Credit Scores and Inequality Across the Life Cycle: A Discussion

Stefania Albanesi, University of Miami, NBER and CEPR

June 12, 2025

1 Introduction

Credit scores play a key role in the allocation of consumer credit in the United States. Lenders report on the status of consumer loans to the three credit bureaus, Equifax, Experian and Transunion. The purpose of this reporting is to address issues of asymmetric information and the lack of exclusivity in consumer lending. Credit scores are then utilized to rank consumers according to their predicted default risk. These scores are derived from proprietary prediction models based on credit report data. The purpose of these credit scores is to provide a single, convenient proxy for the detailed information contained within a credit report. When a consumer applies for a new loan, the lender checks their credit score. Consequently, the size of the loan and the interest rates offered largely depend on this credit score, though lenders will typically also request income and asset information for larger loans, such as mortgages and vehicle loans.

Despite the importance of credit score in consumer credit, scholarly research on the role, properties and consequences of credit scores has been quite limited. The authors have sought to fill this gap in prior work (Chatterjee et al. (2023)) and with the current contribution, by developing quantitative equilibrium models of credit scores. In the models, credit scores are conceptualized as an informative and unbiased signal of each individual borrower's default risk, which depends on a fixed individual characteristic, the consumer's impatience proxied by their discount factor, as well as idiosyncratic income and expenditure shocks. It is assumed that lenders and borrowers rationally understand these scores, which are subject to Bayesian updating. The presence of these scores allows for a separating equilibrium in the credit market.

There is a notable gap between the properties of credit scores in their model and in reality. Empirical credit scores lack transparency in inputs and methods, and there are few publicly available performance metrics to evaluate their accuracy or fairness. Approximately 8% of consumers are unscored Brevoort, Grimm, and Kambara (2016), a group that disproportionately includes low-income individuals, minorities, and young people. There is substantial evidence indicating that current credit scoring systems display poor statistical performance and that the lack of performance is not evenly distributed in the population, as highlighted in research by Albanesi and Vamossy (2019), Albanesi and Vamossy (2024), Albanesi, DeGiorgi, and Nosal (2022), Blattner and Nelson (2021), and Di Maggio, Ratnadiwakara, and Carmichael (2022).

The current paper is based on the observation that credit scores display a life cycle pattern, that shows them increasing with age and income. A model is then developed to replicate these properties. I will discuss the empirical link between age, income and credit scores and review evidence suggesting that that link may be a function of poor performance of current credit scoring algorithms, rather than a desirable feature to emulate. Additionally, I will review some of the evidence that, including additional information, such as medical debt, does not add to the ability of credit scores to proxy default risk. I will conclude with some directions for further research.

2 What Do Credit Scores Measure?

Credit scores provide an ordinal ranking of consumers based on the probability that they will default on a debt obligation within a certain future horizon. In this context, a "default" is typically defined as being 90 or more days past due on any debt, and the "horizon" for this prediction is usually the next six to eight quarters.

The main credit scoring products in the market are FICO, which has been available since 1989, and VantageScore, introduced in 2006. It's noteworthy that the FICO score was exclusively allowed for Government-Sponsored Enterprise (GSE) backed mortgages from 1996 until 2019, while VantageScore has also been permitted for this purpose starting only in 2019.

Credit scoring models change over time. For example, FICO 10 and VantageScore 4.0 were current versions noted in 2023, though many financial institutions may not use the most recent model. Regarding what is publicly known about these proprietary models, the law mandates the disclosure of the four to five most important factors that drive variations in an individual's credit score. These factors generally include the length of credit history, the mix of credit types used, credit utilization ratios, payment history, outstanding balances, and recent credit demand.

2.1 Information and Credit Scores

There are three pieces of legislation that affect credit scores.[1] The Fair Credit Reporting Act (FCRA), enacted in 1970, determines what information can be included in a consumer's credit report, which is the source of information for credit scoring models. According to this legislation, credit reports should generally only contain information pertaining to an individual's debt history. Consequently, information regarding income, wealth, or occupation is not included, primarily due to concerns about potential inaccuracies, as this information

[1] Lauer (2017) provides a detailed account of the history of the development of consumer credit reports and credit scoring systems in the United States.

would be collected by lenders at the time of loan origination and would typically become stale over time. A central tenet of this act is the seven year time limit for which most adverse information can be included in a credit report.[2] Legislators aimed to strike a balance between lenders' need to assess risk and an individual's ability to recover from past financial distress. For consumers, knowing that negative marks will eventually be removed from their credit reports can incentivize individuals to work towards better financial habits and acknowledges that people's financial situations can change. For the credit reporting system, it helps ensure that the information used to evaluate creditworthiness is reasonably current and predictive of future behavior, since a person's financial habits from over a decade prior may have little bearing on their current ability to manage debt. Finally, the FCRA was part of a wave of civil rights legislation starting in the mid-1960s and one of its key goals was to prevent discrimination. For this reason, the act prohibits the inclusion of information on protected categories such as race, gender, marital status, religion, and age.

The Equal Credit Opportunity Act (ECOA) of 1974 primarily focuses on preventing discrimination in credit applications. This act introduced the legal notion of "disparate impact," which occurs if a lending practice has a disproportionate adverse effect on a protected population and lacks a sufficient, legitimate business justification. Additionally, the ECOA mandates that lenders must specify three reasons to an applicant explaining why their request for credit was denied.

The FCRA and the ECOA predate the widespread use of credit scores that started in the 1990s (Livshits, Mac Gee, and Tertilt (2016)), but in large degree shape the credit scoring system. As a result of the diffusion of credit scores in loan originations, the U.S. Congress passed the Fair and Accurate Credit Transactions Act (FACT) in 2003. This legislation deputized the Federal Reserve to study and report to Congress the effects of use of credit scoring on the availability and affordability of credit, the statistical relationship between

[2]Certain negative events are excepted from the seven year rule, including Chapter 7 bankruptcy, which can remain on a credit report for up to 10 years from the filing date, unpaid tax liens, and information on certain high-stakes decisions.

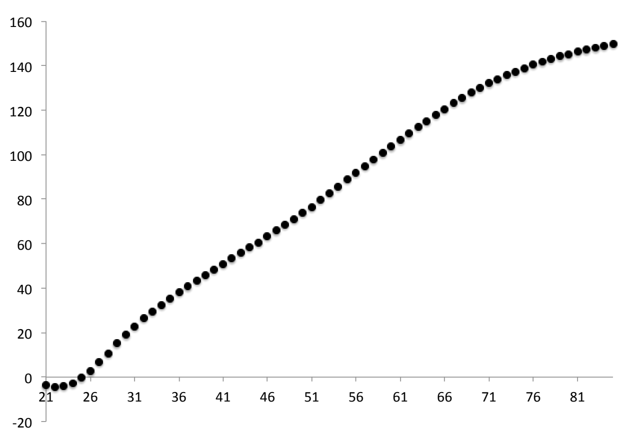
credit scores and quantifiable risks, and the extent to which the use of credit scores affects availability and affordability of credit to protected populations under ECOA. This led to the first comprehensive report on the role and performance of credit scores (FRB (2007)).

I will now discuss the evidence on the relation between credit score, age and income.

2.2 Variation of Credit Scores by Age and Income

Age is a protected category under the FCRA, similarly to gender or race. Hence, credit scores cannot depend on age, even if the date of birth is included in the credit report as a personal identifier. Yet, there is a strong positive relation between age and credit scores (Albanesi, DeGiorgi, and Nosal (2022)), as can be seen in figure 1, which plots estimated credit score age effects. Credit scores increase by more than 40 points between age 26 and age 41. Age effects are somewhat stronger for borrowers younger than 50 years old, though they are positive for the entire age range considered (21-85 years old). This pattern is a consequence of the positive relation between length of the credit history, which is an important factor in credit score variation and positively correlated with age, and credit scores. Age proxies such as length of the credit history are admissible under ECOA, since they are predictive (FRB (2007)), and therefore serve a legitimate business justification.

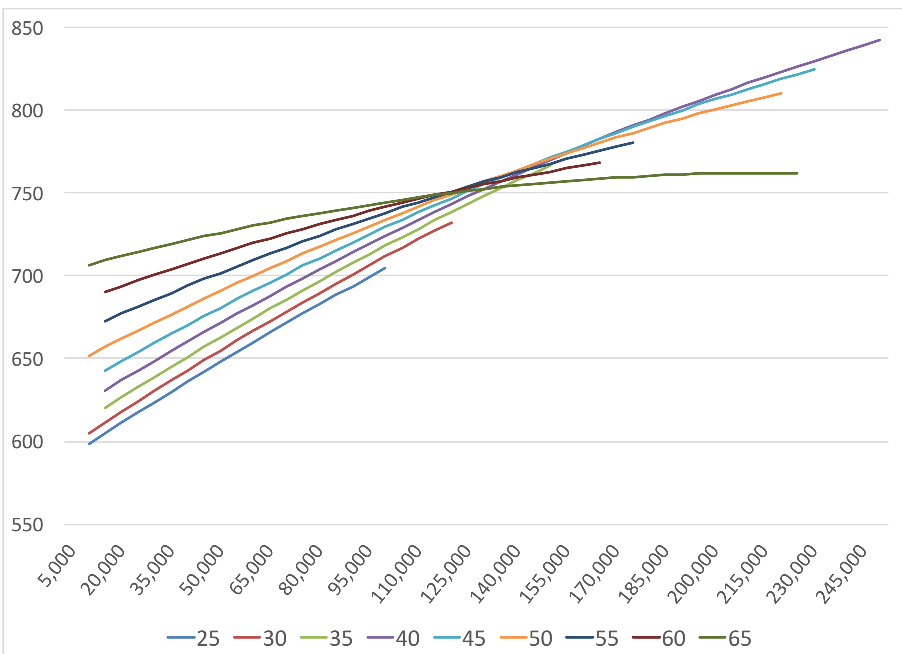
Figure 1: Age Effects in Credit Scores



The horizontal axis measures age in years. The vertical axis measures credit score points. The values are estimated age effects. Sample period 2001-2012. Source: Albanesi, DeGiorgi, and Nosal (2022).

Income is also positively related to credit scores (Albanesi, DeGiorgi, and Nosal (2022)). This can be clearly seen in figure 2, which plots the relation between credit score and income for different five year age groups. The relation between income and credit scores is positive for all age groups, but tends to be steeper for younger borrowers, with an increase in income from \$50,000 to \$65,000 associated with a rise in the credit score from 645 to 660 for consumers in the 25-29 age range and from 700 to 710 for those in the 55-59 age range. While income is not in the credit report, it is used by lenders in loan applications, particularly for vehicle loans, personal loans and mortgages. Credit mix, which is positively associated with credit scores and grows with income (Blattner and Nelson (2021), Albanesi and Vamossy (2024)), is another important factor in credit score variation, leading to a positive relation between income and credit scores.

Figure 2: Variation of Credit Scores By Income and Age



The horizontal axis measures income in 2009 USD. The vertical axis measures credit score points. The table reports estimates from a regression of credit score on income for 5 year age groups 25-29 year olds (25), 30-34 year olds (30) and so on. Source: Albanesi, DeGiorgi, and Nosal (2022).

The systematic variation of credit scores by age and income suggests that credit scores do not reflect an immutable trait, such as character, but rather the economic circumstances of the borrower, as captured by the information in their credit report. This is consistent with the seven year cut-off for information in the credit report, established by the FCRA, which was intended to prevent negative information to have overly persistent effects on borrowers' access to credit.

3 How Good Are Credit Scores?

The main goal of the quantitative model developed by the authors is to replicate the positive relationship between age, income, and credit scores. However, a key question is whether this positive relation is a desirable property, that is, whether it correctly reflects the distribution of default risk. Research by Albanesi and Vamossy (2024) evaluates credit score performance by building a machine learning model that substantially outperforms widely used credit scoring algorithms in ranking consumers by default risk. They use this model as a benchmark against which to evaluate certain properties of credit scores, shedding light on such observed correlations.

One of their main findings is that credit scores misclassify a substantial fraction of borrowers, that is they place them in a risk profile that is different from the one corresponding to their default risk. This is illustrated in Table 1, which stratifies borrowers by their credit score risk profile, using the standard industry categories– Deep Subprime, Subprime, Near Prime, Prime, and Super Prime– that are critical for the determination of credit limits and interest rates on consumer loans. The table reports in which risk profile the high performing machine learning model of default risk would place them. Values on the diagonal show the fraction of borrowers that the machine learning model would place in the same risk profile as the credit score, and off-diagonal items show the fraction of borrowers in each credit score risk profile who are misclassified. The last column shows the overall percentage of borrowers who are misclassified by the credit score for each credit score risk profile. For example, among borrowers with a Subprime credit score, the machine learning model places 53% in the Subprime profile, 15% in the Deep Subprime profile and 22% as Near Prime profile. Among borrowers with a Prime credit scores, only 66.6% should be in the Prime profiles, while 16% display default risk consistent with a Super Prime profiles, 13% with a Near Prime profile and 4% with a Subprime profile.

Table 1: Credit Score and Model Risk Profile Comparison

	Model Based Risk Profile					
Credit Score	Deep Subprime	Subprime	Near Prime	Prime	Super Prime	Misclassified
Deep Subprime	45.06	43.57	8.89	2.44	0.03	54.94
Subprime	14.67	52.71	22.46	10.07	0.08	47.29
Near Prime	1.66	38.05	30.35	28.08	1.86	69.65
Prime	0.17	4.43	12.83	66.50	16.07	33.50
Super Prime	0.02	0.05	0.15	25.71	74.08	25.92

Notes: Rows correspond to industry-defined credit score risk profiles. Columns are fractions in each corresponding model-based risk profile. Column "Misclassified" reports the fraction of borrowers in each credit score risk profile that would be placed in a different risk profile by ML model. All values in percentage. Industry defined risk profiles are determined by the following credit score ranges: Deep Subprime 300-499, Subprime 500-600, Near Prime 601-660, Prime 661-780, Super Prime 781-850. Time period 2006Q1-2016Q2. Source: Albanesi and Vamossy (2024)).

The differences in assigned risk profiles between the credit score and the machine learning model are significant, reflecting substantial variation in default risk, as reported in Table 2 for Subprime and Prime borrowers, who comprise the largest fraction of consumers. For example, Subprime borrowers have an average default rate of 44%. However, borrowers who are Subprime based on the credit score but Near Prime based on the higher performing machine learning model have a 17% realized default rate, whereas those who are Subprime according to the model have a 48% default rate. These findings indicate that credit scores' ability to separate borrowers based on their true likelihood of default is limited.

3.1 Explanatory Variables

Albanesi and Vamossy (2024) show that, among various factors highlighted by the credit scoring industry, payment history and outstanding balances are those most strongly associ-

Table 2: Default Risk Variation by Credit Score Profile

Risk Profile	Default Rate		Risk Profile	Default Rate	
Credit Score	Realized	Predicted	Model Based	Realized	Predicted
Subprime	43.68	43.10	Deep Subprime	92.48	92.46
			Subprime	48.25	47.59
			Near Prime	17.09	16.61
			Prime	8.27	7.20
			Super Prime	4.00	1.00
Prime	6.05	6.62	Deep Subprime	90.87	91.53
			Subprime	25.88	32.72
			Near Prime	13.17	15.17
			Prime	4.38	4.40
			Super Prime	0.92	0.91

Notes: Realized and model predicted default rates by credit score and model-based risk profile. Time period 2006Q1-2016Q2. Source: Albanesi and Vamossy (2024).

ated with the likelihood of default, and they account for 35% and 21%, respectively, in the variation in default risk rankings based on a high performing machine learning model, as reported in Table 3. For credit scores, payment history is also the most important factor, accounting for 40% of the variation in credit scores. However, current credit scoring models tend to overweigh credit utilization and age of accounts and credit mix, which are not strongly associated with default outcomes in the machine learning model. I will now discuss how this property of credit scores introduces a potential for *data bias* in their predictions.

Table 3: Relevance of Factors associated with Default Risk Rankings

Feature Group	Machine Learning Model	Credit Score
Payment History	0.35	0.40
Account Age and Credit Mix	0.12	0.21
Credit Utilization	0.12	0.20
Balances	0.21	0.11
Recent Credit	0.05	0.05
Available Credit	0.15	0.03

Notes: Fraction of the variation in default risk rankings explained for machine learning model and for credit score. Source: Albanesi and Vamossy (2024).

4 Performance and Standing by Age and Income

Data bias occurs when there are sources of noise that come from the underlying data that negatively impact the performance of credit scoring models (Blattner and Nelson (2021)). Compositional differences can be a source of this bias. The high weight placed on the length of the credit history and credit mix in credit scoring models contributes to this bias, given systematic differences in these attributes across the population.

This can be seen in Table 4, which reports the feature composition by demographic group, focussing on two dimensions: length of the credit history, where a short credit history corresponds to the category *Thin file*, and credit mix. Borrowers in the first quantile of the income distribution and borrowers younger than 30 years old are disproportionately likely to have thin files and a small credit mix. This unfavorable feature composition generates substantially lower performance of standard credit scoring models for these groups of borrowers.

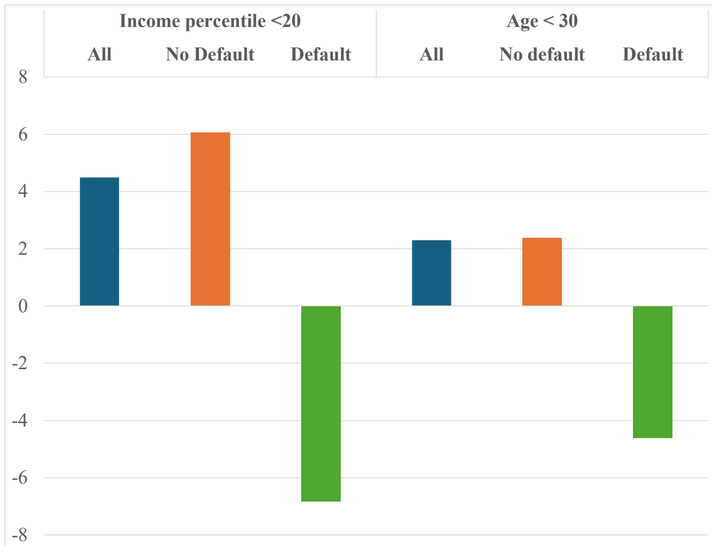
Table 4: Feature Composition by Demographic Group

		Demographic Group			
Feature Category	Status	$Income_{p20}$	$Income_{\geq p20}$	Age < 30	Age ≥ 30
Credit History	Thin file	0.93	0.39	0.89	0.39
	Thick file	0.07	0.61	0.11	0.61
Credit Mix	Small	0.95	0.61	0.91	0.62
	Large	0.05	0.39	0.09	0.38

Notes: Share of borrowers in each feature composition category by demographic group. Feature composition categories, with shares within each category adding up to 1. Time period 2006Q1-2016Q2. Source: Albanesi and Vamossy (2024).

Albanesi and Vamossy (2024) show that a high performing machine learning model substantially reduces the performance gaps for demographic groups with unfavorable feature composition, such as low income and young borrowers, thus reducing data bias. Additionally, they show that the default risk rankings based on the machine learning model improve

Figure 3: Differences in Default Risk Rankings by Demographic Group



Notes: The vertical axis measures the difference in percentile rankings if default risk for a high performance machine learning model relative to the credit score. Demographic groups are: consumers with income in the first quintile (*Income percentile* < 20), consumers younger than 30 years old (*Age* < 30). Source: Albanesi and Vamossy (2024).

the standing of young and low income borrowers, so that their rankings rise. This can be seen in Figure 3, which reports the differences in rankings between the machine learning model and the credit score by demographic group and by default status, within each demographic group. Borrowers in the lowest income quantile would gain on average 4.5 percentiles in ranking over the credit score, which roughly translates to 25 credit score points. Among those in this group who do not default, the increase in ranking relative to the credit score is 6 percentiles, while among those who do default, there is a decline in ranking of approximately 7 percentiles. For borrowers younger than 30 years old, the average increase in ranking is only 1 percentile. But for those who do not default it is two percentiles, or approximately 12 credit score points, while there is a decline in ranking relative to the credit score of 6 percentiles.

These results suggest two important lessons. First, the variation in credit scores by income and age in part reflects poor statistical performance and data bias in credit scoring models. Second, more data may not be needed to improve default predictions with a sufficiently sophisticated statistical technology. This observation is key for understanding the role of medical debt in credit scores.

4.1 Medical Debt

Medical debt was included in credit reports via medical collections. The rule to remove medical collections from credit reporting data took place in phases. The first phase, effective July 1, 2022, removed paid medical collections and unpaid medical collections that are younger than one year, to allow time for insurance payment processing. The second phase, implemented in April 2023, removed unpaid medical collections lower than $500. The third phase would eliminate all medical collections altogether, and is planned under a January 2025 decision by the Consumer Financial Protection Bureau.[3]

The authors argue that excluding medical debt from credit reports and therefore from credit scores lowers welfares and generates further inefficiencies in the allocation of credit, because it reduces the informativeness of credit scores in their model, and therefore their ability to separate between different types of consumers. There are both empirical and conceptual grounds to be skeptical of this result.

From an empirical perspective, there is overwhelming evidence that eliminating medical collections from credit reports will not have any practical effects on consumers. First, the Consumer Financial Protection Bureau came the conclusion that medical collections are often inaccurately reported (CFPB (2014)), and they could add noise to credit reports, eventually amplifying data bias.[4] VantageScore (2022) tests the impact on Vantage Score of removing

[3]The full explanation to the changes in medical collections included in credit reports is available here `https://www.consumerfinance.gov/about-us/newsroom/cfpb-finalizes-rule-to-remove-medical-bills-from-credit-reports/`.

[4]This is not the first initiative in which information has been removed from credit reports to improve reporting accuracy. In June 2016, Equifax, Experian and TransUnion announced a series of initiatives intended to enhance credit data reporting accuracy with respect to collection items. This led to the removal of civil judgments, substantial reduction in the number of tax liens, and a reduction in medical-related agency collections, specifically those less than 180 days old. VantageScore (2017) found that the performance and composition of models developed without this information delivers equivalent predictive insight.

medical collections, finding a negligible impact, since medical collections are not predictive of default. Finally, Duarte et al. (2025) conduct a comprehensive analysis of the impact of removing medical collections, both smaller and larger than $500, and find that medical debts are not meaningfully predictive of defaults, and conclude that their removal will not affect the allocation of credit.

From a conceptual standpoint, there are three main flaws in the argument around the role of medical debt in credit scores. First, the model does not consider the measurement error in medical debts that is one of the key reasons that lead to its exclusion from credit reports. Allowing for measurement error, the results on the role of the exclusion of medical debt would be much weaker in the model. Second, the paper emphasizes the distinction between medical debt and medical delinquency, where debt is not chosen, while delinquency is chosen and therefore informative. But this fails to be true in practice, since medical debts appeared in credit reports only as collections, and therefore already delinquent. Moreover, the collection amounts associated with a single medical bill tended to vary over time as insurance payments and cost changes were applied. This implies that the amount of the delinquency is not something that the borrower chooses or controls. Moreover, borrowers had the incentive to delay payment in order for the process of bill finalization to complete to avoid overpayment. In other words, the distinction between medical debt and medical delinquency while important in the model, is moot in practice. Finally, Immorlica (2025) show that with adverse selection, more information does not necessarily increase welfare when signals are imperfect.

5 Conclusions

The life cycle pattern in credit scores reflects the positive association between length of credit history and credit mix, two important factors in credit score variation that rise with age and income, and credit scores. However, the evidence summarized in my discussion suggests that these factors have weaker association with default than implied by widely used credit scoring algorithms. The excessive weight on length of credit history and credit mix in credit scores reduces the standing of young and low income borrowers below what would be justified by their default behavior. However, improvements in predictive technologies can ameliorate these limitations and also increase the standing of young and low income borrowers in default rankings. There is substantial evidence suggesting that medical debt is not predictive of default and its exclusion from credit reports will not affect default rankings or credit allocation.

The gap between the empirical behavior and salience of credit scores and their notional behavior in the model generates a tension that needs to be resolved in order to make progress. Given the additional complexities and frictions that credit scoring systems face in the real world and that are not present in the model, one cannot simply take the empirical properties of credit scores and assume that they are optimal and construct models that rationalize them, if those models do not incorporate realistic constraints on affecting the properties of credit scores.

From a theoretical standpoint it is important, and indeed critical, to study the information structure on consumer credit markets from a normative perspective, with the goal of generating an efficient and equitable allocation of credit. For example, in our current setting, credit bureaus store troves of sensitive, identifiable personal information, with the seven year rule and the other restrictions in the Fair Credit Reporting Act as the only limitation to this data hoarding. Given the emerging concerns over data privacy and safety[5], an important theoretical question is: what is the minimum amount of data that is needed for reliable default risk rankings? Second, I have noted an emerging body of evidence on the statistical underperformance of credit scoring algorithms. However, the credit scoring industry has been dominated for decades by just one major player with little oversight and limited incen-

[5] A major data breach at Equifax, one of the three credit reporting agencies in the United States, was publicly disclosed on September 7, 2017, revealing that the personal and financial data of an estimated 147 million Americans had been compromised. The sensitive information stolen by hackers included names, Social Security numbers, dates of birth, addresses and more.

tives to improve its predictive technology. The optimal industrial structure for information aggregators and credit scoring firms on consumer credit markets is another crucial research area, as well as optimal regulation in terms of disclosure and innovation incentives for these important players.

Normative questions on the information structures on consumer credit markets should be coupled with a positive analysis of the current system, including credit scores. For quantitative work specifically, properties of credit scores in models should reflect as much as possible their empirical counterparts. Measurement error is an important feature of consumer credit data and credit scores are plagued with data bias, which means that the distribution of perceived default risk systematically differs from the actual. This affects the allocation and price of credit across the population and may affect default outcomes and generate equilibrium effects. Such feedback mechanisms need to be incorporated into the modeling.

References

Albanesi, Stefania, Giacomo DeGiorgi, and Jaromir Nosal. 2022. "Credit growth and the financial crisis: A new narrative." *Journal of Monetary Economics* 132:118–139.

Albanesi, Stefania, and Domonkos F Vamossy. 2019, August. "Predicting Consumer Default: A Deep Learning Approach." Working paper 26165, National Bureau of Economic Research.

———. 2024, September. "Credit Scores: Performance and Equity." Working paper 32917, National Bureau of Economic Research.

Blattner, Laura, and Scott Nelson. 2021. "How costly is noise? Data and disparities in consumer credit." *arXiv preprint arXiv:2105.07554*.

Brevoort, Kenneth P., Philipp Grimm, and Michelle Kambara. 2016. "Credit Invisibles and the Unscored." *Cityscape* 18 (2): 9–34.

CFPB. 2014. "Consumer credit reports: A study of medical and non-medical collections." Technical Report, Consumer Financial Protection Bureau.

Chatterjee, Satyajit, Dean Corbae, Kyle Dempsey, and José-Víctor Ríos-Rull. 2023. "A quantitative theory of the credit score." *Econometrica* 91 (5): 1803–1840.

———. 2025. "Credit Scores and Inequality Across the Lifecyle." Technical Report, National Bureau of Economic Research.

Di Maggio, Marco, Dimuthu Ratnadiwakara, and Don Carmichael. 2022, March. "Invisible Primes: Fintech Lending with Alternative Data." Working paper 29840, National Bureau of Economic Research.

Duarte, Victor, Julia Fonseca, Divij Kohli, and Julian Reif. 2025, April. "The Effects of Deleting Medical Debt from Consumer Credit Reports." Working paper 33644, National Bureau of Economic Research.

FRB. 2007. "Report to the congress on credit scoring and its effects on the availability and affordability of credit." *Board of Governors of the Federal Reserve System.*

Immorlica, Nicole, Sztutman-Andre Townsend Robert. 2025. "Optimal Credit Scores Under Adverse Selection." Unpublished manuscript, MIT.

Lauer, Josh. 2017. *Creditworthy: A history of consumer surveillance and financial identity in America.* Columbia University Press.

Livshits, Igor, James C Mac Gee, and Michele Tertilt. 2016. "The democratization of credit and the rise in consumer bankruptcies." *The Review of Economic Studies* 83 (4): 1673–1710.

VantageScore. 2017. "Negative Data Suppression and Impacts on Credit Score Models." Technical Report, Vantage Score Solutions.

———. 2022. "Impact to VantageScore Credit Scores Due to Changes in Medical Collection Information Reporting." Technical Report, Vantage Score Solutions.