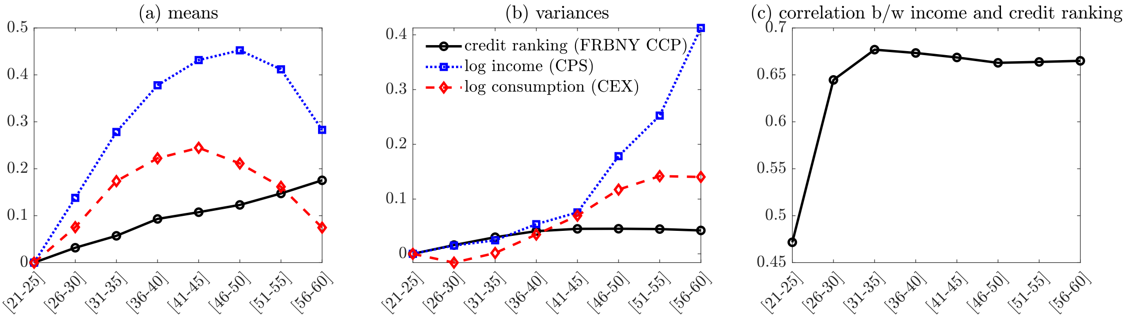
# Credit Scores and Inequality Across the Life Cycle

Satyajit Chatterjee (FRB Philadelphia)
Dean Corbae (Wisconsin and NBER)
Kyle Dempsey (Ohio State)
José-Víctor Ríos-Rull (Penn, UCL, and CAERP)

April 11, 2025
NBER Macro Annual

# Motivation: Credit Rankings, Income, and Consumption by Age




(a) credit rankings increase monotonically (vs. hump shapes in income, consumption)

(b) dispersion in credit rankings rises early in life alongside income, consumption

(c) correlation between income and credit ranking weakest early in life

# Our Questions

How does *reputational inequality* contribute to consumption inequality?

- how is reputational inequality (meas. by dispersion in credit rankings) affected by income inequality?
- how do signals about unobservable type contribute to reputational inequality?

What unobservables correlated with borrower risk are credit scores trying to capture?

- widely viewed as proxy for character or conscientiousness (the “C” in OCEAN).
- *challenges:* how to identify causal link between reputation and income?

How do changes in information technologies (e.g. big data) and / or changes in regulation of individual’s credit history contribute to consumption inequality?

- who are the winners and losers from pooling / separation of unobservable types?

# Framing the Question

What role does income and unobservable type play in repayment prob / credit scores?

Suppose, for individual $i$ in age bracket $n$:

$$S_{i,n} = \alpha_n Y_{i,n} + U_{i,n} \tag{1}$$

- $S_{i,n}$ is observable credit ranking
- $Y_{i,n}$ is observable income (i.e. higher income more likely to repay)
- $U_{i,n}$ is a noisy public signal of $i$'s unobservable type (i.e. an assessment or "type score" where higher $U_{i,n}$ is assessed to be the less risky or more conscientious type)

How might we explain the data patterns in Figure 1 using (1)?

- Is there correlation between $Y_{i,n}$ and $U_{i,n}$?

# Is it all income?

If so, how to explain strictly increasing mean credit rankings $\overline{S}_n$ versus hump shape of mean income $\overline{Y}_n$?

- can learning help explain rising $\overline{S}_n$ despite declining $\overline{Y}_n$ post 45?

# Is it all income?

If so, how to explain strictly increasing mean credit rankings $\overline{S}_n$ versus hump shape of mean income $\overline{Y}_n$?

- can learning help explain rising $\overline{S}_n$ despite declining $\overline{Y}_n$ post 45?

How to explain low correlation between credit rankings and income early in life?

$$\text{corr}(S_{i,n}, Y_{i,n}) \equiv \rho_{S_n,Y_n} = \frac{\alpha_n \sigma^2_{Y_n} + \sigma_{Y_n,U_n}}{\sigma_{S\,n}\sigma_{Y\,n}} \tag{2}$$

where $\sigma_{S_n} = [\alpha_n^2 \sigma^2_{Y_n} + 2\alpha_n \sigma_{Y_n,U_n} + \sigma^2_{U_n}]^{1/2}$.

- when $\sigma_{Y_n,U_n} = 0$ (our benchmark), correlation rises if income inequality rises more than learning early in life followed by more learning/separation later in life.

# Is it all income?

If so, how to explain strictly increasing mean credit rankings $\overline{S}_n$ versus hump shape of mean income $\overline{Y}_n$?

- can learning help explain rising $\overline{S}_n$ despite declining $\overline{Y}_n$ post 45?

How to explain low correlation between credit rankings and income early in life?

$$\mathrm{corr}(S_{i,n}, Y_{i,n}) \equiv \rho_{S_n,Y_n} = \frac{\alpha_n \sigma^2_{Y_n} + \sigma_{Y_n,U_n}}{\sigma_{S\,n}\sigma_{Y\,n}} \tag{2}$$

where $\sigma_{S_n} = [\alpha_n^2 \sigma^2_{Y_n} + 2\alpha_n \sigma_{Y_n,U_n} + \sigma^2_{U_n}]^{1/2}$.

- when $\sigma_{Y_n,U_n} = 0$ (our benchmark), correlation rises if income inequality rises more than learning early in life followed by more learning/separation later in life.

Learning is potentially an important factor in explaining the data.

# Outline

▸ literature

2 period adverse selection model (simple version of Chatterjee et al. (2023))

- 2 unobservable types: patient / conscientious (low risk) versus impatient (high risk).
- scorers use credit market actions to learn/ revise posterior about type
- unobservable discrete choice (EV) shocks avoid oep beliefs & cloud separation

# Outline

▸ literature

2 period adverse selection model (simple version of Chatterjee et al. (2023))

- 2 unobservable types: patient / conscientious (low risk) versus impatient (high risk).
- scorers use credit market actions to learn/ revise posterior about type
- unobservable discrete choice (EV) shocks avoid oep beliefs & cloud separation

Add moral hazard: unobservable effort choice allows for correlation between unobservable type and income growth

# Outline

▸ literature

2 period adverse selection model (simple version of Chatterjee et al. (2023))

- 2 unobservable types: patient / conscientious (low risk) versus impatient (high risk).
- scorers use credit market actions to learn/ revise posterior about type
- unobservable discrete choice (EV) shocks avoid oep beliefs & cloud separation

Add moral hazard: unobservable effort choice allows for correlation between unobservable type and income growth

Add medical shocks to illustrate the role of recent regulation excluding medical debt and delinquency from credit records

- quantifies positive and normative implications (pooling and reputational incentives)

# Baseline Adverse Selection Environment

$N = 2$-period lived, risk-averse households, heterogeneous with respect to

- unobservable type $\tau$: $\tau_n \in \{L, H\}, \mathbb{P}(\tau_1 = H) = \rho, \tau_2 \sim Q^\tau(\tau_2 \mid \tau_1)$
  - implement as **patience** ($\beta_n = \beta_{\tau_n}$) to capture "conscientiousness" (in $N = 2$, $Q^\tau = I$)
  - $\beta_H > \beta_L \implies$ $H$-type less likely to borrow and default (i.e. less risky)
- income $y_n$: 2-point process $\in \{y_L, y_H\}$, $y_2 \sim Q^y(y_2)$, $y_1 = y_L$ independent of type
- wealth $a_n$: short term risky asset, $a_1 = 0$ for everyone, $a_2 \in A$ a discrete set
- type score $s_n$: public noisy assessment of probability that $\tau_n = H$ ($s_1 = \rho$ for all)
- terminal reputation $\phi$: in $N = 2$ receive utility $\phi s_3$

Risk-neutral / deep-pocketed creditors, take risk free rate $r$ as given, make risky loans at price $q_n$

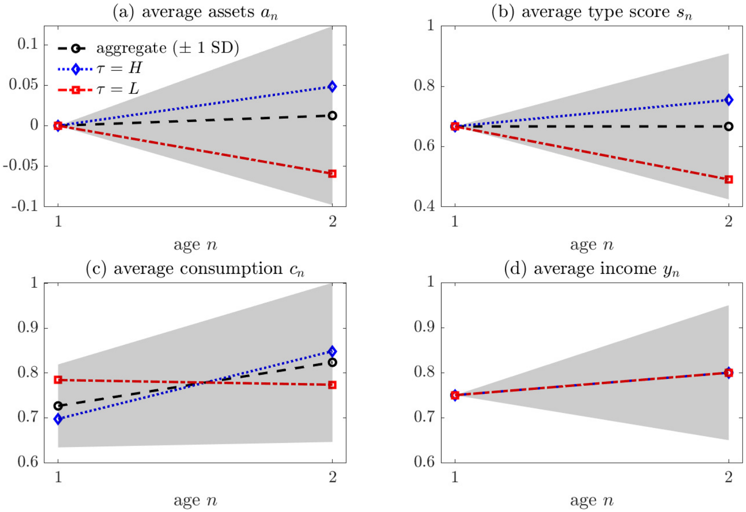
- $q_n$ depends on debt choice $a_{n+1} < 0$ and Bayesian risk assessments $s_n = \rho$.

Exogenous: $(\tau_n, y_n)$, Endogenous: $(a_n, s_n, q_n)$

# $N = 2$ Choices and Timing

1. individuals choose $a_2$ given price schedule (no recovery): $q_1(a_2) \equiv \frac{\mathbb{P}(d_2=0 \mid a_2)}{1+r}$
2. type assessment is updated according to Bayes' Rule: $s_2 \equiv \mathbb{P}(\tau = H \mid a_2)$
3. consumption: $c_1 = y_1 - q_1(a_2)$ yields utility $\mathbb{E}_{\epsilon^{a_2}}[U(c_1) + \epsilon^{a_2}]$
4. at the beginning of age 2, $y_2$ is realized: $y_2 \sim Q^y(y_2)$
5. if $a_2 < 0$, individual makes a default choice: $d_2 \in \{0, 1\}$
6. type assessment is updated via Bayes' Rule: $s_3 \equiv \mathbb{P}(\tau = H \mid d_2, y_2, a_2, s_2)$
7. consumption: $c_2 = y_2 + (1 - d_2)a_2$ yields utility $\mathbb{E}_{\epsilon^{d_2}}[U(c_2) + \epsilon^{d_2} + \phi s_3]$

# Adverse Selection Model Predictions




Divergence in patience yields:

(a) divergence in savings decisions at age 1

(b) leads to dispersion in type scores (mean differences and total dispersion)

(c) divergence in consumption paths

(d) despite no separation on income (to be relaxed)

# Information, Reputational Incentives and Inequality

What are the consequences of restricting information in credit records?

**Experiment:** assume type assessments $s_n$ are not tracked over time ("no tracking")

| | **baseline** | | **no tracking** | |
|---|---|---|---|---|
| | $H$ mean $-$ $L$ mean | std. dev. | $H$ mean $-$ $L$ mean | std. dev. |
| age 2 wealth | 0.11 | 0.11 | 0.02 | 0.09 |
| age 1 consumption | -0.09 | 0.09 | -0.01 | 0.05 |
| age 2 consumption | 0.07 | 0.18 | 0.01 | 0.15 |

- all differences due to reputational incentives
- tracking raises incentive to separate $\implies$ more wealth and consumption inequality

# Adding Moral Hazard to the Baseline

*How might unobservable type affect income?*

Replace exogenous earnings process $\mathbb{P}(y_2)$ with endogenous effort-driven earnings process:

- unobservable effort $e_1 \in \{0, 1\}$ at start of age 1 raises likelihood of $y_h$ at 2: $\mathbb{P}(y_h \mid e_1 = 1) > \mathbb{P}(y_h \mid e_1 = 0)$
- utility *cost* $\kappa$ incurred today
- *benefit:* higher future expected income

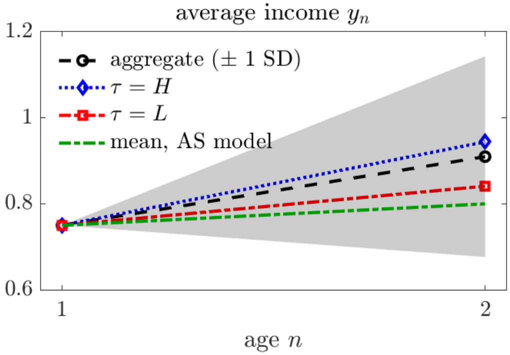
# Adding Moral Hazard to the Baseline

*How might unobservable type affect income?*

Replace exogenous earnings process $\mathbb{P}(y_2)$ with endogenous effort-driven earnings process:

- unobservable effort $e_1 \in \{0, 1\}$ at start of age 1 raises likelihood of $y_h$ at 2: $\mathbb{P}(y_h \mid e_1 = 1) > \mathbb{P}(y_h \mid e_1 = 0)$
- utility *cost* $\kappa$ incurred today
- *benefit:* higher future expected income

$\Longrightarrow$ $H$-types experience higher earnings growth, feedback from income to type assessment



# Adding Medical Expenditures to the Baseline

Survey evidence: unexpected medical expense shocks a major reason for delinquency

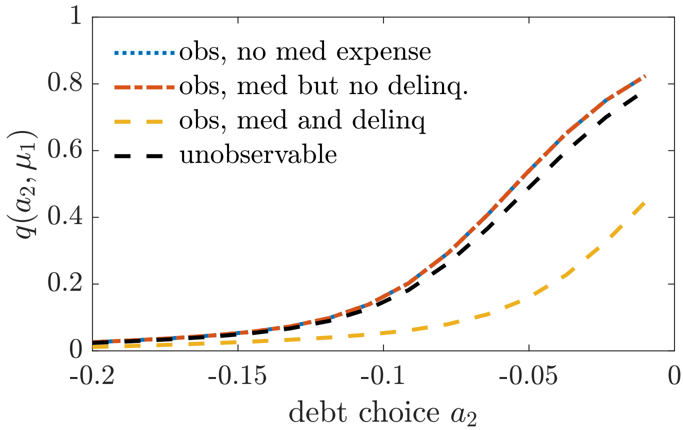
1. medical debts have little predictive value about borrowers' ability to repay others
2. consumers frequently report receiving inaccurate bills

$\Longrightarrow$ rationale for recent CFPB ruling to exclude medical bills from credit reports used and prohibit lenders from using medical information in their decisions.

Using our laboratory, we introduce medical expense shocks $m_1 \in \{0, \overline{m}\}$ with probability $\chi = \mathbb{P}(m_1 = \overline{m})$ independent of type (i.e. pure bad luck).

- delinquency decision $\delta_1 \in \{0, 1\}$ conditional on medical shock may signal one's type
- this in turn may affect terms of credit $\Longrightarrow$ feedback to reputational incentives

# Credit prices and Medical Info Restrictions




First order effect of removing medical records: increased ***pooling*** in credit market.

- helps those who have gone delinquent
- hurts the rest

# Welfare Effects of Restricting Information about Medical Shocks

Ex ante the welfare gain from going from the world of observable medical records to unobservable medical records is 5.6% for $H$-types and 2.4% for $L$-types

$\Longrightarrow$ both types prefer the world *with* information restrictions *despite pooling*

Why is restricting information preferable, even to type $H$?

- life cycle income growth $\Longrightarrow$ want to push consumption forward
- delinquency revelatory of being $L$-type *when its observable*
- $H$-types value of reputation at odds with desire to delay repayment
- medical DQ rate for $H$ / $L$ types: 0% / 4% observable vs 32% / 39% unobservable

# Summary: Information versus Reputational Incentives

Moving from "baseline" to "no tracking" and from observable to unobservable medical records ***both restrict information***.

But these two cases are very different in terms of ***reputational incentives***:

- baseline to no tracking *kills* them. Consider marginal benefit of saving more:

$$\text{saving: } \frac{\partial V_n}{\partial a_{n+1}} + \frac{\partial V_n}{\partial s_{n+1}} \frac{\partial s_{n+1}}{\partial a_{n+1}} \qquad \text{effort: } \left( \frac{\partial V_n}{\partial y_{n+1}} + \frac{\partial V_n}{\partial s_{n+1}} \frac{\partial s_{n+1}}{\partial y_{n+1}} \right) \frac{\partial y_{n+1}}{\partial e_n}$$

- red term absent in no tracking (also in full information / "big data" economy!)
- by contrast, observable to unobservable medical expenses merely *alters the inference*

***Key quantitative questions:*** how big is the marginal value of a better reputation? How elastic is reputation with respect to actions, income? Much data work left to do!

# Conclusions and Future Directions

**Framework** extends to $N > 2$ to match data, relax exogenous reputation (i.e. $\phi = 0$).

- can be used to evaluate the role of reputational inequality in income and consumption inequality over the life cycle
- provides a laboratory to evaluate the role that information technologies and restrictions on scoring play in welfare analysis

# Conclusions and Future Directions

**Framework** extends to $N > 2$ to match data, relax exogenous reputation (i.e. $\phi = 0$).

- can be used to evaluate the role of reputational inequality in income and consumption inequality over the life cycle
- provides a laboratory to evaluate the role that information technologies and restrictions on scoring play in welfare analysis

**Methodology:** simple, quantitative framework for dynamic hidden information problems

- *extreme value shocks*: no need for exogenous oep beliefs in Perfect Bayesian Eqm
- *block recursive*: don't need entire cross-sectional distribution to price risky debt
  - given observable default relevant state vector $(y_n, a_n, s_n)$; restricting information in a credit record can break it.

# APPENDIX

# Related literature

▸ back

**Unsecured credit:** Athreya (2002), Chatterjee et al (2007), Livshits et al (2007), Athreya et al (2008), Raveendranathan (2020), Herkenhoff and Raveendranathan (2024), many others

- *hidden info:* Chatterjee et al (2023), Livshits et al (2016), Sanchez (2018), Narajabad (2012)
- We add: moral hazard (feedback from income to type score)
- We add: observability restrictions (e.g. medical, Fulford and Low (2024), Lauer (2017))

**Credit Scores and Reputational inequality:** Poon (2007), Pasquale (2015), Golden et al (2016), Israel et al (2014), Albanesi and Vamossy (2019, 2024)

- We add: quantitative framework to link to credit scores, income, and consumption
- *what is a type?* OCEAN, personality economics (Heckman et al (2023), Ameriks et al (2017))

**Methods:** use dynamic discrete choice (Rust (1987)) to solve oep beliefs (Kreps and Sobel (1994))