be relatively high. These two countervailing forces lead us into a sweet spot of hidden information consistent with the mean and standard deviation of rankings of credit scores across the age profiles in the data.

Fourth, we use our estimates to explore the role of hidden information in the U.S. unsecured credit market. We start by considering a policy counterfactual in which lenders are prohibited from keeping track of the history of an individual's asset market actions but can condition on the observable length of individuals' credit history (effectively their age). In this case, impatient types are pooled with patient types without having to bear the costs of imitating them in order to obtain better borrowing terms. Since young adults wish to borrow against their higher expected future income, and most start their adult life impatient, the policy has the possibility of improving the welfare of those young adults. However, the policy removes the incentives to maintain a good reputation, which leads to individuals facing higher interest rate offerings. We find the negative incentive effects roughly offset the potential pooling benefits except for young, poor impatient adults, who are made substantially better off.

Our second counterfactual considers an economy in which one's type is perfectly observable. The findings are intuitive. Since the impatient are known in this economy, they face a more adverse situation; their interest rates are higher and they borrow less. The opposite is true for the patient. As people age, this knowledge becomes less relevant because people accumulate precautionary balances and rarely borrow. The benefits to the patient outweigh the costs to the impatient, resulting in a relatively large welfare improvement associated with full information. Furthermore, individual-level allocations in

[2]There is an extensive empirical literature finding evidence of adverse selection in credit markets which includes Ausubel (1999), Agarwal, Chomsisengphet, and Liu (2010), Einav, Jenkins, and Levin (2013), and Hertzberg, Liberman, and Paravisini (2018). Related empirical papers which study credit scoring and default include Albanesi, DeGiorgi, and Nosal (2022) and Albanesi and Vamossy (2019).

the full information economy are quite different from individual-level allocations in the base economy, showing that our baseline economy is still far from being a full information economy.

Fifth, we make several methodological contributions. We combine both screening and dynamic signaling where these screening and signaling opportunities are constrained by noise which we introduce via extreme value shocks.[3] The shocks ensure that beliefs held by lenders following any feasible action are determined in equilibrium (reminiscent of Selten (1975) and Myerson (1978)).[4] The shocks also cloud inference about unobservable type; different types may choose the same action analogous to a semi-separating or partial pooling equilibrium. Finally, the shocks eliminate multiplicity of equilibria that can arise in signaling games from variation in off-the-equilibrium-path beliefs and provide tractability.

Sixth, we extend quantitative theory models of default with full information, like that in Livshits, MacGee, and Tertilt (2007) and Chatterjee, Corbae, Nakajima, and Rios-Rull (2007), to include hidden information which requires us to index the pricing of credit to the market assessment of individual types.[5,6] 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. Here, in the presence of hidden information, past asset market actions are informative about an individual's unobservable characteristics that are correlated with their repayment probability encapsulated in a credit score. Related quantitative theory papers with hidden information applied to consumer default include Athreya, Tam, and Young (2012) and Exler, Livshits, MacGee, and Tertilt (2021).[7] The former paper makes an anonymous markets

[3]The microeconomic literature classifies hidden knowledge/adverse selection models as "screening" or "signaling" models (Riley (2001)). As in screening models, in our paper lenders offer loans distinguished by loan characteristics (size of the loan) and observable personal characteristics (income, previous history) that give ample scope for separation (if such separation is desirable from an individual point of view and can be sustained in equilibrium). And, as in signaling models, there are actions that an individual can take (e.g., saving) that have no effect on the payoff to any lender but which convey valuable information to them. In the use of history to condition prices, our model shares a connection to the microeconomic literature that studies the conditioning of prices on customers' purchase histories (see, for instance, Acquisti and Varian (2005)).

[4]Even if there was no hidden information and no Bayesian updating of beliefs, a continuous support of shocks would be needed to ensure the existence of a *pure* strategy equilibrium; otherwise, existence would require that people be allowed to play mixed strategies. Despite individuals playing pure strategies, the shocks ensure different types may choose the same action analogous to a "semi-separating" equilibrium. Further, the assumption that the shocks are drawn from a type I extreme value distribution delivers the tractability as in Rust (1987).

[5]In full information environments, the observation that the competitive pricing of defaultable debt requires indexing the price of the loan to some observable characteristics like its size appeared in a clear form in Jaffee and Russell (1976) and Eaton and Gersovitz (1981). A large literature on quantitative models of defaultable consumer and sovereign debt now exists (see Exler and Tertilt (2020) and Aguiar, Chatterjee, Cole, and Stangebye (2016) for recent surveys).

[6]Gale (1992) and Dubey and Geanokoplos (2002) proved existence of competitive equilibrium in environments with hidden information. In contrast to us, they adopted the *anonymous markets assumption* of classical GE theory which does not permit prices to depend on personal characteristics of buyers or sellers (such as a credit score). Prescott and Townsend (1984) characterized constrained efficient allocation in an adverse selection environment but showed that there is no natural decentralization of it via a price system. Guerreri, Shimer, and Wright (2010) proved existence and uniqueness of separating equilibria in static adverse selection models by expanding the contract space to include competitive search over submarkets, which helps sustain separation. Our framework expands the contract space to include dynamic type scores which are used to help separate borrowers.

[7]Other related papers which include an information problem of some sort are Luo (2017), Kovrijnykh, Livshits, and Zetlin-Jones (2019), Nelson (2022), and Blattner, Hartwig, and Nelson (2022).

assumption where only current asset choices are observed but no prior information about an individual's asset market behavior can be used to infer their unobservable type to price credit, while the latter paper makes assumptions on types that effectively eliminate the adverse selection problem for lenders. Closely related quantitative theory papers with hidden information applied to sovereign default include D'erasmo (2011) and Fourakis (2021).

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 an improvement in technology is an increase in signal precision. These include Narajabad (2012), Livshits, MacGee, and Tertilt (2016), Drozd and Serrano-Padial (2017), and Sanchez (2018). For instance, Narajabad (2012) examined the polar cases where the credit market lacks information on borrower's riskiness and rating technologies do not work well resulting in a pooling equilibrium versus the case where there is sufficient information to separate borrowers according to their unobservable cost of default. Livshits, MacGee, and Tertilt (2016) considered 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 partial pooling of borrowers in a given contract falls.

While previous quantitative theory models imposed exogenous punishment, we incorporate dynamic reputation as a means of disciplining borrowers along the lines of Diamond (1989) and more recently Elul and Gottardi (2015). Our reputational environment, where everyone optimizes but people have hidden knowledge about their preferences, is closely linked to repeated games with incomplete information (see Peski (2014) for a discussion of this literature). Reputation in debt markets in which one player is a commitment type has 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, Corbae, and Ríos-Rull (2008), Corbae and Glover (2018), and Braxton, Herkenhoff, and Phillips (2020).

Section 2 describes our baseline economy with hidden information. Section 3 describes the equilibrium problems faced by our agents. Section 4 describes how we map the model to data and Section 5 studies the properties of the estimated model. Section 6 compares our baseline economy to alternative economies with different information structures. Section 7 concludes. There is an accompanying Supplemental Material (Chatterjee, Corbae, Dempsey, and Ríos-Rull (2023)), where we provide additional theoretical (Appendix A) and computational (Appendix B) results, description of data (Appendix C), and an extension of the baseline model to delinquency (Appendix D). There is also a file with Additional Materials posted in a working paper posted on the author's website.

## 2. ENVIRONMENT

We pose a model of perpetual youth as in Blanchard (1985) and Yaari (1965) with constant population. Agents die with probability $1-\rho$ at the end of the period and those who die are replaced by newborns so that there is always a unit measure of agents. An individual's persistent log earnings, denoted $e_t \in \mathcal{E} = \{e_1, e_2, \dots, e_E\} \subset \mathbb{R}_{++}$, are exogenously drawn from a stationary finite state Markov process $Q^e(e_{t+1}|e_t)$. In addition, there are purely transitory (log) earnings, denoted $z_t \in \mathcal{Z} = \{z_1, z_2, \dots, z_Z\} \subset \mathbb{R}_{++}$, which are exogenously drawn from a stationary probability distribution $H(z_t)$. All earnings draws are independent across individuals and we denote individual total earnings as $y_t(e_t, z_t) = \exp(e_t + z_t)$. A newborn's earnings are drawn from initial distribution $F_e$.[8]

At time $t$, individuals can choose assets $a_{t+1} \in \mathcal{A} = \{a_1, a_2, \dots, a_N\} \subset \mathbb{R}$, where $a_1 < a_2 < \cdots < a_N$, at discount price $q_t$ determined in a competitive market. We assume the finite set $\mathcal{A}$ includes 0 with $a_1 < 0$ and $a_N > 0$. If an agent holds debt (i.e., $a_t < 0$), she can choose whether or not to file for bankruptcy $d_t \in \mathcal{D} = \{0, 1\}$. If she files (i.e., $d_t = 1$), then in the period of filing she cannot borrow or save (i.e., $a_{t+1} = 0$) and her earnings net of the costs of bankruptcy become $y_t(e_t, z_t)(1 - \kappa_1) - \kappa$, where $\kappa > 0$ is the bankruptcy filing fee and $\kappa_1 \in (0, 1)$ proxies for the negative consequences of bankruptcy on one's earnings.[9]

In each period $t$, the individual values consumption $c_t$ using a utility function $u(c_t) : \mathbb{R}_{++} \to \mathbb{R}$ which is is continuous, increasing, and concave. At time $t$, an individual discounts her future utility at rate $\beta_t \in \mathcal{B} = \{\beta_1, \beta_2, \dots, \beta_B\}$ if she survives. Her discount factor varies stochastically over time drawn from a finite state Markov process $Q^\beta(\beta_{t+1}|\beta_t)$. The $\beta_t$ are drawn independently across individuals and are unobservable to others. We call $\beta_t \in [0, 1)$ a household's *type*.

In addition, households receive action-specific, additively separable extreme value preference shocks which enter households' flow utility each period. The first set of shocks attach to the bankruptcy/no bankruptcy choice ($d_t \in \{0, 1\}$) and is therefore drawn only by borrowers:

$$\nu_t = \left(\nu_t^{d=0}, \nu_t^{d=1}\right). \tag{1}$$

The second set of shocks of length $N$ attaches to $a_{t+1}$ choices in the event of no default:

$$\epsilon_t = \left(\epsilon_t^{a_1}, \epsilon_t^{a_2}, \dots, \epsilon_t^{a_N}\right). \tag{2}$$

The vectors $\nu_t$ and $\epsilon_t$ are drawn independently across individuals. Each element $\nu_t^d$ and $\epsilon_t^{a_n}$ is drawn from type I extreme value distributions $F_\nu$ and $F_\epsilon$ with scale parameters $\alpha$ and $\lambda$, respectively.[10]

Intermediaries can observe individuals' persistent earnings (i.e., $e_t$) and asset market behavior (i.e., $a_t$, $d_t$, and $a_{t+1}$), but cannot observe their preferences (i.e., $\nu_t$, $\epsilon_t$, and $\beta_t$) nor the transitory component of earnings (i.e., $z_t$). Since $\nu_t$, $\epsilon_t$, and $z_t$ are i.i.d. over time and individuals, nothing can be learned about their future values from their current values. In contrast, since $\beta_t$ is drawn from a persistent Markov process, the probability distribution of its future values depends on its current (unobservable) value. We denote the creditor's probability assessment that an individual is of type $\beta_i$ at the beginning of period $t$ before any actions are taken as $s_t(\beta_i) = \Pr(\beta_t = \beta_i)$. We call $s_t = (s_t(\beta_1), \dots, s_t(\beta_B))$ an individual's *type score* with $\sum_{i=1}^B s_t(\beta_i) = 1$.[11]

Given an individual's observable characteristics $\omega_t = (e_t, a_t, s_t)$ as well as their credit market actions $(d_t, a_{t+1})$, financial intermediaries revise their assessments of an individ-

[8]By setting $F_e$ to the degenerate distribution that has all mass on the lowest persistent earnings level, we can parsimoniously achieve a rising earnings profile over an individual's working life as an approximation to the full life cycle model in Livshits, MacGee, and Tertilt (2007).

[9]See Corbae and Glover (2018) for a model in which a poor credit record adversely affects an individual's earnings.

[10]For reasons that we explain when computing the model in Section 4.1, we permit the location parameters of the extreme value shocks to depend on the options available to households.

[11]Of course, the framework is rich enough to add more unobservables. For instance, if the persistent component of earnings is unobservable, then $s_t = (s_t(\beta_1, e_1), s_t(\beta_1, e_2), \dots, s_t(\beta_B, e_E))$.

ual's type from $s_t$ via Bayes's rule.[12] We denote this update as $\psi_t^{(d_t,a_{t+1})}(\omega_t) \in [0,1]^B$. As a result of this assessment, the prices faced by an individual in the credit market will also depend on her observable state and her credit market actions. Thus, we denote the price function for an individual with observable characteristics $\omega_t$ who chooses assets $a_{t+1}$ by $q_t^{a_{t+1}}(\omega_t)$. The arguments of $q_t$ influence the price because they can directly affect the likelihood of repayment on a loan (as in standard debt and default models) and indirectly by revealing information about the individual's current type (this is encoded in the update $\psi_t^{(0,a_{t+1})}(\omega_t)$).[13] Importantly, note that all the *future* implications of current credit market actions are encapsulated in the update $\psi_t^{(d_t,a_{t+1})}$. In particular, the only punishment to bankruptcy in future periods (aside from those that follow from the requirement that saving is not permitted in the filing period) is the possible loss of reputation stemming from intermediaries' adverse assessments of her unobservable type.

For technical reasons, we assume $s_t \in \mathcal{S}$, a finite subset of $[0,1]^B$. This assumption makes it possible to apply standard methods to prove existence of equilibrium. Since the posterior $\psi_t$ may not lie on one of the finite points in $\mathcal{S}$, we assign it randomly to nearby points in $\mathcal{S}$. We denote the probability mass function implied by our random assignment as $Q^s(s_{t+1}|\psi_t)$.[14] We can show that the following holds:

LEMMA 1: *There exists an assignment rule satisfying*: (*i*) $\mathbb{E}_{s_{t+1}\in\mathcal{S}}[Q^s(s_{t+1}|\psi_t)] = \psi_t$ (*i.e., consistency*), (*ii*) *the variance of the approximation error* (*i.e., of* $s_{t+1}$ *from* $\psi_t$) *is arbitrarily small*, *and* (*iii*) $Q^s(s_{t+1}|\psi_t)$ *is continuous in* $\psi_t$.

DEFINITION 1: The *timing* in any given period is as follows:

1. All individuals (survivors and newborns) begin with the vector $(\beta_t, e_t, a_t, s_t)$ and receive a transitory earnings shock $z_t$.
2. Individuals who have $a < 0$ receive the random utility vector $\nu_t$ and decide whether to file for bankruptcy ($d_t = 1$) or not ($d_t = 0$).
3. Individuals who have not filed for bankruptcy receive the random utility vector $\epsilon_t$ and choose a feasible action given prices $q_t^{a_{t+1}}(\omega_t)$.
4. Based on each individual's actions $(d_t, a_{t+1})$ and observable characteristics $\omega_t$, intermediaries revise their assessments of an individual's type via Bayes's rule, updating $s_t$ to $\psi_t$.
   (a) Individuals who survive draw beginning-of-next-period realizations of $\beta_{t+1}$ and $e_{t+1}$ from the exogenous transition functions $Q^\beta(\cdot|\beta_t)$ and $Q^e(\cdot|e_t)$. The beginning of next-period type score $s_{t+1}$ is drawn from the probability mass function $Q^s(\cdot|\psi_t)$.
   (b) Newborns begin life with $\beta_{t+1}$ drawn from initial distribution $F_\beta$, earnings class $e_{t+1}$ drawn from initial distribution $F_e$, zero assets, and a type score $s_{t+1}$ equal to $F_\beta$ for consistency. We assume $F_\beta \in \mathcal{S}$.

[12]As in the original econometric use of extreme value shocks in the discrete choice literature, $(\nu_t, \epsilon_t)$ provide a parsimonious way to capture how a type scorer may observe different choices $(d_t, a_{t+1})$ by two individuals in the same observable starting state $(\omega_t)$ due to, for instance, unobserved preference heterogeneity.

[13]Note that in the absence of hidden information regarding type, the pricing function would be independent of $a_t$ (as is the case in Chatterjee et al. (2007)) since past debts have no bearing on the repayment probability of newly incurred debt.

[14]This rule is specified in equation (28) in Appendix A of the Supplemental Material. There is nothing of substance in this randomization over contiguous elements of $\mathcal{S}$ since $\mathcal{S}$ is finely gridded when we estimate the model.

## 3. EQUILIBRIUM

### 3.1. *Individuals' Problem*

Let $x_t$ be denoted $x$ and $x_{t+1}$ be denoted $x'$. Denote the part of the state space observable to creditors by $\Omega = \{\mathcal{E} \times \mathcal{A} \times \mathcal{S}\}$ with typical element $\omega$. An individual takes as given:

- the price function $q^{a'}(\omega) : \mathcal{A} \times \Omega \to [0, 1]$,
- the type scoring functions $\psi^{(0,a')}(\omega) : \mathcal{A} \times \Omega \to [0, 1]^B$ and $\psi^{(1,0)}(\omega) : \Omega \to [0, 1]^B$ which perform Bayesian updating of an individual's type based on all observables following asset choice and bankruptcy, respectively.

For ease of notation, we will denote the triplet of functions $\{q^{a'}(\omega), \psi^{(0,a')}(\omega), \psi^{(1,0)}(\omega)\}$ by $f \in F$, where $F = \{(f_1, f_2, f_3) | f_1 : \mathcal{A} \times \Omega \to [0, 1], f_2 : \mathcal{A} \times \Omega \to [0, 1]^B$ and $f_3 : \Omega \to [0, 1]^B\}$.

DEFINITION 2: Given $(z, \omega)$ and $f \in F$, the *set of feasible actions* is a finite set $\mathcal{F}(z, \omega|f)$ that contains all actions $(d, a')$ such that consumption $c^{(d,a')}(z, \omega|f)$ is strictly positive where:

$$c^{(d,a')}(z, \omega|f) = \begin{cases} y\big(e(\omega), z\big) + a(\omega) - q^{a'}(\omega) \cdot a' & \text{if } \big(d, a'\big) = \big(0, a'\big), \\ y\big(e(\omega), z\big)(1 - \kappa_1) - \kappa & \text{if } a(\omega) < 0 \text{ and } \big(d, a'\big) = (1, 0), \end{cases} \tag{3}$$

where we use $a(\omega)$, $e(\omega)$, and $s(\omega)$ to denote the corresponding elements of $\omega$.

ASSUMPTION 1: $y(e_1, z_1) + \min\{a_1, -\kappa - \kappa_1 y(e_1, z_1)\} > 0$.

We make this assumption to ensure that it is always feasible for an indebted individual to file for bankruptcy and always feasible for her to pay back her debt.

We work backwards from an individual's state at stage 3 in timing. Given the functions $f$, for $(d, a') \in \mathcal{F}(z, \omega|f)$ we denote the *conditional value function*

$$\begin{aligned} &v^{(d,a')}(\beta, z, \omega|f) \\ &\quad = u\big(c^{(d,a')}(z, \omega|f)\big) \\ &\qquad + \beta\rho \cdot \sum_{(\beta',z',e',s')} Q^{\beta}\big(\beta'|\beta\big)Q^{e}\big(e'|e\big)H\big(z'\big)Q^{s}\big(s'|\psi^{(d,a')}(\omega)\big)W\big(\beta', z', \omega'|f\big), \end{aligned} \tag{4}$$

where the *expected value function* $W$ integrates the value function over $\nu$ and is defined below.

The value function $V^{\text{ND}}(\epsilon, \beta, z, \omega|f) : \mathbb{R}^N \times \mathcal{B} \times \mathcal{Z} \times \Omega \to \mathbb{R}$ for an individual who chooses not to file for bankruptcy at stage 2 is then given by

$$V^{\text{ND}}(\epsilon, \beta, z, \omega|f) = \max_{(0,a') \in \mathcal{F}(z,\omega|f)} v^{(0,a')}(\beta, z, \omega|f) + \epsilon^{a'}, \tag{5}$$

$$W^{\text{ND}}(\beta, z, \omega|f) = \int V^{\text{ND}}(\epsilon, \beta, z, \omega|f)\, \mathrm{d}F_{\epsilon}(\epsilon). \tag{6}$$

Given the sequential nature of choices in our timing, the value function at stage 2 is then given by

$$V(\nu,\beta,z,\omega|f) = \begin{cases} W^{\mathrm{ND}}(\beta,z,\omega|f) & \text{if } a(\omega)\geq 0,\\ \max\{v^{(1,0)}(\beta,z,\omega|f)+\nu^{D}, W^{\mathrm{ND}}(\beta,z,\omega|f)+\nu^{\mathrm{ND}}\} & \text{if } a(\omega)<0.\end{cases} \tag{7}$$

$W^{\mathrm{ND}}$ shows up because $\epsilon$ has not yet been drawn at stage 2. Finally, as promised we have

$$W(\beta,z,\omega|f) = \int V(\nu,\beta,z,\omega|f)\,\mathrm{d}F_{\nu}(\nu). \tag{8}$$

Given that $\nu$ and $\epsilon$ are drawn from type I extreme value distributions, there are simple closed form solutions for choice probabilities. Conditional on not filing for bankruptcy, let $\tilde{\sigma}^{(0,a')}(\beta,z,\omega|f)$ be the probability that the individual in state $(\beta,z,\omega)$ chooses action $a'\in\mathcal{F}(z,\omega|f)$:

$$\tilde{\sigma}^{(0,a')}(\beta,z,\omega|f) = \begin{cases} \dfrac{\exp\left\{\dfrac{v^{(0,a')}(\beta,z,\omega|f)}{\lambda}\right\}}{\displaystyle\sum_{(0,\hat{a}')\in\mathcal{F}(z,\omega|f)}\exp\left\{\dfrac{v^{(0,\hat{a}')}(\beta,z,\omega|f)}{\lambda}\right\}} & \text{for } a'\in\mathcal{F}(z,\omega|f),\\ 0 & \text{for } a'\notin\mathcal{F}(z,\omega|f).\end{cases} \tag{9}$$

Note that infeasible actions are assigned zero probability. Similarly, the probability of bankruptcy for an individual with debt ($a(\omega)<0$) is

$$\sigma^{(1,0)}(\beta,z,\omega|f) = \frac{\exp\left\{\dfrac{v^{(1,0)}(\beta,z,\omega|f)}{\alpha}\right\}}{\exp\left\{\dfrac{v^{(1,0)}(\beta,z,\omega|f)}{\alpha}\right\}+\exp\left\{\dfrac{W^{\mathrm{ND}}(\beta,z,\omega|f)}{\alpha}\right\}}. \tag{10}$$

Then, given that $\nu$ and $\epsilon$ are independent, asset choice probabilities are given by

$$\sigma^{(0,a')}(\beta,z,\omega|f) = \tilde{\sigma}^{(0,a')}(\beta,z,\omega|f)\big(1-\sigma^{(1,0)}(\beta,z,\omega|f)\big), \tag{11}$$

noting that an individual with $a(\omega)\geq 0$ has $\sigma^{(1,0)}(\beta,z,\omega|f)=0$ by definition. Furthermore, these choice probability expressions imply simple expressions for $W^{\mathrm{ND}}$ in (6) and $W$ in (8). Specifically,

$$W^{\mathrm{ND}}(\beta,z,\omega|f) = \lambda\ln\Bigg(\sum_{(0,a')\in\mathcal{F}(z,\omega|f)}\exp\left\{\frac{v^{(0,a')}(\beta,z,\omega|f)}{\lambda}\right\}\Bigg)+\lambda\gamma_E+\overline{\epsilon}(\mathcal{A}) \tag{12}$$

and

$$W(\beta,z,\omega|f) = \begin{cases} W^{\mathrm{ND}} & \text{if } a(\omega)\geq 0,\\ \alpha\ln\left(\exp\left\{\dfrac{v^{(1,0)}(\beta,z,\omega|f)}{\alpha}\right\}+\exp\left\{\dfrac{W^{\mathrm{ND}}(\beta,z,\omega|f)}{\alpha}\right\}\right) & \\ \quad +\alpha\gamma_E+\overline{\nu}(\mathcal{D}) & \text{if } a(\omega)<0,\end{cases} \tag{13}$$

where $\gamma_E$ is the Euler–Mascheroni constant.

In Appendix A.5 of the Supplemental Material, we prove the following:

THEOREM 1: *Given $f$, there exists a unique solution $W(\beta, z, \omega|f)$ to the individual's decision problem in* (3)–(8).

### 3.2. *Intermediaries' Problem*

Competitive intermediaries with deep pockets have access to an international credit market where they can borrow or lend at the risk-free interest rate $r \geq 0$. Any given intermediary takes prices $q$ and scoring function $\psi$ (i.e., $f$) as given. We assume that losses and gains resulting from individuals' deaths accrue to the financial intermediary effectively implementing an annuity contract. The profit $\pi^{a'}(\omega|f)$ on a contract of size $a'$ with agents with observables $\omega$ is

$$\pi^{a'}(\omega|f) = \begin{cases} \rho \cdot \dfrac{p^{a'}(\omega|f) \cdot (-a')}{1+r} - q^{a'}(\omega) \cdot (-a') & \text{if } a' < 0, \\ q^{a'} \cdot a' - \rho \cdot \dfrac{a'}{1+r} & \text{if } a' \geq 0, \end{cases} \tag{14}$$

where the probability of repayment on a contract of size $a'$ made to individuals with observable characteristics $\omega$ is $p^{a'}(\omega|f) : (\mathbb{R}_{--} \cap \mathcal{A}) \times \Omega \to [0, 1]$. Given perfect competition and constant returns to scale in lending, if a solution to the intermediary's problem exists, then optimization by the intermediary implies zero profits for strictly positive measures of contracts issued or

$$q^{a'}(\omega|f) = \begin{cases} \dfrac{\rho \cdot p^{a'}(\omega|f)}{1+r} & \text{if } a' < 0 \\ \dfrac{\rho}{1+r} & \text{if } a' \geq 0. \end{cases} \tag{15}$$

Assessing an individual's probability $p^{a'}(\omega|f)$ of repaying a debt next period given her current observable characteristics $\omega$ given unobservable $(\beta, \epsilon, z)$, takes two steps:

1. Assess the probability that an individual in state $\omega$ who takes action $(d, a')$ will be of unobservable type $\beta'$ next period via Bayes's rule (the type scoring function $\psi_{\beta'}^{(d,a')}(\omega)$).
2. For each possible future unobservable type $\beta'$, compute the individual's probability of future repayment conditional on being that type and transitions over observable characteristics and then compute the weighted sum over future types to obtain $p$.

Starting with step 1, an individual's probability of being type $(\beta'_1, \dots, \beta'_B)$ next period is given by the type scoring function $\psi^{(d,a')}(\omega) = (\psi_{\beta'_1}^{(d,a')}(\omega), \dots, \psi_{\beta'_B}^{(d,a')}(\omega))$, where

$$\psi_{\beta'}^{(d,a')}(\omega|f) = \begin{cases} \displaystyle\sum_{\beta} Q^{\beta}(\beta'|\beta) \cdot \frac{\sum_z \sigma^{(d,a')}(\beta, z, \omega|f) \cdot H(z) \cdot s(\beta)}{\sum_{\hat{\beta},z} \sigma^{(d,a')}(\hat{\beta}, z, \omega|f) \cdot H(z) \cdot s(\hat{\beta})} \\ \quad \text{for } (d, a') \in \mathcal{F}(z, \omega|f), \\ \displaystyle\sum_{\beta} Q^{\beta}(\beta'|\beta) \cdot s(\beta) \\ \quad \text{for } (d, a') \notin \mathcal{F}(z, \omega|f). \end{cases} \tag{16}$$

Note that in (16), the assessment uses Bayes's rule to assign the probability of an individual in state $\omega$ taking a feasible action $(d, a')$ being of type $\beta'$ next period. By (9) and (10), the probability of choosing any $(d, a') \in \mathcal{F}(z, \omega|f)$ is strictly positive for every $\beta'$. Hence, $\psi_{\beta'}^{(d,a')}(\omega|f)$ is well-defined in (16) for all feasible actions. Thus, since every feasible action is chosen with some probability due to the presence of extreme value shocks, we avoid having to assign off-the-equilibrium path beliefs for feasible actions. For completeness, without loss of generality, the bottom branch of (16) handles the case of infeasible actions. Turning to step 2, given observable state $\omega$, we obtain the probability of repayment the intermediary uses for pricing debt (i.e., for $a' < 0$) via

$$p^{a'}(\omega|f) = \sum_{\beta', z', e', s'} H(z') \cdot Q^e(e'|e) \cdot Q^s(s'|\psi^{(0,a')}(\omega|f)) \cdot s'(\beta') \cdot \left(1 - \sigma^{(1,0)}(\beta', z', e', a', s'|f)\right). \tag{17}$$

## 3.3. *Evolution*

Let $\mu(\beta, z, \omega|f)$ denote the beginning-of-period measure of individuals in state $(\beta, z, \omega)$ for a given $f$. Then, the cross-sectional distribution evolves according to

$$\mu'(\beta', z', \omega'|f) = \sum_{\beta, z, \omega} T(\beta', z', \omega'|\beta, z, \omega; f) \cdot \mu(\beta, z, \omega|f), \tag{18}$$

where the transition function is

$$\begin{aligned} T(\beta', z', \omega'; \beta, z, \omega|f) = {} & \rho \cdot Q^\beta(\beta'|\beta) \cdot H(z') \cdot Q^e(e'|e) \\ & \cdot \sigma^{(d,a')}(\beta, z, \omega|f) \cdot Q^s(s'|\psi^{(d,a')}(\omega|f)) \\ & + (1-\rho) \cdot F_\beta(\beta') \cdot H(z') \cdot F_e(e') \cdot \mathbf{1}_{\{a'=0\}} \cdot \mathbf{1}_{\{s'=F_\beta\}}. \end{aligned} \tag{19}$$

The first line in equation (19) is the probability of a survivor transitioning to $(\beta', z', e', a', s')$ while the second line is the probability that a newborn arrives in state $(\beta', z', e', a', s')$.

An invariant distribution is a fixed point $\overline{\mu}(\cdot|f) = T\overline{\mu}(\cdot|f)$. In Appendix A.3 of the Supplemental Material, we prove the following:

LEMMA 2: *There exists a unique invariant distribution $\overline{\mu}(\cdot|f)$ and $\{\mu_0 T^n\}$ converges to $\bar{\mu}(\cdot|f)$ at a geometric rate for any initial distribution $\mu_0$.*

Note that although the invariant distribution is critical for computing cross-sectional moments used to map the model to the data, none of the other equilibrium objects (i.e., the set of functions $f$, the value function $V$, or the decision rule $\sigma$) takes $\mu$ as an argument. This simplifies the model and eases the computational burden, but is not necessary. Other specifications in which knowledge of the distribution is required are possible, but we do not consider these in the baseline model.

## 3.4. *Existence*

We can now give the definition of a stationary Recursive Competitive Equilibrium.

DEFINITION 3: A stationary *Recursive Equilibrium* is a pricing function $q^*$, a type scoring function $\psi^*$, a choice probability function $\sigma^*$, and a steady-state distribution $\overline{\mu}^*$ such that:

(i) optimality: $\sigma^{(d,a')*}(\beta, z, \omega|f^*)$ satisfies (9) and (10) for all $(\beta, z, \omega) \in \mathcal{B} \times \mathcal{Z} \times \Omega$ and $(d, a') \in \mathcal{F}(z, \omega|f^*)$,

(ii) zero profits: $q^{a'*}(\omega|f^*)$ satisfies (15) with equality for all $\omega \in \Omega$ with $p^{a'*}(\omega|f^*)$ satisfying (17) for all $\omega \in \Omega$,

(iii) Bayesian updating: $\psi_{\beta'}^{(d,a')*}(\omega|f^*)$ satisfies (16) for all $(\beta', \omega) \in \mathcal{B} \times \Omega$, and

(iv) stationary distribution: $\overline{\mu}^*(\beta, z, \omega|f^*)$ solves (18) for $T(\beta', z', \omega'; \beta, z, \omega|f^*)$.

The key step in proving the existence of a Recursive Competitive Equilibrium is proving that the value function $W(\beta, z, \omega|f)$ is continuous in $f$. In Appendix A.4 of the Supplemental Material, we first prove the following:

LEMMA 3: *$W(\beta, z, \omega|f)$ is continuous in $f$, and for any $(d, a') \in \mathcal{F}(z, \omega|f)$, $\sigma^{(d,a')}(\beta, z, \omega|f)$ is continuous in $f$.*

Using this result, we then prove the following:

THEOREM 2: *There exists a stationary Recursive Equilibrium.*

### 3.5. *Equivalence to an Economy With Credit Scores*

In the economy described thus far, an individual's reputation is her type score. In U.S. credit markets, an important measure of reputation is the *credit score*. A credit score is an index that is positively related to the likelihood of repayment. The goal of this subsection is to show that under certain conditions, the equilibrium described in the previous subsections 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 given other relevant characteristics such as earnings and current assets.

In this paper, we formalize the notion that a credit score depicts a consumer's creditworthiness by defining it to be the probability of repayment on a loan of some standard size $a' = \bar{a} < 0$. According to the timing in Definition 1 part 4(a), since $\omega = (e, a, s)$ is known at the end of $t-1$, an individual's credit score can be calculated at the end of period $t-1$ to be $m = p^{\bar{a}}(e, a, s)$ using equation (17). The credit score of newborns who arrive at the end of period $t-1$ is $m = p^{\bar{a}}(e_1, 0, F_\beta)$.

In the financial arrangement with credit scores, an individual in (observable) state $\hat{\omega} = (e, a, m)$ takes as given a pricing function $q^{a'}(\hat{\omega})$ and a credit-score transition function $Q_m^{(d,a')}(m'|e', \hat{\omega})$ which tells her the probability distribution of her future credit score conditional on her current observable state, current actions, and future earnings. Intermediaries take as given the pricing function (which must satisfy the zero profit condition) and the probability of repayment function $p^{a'}(\hat{\omega})$ (which must be consistent with the individual's objective likelihood of repayment). In Appendix A.5 of the Supplemental Material, we restate the household and financial intermediary problems for this financial arrangement and provide a definition of a Recursive Equilibrium with Credit Scores.

An equivalence between the type-scoring and credit-scoring environments will exist if there is a one-to-one and onto mapping between $s$ and $m$, holding fixed the other factors that affect credit scores, namely, $e$ and $a$. Then, wherever $s$ appears in the theoretical model, it can be replaced by $m$. Thus, the equivalence will hold if the inverse function

$(p^{\bar{a}*})^{-1}(e, a, m)$ exists.[15] Now note that since $\mathcal{S}$ is a finite collection of grid points, the occurrence of distinct grid points in $\mathcal{S}$ mapping to precisely the same probability of repayment on $\bar{a}$, given $e$ and $a$, will be purely coincidental.[16] Thus, barring coincidences, the mapping $m = p^{\bar{a}*}(e, a, s)$ will be one-to-one, and it can be made onto by restricting the range of $p$ to contain only those $m$ that are implied by some $s \in \mathcal{S}$, given $e$ and $a$. In other words, regardless of the number of types, the finite support of $s$ can be used to encode both an individual's type score and her probability of repayment on $\bar{a}$. In our application, we verify that the one-to-one property holds and $(p^{\bar{a}})^{-1}(e, a, m)$ exists.

In Appendix A.5 of the Supplemental Material, we prove the following:

THEOREM 3: *Given a Recursive Equilibrium, let $m = p^{\bar{a}*}(e, a, s)$. Suppose that the inverse function $s = (p^{\bar{a}*})^{-1}(e, a, m)$ exists. Then a Recursive Equilibrium with Credit Scores exists in which the choice probabilities $\sigma^{(d,a')*}(\beta, z, e, a, m) = \sigma^{(d,a')*}(\beta, z, e, a, s)$ for $s = (p^{\bar{a}*})^{-1}(e, a, m)$.*

## 4. MAPPING THE MODEL TO DATA

We now examine the U.S. unsecured credit market through the lens of our model. We rely on the equivalence result between the model with type scores and the model with credit scores described in Theorem 3 when there are two $\beta$ types ($\beta_H > \beta_L$). Specifically, it allows us to use the model with type scores in order to target the joint behavior of earnings, aggregate credit market moments, and credit rankings over their working age. We then verify that the sufficient condition in Theorem 3 is satisfied for the estimated parameters so that the equivalence result holds.

In the data, a credit score is an *ordinal* measure of creditworthiness, typically ranging from around 300 to 850, not a direct estimate of the probability of repayment $m$.[17] To close this gap, we associate with $m = p^{\bar{a}}(\omega)$ a number in the unit interval that gives $p^{\bar{a}}(\omega)$'s position (i.e., ranking) in the overall distribution of $p^{\bar{a}}(\omega)$ in the model economy.

DEFINITION 4: An individual's *credit ranking* in state $\omega$ is given by

$$\chi^{\bar{a}}(\omega) = \sum_{\tilde{\omega} \in J^{\bar{a}}(\omega)} \mu(\tilde{\omega}), \tag{20}$$

where $J^{\bar{a}}(\omega) = \{\tilde{\omega} : p^{\bar{a}}(\tilde{\omega}) \leq p^{\bar{a}}(\omega)\}$ and $\mu(\omega) = \sum_{\beta,z} \mu(\beta, z, \omega)$.

Clearly, $\chi^{\bar{a}}(\omega) \in [0, 1]$. We construct the data analogue of $\chi^{\bar{a}}(\omega)$ by associating with each credit score its percentile position in the overall distribution of credit scores.[18] Simply

[15]If the relationship between $s$ and $m$ is not one-to-one, then two individuals with the same $e$, $a$, and $m$ choosing the same level of debt could face different prices in the economy with type scores because their $s$'s are different.

[16]To see why, suppose that there are three types of individuals and let $s = \{s_1, s_2, s_3\}$ be a specific type score. For concreteness, assume that, in the same circumstances, type 1's probability of repayment is greater than type 2's, and type 2's is greater than type 3's (i.e., $\sigma^{(1,0)}(\beta_1, z, e, \bar{a}, s|f) < \sigma^{(1,0)}(\beta_2, z, e, \bar{a}, s|f) < \sigma^{(1,0)}(\beta_3, z, e, \bar{a}, s|f)$). If we now consider another $\hat{s}$ where the value of $\hat{s}_1$ is higher than $s_1$, the probability of repayment on $\bar{a}$ can remain unchanged if $\hat{s}_2$ is lower than $s_2$ by some specific amount and $\hat{s}_3$ is higher than $s_3$ by some specific amount. However, it will be a pure coincidence if the exact combination $(\hat{s}_1, \hat{s}_2, \hat{s}_3)$ is an element of $\mathcal{S}$.

[17]See, for example, https://www.investopedia.com/terms/c/credit_score.asp.

[18]Thus, while our theory is in terms of type scores which we map to credit scores in [0, 1] via Theorem 3, since real-world credit scores have no interpretation in the model, we convert them to something interpretable

put, after computing credit scores for each individual in the economy, we then line them up and associate each individual with its rank in the credit-score distribution.

Furthermore, real-world credit scores do not mention any specific level of borrowing $\bar{a}$. One way to interpret this fact is to think that the ranking of individuals with respect to probability of repayment holds for *any* level of debt. For this to be true, we need the following property:

DEFINITION 5: Let $\hat{a} < \bar{a} < 0$. Then, $p^{a'}(\omega)$ *preserves order with respect to* $a'$ if $p^{\bar{a}}(\omega) \geq p^{\bar{a}}(\tilde{\omega})$ if and only if $p^{\hat{a}}(\omega) \geq p^{\hat{a}}(\tilde{\omega})$.

If $p^{a'}(\omega)$ preserves order, then $J^{\hat{a}}(\omega) = J^{\bar{a}}(\omega)$ and $\chi^{\bar{a}}(\omega)$ becomes invariant to the choice of $\bar{a}$. This order preserving property holds for a wide range of debt levels for our estimated model in Section 4.

Credit rankings, earnings, and assets all grow with age on average, and we want our model to capture those features. Unfortunately, we do not have access to a panel data set which contains all these dimensions. So we use a version of simulated method of moments to estimate our model. Specifically, we take some non-controversial information from outside the model: the earnings process, the risk-free rate of return, demographics, preferences over risk, a measurement of the costs of bankruptcy filings, and a generic value of debt ($\overline{a}$) to which the credit score is normalized.[19]

Next, we obtain a set of data moments that summarize the properties of the unsecured credit market (bankruptcy filing rates, average interest rates, dispersion of interest ratios, fraction of households in debt, debt-to-income ratio) and we approximate the behavior of credit scores as a function of age (specifically, affine functions of the mean and the standard deviation of credit scores and the autocorrelation of the annual change in individual scores). One can interpret age as the length of an individual's credit history; agents are "born" with no credit history and the length of their credit history grows with age.

We then proceed to estimate the parameters of interest, which are the values of patience for both types, the transition probabilities of types and their frequency at birth, the proportional earnings loss from bankruptcy, as well as measures of noise (the variances of the extreme value shocks) by minimizing the weighted sum of squared differences between the values of the moments in the data and their model counterparts. We have tried various alternative sets of moments with minimal effects on the findings. While earnings, credit, and bankruptcy statistics have been used since Chatterjee et al. (2007) and Livshits, MacGee, and Tertilt (2007), credit scores, and their evolution by age, have not. The evolution of credit scores is crucial for understanding the building of a reputation over the early part of the life-cycle.

## 4.1. *Computation*

Computation of equilibrium requires solving for two endogenous functions: the bond price function and the type-score updating function. The bond price function is standard in unsecured debt models like Chatterjee et al. (2007), except that the endogenous type score is an additional dimension. The type scoring function is new: individuals take as given how feasible actions change the market's perception of their type that is updated

in both the data and the model: credit rankings. Hence we use all three concepts: (1) $s$; (2) $m = p^{\bar{a}}(e, a, s)$; (3) $\chi^{\bar{a}}(\omega) = \sum_{\tilde{\omega} \in J^{\bar{a}}(\omega)} \mu(\tilde{\omega})$.

[19]We verify that the choice of $\overline{a}$ does not matter provided it is higher than the bankruptcy filing costs.

using Bayes's law and this perception has to be consistent with the actions taken by both types.

While we introduced extreme value shocks in order to keep the Bayesian posterior well behaved, it can, however, exacerbate grid sensitivity associated with approximating continuous choices.[20] We deal with this by making the location parameters of the extreme value shock associated with each asset choice depend positively on the measure of consumption points that are associated with that action in the individual's feasible set $\mathcal{F}(z, \omega|f)$.[21] We describe adjustments to mitigate grid sensitivity after introducing the functional forms of the shock distributions in the next section.

## 4.2. *Estimation*

We use a minimum distance estimator to parameterize the model. We discuss which ex ante restrictions we specify (Section 4.2.1), the targets that we use and the data from which they come (Section 4.2.2), the estimation strategy (Section 4.2.3), the estimates (Section 4.2.4), and we finish with a discussion of the robustness of our estimates (Section 4.2.5).

### 4.2.1. *Functional Forms and Parameters Chosen Outside the Model*

A model period is one year. We take the relevant working life span of people to be 40 years, as the bulk of borrowing is by young people, implying a working age survival probability of 0.975. We choose a CRRA utility function with risk aversion parameter 1.5. We pose a risk-free rate of 1%, which implies an effective interest rate on savings of 3.59% in the presence of perfect annuity markets. We take the cost of filing for bankruptcy to be about 1.5% of median earnings taken from Albanesi and Nosal (2018).[22] Since it is a dominant action not to invoke bankruptcy on debts less than the filing cost, we choose $\bar{a}$ (the debt value used to compute the probability of repayment for a credit score) to be 3.5% of median earnings (i.e., well above those costs). Finally, we take the earnings class to be the persistent AR1 process estimated by Floden and Lindé (2001) and assume agents are born with the lowest earnings level to replicate the upward earnings path during one's working age.[23] These parameters chosen outside the model are summarized in Table I.

[20]As an example, imagine that we are approximating the interval [0, 2] with a discrete grid that is log-spaced. This means that there are more grid points in [0, 1] than in [1, 2]. Now assume that the value of the action associated with any grid point $i$ is just $v + \epsilon^i$. Then, any one of these grid points has an equal chance of being selected. But, since there are more points in [0, 1], it is more likely that the choice will be from that interval. In the context of our model, this effect imparts a bias toward actions close to the origin (debt or small levels of assets).

[21]For instance, in the example of Footnote 20, the adjustment to the location parameter of the shocks lowers the mean of the extreme value shocks associated with closely-packed choices. The result is that, with the adjustment, it is equally likely that the best choice is in [0, 1] or [1, 2]. This adjustment has implications for savings behavior explored in Briglia, Chatterjee, Corbae, Dempsey, and Rios-Rull (2021). See also Appendix B.4 of the Supplemental Material.

[22]Albanesi and Nosal (2018) reported a filing fee of \$697 in 2005 pre-BAPCA. Median household income in 2004, adjusted for 3.39% inflation between 2004 and 2005, was \$45,837; the ratio of these numbers yields 1.52%.

[23]Recalling total earnings is given by $y_t(e_t, z_t) = \exp(e_t + z_t)$, we approximate the AR1 process by a five-state Markov chain using the Adda and Cooper (2003) method, which yields support $\mathcal{E} = \{-0.71, -0.27, 0.00, 0.27, 0.71\}$, transition matrix $Q^e(e'|e) = \begin{bmatrix} 0.767 & 0.207 & 0.025 & 0.001 & 0.000 \\ 0.207 & 0.496 & 0.253 & 0.043 & 0.001 \\ 0.025 & 0.253 & 0.446 & 0.253 & 0.025 \\ 0.001 & 0.043 & 0.253 & 0.496 & 0.207 \\ 0.000 & 0.001 & 0.025 & 0.207 & 0.767 \end{bmatrix}$, and a transitory component with a three-point uniform distribution on support $\mathcal{Z} = \{-0.25, 0, 0.25\}$.

TABLE I

PARAMETERS CHOSEN OUTSIDE THE MODEL.

| Parameter | | Value | Notes |
|---|---|---|---|
| | Demographics and preferences | | |
| Survival probability | $\rho$ | 0.975 | avg. life span 40 years |
| Risk aversion | $\gamma$ | 1.5 | CRRA preferences |
| Earnings at birth | $\underline{e}$ | −0.71 | See Footnote 23 |
| | Technology | | |
| Risk-free rate (%) | $r$ | 1.000 | |
| Bankruptcy filing cost | $\kappa$ | 0.0152 | 1.5% of median earnings |
| Debt level for computing credit score | $\bar{a}$ | −0.035 | 2.9% of median earnings |
| | Earnings | | |
| Persistence of $\log(e)$ | $\rho_e$ | 0.9136 | Floden and Lindé (2001) |
| Variance of innovations to $\log(e)$ | $\nu_e^2$ | 0.0426 | Floden and Lindé (2001) |
| Variance of $\log(z)$ | $\nu_z^2$ | 0.0421 | Floden and Lindé (2001) |

We parameterize the cumulative distribution function of the type I extreme value $\nu$ shocks associated with the default choice as

$$F_\nu(\nu^d; \alpha) = \exp\left\{-\exp\left(-\frac{\nu^d - \overline{\nu}}{\alpha}\right)\right\} \quad \text{for } d \in \{0, 1\}. \tag{21}$$

Given $\alpha$, we choose the location parameter $\overline{\nu}$ to eliminate the incentive for a household to choose debt simply in order to obtain favorable draws of the extreme value shock associated with the bankruptcy decision.[24] We parameterize the cumulative distribution function of the $\epsilon$ shocks associated with asset choices as

$$F_\epsilon(\epsilon^{a_n}; \lambda) = \exp\left\{-\exp\left(-\frac{\epsilon^{a_n} - \overline{\epsilon}^{a_n}(z, \omega)}{\lambda}\right)\right\} \quad \text{for } n \in \{1, \dots, \overline{n}(z, \omega)\}, \tag{22}$$

where $\overline{n}(z, \omega)$ is the index of the largest budget feasible $a'$ for an agent with $(z, \omega)$. As discussed in Section 4.1, given $\lambda$, we specify choice and state-specific means for the $\epsilon$ shocks $\overline{\epsilon}^{a_n}(z, \omega)$ to mitigate grid sensitivity.[25] Estimates of the scale parameters $\alpha$ and $\lambda$ are discussed in Section 4.2.4.

### 4.2.2. *Data and Targets*

The set of statistics that we deem important to target pertain to the main aggregate characteristics of the U.S. unsecured credit market: credit usage (the fraction of households in net debt and the debt-to-income ratio, credit terms (average interest rates and

[24] We set $\bar{\nu} = -\alpha \cdot (\gamma_E + \ln(|\mathcal{D}|))$, where $\gamma_E$ is the Euler–Mascheroni constant in equation (35) in Appendix B.1 of the Supplemental Material so that $\mathbb{E}[\max\{\nu^D, \nu^{\text{ND}}\}] = 0$. This correction implies that for an indebted household $a(\omega) < 0$ for whom $v^{(1,0)} = W^{\text{ND}}$ in (7), the ex ante value $W$ in (8) is equal to $v^{(1,0)} = W^{\text{ND}}$. In other words, the presence of the default/no-default shocks do not add any extra utility in expectation.

[25] Specifically, we set $\overline{\epsilon}^{a_n}(z, \omega | f_j) = -\lambda \cdot (\gamma_E - \ln(\eta^{a_n}(z, \omega | f_j)))$ in equation (37) of Appendix B.1 of the Supplemental Material, where $\eta^{a_n}(z, \omega | f_j)$ is the measure of consumption in an agent's budget set accounted for by a given asset choice $a_n$. This maps our exogenous discrete grid over $a'$ into consumption weights that help correct distortions to individual decision making when adding arbitrary points to the $a'$ grid. In particular, the correction down-weights choices on dense portions of the grid.

their dispersion), and the bankruptcy rate. Importantly, we are also interested in matching properties of the age profile of credit rankings. In particular, we match the intercept and slope of the mean and standard deviation of credit rankings across the working age profile, as well as the mean of the autocorrelation of credit ranking changes.[26] This amounts to using 10 moments as targets.

To obtain these data targets, we use three primary sources: the Survey of Consumer Finances (SCF), the administrative records of the U.S. Bankruptcy Courts, and the Federal Reserve Bank of New York Consumer Credit Panel/Equifax (FRBNY CCP/Equifax). The first provides information on individual-level variation in debt and interest rates, the second provides information on aggregate bankruptcy filing rates, and the last contains individual-level information on credit records from an anonymized panel which provides us with moments on the variation and evolution of credit scores. The credit-score measure is the Equifax Risk Score (hereafter Risk Score), which is a proprietary credit score similar to other risk scores used in the industry.

We choose 2004 as our baseline year. This is because this is the latest year for which the bankruptcy filing statistics are unaffected by the Bankruptcy Abuse Prevention and Consumer Protection Act of 2005 (which changed the eligibility requirements for a discharge in ways we do not model in this paper). To align with this choice, we use the 2004 SCF for our credit market moments and the data for 2004 from the CCP for the age profile of credit rankings. For the autocorrelation of year-to-year changes in credit rankings, we use CCP data from 2003, 2004, and 2005.

In the SCF, we focus on the subset of households with heads between the ages of 20 and 60 years excluding the top 5% of the wealth distribution, for whom we think our theory is not relevant. The fraction of indebted households is the fraction of such households with negative net worth. The average debt-to-income ratio is the ratio of total unsecured debt of indebted households to 2004 per-household U.S. GDP. For the mean and standard deviation of interest rates, we used the interest rates reported on unsecured debt by all households with negative net worth.

The bankruptcy rate is the ratio of the total number of nonbusiness Chapter 7 filings in 2004 reported by the U.S. Bankruptcy Courts, scaled by the total number of U.S. households in 2004.

While the previous credit market data targets have been used in numerous quantitative theory papers on bankruptcy, what is novel is our use of the age profile of Risk Score moments. 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 Supplemental Material). With these 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. With these age bin data values, we estimate affine age profiles for means, standard deviations. To compute autocorrelations of year-to-year changes in credit rankings, we create credit rankings for 2003 and 2005 for each individual. We place individuals in the same 5-year age bins and compute the correlation between the change in rankings between 2003 and 2004 and between 2004 and 2005 for each age bin.[27] Figure 1 shows the data and affine approximations to the data as well as the model-generated data and approximation.

[26]The age profile of the autocorrelation of credit ranking changes showed no significant slope in the data, so we only target the intercept.

[27]We do not use the slope of the age profile of the autocorrelation of year-to-year credit-score changes as a target because it is zero, which makes matching the relative deviation between model and data a problem.

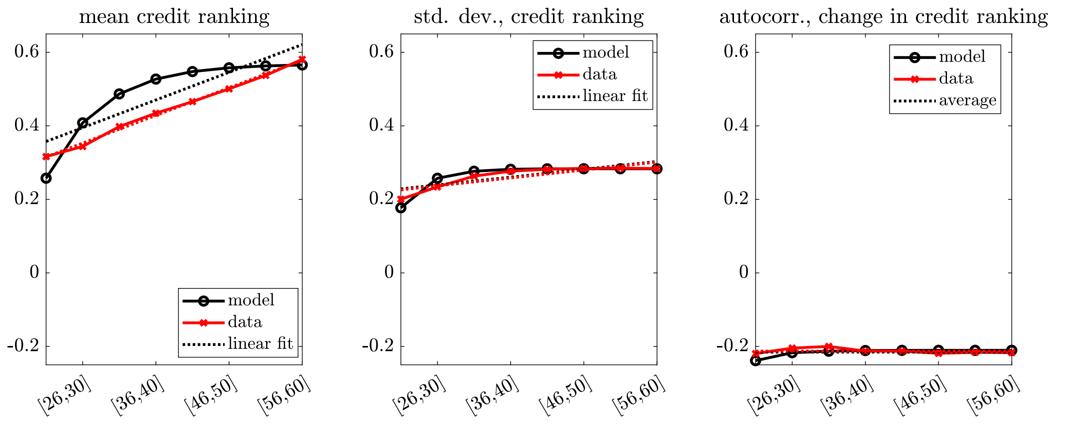


FIGURE 1.—Credit ranking age profile: model versus data. *Notes*: The credit ranking data are based on author calculations using FRBNY CCP/Equifax data, whose construction is detailed in Appendix C of the Supplemental Material. The linear approximation to the model-generated credit ranking age profiles uses the regression coefficients from Table II.

### 4.2.3. *Estimation Strategy*

To estimate the parameters, we proceed with a mixture of simulated method of moments for the aggregate statistics and indirect inference for the affine fit (intercept and slope coefficients) of the age profile of credit-score means, standard deviations, and autocorrelations. Our system is overidentified (we have 8 parameters $\theta = (\beta_H, \beta_L, Q^\beta(L'|H), Q^\beta(H'|L), F_{\beta_H}, \kappa_1, \alpha, \lambda)$ and 10 moments (five from the credit market and five of the age profile of credit rankings) so not all moments will be exactly replicated. Specifically, the consistent estimated parameter values in Table III solve

$$\widehat{\theta} = \arg\min_{\theta} \widehat{g}'(\theta)\widehat{W}\widehat{g}(\theta), \tag{23}$$

where $\widehat{g}(\theta) = (\widehat{S} - S(\theta))$ is the (percentage) difference between data $\widehat{S}$ and model $S(\theta)$ moments and $\widehat{W}$ is a weighting matrix. In our base estimation, we use an identity weighting matrix which we discuss further in Section 4.2.5.[28]

### 4.2.4. *Parameter Estimates*

The values of the data moments and their model counterparts, as well as the average mean squared errors (both total and for each of the two blocks of moments), are reported in Table II and Figure 1.

The estimated values for patience are $\beta \in \{0.809, 0.930\}$, so that low types have a 13% lower discount factor than high types.[29] This differential allows reputation acquisition to play a role in equilibrium: types are far enough apart to want to behave differently, but close enough that imitation is not too costly when individuals start with low earnings and zero assets, that is, when they are young. This is explained in more detail in Section 5.3.

The estimated transition matrix $Q^\beta$ implies an ergodic distribution featuring 48% high types, but the life-cycle demographic structure implies a slightly lower stationary fraction

[28]For more details on computation and estimation, see Appendix B.1 of the Supplemental Material.
[29]Our annual estimates of discount factors translate into quarterly values $\beta_H = 0.982$ and $\beta_L = 0.948$.

TABLE II

ESTIMATION TARGETS.

| Moment (%) | | Data | Model |
|---|---|---|---|
| Aggregate credit market moments | | | |
| Bankruptcy rate | BR | 1.00 | 1.02 |
| Average interest rate | AI | 11.9 | 11.5 |
| Interest rate dispersion | SDI | 7.00 | 7.08 |
| Fraction of HH in debt | FID | 7.92 | 9.16 |
| Debt-to-income ratio | DTY | 0.40 | 0.26 |
| Credit ranking age profile moments | | | |
| Intercept, mean credit ranking | I:MCR | 0.278 | 0.320 |
| Slope, mean credit ranking | S:MCR | 0.038 | 0.038 |
| Intercept, std. dev. credit ranking | I:SDCR | 0.215 | 0.218 |
| Slope, std. dev. credit ranking | S:SDCR | 0.011 | 0.011 |
| Average autocorrelation of credit ranking changes | AUTO | −0.220 | −0.215 |
| Sum of squared errors | | | |
| Aggregate credit market moments | | | 0.144 |
| Credit ranking age profile moments | | | 0.023 |
| Total | | | 0.167 |

*Note*: The credit ranking data are based on author calculations using FRBNY CCP/Equifax data. The sum of squared errors is computed in percentage deviation terms to control for relative magnitudes of moments, each receiving equal weight.

equal to 47%.[30] There is demographic improvement in the average assessment of an individual's type (and hence creditworthiness) over one's lifetime: agents' initial type scores are consistent with the estimated initial fraction of high types, $F_\beta = 32\%$, with the average assessment updating via $\overline{s}' = \overline{s} \cdot Q^\beta(H|H) + (1-\overline{s}) \cdot Q^\beta(H|L)$ which converges to 48%. This is independent of alternative credit arrangements we consider. The fact that $F_\beta$ is

TABLE III

PARAMETERS CHOSEN WITHIN THE MODEL.

| Parameter | | Value |
|---|---|---|
| Evolution of types | | |
| High discount factor | $\beta_H$ | 0.930 |
| Low discount factor | $\beta_L$ | 0.809 |
| High to low $\beta$ transition | $Q^\beta(L'\|H)$ | 0.226 |
| Low to high $\beta$ transition | $Q^\beta(H'\|L)$ | 0.205 |
| Fraction high $\beta$ at birth | $G_{\beta_H}$ | 0.318 |
| Proportional default cost | $\kappa_1$ | 0.067 |
| Extreme value parameters | | |
| Scale in $F_\nu(\nu; \alpha)$ | $\alpha$ | 0.029 |
| Scale in $F_\epsilon(\epsilon; \lambda)$ | $\lambda$ | 0.002 |

[30]The fraction of type $H$ in the stationary distribution (call it $\mu_H$) solves $\mu_H = \rho \cdot [(1 - Q^\beta(L'|H))\mu_H + Q^\beta(H'|L)(1-\mu_H)] + (1-\rho) \cdot F_{\beta_H}$ or $\mu_H = \frac{\rho Q^\beta(H'|L)+(1-\rho)F_{\beta_H}}{1-\rho(1-Q^\beta(L'|H))+\rho Q^\beta(H'|L)}$. For our estimated parameter values in Table III, $\mu_H = 0.47$. In a cohort, the fraction of type $H$ asymptotes to 0.48 (the value of $\mu_H$ corresponding to $\rho = 1$).

estimated to be below the stationary fraction of high types is a robust consequence of the rising average credit score over the age profile.

The estimated earnings loss from filing for bankruptcy is 6.7 percent of the persistent component of earnings. This relatively large estimated cost indicates that there are costs to bankruptcy that are not captured by the loss of reputation in unsecured credit markets only. For example, reputation loss can impact one's job finding prospects, secured borrowing costs like mortgages, and even insurance premia. For reasons of parsimony, these other channels are captured by our estimate of $\kappa_1$.

Our estimates of the parameters $\alpha$ and $\lambda$ of the extreme value distributions imply that there is more noise in the bankruptcy decision than in asset choices, but not so much that fundamentals are overridden (i.e., fundamental heterogeneity in unobservable type and earnings are the key drivers of default and asset choice). One measure of the size of the extreme value shocks is how noisy consumption decisions are; at an individual level, the variance of consumption decisions, conditional on state $(\beta, z, \omega)$, is zero without the extreme value shocks. In our model, the average coefficient of variation of consumption across all agents is only modestly higher at 2.03%.[31] This is especially true for the $\epsilon$ shock process associated with the asset choice decision where the share of total borrowing and saving actions by "modal agents" (i.e., those for whom an action in the set under consideration is the mode) is 85.6% and 99.9%, respectively.[32] On the other hand, the default/no-default action is associated with more variability in the shock process $\nu$ where the share of modal defaulters is only 5.25%. This can be explained by the fact that our parsimonious $\nu$ shock process is capturing other unobservable factors behind the default decision not included in our model (e.g., Chatterjee et al. (2007) included other shocks to capture medical expenses and lawsuits which survey respondents cited as reasons for filing for bankruptcy).

### 4.2.5. *Sensitivity Analysis*

In lieu of standard errors, we provide a measure of the sensitivity of our parameter estimates to the moments of the data using the local methods in Andrews, Gentzkow, and Shapiro (2017). Table IV presents a version of their sensitivity measure $\Lambda$ applied to classical minimum distance estimation:

$$\Lambda = -\left(G'WG\right)^{-1}G'W, \tag{24}$$

where $G = \mathbb{E}[\nabla_\theta \widehat{g}(\theta)]$ is the $10 \times 8$ probability limit of the Jacobian and $W$ is the probability limit of the weighting matrix, which we have simply taken to be the identity matrix. $\Lambda$ measures how sensitive the parameter estimates in Table III are to local perturbations of the data moments. Further, there is a tight connection between $\Lambda$ and standard errors in GMM/SMM. Specifically, given (24), the limiting distribution of the estimates can be written

$$\sqrt{T}(\widehat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}\left[0, \Lambda\Omega\Lambda'\right], \tag{25}$$

where $\Omega = \mathbb{E}[g(\theta)g(\theta)']$ is the limiting variance-covariance matrix of the data moments, $\theta_0$ is the true parameter value, and $T$ is sample size. For a given $\Omega$, (25) makes clear

[31]At each point in the state space, we compute the standard deviation and mean of consumption implied by the decision rule $\sigma$. We then take the ratio of these numbers at each point and average over the stationary distribution.

[32]See Appendix B.5 of the Supplemental Material for a description of these calculations (in particular equation (41)).

TABLE IV

SENSITIVITY ANALYSIS: IMPLIED PERCENTAGE CHANGE IN PARAMETER GIVEN 1% CHANGE IN EMPIRICAL MOMENT.

| | BR | AI | SDI | FID | DTY | I:MCR | S:MCR | I:SDCR | S:SDCR | AUTO |
|---|---|---|---|---|---|---|---|---|---|---|
| $\beta_H$ | 0.00 | −0.08 | −0.06 | −0.08 | 0.09 | −0.12 | −0.12 | −0.11 | −0.04 | 0.02 |
| $\beta_L$ | 0.34 | −0.11 | 0.02 | −0.04 | −0.21 | 0.51 | 0.44 | 0.47 | 0.13 | −0.02 |
| $Q^\beta(L'\|H)$ | −0.47 | 0.93 | 0.36 | 2.33 | −2.67 | 23.42 | 14.19 | 19.71 | 4.29 | 0.65 |
| $Q^\beta(H'\|L)$ | −5.83 | 6.06 | 0.54 | 3.12 | 2.44 | 5.57 | −0.03 | 2.60 | 0.00 | 0.71 |
| $G_{\beta_H}$ | −2.49 | 2.45 | 0.06 | 1.24 | 1.39 | −10.66 | −0.69 | 7.17 | −0.29 | 0.39 |
| $\kappa_1$ | 0.17 | −0.19 | −0.03 | −0.06 | 0.01 | −0.64 | −0.43 | −0.88 | −0.13 | 0.01 |
| $\alpha$ | −0.35 | −0.22 | −0.03 | 0.48 | 0.03 | −0.59 | −0.43 | −0.77 | −0.13 | 0.02 |
| $\lambda$ | 7.22 | −6.67 | −0.22 | −7.07 | 2.02 | −3.32 | −3.46 | −5.06 | −1.03 | 0.08 |

*Note*: Each entry corresponds to the implied percentage change in the estimated parameter in the row associated with a 1% change in the indicated empirical moment in the column. Abbreviations for moments are available in Table II. All numbers are reported in percentage points, that is, the 0.00 (which when expanded is actually 0.0011%) in the top left cell implies that if the bankruptcy rate were 1% higher, our estimate of $\beta_H$ would increase by 0.0011%, from 0.930 to 0.931.
BR: bankruptcy rate; AI: interest rate average; SDI: interest rate dispersion; FID: fraction of households in net debt; DTY: debt-to-income ratio; I:MCR, S:MCR: intercept and slope of the mean of credit rankings across the age profile; I:SDCR and S:SDCR: intercept and slope of the standard deviation of credit rankings across the age profile; AUTO: mean of the autocorrelation of credit ranking changes.

that small values of $\Lambda$ are associated with more precise parameter estimates. Since there are several moments where there is effectively no sample variation in the data (i.e., the 2004 bankruptcy rate is the population moment and the credit ranking statistics from the 2004 FRBNY CCP/Equifax essentially comprise the population moment due to its large sample size) and we are drawing from very different data sources, some elements (e.g., zeros) of the estimate of $\Omega$ are hard to interpret. Hence, in lieu of standard errors in Table III, we instead focus on $\Lambda$ in Table IV.[33]

Looking across the rows of Table IV, there are several parameters ($\beta_H$, $\beta_L$, $\kappa_1$, and $\alpha$) which appear to be relatively insensitive to a 1% change in our moments.[34] In contrast, parameters like $Q^\beta(L'|H)$, $Q^\beta(H'|L)$, and $F_{\beta_H}$ appear to be sensitive to moments like the intercept of the age profile of mean credit rankings (I:MCR). This suggests that if we had parameterized the age profile of credit rankings differently (say with a quadratic instead of affine function), the estimate of these parameters might be affected. The fact that these parameters are all jointly sensitive to the intercept of the mean credit ranking profile is related to the fact that the affine function (intercept and slope) depends on the transition of high types across age bins which depends explicitly on all those parameters (see Footnote 30). There also appears to be sensitivity in the estimate of the variance of extreme value shocks associated with asset choices with respect to several of the moments, but this may be related to our very small estimate of $\lambda$ in Table III.

### 4.2.6. *Jacobian Matrix*

The transformation of the estimated sensitivity matrix $\hat{\Lambda}$ presented in Table IV is useful for thinking about which moments of the data drive our parameter estimates. The estimated Jacobian matrix $\hat{G}$ in Table V is an essential input into that analysis (per (24)),

[33] Appendix B.3 of the Supplemental Material provides a detailed explanation of our implementation of Andrews, Gentzkow, and Shapiro (2017).

[34] Our sensitivity numbers are not invariant to the scaling of the parameters. The most natural scaling would be with respect to $\Omega$, but we do not pursue this here given the issues with estimation of $\Omega$ discussed above.

TABLE V

JACOBIAN ANALYSIS: (NUMERICAL) DERIVATIVE OF TARGET MOMENTS WITH RESPECT TO ESTIMATED PARAMETERS, $\hat{G}'$.

| | BR | AI | SDI | FID | DTY | I:MCR | S:MCR | I:SDCR | S:SDCR | AUTO |
|---|---|---|---|---|---|---|---|---|---|---|
| $\beta_H$ | −3.03 | −19.6 | 151 | −20.8 | −0.65 | −0.33 | 0.07 | −0.33 | 0.05 | 1.02 |
| $\beta_L$ | −8.80 | −30.9 | 59.7 | −64.1 | −2.05 | −0.27 | 0.05 | −0.32 | 0.05 | 2.09 |
| $Q^\beta(L'\|H)$ | 0.81 | 4.29 | −6.47 | 5.79 | 0.18 | 0.06 | −0.02 | 0.07 | −0.02 | −0.71 |
| $Q^\beta(H'\|L)$ | −1.11 | −6.24 | −1.53 | −7.59 | −0.24 | −0.10 | 0.02 | −0.07 | 0.02 | −0.22 |
| $G_{\beta_H}$ | −0.05 | 0.11 | −2.49 | −0.24 | −0.01 | 0.06 | −0.01 | −0.02 | 0.00 | −0.28 |
| $\kappa_1$ | 7.90 | −209 | −195 | 215 | 10.1 | 5.74 | −1.15 | 7.83 | −1.35 | −58.6 |
| $\alpha$ | −24.2 | 536 | −104 | −628 | −20.6 | −2.66 | 0.54 | −3.31 | 0.54 | 19.0 |
| $\lambda$ | −15.5 | 846 | −1802 | 8.38 | −19.5 | −5.07 | 1.09 | −6.66 | 1.36 | −25.5 |

*Note*: Each entry is the numerical derivative of the target moment (column) with respect to a 0.1% change in the corresponding parameter (row). BR: bankruptcy rate; AI: interest rate average; SDI: interest rate dispersion; FID: fraction of households in net debt; DTY: debt-to-income ratio; I:MCR, S:MCR: intercept and slope of the mean of credit rankings across the age profile; I:SDCR and S:SDCR: intercept and slope of the standard deviation of credit rankings across the age profile; AUTO: mean of the autocorrelation of credit ranking changes.

while also containing useful information on its own for understanding how model parameters drive model moments.

We start by thinking about implications for the aggregate credit market moments (five leftmost columns of the table). If either the high or low $\beta$ type becomes more patient, there is less borrowing (both the fraction in default and the debt-to-income ratio drop), which leads to less bankruptcy and lower average interest rates. If the transition from type $H$ to $L$ ($L$ to $H$) rises, leading to more (less) $L$ types on average, there is more (less) borrowing, higher (lower) interest rates, and higher (lower) bankruptcy rates. A higher initial fraction of high types yields less borrowing and fewer bankruptcies. Finally, a higher variable filing cost (i.e., a higher punishment to filing for bankruptcy) yields much lower interest rates, inducing more borrowing and ultimately leading to more bankruptcy in equilibrium.

Increasing either $\alpha$ or $\lambda$ increases the noisiness of the associated decision: bankruptcy versus no bankruptcy and the choice of $a'$, respectively. Increasing either parameter clouds lenders' ability to infer types based on actions. Consider bankruptcy ($\alpha$) first. Since bankruptcy is generally not optimal for most borrowers, raising $\alpha$ increases the rate at which borrowers file, all else equal. In equilibrium, though, this leads to a surge in interest rates across all loans and a sharp decline in both the share of borrowers and the overall amount of borrowing, ultimately lowering the bankruptcy rate. Next, consider borrowing and saving ($\lambda$). Raising $\lambda$ increases the likelihood with which households will deviate from their "optimal choices" or, put differently, be "off their Euler equations." The first-order effect of this is that households are much more willing to take on small debts which carry non-trivial interest rates due to default risk, driving up the fraction in debt and the average interest rate. At the same time, though, since debt to income falls, the total volume of debt decreases, lowering the bankruptcy rate in equilibrium.

Finally, consider the credit ranking life-cycle moments (five rightmost columns of the table). In general, these moments are less sensitive to our target parameters than the aggregate credit market moments because: (i) a sizable portion of the life-cycle of credit rankings is driven by the exogenous life-cycle of earnings; and (ii) the endogenous upward trend in types is not changed much in a neighborhood of our initial parameters. This second point, in particular, explains the relatively small magnitudes in the first five rows and

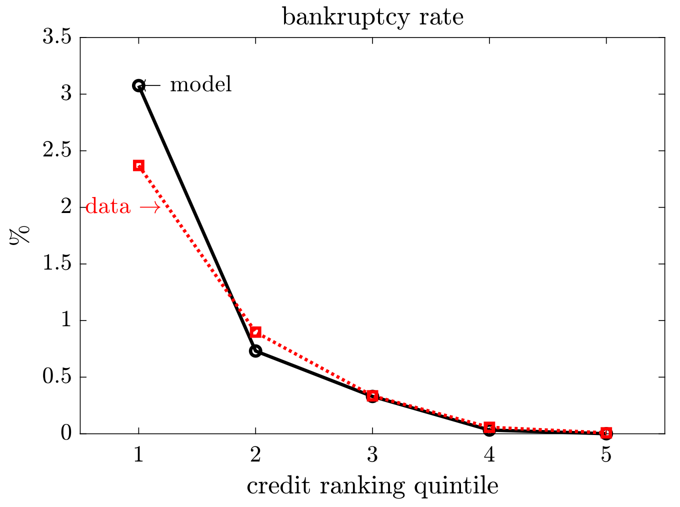


FIGURE 2.—Bankruptcy rate by credit ranking quintiles. *Notes*: The credit ranking data are based on author calculations using FRBNY CCP/Equifax data.

last five columns of Table V. Turning to the final three rows, then, we can highlight several intuitive patterns. Increasing the variable filing cost makes bankruptcy less attractive for the young, increasing their credit ranking at birth (i.e., I:MCR increases) but lowering the upward trend in credit ranking (i.e., S:MCR decreases). Since default becomes a clearer signal of type, then, we see a similar pattern for the cross-sectional variation of credit ranking. Raising either $\alpha$ or $\lambda$ has the opposite effect: by driving up the incentive to file or borrow when young, it lowers credit rankings at birth and increases the upward trend in mean credit rankings, and by lowering the informational content of decisions, it lowers variation in credit rankings at birth (consistent with more pooling).

### 4.3. *Model Fit: Credit Rankings and Bankruptcy Filing*

We next assess how the model performs relative to certain non-targeted properties in the data. Figure 2 shows the non-targeted bankruptcy rate by credit ranking quintiles in the data and in the model. As in the data, the model generates high filing rates among individuals with low credit rankings and low filing rates among individuals with high credit rankings. The model replicates the decreasing pattern in the data.

Another key property of real-world risk scores is that they fall upon bankruptcy and mean revert. We illustrate this property for credit rankings in both the data and the model in Figure 3. Specifically, we conduct an event study of the average change in credit rankings around a bankruptcy filing for various age bins. While the model underpredicts the fall in credit rankings for younger cohorts and overpredicts the long-run recovery in credit rankings, it does remarkably well in matching the rank at the time of filing and the patterns we see in the data despite not being targeted in our estimation.

## 5. MODEL MECHANICS

### 5.1. *Choice Mechanics*

The workings of our model depend on differences in patience among types, as well as their earnings. Figure 4 uses likelihood ratios to illustrate the fact that patience matters

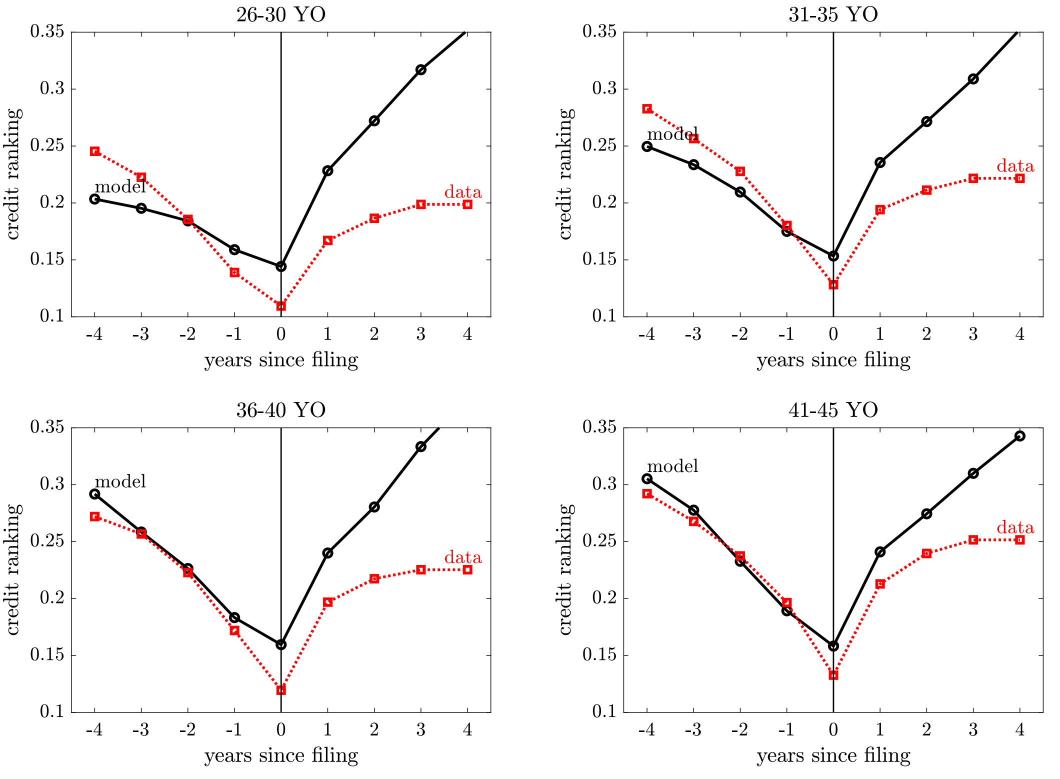


FIGURE 3.—Event study: credit rankings around bankruptcy filings by age. *Notes*: The credit ranking data are based on author calculations using FRBNY CCP/Equifax data. Model results are obtained by simulating a panel of 10,000 individuals for 1000 periods and dropping the first 100 periods. Bankruptcies are then isolated, and each data point reported represents the mean of credit rankings across all bankruptcies for the given lead or lag from the date of the bankruptcy (normalized to 0). The results are binned by 5-year age groups consistent with our earlier results.

in the simple sense that agents of different types take different actions.[35] The left panel shows the bankruptcy likelihood ratio across different persistent earnings levels as a function of debt. While the low type is more likely than the high type to file for bankruptcy for all earnings levels, the fact that the likelihood ratio increases in earnings indicates the importance of differences in type. While both types have the same current gain from default, type $\beta_H$ cares more about the future consequences of a drop in type scores which disincentivizes her from filing for bankruptcy. Further, for any given earnings level, the difference in bankruptcy probability across types declines as debt increases: as debt increases, the current gain from bankruptcy rises enough to offset the future consequences of a drop in one's type score.

[35]Here we define the likelihood ratio of an action $(d, a')$ as the type $\beta_L$ choice probability relative to the sum of the two choice probabilities. That is,

$$\frac{\sigma^{(d,a')}(\beta_L, z, \omega)}{\sigma^{(d,a')}(\beta_L, z, \omega) + \sigma^{(d,a')}(\beta_H, z, \omega)},$$

which lies in $[0, 1]$. For an action which is uninformative about an agent's type, this ratio is 0.5.

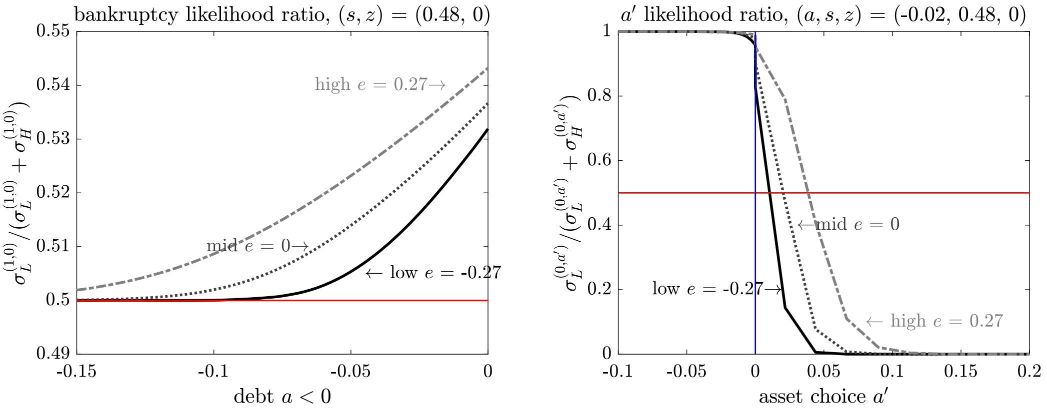


FIGURE 4.—Likelihood ratios of default and borrowing/saving decisions.

The right panel of Figure 4 shows likelihood ratios associated with asset choices. Type $\beta_L$ borrows much more frequently than type $\beta_H$, regardless of earnings. Regarding savings, type $\beta_L$ saves more frequently than type $\beta_H$ (the likelihood ratio is greater than 0.5) for small levels of savings but saves less frequently for higher levels of savings. Furthermore, the level of savings beyond which type $\beta_L$ saves less frequently than type $\beta_H$ is increasing in earnings. These properties are consistent with type $\beta_L$ tending to choose lower values of $a'$ (i.e., higher values of current consumption) than type $\beta_H$ in the same circumstances.

As is a feature of many default models, the probability of bankruptcy is increasing in debt and decreasing in earnings for those with sufficiently large debt.[36] The latter implies that, all else equal, realizations of an individual's earnings have important implications for credit rankings as evident in equation (17). This highlights the composite nature of credit scores: they depend on earnings, asset positions, and type scores.

While age is not a state variable in the decision problems of individuals and lenders, demographics play a role through how we model the arrival of newborns and the Markov process for hidden type. Specifically, since all newborns begin with the lowest earnings class, our Markov process for earnings implies that earnings are expected to rise through an individual's life as shown in the top left panel of Figure 5. The earnings profile induces an increasing wealth profile in the top center panel. Given the rising earnings profile, the young do the lion's share of borrowing as evident in the top right panel.[37] Our estimates of $F_\beta$ and $Q^\beta$ from Section 4 imply that the fraction of type $H$ newborns is lower than the long-run fraction of type $H$. This implies that average type score rises with age according to $\overline{s}' = \overline{s} \cdot Q^\beta(H|H) + (1 - \overline{s}) \cdot Q^\beta(H|L)$, documented in the bottom left panel of Figure 5. The age profile for types scores induces a similar ordering for credit rankings in the bottom left panel. Finally, the bottom right panel shows that the "within-group" variance of consumption is higher for type $L$ than type $H$, consistent with more precautionary saving

[36]We document these facts in the Additional Material to this article. For very small debt, however, the lowest earners (who have the highest marginal utility of consumption) are least likely to default in order to avoid bearing the costs ($\kappa_0$ and $\kappa_1 \times \exp(e)$) of bankruptcy.

[37]Given that economy-wide debt across all ages by type $H$ is lower than that for type $L$ (i.e., the denominator of the share), type $H$ have a higher share than type $L$ when each is poor, which reverses as type $H$ accumulate more savings through time.

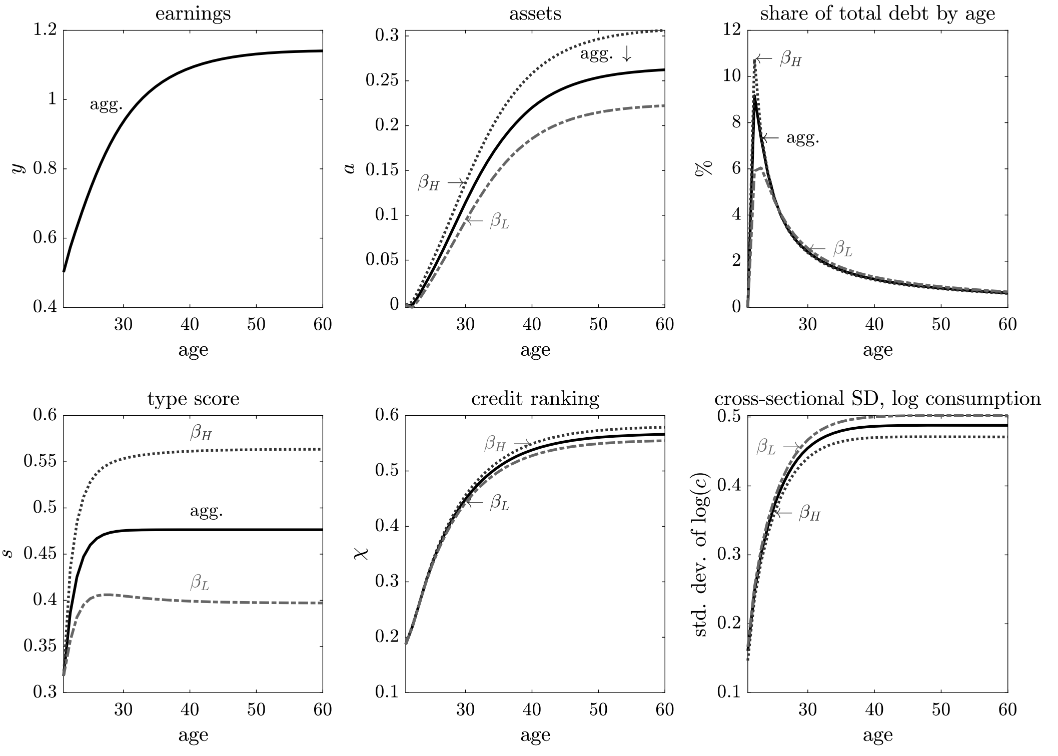


FIGURE 5.—Average moments by age and type in baseline model. *Notes*: In each panel, each line corresponds to the average moment indicated at the specified age in the baseline model. For the type-specific measures, the average is computed conditional on type. The share of total debt by age is the share of economy-wide debt across all ages for the indicated type accounted for by agents of that type at that age. For example, high-$\beta$ 22-year-olds account for 10.6% of the debt held by high-$\beta$ agents.

by (top right panel) and greater credit access (bottom center panel) for type $H$.[38] A notable feature of the bottom right panel of Figure 5 is that the cross-sectional variance of consumption grows in early life, consistent with the empirical evidence in Figure 14 of Heathcote, Perri, and Violante (2010).

### 5.2. *Scoring Mechanics*

Figure 6 plots the change in the public assessment of an individual's type resulting from Bayesian updating given her current type score and actions (i.e., $\psi^{(d,a')}(e, a, s)$ in (16)). Because our estimates exhibit non-zero off-diagonal elements of $Q^{\beta}$, an individual's type can switch from one period to the next even if their action reveals themselves to be one type or another. Thus, the domain of the type scoring function in Figure 6 lies between $s = 0 + Q^{\beta}(H'|L) = 0.205$ and $s = 1 - Q^{\beta}(L'|H) = 0.774$. The left plot shows the different updates for bankruptcy filers and non-filers for $a = -0.02$ integrated over earnings and all $a'$ choices in the case of non-filers. It also plots the posterior of an agent's type even if their

[38] Krueger and Perri (2006) termed "across-group" variation owing to observable differences like education and "within-group" variation the residual which includes idiosyncratic income. Here, we are grouping people on observables like age and also unobservables like type.

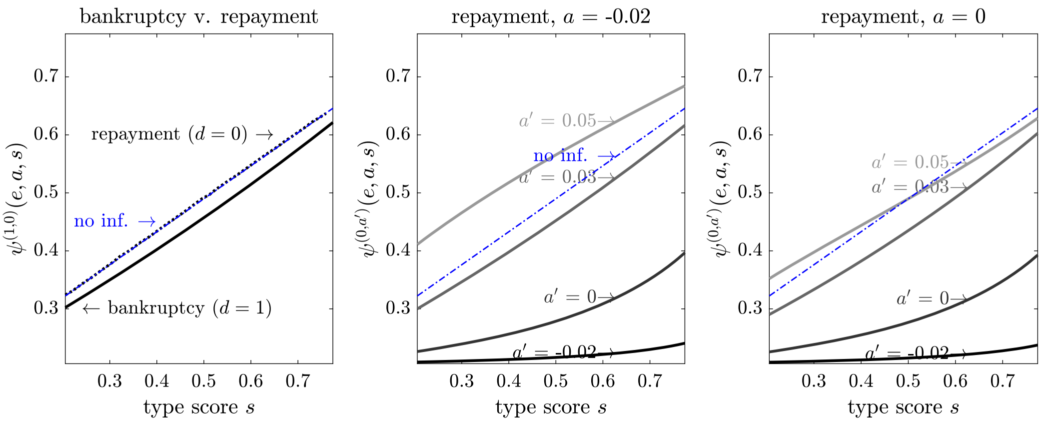


FIGURE 6.—Type-score responses. *Notes*: In each panel, each line is the implied type-score update ($\psi$, $y$-axis) given the current type score ($s$, $x$-axis) for the indicated action. For the left panel, the actions considered are repayment and bankruptcy, conditional on having debt $a = -0.02$. For the middle and right panels, we consider four non-bankruptcy actions, $a' \in \{-0.02, 0, 0.03, 0.05\}$. This shows how different choices affect the type-score update for two different levels of wealth, $a = -0.02$ (middle panel) and $a = 0$ (right panel). In each panel, the blue "no inference" line corresponds to the type-score update the borrower would receive just based on the upward drift in $\beta$ over age.

actions are not observed, which we call "no inference."[39] The mean reversion in type score accounts for why a person's score falls upon repayment if the current score is sufficiently high or rises upon default if it is sufficiently low. Still, it remains true that repaying leads to a higher type score than filing for bankruptcy. Since the $\nu$ shock is noisy, the choice to not file does not reveal much; thus, the no bankruptcy line is only imperceptibly higher than the no inference line in the left panel of Figure 6.

The center and right plots of Figure 6 show the Bayesian updates that result from different actions taken by an individual either already in debt (center at $a = -0.02$) or with zero assets (right) for the median earner ($e = z = 0$). For an individual already in debt, staying in debt (i.e., $a' = -0.02$) signals the individual is likely to be type $\beta_L$ as we saw in Figure 4, leading to a drop in their posterior score. As the individual chooses higher $a'$, their posterior rises. It is not until sufficiently high savings choices (e.g., $a' = 0.05$) that there is enough separation to raise an individual's posterior higher than what would be associated with mean reversion only (i.e., no inference). The right panel documents that starting from a higher asset position ($a = 0$), all the assessments shift down; that is, the smaller net change in asset position makes the inference less likely to be a high type.

Figure 7 illustrates some important points. First, the figure verifies a form of the sufficient condition (one-to-one mapping between type scores and credit scores conditional on persistent earnings and assets) in Theorem 3 that establishes the equivalance between the fundamental type-score equilibrium (RCE) and the credit-score equilibrium (RCECS). Specifically, we graph the inverse function since Theorem 3 assumes that the inverse function $s = (p^{\bar{a}*})^{-1}(e, a, m)$ exists. This graph is indeed one-to-one, and we have verified that the function is one-to-one conditional on observables across the entire state

[39]One might expect to compare the type scoring function to the 45 degree line to see whether the agent's reputation improves or deteriorates. However, given the upward trend in mean type scores implied by the discount factor process, it is more natural to compare to the no inference line.

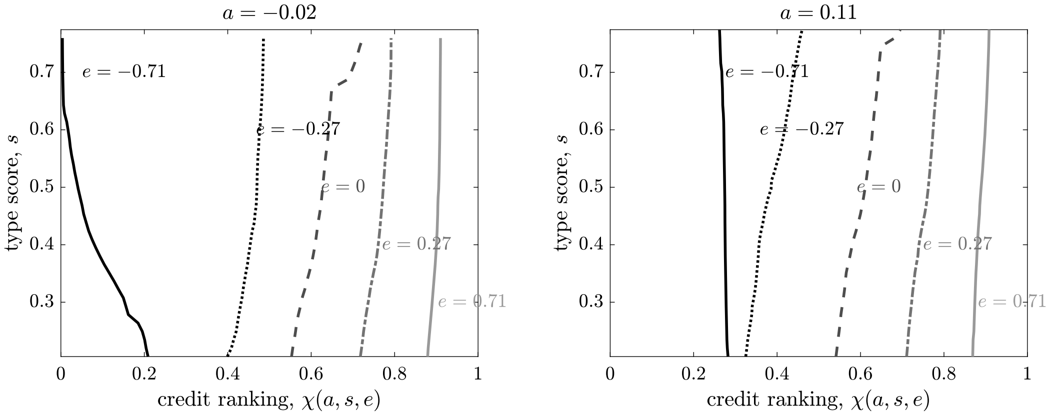


FIGURE 7.—Credit-scoring function.

space. Second, it shows the important effect of earnings on credit rankings; higher earnings are correlated with higher rankings. Third, higher beginning of period assets are not necessarily correlated with higher rankings (e.g., for $e = 0.71$, borrowing $\overline{a}$ from a higher initial asset holding lowers one's credit ranking). Finally, while higher type scores are associated with higher credit rankings for most earnings levels, it is not true when earnings are very low. In this case, type $\beta_H$ actually file for bankruptcy slightly more frequently than type $\beta_L$, generating the negative relation between type score and credit ranking. At low earnings, it may actually be optimal to borrow rather than default (see Figure 1 in Chatterjee et al. (2007)). In our current case where we have different types, the gain from borrowing is stronger for type $\beta_L$ than for $\beta_H$ since they care less about the relative drop in their future type score.

In Figure 8, we plot the cross-section of debt choices across credit ranking quintiles (behind their aggregate counterparts in Table II). The figure illustrates that borrowers with low credit rankings are more likely to be in debt and have high debt-to-income ratios. It

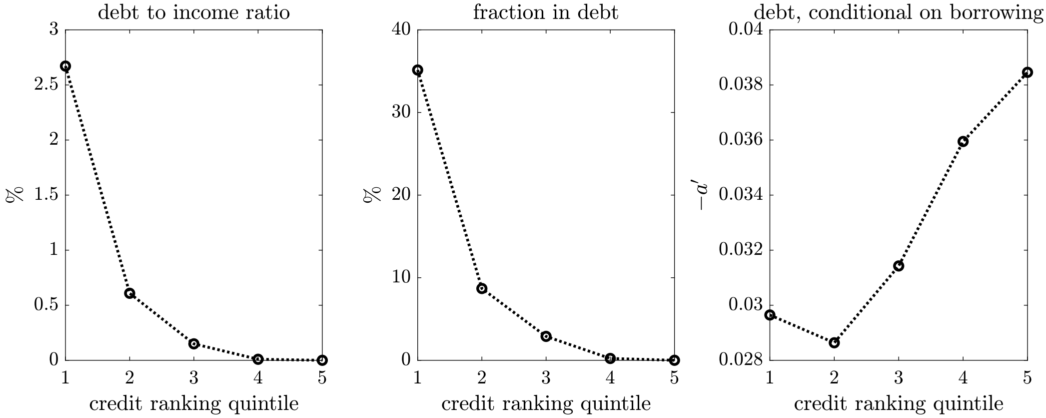


FIGURE 8.—Outcomes by credit ranking quintiles. *Notes*: Average moments are computed as the average conditional on credit ranking quintile. The debt-to-income ratio is the average of individual debt to individual income across the population. The size of debt conditional on borrowing averages across all choices made by each agent in a given state.

also illustrates that conditional on actually borrowing, those with high credit rankings tend to borrow more (since they can do so at lower interest rates). These relations about debt, income, and credit status are consistent with the empirical findings in Diaz-Gimenez, Glover, and Rios-Rull (2011, Table 17, p. 19).

### 5.3. *Selection and Reputation Effects*

We now consider how one's current asset choice affects the price they face today via revelation about the individual's unobservable type. If an individual of observable type $\omega = (e, a, s)$ were to borrow $a'$, she would be facing a price that depends on the default probabilities of her type tomorrow $\omega' = (e', a', s')$. Since $a'$ is given at the beginning of the next period and does not affect the probability distribution of $e'$, what matters is how $a'$ affects the update $s'$. To isolate the contribution of the asset choice on the update, we compare the baseline equilibrium price which depends on the update through $Q^s(s'|\psi^{(d,a')})$ with a price schedule that results from excluding $a'$ from the Bayesian updating formula. This price schedule is now given by $\widetilde{q}^{a'}(a, s, e) = \rho \widetilde{p}^{a'}(a, s, e)/(1 + r)$, where

$$\widetilde{p}^{a'}(a, s, e) = \sum_{\beta', e', z', s'} H(z')Q^e(e'|e)Q^s(s'(\beta')|\widetilde{s}')s'(\beta')\left[1 - \sigma^{(1,0)}(\beta', e', z', a', s')\right], \quad (26)$$

where $\widetilde{s}' = s \cdot Q^\beta(H|H) + (1 - s) \cdot Q^\beta(H|L)$ updates the prior $s$ using only $Q^\beta$ and ignores the information in the borrowing choice $a'$.

Figure 9 shows the percentage increase in $\widetilde{q}$ relative to the baseline equilibrium price $q$. Recall from Figure 4 that borrowing reveals oneself to very likely be type $\beta_L$ and therefore when lenders cannot take this information into account, ignoring what can be learned from selection effects induces the price $\widetilde{q}$ to exceed $q$. The figure shows that if an individual starts with a low prior $s$ (here we take the 10th percentile $s = 0.21$), the price effect is smaller than if the individual starts with a high prior $s$ (here we take the 90th percentile $s = 0.76$). Further, the positive price effects for the two priors are amplified the more

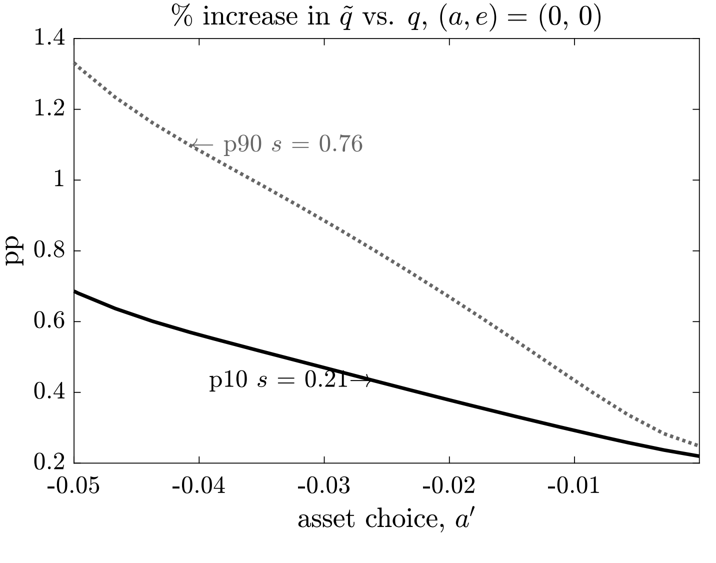


FIGURE 9.—Static effect of borrowing choice.

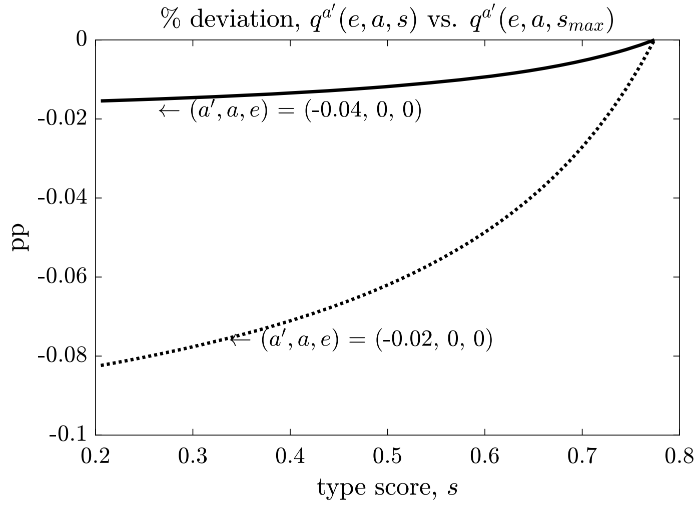


FIGURE 10.—Reputation and prices.

borrowing is undertaken. The latter effect arises since, for higher debt levels, the pool of borrowers will tend to contain less creditworthy type $\beta_L$—that is, more adverse selection.

In addition to the contemporaneous effects discussed above, asset choices can also have long-lasting effects. This requires that: (i) prices depend on an individual's current reputation (i.e., type score) for a given current action; and (ii) her choice today affects her future reputation (and hence future prices). Having established the second condition in Figure 6, we next establish the first in Figure 10.

Specifically, Figure 10 plots the percentage change in debt prices that an individual with current type score $s$ can obtain relative to a person with the highest type score $s = 0.77$ for two different debt choices. The fact that both lines are downward sloping establishes that type scores matter; for a given debt choice, a higher $s$ fetches a higher price. The reason is simple; since $\beta_H$ types repay with higher probability than the $\beta_L$ types, there is information about the probability of repayment in the current type score (which, in turn, reflects the history of the individual's past actions). The fact that the line in Figure 10 is less steep for large debts is consistent with Figure 4: type $\beta_L$ are almost perfectly separated from type $\beta_H$ at those debt levels, and so the prior $s$ does not matter as much for assessing repayment probability.

Finally, we examine signaling costs. In models with hidden types, the "bad" types have an incentive to imitate the "good" types in order to pool with them and obtain better terms of trade, while the "good" types have an incentive to separate themselves from the "bad" types to get even better terms. Here we assess the costliness for an impatient type $L$ to imitate the actions of a patient type $H$.

Changing one's action has three effects: (i) a change in today's consumption; (ii) a change in tomorrow's net wealth; and (iii) a change in tomorrow's reputation. To explore these three effects for a type $L$ to imitate a type $H$, we assume the type $L$ follows the choice probability function $\sigma(\beta_H, z, \omega)$ instead of $\sigma(\beta_L, z, \omega)$. One measure of the consumption cost from a type $L$ individual mimicking a type $H$ is the average difference in consumption between $H$ types and $L$ types implied by the differences in their choice probabilities relative to the average consumption of a type $L$ individual. Similar measures can be computed for next-period net wealth and credit ranking.

TABLE VI

SIGNALING COSTS AND BENEFITS.

| % Average Gain in: | Consumption ($\widehat{C}$) | Wealth ($\widehat{A}$) | Credit Ranking($\widehat{\chi}$) |
|---|---|---|---|
| All | −3.65 | 3.80 | 1.31 |
| Newborns | −0.77 | 0.81 | 0.37 |

*Note*: The first column measure is $\hat{C} = \frac{\sum_{z,\omega} \mu(\beta_L, z, \omega)[\sum_{(d,a')\in F(z,\omega)} (\sigma^{(d,a')}(\beta_H, z, \omega) - \sigma^{(d,a')}(\beta_L, z, \omega)) c^{(d,a')}(z, \omega)]}{\sum_{z,\omega} \mu(\beta_L, z, \omega)[\sum_{(d,a')\in F(z,\omega)} \sigma^{(d,a')}(\beta_L, z, \omega) c^{(d,a')}(z, \omega)]}$, while the second and third columns substitute $a'(z, \omega)$ and $\chi^{(0,\overline{a})}(\omega)$ for $c^{(d,a')}(z, \omega)$.

Table VI provides these calculations for our calibrated parameters (i.e., where $(\beta_H - \beta_L)/\beta_H = 13\%$). The table illustrates an important point. Type $L$ newborns have a much lower consumption loss to mimicking a type $H$ individual than their older counterparts. This is because the imitation costs are increasing in earnings and assets, both of which rise on average through one's life as evident in Figure 5. Since type $H$ choose to save more, this imposes a bigger consumption loss to type $L$ from mimicking later in life. Alternatively, it is easier to mimic when young, as the dispersion in assets and scores is lower in youth. The fact that it is less costly to mimic when young implies there is more pooling among the young, and the fact that it is more costly to mimic when old implies there will be more separation among the old. The consequence is that while there is a bigger jump in credit ranking of $\beta_L$ type when mimicking in old age, it is more costly to do so.

## 6. IMPACT OF ALTERNATIVE INFORMATION STRUCTURES

How important is the information structure for allocations? What would happen if society outlawed tracking of individual credit histories and with it the incentives to build a good reputation? Would the credit market shrink dramatically as the usefulness of maintaining a good reputation disappears? These are natural questions that we can answer quantitatively by using our model to compare outcomes in economies that differ from our baseline only in their information structures.

Before we get into the details of our answers to these questions, it is important to keep in mind certain features of hidden information that are present in our model. First, because of imperfect separation, low types are subsidized by high types. Second, there are incentives to repay debt and to save more to imitate a high type. Third, there can be important interactions of hidden information across the age profile. Specific to our model, all newborns are low earners and face an (expected) upward sloping age-earnings profile. Thus, newborns and young have *a life-cycle reason to borrow* and so are more impacted by hidden information. Finally, individuals face idiosyncratic earnings shocks against which direct insurance is unavailable. Since borrowing to smooth consumption is costly in all the economies that we explore, all individuals have *a precautionary savings motive*. Differences between the economies imply not only that individuals behave differently on account of the incentives that they face, but also that the equilibrium prices reflect these changes. Accordingly, we have to look at both aspects simultaneously.

### 6.1. *Description of Alternative Economies*

We now consider two stark alternative information structures in which reputation plays a limited role: one where past actions cannot be used to price-discriminate but demographic drift can be used to infer type, and another where type is public information.[40]

Our first alternative economy poses hidden information as in the baseline model (hereafter termed BASE), but prohibits creditors from using a person's past to price loans. We assume that the length of one's credit history (proxied here by age) is both publicly observable and legally used to price debt. This assumption isolates the role of reputation from the role of the demographic drift in the credit market. In this economy, some information about the individual's type is learned contemporaneously from her asset choice, but this information is not carried across periods. We refer to this economy as the *no-tracking* (NT) economy since an individual's assets cannot be tracked over time. To be concrete, in this alternative economy, individuals' type scores initially equal the fraction of high types among newborns and evolve thereafter according to the demographic drift.[41] This implies a one-to-one mapping between an individual's age and the prior that she is a high type (her type score in this alternative economy). Consistent with no-tracking, lenders are also not allowed to use information about an individual's beginning-of-period asset holdings when pricing loans since this also contains information about her past actions. However, lenders are able to use the current action and the cross-sectional distribution of agents in the NT economy when forming a posterior about the likelihood of repayment necessary to price loans. The NT economy has hidden information and cross-subsidization, but there are no dynamic reputational incentives, as actions cannot be used to impute type as in the BASE economy.

Our second alternative features *full information* (hereafter termed FI) where the type is directly observed by lenders that use it to price-discriminate. Except for our extreme value shocks, this alternative economy is similar to Chatterjee et al. (2007). There is no need to infer a person's type, and the price for a loan of size $a'$ depends directly on $\beta$ and all other relevant observables. Importantly, prices do not depend on $s$ nor $a$ because they are not directly payoff relevant. Comparing the FI economy to the BASE reveals the full impact of hidden information: in the FI, there is no cross-subsidization nor are there any incentives to imitate or separate.

Figure 11 highlights some of the main differences among the three economies. The left panel shows the CDFs of type scores for each type. This indicates the degree to which information about type is revealed to creditors. In the BASE economy, the $L$-type CDF rises steeply and fast, indicating that most type $L$ individuals have low scores. In contrast, the $H$-type CDF rises more gradually, indicating that type $H$ individuals have more dispersed type scores. In the NT economy, type scores (priors) are trapped between 0.32 (score at birth) and 0.48 (score at age infinity). Although people of all types share the same age-specific type score at each age, there are more type $L$ individuals at each age than type $H$ and, consequently, the CDF of low types rises somewhat faster. Most importantly, the CDFs are closest to each other for the NT economy, indicating that less is being learned about an individual's type as she ages, compared to the BASE and FI economies.

The right panel of Figure 11 plots the mean and standard deviation of type scores for each age across the alternative economies. Importantly, the mean type score at each age

[40]For the formal specification of these alternative economies, see Appendix B.6 of the Supplemental Material.

[41]As before, the evolution is simply given by $\bar{s}' = \bar{s} \cdot Q^{\beta}(H|H) + (1 - \bar{s}) \cdot Q^{\beta}(H|L)$ with initial condition $\bar{s} = 0.32$.

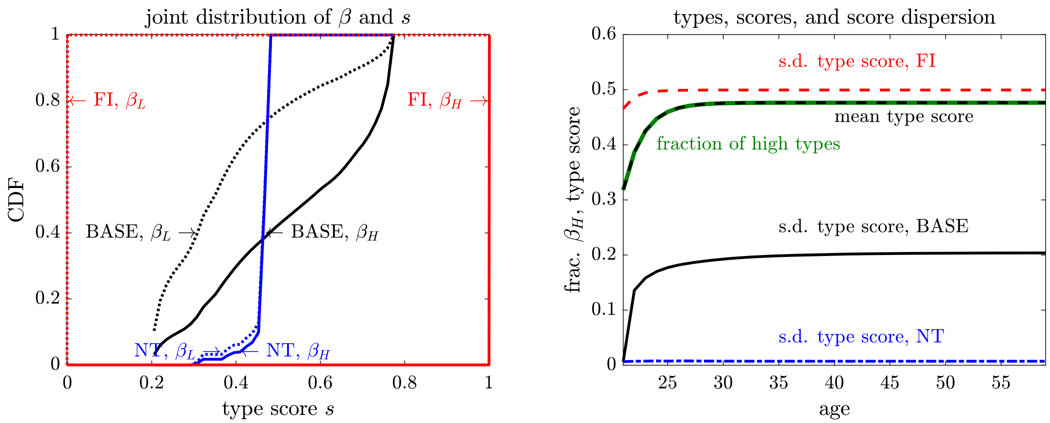


FIGURE 11.—Evolution of types and type scores in alternative economies. *Notes*: The left panel plots the type-specific CDF of type scores in each model economy. Black/blue/red refer to BASE/NT/FI model economies, and solid (dashed) lines refer to high (low) types. The green and black dashed lines of the right panel correspond to the fraction of high types at the indicated age and the mean type score at each age in each economy. The three model variants considered have different type-score standard deviations.

is correctly assessed in all economies to be equal to the fraction of high types. In the NT economy, the standard deviation of type scores at all ages is zero, as nothing is learned.[42] For the FI economy, the dispersion in type scores at any age is the dispersion of types themselves in the economy. In the BASE economy, it is increasing at a faster rate than in the FI economy, consistent with learning.

### 6.2. *No-Tracking Economy Results*

The key feature of NT is that the only exogenous information (earnings class and age-implied type score) can be used by lenders in the future. Thus, there are no incentives to maintain one's reputation in asset markets. This can cause equilibrium price menus to drop, as is evident from the fact that average interest rates rise in Table VII.[43] In response to the rise in interest rates, the fraction of the population in debt falls. Since reputation effects are absent, though, those who choose to borrow are willing to do so at higher interest rates. Consequently, Table VII shows that, in equilibrium, both the debt-to-income ratio and average interest rate rise. Higher debt in turn leads to an increase in the bankruptcy rate.

Table VII also decomposes the aggregate statistics by unobservable type. It shows that type $\beta_H$ borrowing and default is more sensitive to the change in incentives compared to type $\beta_L$. This arises due to the lack of persistent reputational costs incurred through borrowing and default, which were present in the baseline economy.

Regarding welfare, we focus on newborns.[44] Rising interest rates associated with the effect of eliminating incentives that rely on credit histories generally make newborns worse

[42] Strictly speaking, because the evolution of $s$ implied by the Markov type transition function for the NT economy typically does not yield scores which fall on the grid points in $\mathcal{S}$, there is some negligible dispersion in type scores.

[43] For an example of such a price menu, see Figure 13 in Appendix B.6.1 of the Supplemental Material.

[44] One consequence of focusing on newborns is that we do not need to compute a transition. At the moment of the policy switch, the average asset holdings of older cohorts are potentially different from those of the same age group in the steady state of the NT economy. Hence, even in a small open economy, all except the

TABLE VII

COMPARISON OF BASELINE, NO TRACKING, AND FULL INFORMATION ECONOMIES.

| Economy Discount Factor Type | No Tracking (NT) | | | Full Information (FI) | | |
|---|---|---|---|---|---|---|
| | High | Low | All | High | Low | All |
| | | | Panel A: % difference from BASE | | | |
| bankruptcy rate | 1.40 | 0.95 | 1.12 | −0.98 | 0.36 | −0.13 |
| average interest rate | 2.02 | 1.02 | 1.44 | −7.15 | 0.83 | −2.52 |
| interest rate dispersion | 10.7 | 0.75 | 7.57 | −5.12 | 0.71 | −2.15 |
| fraction in debt | −0.21 | −0.13 | −0.15 | 0.23 | −0.10 | 0.00 |
| debt-to-income ratio | 0.45 | 0.34 | 0.39 | −0.30 | 0.12 | −0.04 |
| | | | Panel B: wealth equivalent welfare measure, newborns (% of mean wealth) | | | |
| low $z$ | −0.001 | 0.089 | 0.060 | 0.187 | 0.089 | 0.121 |
| median $z$ | −0.000 | −0.000 | −0.000 | 0.089 | 0.044 | 0.058 |
| high $z$ | −0.000 | −0.000 | −0.000 | 0.135 | 0.089 | 0.104 |
| mean | −0.001 | 0.030 | 0.020 | 0.137 | 0.074 | 0.094 |

*Note*: Each entry in Panel A is the difference, in percentage points of the BASE moment, of the moment in the indicated alternative economy (FI or NT) relative to the BASE economy. Panel B reports the amount of additional wealth an agent would have to be given in the baseline economy in order to be indifferent between being born into the indicated alternative economy in the indicated state and being born in the baseline economy. The units for Panel B are percentages of mean wealth. Table 10 in Section C of the Additional Material, Chatterjee et al. (2023), explores how variations in $\alpha$ and $\lambda$ affect the "all" columns in this table.

off. Only the type $\beta_L$ newborns with lowest transitory earnings shock $z$ benefit from the cross-subsidization that comes with no tracking. Since their gain is large (0.089) relative to the small losses in all other cases, the mean overall gain for newborns is positive (0.020).[45]

## 6.3. *Full Information Results*

Under full information (FI), types are observed and cross-subsidization ends, as do incentives for a type $\beta_L$ to imitate a type $\beta_H$. Therefore, equilibrium debt price menus change fundamentally for each type.[46] Specifically, as one might expect, type $\beta_L$ in the FI economy face lower loan prices (higher interest rates) and type $H$ face higher prices (lower interest rates) than the BASE economy where there is some cross-subsidization. Price differences also change with age. Interestingly, as agents accumulate assets through time, the act of borrowing is assessed to be even more likely to come from a low type so there is little difference between prices in FI and BASE for type $\beta_L$ but large differences for type $\beta_H$. These differences in interest rates faced by the two types are clearly illustrated in Table VII.

In response to the changes in the menu of interest rates, Table VII documents that the fraction of type $H$ (type $L$) who borrow rises (falls) as one would expect. While there are

newborns face a transition of prices as the cross-sectional distribution used to infer future default probabilities evolves to the invariant distribution.

Besides this, there are multiple ways to think of how the switch from the BASE to NT would be implemented for people already alive. One possibility is to immediately outlaw the use of personal asset market history beyond the length of one's credit history (i.e., age). Alternatively, one could treat older individuals just like newborns, using information on their asset holdings and type score for the period of the policy switch, but then knowledge about subsequent savings or defaults cannot be used. Hence, rather than make a choice on implementation, we focus on newborns.

[45]Our wealth equivalent welfare measure is standard; details are in the Additional Material.

[46]For an example, see Figure 14 in Appendix B.6.2 of the Supplemental Material.

also no reputation effects in the FI case, the changes in debt-to-income ratio come about for different reasons for the two types. Debt-to-income for type $L$ increases despite the rise in interest rates because they were holding so little debt in the BASE economy in order to raise their reputation by mimicking high types (i.e., the rise is not that they are holding more debt in FI but that they were holding so little debt in BASE). The lower debt-to-income ratio for type $H$ arises from the large increase in $q^{a'}_{\text{BASE}}$ which makes it cheaper to achieve a desired inflow for consumption (i.e., since $q^{a'}_{\text{BASE}}$ rises, one can achieve the desired inflow $q^{a'}_{\text{BASE}}a'$ with a smaller $a'$). Lower (higher) debt-to-income for type $H$ ($L$) explains the fall (rise) in bankruptcy across type in Table VII. One important takeaway from the differences across type is that they tend to cancel out in the aggregate, leading to only slight differences in aggregate statistics except for the impact on equilibrium interest rates.

As Table VII documents, all newborns are better off in the FI economy. While it is clear that the newborn high types would rather live with full information where they do not subsidize the low types, even low types prefer (albeit less so) full information since they transit to type $H$ with a relatively high probability $Q^{\beta}(H'|L) = 0.205$. The aggregate welfare gains from eliminating cross-subsidization are quite high (0.094) in the FI economy relative to the gain (0.020) in the NT economy. Thus, our "big data" BASE economy yields welfare properties for newborns which are very close to the "small data" no-tracking economy. This is in contrast to the relatively large welfare gains that can come from eliminating hidden information.

## 7. CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

In this paper, we present a hidden information model of unsecured consumer credit with risk of default. People are subject to unobserved persistent and transitory shocks, and the history of people's asset market actions helps forecast future defaults. The setup is possibly the simplest environment to quantitatively study the role of credit scores in regulating consumer credit. We showed how this can be done using shocks drawn from an extreme value distribution and recursive updating of beliefs.

Our quantitative model not only accounts for aggregate credit market moments, but also the age profile of credit rankings observed in U.S. data. In this sense, our model provides a quantitative theory of the credit score.

Two implications of our theory are worth highlighting. First, we found that restricting lenders' access to an individual's history of asset market actions (no tracking) leads to an overall welfare gain for young adults. Since the young tend to borrow against their future income, the insurance afforded poor young adults of low type who are cross-subsidized by others in better standing outweighs the costs of higher interest rates associated with negative incentive effects from not having to maintain a good reputation. Our "big data" baseline model suggests that the intratemporal insurance for a subset of the population in a "small data" economy can outweigh the incentive effects worsening intertemporal insurance.

Second, even though our model allowed lenders unrestricted access to the history of all actions relevant for inferring an individual's type, the equilibrium allocations at an individual level remain far removed from those of a full information economy. This stems from the fact that individuals *select* actions that only partially reveal their type, while in the full information economy they get that revelation for free. Despite big differences at the micro level, the macro (aggregate) differences can be small.

For simplicity, we have assumed that the only possible actions for an indebted agent were to either pay back its debt completely and choose another asset position facing prices that are based on its observables or to file for bankruptcy at a cost and have its debts discharged. In the real world, indebted agents can also go delinquent. In Appendix D of the Supplemental Material, we modify our model along the lines of Athreya, Mustre-del-Rio, and Sanchez (2019) to include a delinquency option and quantitatively assess its implications for credit market outcomes. Notably, the following basic results from the BASE model hold in the extended model: (i) type $\beta_L$ are more likely to go bankrupt; (ii) each type is more likely to file for bankruptcy at higher levels of debt; and (iii) bankruptcy on average leads to a downward revision of one's type score.

Where next? First, type does not have to correspond to an individual's hidden time preference. Alternatively, it could correspond to hidden ability differences that exogenously affect earnings. Hidden time preference can affect a hidden human capital decision (i.e., moral hazard) to endogenously affect earnings or a variety of other personal traits.

Second, reputation in the unsecured credit market can spill over to other markets, reinforcing reputation effects. A person's reputation (or type score) in the unsecured credit market may have implications for other markets (e.g., insurance, labor, housing) and other interactions (marriage) that are worth exploring. Finally, considering the interaction of financial literacy and imperfect competition in the unsecured consumer credit market are important directions for future research.

## REFERENCES

ACQUISTI, ALESSANDRO, AND HAL R. VARIAN (2005): "Conditioning Prices on Purchase History," *Marketing Science*, 24, 367–381. [1806]

ADDA, JEROME, AND RUSSELL COOPER (2003): *Dynamic Economics: Quantitative Methods and Applications*. Cambridge, MA: MIT Press. [1817]

AGARWAL, SUMIT, SOUPHALA CHOMSISENGPHET, AND CHUNLIN LIU (2010): "The Importance of Adverse Selection in the Credit Card Market: Evidence From Randomized Trials of Credit Card Solicitations," *Journal of Money, Credit and Banking*, 42, 743–754. [1805]

AGUIAR, MARK, SATYAJIT CHATTERJEE, HAROLD L. COLE, AND ZACHARY R. STANGEBYE (2016): "Quantitative Models of Sovereign Debt Crises," in *Handbook of Macroeconomics*, Vol. 2B, ed. by Harald Uhlig and John B. Taylor. Elsevier B.V., 1697–1755. [1806]

ALBANESI, STEFANIA, AND JAROMIR NOSAL (2018): "Insolvency After the 2005 Bankruptcy Reform," Tech. Rep. Working Paper No. 24934, National Bureau of Economic Research. [1817]

ALBANESI, STEFANIA, AND DOMONKOS F. VAMOSSY (2019): "Predicting Consumer Default: A Deep Learning Approach," Working Paper No. 26165, National Bureau of Economic Research. [1805]

ALBANESI, STEFANIA, GIACOMO DEGIORGI, AND JAROMIR NOSAL (2022): "Credit Growth and the Financial Crisis: A New Narrative," *Journal of Monetary Economics*, 132, 118–139. [1805]

AMADOR, MANUEL, AND CHRISTOPHER PHELAN (2021): "Reputation and Sovereign Default," *Econometrica*, 89, 1979–2010. [1807]

ANDREWS, ISAIAH, MATTHEW GENTZKOW, AND JESSE SHAPIRO (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, 132, 55–70. [1822,1823]

ATHREYA, KARTIK, JOSE MUSTRE-DEL-RIO, AND JUAN M. SÁNCHEZ (2019): "The Persistence of Financial Distress," *Review of Financial Studies*, 32, 3851–3883. [1838]

ATHREYA, KARTIK, XUAN S. TAM, AND ERIC R. YOUNG (2012): "A Quantitative Theory of Information and Unsecured Credit," *American Economic Journal: Macroeconomics*, 4, 153–183. [1806]

AUSUBEL, LAWRENCE (1999): "Adverse Selection in the Credit Card Market," Unpublished, University of Maryland. [1805]

BLANCHARD, OLIVIER J. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93, 223–247. [1807]

BLATTNER, LAURA, JACOB HARTWIG, AND SCOTT NELSON (2022): "Information Design in Consumer Credit Markets," Working paper. [1806]

BRAXTON, J. CARTER, KYLE F. HERKENHOFF, AND GORDON M. PHILLIPS (2020): "Can the Unemployed Borrow? Implications for Public Insurance," Working Paper 27026, National Bureau of Economic Research. [1807]

BRIGLIA, LUIGI, SATYAJIT CHATTERJEE, DEAN CORBAE, KYLE DEMPSEY, AND JOSÉ-VÍCTOR RÍOS-RULL (2021): "Saving for a Sunny Day: An Alternative Theory of Precautionary Savings," Unpublished. [1817]

CHATTERJEE, SATYAJIT, DEAN CORBAE, KYLE DEMPSEY, AND JOSÉ-VÍCTOR RÍOS-RULL (2023): "Additional Materials to "A Quantitative Theory of the Credit Score"," Working Paper, Ohio State University. [1836]

CHATTERJEE, SATYAJIT, DEAN CORBAE, MAKOTO NAKAJIMA, AND JOSÉ-VÍCTOR RÍOS-RULL (2007): "A Quantitative Theory of Unsecured Consumer Credit With Risk of Default," *Econometrica*, 75, 1525–1589. [1806,1809,1816,1822,1830,1834]

CHATTERJEE, SATYAJIT, DEAN CORBAE, AND JOSÉ-VÍCTOR RÍOS-RULL (2008): "A Finite-Life Private-Information Theory of Unsecured Consumer Credit," *Journal of Economic Theory*, 142, 149–177. [1807]

CHATTERJEE, SATYAJIT, DEAN CORBAE, KYLE DEMPSEY, AND JOSÉ-VÍCTOR RÍOS-RULL (2023): "Supplement to 'A Quantitative Theory of the Credit Score'," *Econometrica Supplemental Material*, 91, https://doi.org/10.3982/ECTA18771. [1807]

COLE, HAROLD L., AND PATRICK KEHOE (1998): "Models of Sovereign Debt: Partial versus General Reputations," *International Economic Review*, 39, 55–70. [1807]

CORBAE, DEAN, AND ANDREW GLOVER (2018): "Employer Credit Checks: Poverty Traps versus Matching Efficiency," NBER Working Paper 25005. [1807,1808]

D'ERASMO, PABLO (2011): "Government Reputation and Debt Repayment in Emerging Economies," Unpublished. [1807]

DIAMOND, DOUGLAS W. (1989): "Reputation Acquisition in Debt Markets," *Journal of Political Economy*, 97, 828–862. [1807]

DIAZ-GIMENEZ, JAVIER, ANDREW GLOVER, AND JOSÉ-VÍCTOR RÍOS-RULL (2011): "Facts on the Distribution of Earnings, Income, and Wealth in the United States: 2007 Update," *Quarterly Review, Federal Reserve Bank of Minneapolis*, 34, 2–32. [1831]

DROZD, LUKASZ, AND RICARDO SERRANO-PADIAL (2017): "Modeling the Revolving Revolution: The Debt Collection Channel," *American Economic Review*, 107, 897–930. [1807]

DUBEY, PRADEEP, AND JOHN GEANOKOPLOS (2002): "Competitive Pooling: Rothschild-Stiglitz Reconsidered," *The Quarterly Journal of Economics*, 117, 1529–1570. [1806]

EATON, JONATHAN, AND MARK GERSOVITZ (1981): "Debt With Potential Repudiation: Theoretical and Empirical Analysis," *Review of Economic Studies*, 48, 289–309. [1806]

EINAV, LIRAN, MARK JENKINS, AND JONATHAN LEVIN (2013): "The Impact of Credit Scoring on Consumer Lending," *The RAND Journal of Economics*, 44, 249–274. [1805]

ELUL, RONEL, AND PIERO GOTTARDI (2015): "Bankruptcy: Is It Enough to Forgive or Must We Also Forget," *American Economic Journal: Microeconomics*, 7, 294–338. [1807]

EXLER, FLORIAN, AND MICHÈLE TERTILT (2020): "Consumer Debt and Default: A Macro Perspective," *Oxford Research Encyclopedia of Economics and Finance*. [1806]

EXLER, FLORIAN, IGOR LIVSHITS, JAMES MACGEE, AND MICHÈLE TERTILT (2021): "Consumer Credit With Over-Optimistic Borrowers," Working Paper 21-42, Federal Reserve Bank of Philadelphia. [1806]

FLODEN, MARTIN, AND JASPER LINDÉ (2001): "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?" *Review of Economic Dynamics*, 4, 406–437. [1817,1818]

FOURAKIS, STELIOS (2021): "Sovereign Default and Government Reputation," Unpublished. [1807]

GALE, DOUGLAS (1992): "Walrasian Theory of Markets With Adverse Selection," *The Review of Economic Studies*, 59, 229–255. [1806]

GUERRERI, VERONICA, ROBERT SHIMER, AND RANDALL WRIGHT (2010): "Adverse Selection in Competitive Search," *Econometrica*, 78, 1823–1862. [1806]

HEATHCOTE, JONATHAN, FABRIZIO PERRI, AND GIAN LUCA VIOLANTE (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13, 15–51. [1828]

HERTZBERG, ANDREW, ANDRES LIBERMAN, AND DANIEL PARAVISINI (2018): "Screening on Loan Terms: Evidence From Maturity Choice in Consumer Credit," *The Review of Financial Studies*, 31, 3532–3567. [1805]

JAFFEE, DWIGHT M., AND THOMAS RUSSELL (1976): "Imperfect Information, Uncertainty, and Credit Rationing," *The Quarterly Journal of Economics*, 90, 651–666. [1806]

KOVRIJNYKH, NATALIA, IGOR LIVSHITS, AND ARIEL ZETLIN-JONES (2019): "Building Credit Histories With Heterogeneously Informed Lenders," Working Paper 19-17, Federal Reserve Bank of Philadelphia. [1806]

KRUEGER, DIRK, AND FABRIZIO PERRI (2006): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 73, 163–193. [1828]

LIVSHITS, IGOR, JAMES C. MACGEE, AND MICHÈLE TERTILT (2007): "Consumer Bankruptcy: A Fresh Start," *American Economic Review*, 97, 402–418. [1806,1808,1816]

——— (2016): "The Democratization of Credit and the Rise in Consumer Bankruptcies," *Review of Economic Studies*, 83, 1673–1710. [1807]

LUO, WENLAN (2017): "The Consumer Credit Channel of Monetary Policy,", Georgetown University. [1806]
MYERSON, ROGER B. (1978): "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 73–80. [1806]
NARAJABAD, BORGHAN N. (2012): "Information Technology and the Rise of Household Bankruptcy," *Review of Economic Dynamics*, 15, 526–550. [1807]
NELSON, SCOTT (2022): "Private Information and Price Regulation in the US Credit Card Market," Unpublished. [1806]
PESKI, MARCIN (2014): "Repeated Games With Incomplete Information and Discounting," *Theoretical Economics*, 9, 651–694. [1807]
PRESCOTT, EDWARD C., AND R. M. TOWNSEND (1984): "Pareto Optima and Competitive Equilibria With Adverse Selection and Moral Hazard," *Econometrica*, 52, 21–46. [1806]
RILEY, JOHN G. (2001): "Silver Signals: Twenty-Five Years of Screening and Signaling," *Journal of Economic Literature*, 39, 432–478. [1806]
RUST, JOHN (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033. [1806]
SÁNCHEZ, JUAN M. (2018): "The Information Technology Revolution and the Unsecured Credit Market," *Economic Inquiry*, 56, 914–930. [1807]
SELTEN, REINHARD (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 25–55. [1806]
YAARI, MENAHEM (1965): "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, 32, 137–150. [1807]

---

*Co-editor Charles I. Jones handled this manuscript.*

*Manuscript received 31 July, 2020; final version accepted 17 February, 2023; available online 24 February, 2023.*