NBER WORKING PAPER SERIES

DATA AND WELFARE IN CREDIT MARKETS

Mark Jansen
Fabian Nagel
Constantine Yannelis
Anthony Lee Zhang

Working Paper 30235
http://www.nber.org/papers/w30235

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2023, Revised June 2023

We are grateful to Jacelly Cespedes, Tony Cookson, Eduardo Davila, Anthony DeFusco, Andreas Fuster, Andra Ghent, Zhiguo He, Kyle Herkenhoff, Ray Kluender, Christian Leuz, Alberto Manconi, Sanjog Misra, Matt Notowidigdo, Nicola Pavanini, Nathan Seegert, Alp Simsek, Huan Tang, Ansgar Walther, Jialan Wang, Wenyu Wang, Alexander Zentefis, and Wanran Zhao for thoughtful discussions and comments, as well as seminar participants at the University of Chicago, Yale School of Management, Cambridge University Judge School of Business, HEC Paris, INSEAD, Washington University in St. Louis, the University of Kentucky Gatton College of Business, the Swiss Finance Institute, the Swiss Federal Institute of Technology Lausanne, Durham University Business School, SAIF, HEC Lausanne, FIRS, the London Business School Summer Finance Symposium, the Swiss Society for Financial Market Research Conference, the Frankfurt Conference on Regulating Financial Markets, the Young Scholars in Finance Consortium, the Midwest Finance Association Conference, and the Minnesota Corporate Finance Conference. Greg Tracey provided superb research assistance. Calculated (or derived) based on credit data provided by TransUnion, a global information solutions company, through a relationship with the Kilts Center for Marketing at the University of Chicago Booth School of Business. TransUnion (the data provider) has the right to review the research before dissemination to ensure that it accurately describes TransUnion data, does not disclose confidential information, and does not contain material it deems to be misleading or false regarding TransUnion, TransUnion's partners, affiliates or customer base, or the consumer lending industry. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Mark Jansen, Fabian Nagel, Constantine Yannelis, and Anthony Lee Zhang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Data and Welfare in Credit Markets
Mark Jansen, Fabian Nagel, Constantine Yannelis, and Anthony Lee Zhang
NBER Working Paper No. 30235
July 2022, Revised June 2023
JEL No. D6,G20,G21,G28,G5,H81

ABSTRACT

We show how to measure the welfare effects arising from increased data availability. When lenders have more data on prospective borrower costs, they can charge prices that are more aligned with these costs. This increases total social welfare and transfers surplus from borrowers to lenders. We show that the magnitudes of the welfare changes can be estimated using only quantity data and variation in prices. We apply the methodology on bankruptcy flag removals and find that removing prior bankruptcy information substantially increases the social surplus of previously bankrupt consumers, at the cost of a modest decrease in total social welfare. This suggests that flag removals have low efficiency costs for redistributing surplus to previously bankrupt borrowers. We show how the framework can be extended to incorporate adverse selection and imperfect competition.

Mark Jansen
University of Utah
1655 E. Campus Drive
Salt Lake City, UT84112
mark.jansen@utah.edu

Fabian Nagel
University of Chicago
Booth School of Business
5807 S Woodlawn Ave
Chicago, IL 60637
fnagel0@chicagobooth.edu

Constantine Yannelis
Booth School of Business
University of Chicago
5807 S. Woodlawn Avenue
Chicago, IL 60637
and NBER
constantine.yannelis@chicagobooth.edu

Anthony Lee Zhang
University of Chicago
Booth School of Business
5807 S Woodlawn Ave
Chicago, IL 60637
anthony.zhang@chicagobooth.edu

1 Introduction

The past half century saw an explosion of available data to screen and score borrowers. In principle, increasing data availability should allow lenders to charge interest rates that are more aligned with borrowers' true risk, increasing the efficiency of credit allocation and social welfare. Yet changes in interest rates also shift the distribution of social surplus between lenders and borrowers of different risk levels. How can we quantify the effects of increased data availability on social welfare and the distribution of surplus between lenders and borrowers?

This paper builds a tractable framework that measures the welfare effects of increased data availability by treating changes in data availability as a form of *third-degree price discrimination*. The framework provides sufficient statistics to measure the effects of data availability on social welfare and the division of surplus. We then apply the framework in consumer lending markets and show that the removal of prior bankruptcy information from consumer credit reports substantially increases the social surplus of affected consumers at a low cost to social welfare.

Whether to allow data to be priced by lenders has been at the forefront of recent and past policy debates and actions.[1] For example, in 2022 the CFPB prohibited credit bureaus from reporting medical debt. Laws such as the General Data Protection Regulation (GDPR) in the EU and the Fair Credit Reporting Act (FCRA) in the US prohibit lenders from pricing on certain characteristics or sharing information without borrowers' consent. Many countries also require lenders to report to credit registries. Regulators have noted the tradeoffs between personalized pricing and redistribution. For example, while the European Data Protection Supervisor (2021) recommended allowing lenders to use some data for personalized pricing and "*clearly delineating the categories and sources of personal data that may be used for the purpose of creditworthiness assessment*," the regulator also noted that "*limiting ex ante the types of personal data that can be used for creditworthiness assessment, and consumer lending more broadly, to what is necessary and proportionate...would...help protect consumers from being targeted at moments*

[1]Lenders have screened borrowers based on informal data since antiquity (Calomiris and Neal, 2013). In modern times, Fair Isaac and Company (FICO) introduced credit scoring in 1958, and credit bureaus became prominent in the United States beginning in the 1960s. In the 2000s a number of more sophisticated machine learning techniques and alternative credit data were introduced including VantageScore. These advances likely had important welfare consequences, which can potentially be measured in our framework.

of vulnerability with unfair credit offers." Such regulatory scrutiny of personalized loan pricing came as banks increased their spending to $74 billion in 2022, an increase of 37% from 2017. The International Data Corporation estimates 2021 worldwide spending on consumer data and analytic services was $130 billion, growing at 12% through 2022.

To explore the welfare consequences of data provision, we analyze a model of household lending in which lenders acquire data on prospective borrowers. In the absence of data, lenders lose money on high-cost borrowers (i.e., riskier borrowers, who are more likely to default) and make money on low-cost borrowers. When lenders cannot differentiate between borrower types, some high-cost borrowers receive credit even though their willingness to pay is lower than the social cost of lending to them, and some low-cost borrowers do not receive credit even though their willingness to pay is higher than the social cost of lending to them. Acquiring data allows lenders to differentiate between and thus charge different prices to high- and low-cost borrowers.

The framework allows us to quantify the effects of data availability on social surplus, and the distribution of surplus between different kinds of prospective borrowers. Data availability has two effects. First, it increases total social welfare by making credit allocation more efficient: low-cost borrowers that are willing only to pay above their cost but below the pooling rate receive credit, while high-cost borrowers that are willing to pay below their cost do not receive credit. Second, it changes the distribution of surplus between borrowers: low-cost borrowers fare better, and high-cost borrowers fare worse. We show that, when the price changes induced by data availability are small, data removal is a fairly efficient transfer technology, in the sense that the social welfare effects of data removal tend to be small relative to the size of the surplus redistributed from low-cost to high-cost consumers. This is because the social deadweight loss scales quadratically with small changes in prices, whereas the consumer surplus scales linearly.

Our framework is derived from a realistic model of consumer lending markets, which builds on the cost curve approach developed in Einav, Finkelstein and Cullen (2010) and applied in DeFusco, Tang and Yannelis (2022). Unlike earlier cost curve approaches, our baseline model allows us to estimate the welfare consequences of data acquisition or removal using only variation in price and quantity data. The framework allows for (1) multiple periods, (2) defaults which may be costly to consumers, (3) arbitrary structures of lenders' costs, incorporating

for example default, collateral, funding and operational costs, and (4) intensive and extensive margins of loan demand – consumers can adjust loan size, as well as opt out of the loan market entirely.

In the framework, our core notion of social welfare is the total *partial equilibrium money-metric utility* of consumers. The standard justification for this notion of welfare, which holds in our setting, is that any allocation which does not maximize the money-metric surplus is *Pareto dominated* by an allocation which does, combined with lump-sum transfers redistributing surplus across agents. In practice, of course, regulators are often unable to implement these lump-sum transfers, for logistical and political-economy reasons. Taking into account these frictions, a regulator may be willing to choose a Pareto-dominated outcome, such as removing data from markets, in order to redistribute surplus across consumer groups. Analyzing money-metric surplus is still a useful benchmark, for such a regulator to understand quantitatively how much data removal deviates from the first-best outcome of allowing the market to fully use data in pricing, and redistributing surplus through lump-sum transfers. An appealing feature of this framework is that researchers can defer policy decisions to the regulator and need not impose an objective function for the regulator. Instead, the framework provides an intuitive trade-off in USD to the regulator who can decide how much surplus taken from one group is acceptable to redistribute one dollar to another. In contrast to typical welfare weights, regulators can thereby apply their own weights to surpluses in different groups.

While our framework makes a number of strong assumptions in the baseline model, its primary strength is that the welfare and transfer effects of data availability can be estimated with very limited data observed by the econometrician. All that is needed is to observe pre- and post-data availability prices and loan quantities. Intuitively, since we assumed markets are competitive, prices faced by consumers post-data availability fully reflect the social costs of lending to these consumers, taking into account default rates, imperfect collateral recovery rates, operational and funding costs of lending, and any other costs facing lenders. Using pre- and post-data prices and quantities allows us to estimate the slope of consumers' loan demand.[2] With marginal costs, demand slopes, and the size of the price change in hand, we can then compute the size of the consumer surplus trapezoids, and deadweight loss triangles,

[2]We can include more realistic features such as imperfect competition and adverse selection by adding only slightly more data, on loan charge-offs. Additional features can also be added to the simple baseline framework.

under the inefficient pre-data prices.

The framework is broadly applicable, and in an application we apply our methodology to a commonly studied event that leads to information removal under the Fair Credit Reporting Act (FCRA). FCRA requires that flags indicating the occurrence of consumer bankruptcy be removed after seven (ten) years for a Chapter 13 (Chapter 7) bankruptcy. Using administrative data from TransUnion, we find that flag removal leads to discontinuous increases in credit scores, a corresponding drop in interest rates on new loans, and an increase in loan volume. We focus on auto lending in our application.[3]

We use the framework to study social welfare loss and transfers in auto lending. Our theory suggests that data which creates "small" changes in prices produce large transfers relative to deadweight losses. How large or small are these price changes in the setting of bankruptcy flag removal, and how many dollars of social surplus are destroyed for each dollar transferred across groups? Flag removal results in a 17-point increase in credit scores, a 22.6 basis point reduction in interest rates, and an $18 increase in borrowing. Through the lens of our model, we find that bankruptcy flag removals transfer approximately $19 million to previously bankrupt consumers each year, at the cost of roughly $598,000 in social welfare. Thus, for each dollar of surplus transferred to previously bankrupt consumers, only $0.03 of social surplus is destroyed. Our results imply that flag removal is a relatively inexpensive and social efficient way to transfer surplus to previously bankrupt consumers.

Our baseline model rules out imperfect competition and adverse selection, resulting in particularly simple expressions with low data requirements. We show that both forces can be accommodated at the cost of increased data requirement to apply our framework. With imperfect competition, welfare can be quantified if the econometrician has data on costs, which is readily available in many settings. Surprisingly, there is a set of assumptions, which may be approximately satisfied in some settings, under which markups do not affect the results. In particular, if the elasticities of demand are equal across consumer groups, and if the markups charged to consumer groups are constant and unaffected by markups, then the welfare effects

[3]While our methodology is widely applicable to consumer credit markets, we focus on auto loans for two reasons. First, borrowing for an auto purchase is a more common household event than incurring student or mortgage debt. Second, data limitations prevent the measurement of interest rates or institutional details for credit card debt and student loans.

of data availability with imperfect competition are in fact identical to the baseline setting of perfectly competitive markets, so measurements assuming perfect competition produce correct welfare results.

We also show that adverse selection can also be accommodated in our measurement framework, if the econometrician observes data on costs, by estimating the relationship between prices and costs, in a manner similar to Einav et al. (2010) and DeFusco et al. (2022). However, we find little evidence of adverse selection in our setting, and we cannot reject that marginal costs are non-increasing in price. We also discuss how our results are affected by considering consumers' strategic decisions to declare bankruptcy, and by the informativeness of data on default rates.

This paper joins a growing body of work studying the impact of data in financial markets. Begenau, Farboodi and Veldkamp (2018), Farboodi and Veldkamp (2020) and Farboodi et al. (2019), He, Huang and Zhou (2020), and Liu et al. (2023) construct models of data and efficiency in lending markets. Nelson (2018) studies information and price regulation while Tang (2019) uses data from a Fintech to value privacy. Jones and Tonetti (2020) and Acemoglu et al. (2019) more generally study the economics of data. Fuster et al. (2022) analyze the distributional effects of increasing model complexity in credit markets. Our contribution to this literature is to develop a methodology to quantify the welfare benefits of data availability, and to show theoretically and empirically that – somewhat surprisingly – when data is not very informative about default rates, the transfer effects of data availability are large relative to the effects on aggregate welfare.

This paper also relates to a growing literature on the measurement of welfare in credit markets. DeFusco, Tang and Yannelis (2022) build upon Harberger (1964) and Einav, Finkelstein and Cullen (2010) and adapt a methodology from the insurance literature to estimate welfare losses from asymmetric information in consumer credit markets. Much of the existing literature (e.g., Herkenhoff, 2019) estimates or calibrates structural models. Dávila and Walther (2021) study corrective regulation of credit markets with imperfect instruments. This paper contributes to both the aforementioned strands of literature by building a novel theoretical framework and provides a tractable and easy-to-use methodology to measure the welfare effects of data acquisition in credit markets.

This paper also joins an empirical literature on the consequences of bankruptcy, and specifically bankruptcy flag removal. Musto (2004) is the first paper to study the empirical effects of flag removal, studying the impact on credit scores and default. Several studies including Herkenhoff, Phillips and Cohen-Cole (2021), Bos et al. (2018), Dobbie et al. (2020), and Gross, Notowidigdo and Wang (2020) study how bankruptcy flag removal affects credit access, employment, entrepreneurship, consumption, and other outcomes. A related paper is Liberman et al. (2019), who analyze the effect of information removal in Chile, applying a similar methodology based on Einav et al. (2010) to estimate the surplus effects of bankruptcy flag removal. Our results are complementary to Liberman et al. (2019). We build a microfounded model, illustrating the assumptions under which demand and supply curves can be interpreted in terms of consumer utility in the setting of credit markets. Our model also uses variation in interest rates, rather than quantity and default data. We also demonstrate that, while the welfare gains are positive, they are quantitatively small relative to the transfer effects, implying that flag removal may be an efficient tool for redistribution if the regulator has strong preferences for transferring surplus to previously bankrupt consumers.[4] A recent literature on household bankruptcy also links theory to empirics. Gross et al. (2021) study the economic consequences of bankruptcy, Indarte (2021) studies moral hazard and liquidity in bankruptcy, Argyle et al. (2022) study disparities, and Dávila (2020) provides a theoretical framework for optimal bankruptcy exemptions. Our main contribution here is to provide a new and tractable approach to measure the welfare effects of information acquisition and removal. This study applies a new methodology using bankruptcy flag removal as an application, and measures the welfare effects stemming from information removal. We find that there are relatively low social costs in terms of giving borrowers a fresh start in terms of removing the stigma of past

[4]There is also a largely theoretical literature studying the welfare effects of data sharing policies in credit markets, which mostly emphasizes different channels. Pagano and Jappelli (1993) analyze lenders' incentives to share information, and how information sharing can alleviate adverse selection. Vercammen (1995), Padilla and Pagano (2000), and Elul and Gottardi (2015) focus on how the threat of updating credit scores and thus interest rates dynamically affect borrowers' default incentives. Kovbasyuk and Spagnolo (2021) and Blattner et al. (2022) apply Bayesian persuasion methods to analyze optimal information structures in credit markets, taking into account adverse selection and dynamically evolving consumer types. Chatterjee et al. (2020) emphasize how credit scores are a substitute for dynamic reputation in incentivizing repayment. Relative to the literature, our paper emphasizes exogeneneous heterogeneity in the willingness-to-pay for credit; this is key to creating allocative efficiency gains from information revelation, which is a core focus of our analysis. In the baseline model, we abstract away from forces such as dynamic types, reputation, moral hazard and adverse selection, and deal with many of these effects in extensions and appendices.

bankruptcy.

This paper is further related to classic findings that third-degree price discrimination has ambiguous effects on social welfare (Schmalensee, 1981; Varian, 1985, 1989). Chen and Schwartz (2015) theoretically analyze price discrimination for a monopolist with information about costs, and also find that differential pricing tends to improve pricing more generally than in the classic case. This paper also relates to Dávila and Schaab (2021), which analyzes a general framework for welfare analysis of policy when agents are heterogeneous, decomposing welfare effects of policies into four components. In the language of Dávila and Schaab, our policies affect welfare mainly through the intertemporal-sharing and redistribution channels.

The remainder of this paper is structured as follows. Section 2 presents a model of data and consumer welfare, and connects this to empirical analysis. Section 3 presents an empirical application of the framework, using administrative data and the removal of bankruptcy flags. Section 4 extends the main framework to incorporate adverse selection and imperfect competition. Section 5 concludes.

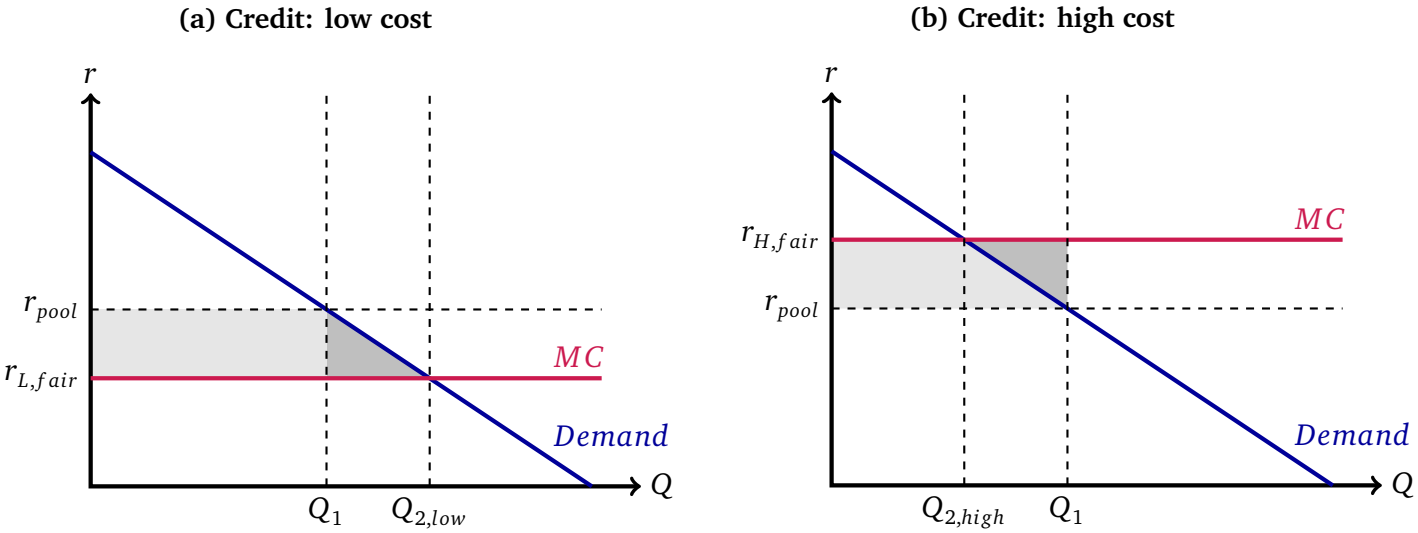
2 Data and Welfare

The model has two components. In Subsection 2.1, we discuss a simple intuition behind the welfare effects of price discrimination in credit markets, and how it differs from conclusions in classic markets. In Subsection 2.2, we build a more detailed model which can be mapped to data on consumer auto loan markets.

2.1 Intuition: Welfare Effects of Price Discrimination in Credit Markets

We think of changes in credit markets resulting from increased data availability as a form of *third-degree price discrimination*. Figure 1 shows a stylized example. Suppose there are two kinds of consumers, each with a different likelihood of defaulting and thus a different cost of lending. Suppose the markets are competitive, so firms set their prices equal to average cost. When firms do not have sufficient data to distinguish the two kinds of consumers, they set interest rates so that they break even on the average borrower cost, representing the rate r_{pool}. Firms thus lose money on the high-cost consumers in the right panel but make money on the

Figure 1: Price Discrimination in Credit Markets



This figure illustrates how third-degree price discrimination affects welfare in credit markets. Suppose there are two groups of prospective borrowers, with low costs (panel a) and high costs (panel b). The red lines show the cost of serving borrowers in each group, and the blue lines show borrowers' demand curves. Lenders are initially unable to distinguish between these borrowers, so set the pooled price r_{pool}. Once lenders are able to distinguish these prospective borrowers, they set $r_{L,fair}$ for the low cost group (a) and $r_{H,fair}$ for the high cost group (b). The dark gray shaded triangles illustrate the increase in social welfare for each group after the price change. In panel (a), the sum of the light gray shaded rectangle and the dark gray shaded triangle represents the increase in consumer welfare after the price change. In panel (b), the light gray shaded area shows the decrease in consumer welfare from the price change.

low-cost consumers in the left panel. This implies that there are two deadweight loss triangles, represented by the dark gray-shaded areas. The left panel shows that credit is underprovided to low-cost consumers: prices are higher than the costs of serving these consumers, so there exist some consumers who would receive credit in the social planner's optimum but not in equilibrium. The right panel shows that credit is overprovided to high-cost consumers. Prices are lower than the costs of serving these consumers, so credit is actually overprovided to them. They should not receive credit in the social planner's optimum, yet they are able to borrow in equilibrium, essentially because the cross-subsidy from low-cost consumers pushes prices sufficiently low.

Suppose now that lenders receive data that allows them to distinguish between the two kinds of consumers.[5] Thus, lenders can make the prices equal to the costs for each group:

[5] In the baseline model, we assume that there is perfect competition and that all gains are consumer surplus.

prices for the left group are $r_{L,fair}$ and costs for the right group are $r_{H,fair}$. Social surplus thus increases for both groups. Low-cost types that are willing to pay higher than their cost, but lower than r_{pool}, receive credit, and high-cost types that are only willing to pay below their cost do not receive credit. Both effects increase aggregate social surplus.

Thus, in settings where data allows lenders to distinguish the costs across borrower types, the intuition in Figure 1 suggests that data and increasing price discrimination will tend to increase social welfare. This is true as long as competitive markets ensure that prices are equal to the cost of serving borrowers, and as long as data is informative about costs. Although in Figure 1, we depict the two groups of borrowers as having the same demand for loans; however, this assumption is not important for the results, because data about demand does not affect prices in competitive markets. In our empirical application, we also do not impose the restriction that demand curves are equal across groups in our empirical application.

We contrast this case with the classic literature on third-degree price discrimination. There are two differences between the classic literature and our case: sellers have *market power*, and data is informative only about consumers' *demand*, and not about costs.[6] In this case, it is well-known that the welfare effects of third-degree price discrimination are ambiguous. Figure 2 shows the intuition in the classic case: we consider two submarkets in which consumers have the same cost, but different demand curves. When a monopolist firm is able to price discriminate, it will tend to lower prices for the low-demand group, and raise prices for the high-demand group. Since prices are above marginal costs for both groups, social welfare increases for the low-demand consumers and decreases for the high-demand consumers; the effect on total welfare is thus ambiguous.

In addition to changing social welfare, price discrimination also changes the distribution of surplus between different kinds of borrowers. Surplus increases for low-cost borrowers, but decreases for high-cost borrowers. The increase in consumer surplus on the left panel of Figure 1 is the sum of the light gray and dark gray areas, and the decrease in consumer surplus on the right plot is the light gray area.

A final point we emphasize is that, for small changes in prices, price discrimination tends to generate larger distributional effects than social welfare effects. The reason for this is that

In Section 4 we extend the framework to imperfect competition and discuss producer surplus.

[6] See, for example, Schmalensee (1981), Varian (1985), and Varian (1989).

the areas of the dark gray social surplus triangles in Figure 1 scale quadratically with price changes, whereas the consumer surplus areas, which are trapezoids, scale approximately linearly with price changes. That is, as r_{pool} approaches $r_{L,fair}$, the dark gray triangle and light gray rectangles both become small; however, the dark gray triangle shrinks faster than the light gray rectangle, so the *ratio* of the triangle area to the rectangle area – the amount of deadweight loss generated, per dollar of social welfare transferred to L-types – decreases towards 0. Thus, small changes in prices redistribute large amounts of surplus between low- and high-cost consumers, with relatively small effects on social welfare.

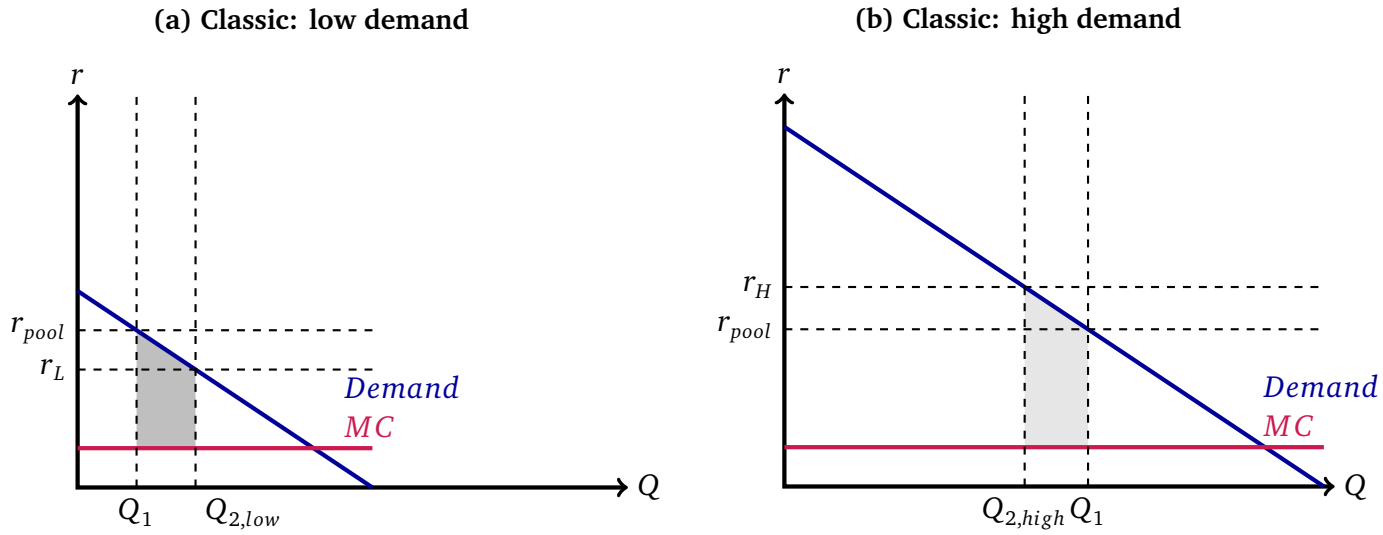
This finding has implications for the effects of data removal, such as the erasure of prior bankruptcy information from consumers' credit files. We find that the data removal policies have two effects: they transfer surplus from low-cost to high-cost consumers, and they lower social welfare, since consumers are pooled and credit allocation becomes less efficient. Our arguments imply that if the data removal-induced price changes are relatively small, the welfare losses will be small relative to the amount of consumer surplus redistribution. Thus, even though data removal decreases social efficiency in lending markets, they and other related policies may be a relatively cheap way, in terms of social costs, to transfer surplus to consumers at the margins of access to credit markets.

Before proceeding, we note that our partial-equilibrium approach of estimating consumer surplus and producer surplus, illustrated in Figure 1, corresponds to a particular way of weighting utility across different consumers: a dollar of surplus has equal weight for each consumer. In partial competitive equilibrium, social surplus is maximized. At any other point, it is possible to create a Pareto improvement by moving to the competitive equilibrium outcome, and redistributing money across consumers.[7] In practice, however, it is often practically and politically infeasible to implement lump-sum transfers around the competitive equilibrium outcome. Policymakers may thus be willing to implement policy interventions that transfer surplus across consumers, even if these interventions deviate from outcomes that maximize money metric social surplus. Bankruptcy flag removals are an example of such a policy.

Our framework can be used to quantify the policymakers' tradeoff between money-metric efficiency and redistribution. By comparing the size of the transfer to the size of the efficiency

[7]For a discussion of how partial-equilibrium consumer and producer surplus are derived from utility theory, see Chapter 4.3 of Jehle and Reny (2011).

Figure 2: Price Discrimination in Classic Markets



This figure illustrates how third-degree price discrimination affects welfare in classic markets. Suppose there are two groups of prospective borrowers—low demand (panel a) and high demand (panel b). The red lines show the cost of serving these borrowers, and the blue lines show borrowers' demand curve. Lenders are initially unable to distinguish between these prospective borrowers, so set price r_{pool}. After lenders are able to distinguish the two groups of borrowers, they set r_L for the low-demand group (as shown in panel a) and r_H for the high-demand group (panel b). The dark gray shaded area in panel (a) shows the welfare gain for the low-demand group, where prices decrease, and the light gray shaded area in panel (b) shows the welfare loss for the high-demand group, where prices increase.

loss in Figure 2, policymakers can measure how much money-metric surplus is destroyed, per unit of surplus that is transferred between the two groups of consumers.

Our approach has the added benefit of allowing the economist to be relatively hands-off as to the weights that are applied to welfare in any given setting. Given the results in any setting, policymakers can decide on the level of deadweight loss they are willing to accommodate in order to transfer surplus between groups. For example, if a policymaker took a stance that a dollar is worth 10% more to previously bankrupt consumers than other consumers, policymakers may be willing to adopt data removal policies as long as less than 10 cents of social surplus is lost, for each dollar transferred between groups.

2.2 Main Model

We now build a more detailed model, which allows for a number of realistic features: loans that last multiple periods; borrower default, which may be harmful; and intensive as well as extensive margins of demand: borrowers' loan size can adjust to interest rates, and borrowers can also choose to enter or exit the market entirely.

In the most general model, we assume that there are j groups of borrowers with different default rates, and that, in each group there are a number of borrowers whose demands for borrowing are different. We will build up to this case in three steps: first, we analyze the case with a single borrower; second, we show how to aggregate welfare across different borrowers with equal default rates; and third, we discuss how to compare welfare across borrower groups with different default rates.

2.2.1 Single-Borrower Case

Suppose there is a borrower who wishes to borrow money to finance consumption. The agent borrows L and spends in period $t = 0$, then pays back the loan in equal nominal payments over periods $t = 1 \dots T$. The prospective borrower's utility is:

$$\underbrace{u_0(c_0)}_{Purchase} + \underbrace{\sum_{t=1}^{T} \beta^t (1-\delta)^t u(c_t)}_{Payment} + \underbrace{\sum_{t=1}^{T} (1-\delta)^{t-1} \delta \sum_{\tilde{t}=t}^{T} \beta^{\tilde{t}} u(c_D)}_{Default} \tag{1}$$

In words, the borrower gets $u_0(c_0)$ upfront from consumption in period 0. We can think of $u_0(c_0)$ as representing utility from increased consumption in period 0. We assume $u_0(c_0)$ is concave: the marginal utility of consumption is decreasing.

Default. Default happens in each period at exogenous rate δ, independently across periods. We assume default can affect consumption in future periods: after default, the borrower receives c_D in each remaining period from t to T. Default might affect consumption, for example, because the borrower faces repossession, or because the borrower's credit score decreases, affecting the ability to borrow in the future.

The borrower's utility thus has three components. In period 0, the borrower gets $u_0(c_0)$. With probability $(1-\delta)^t$, the borrower reaches period t without defaulting, and receives $\beta^t u(c_t)$, where c_t is period-t consumption. With probability $(1-\delta)^{t-1}\delta$, the borrower defaults in period t; in this case, she receives $u(c_D)$ in each period from t to T, which is represented by the rightmost term in (1).

Payments. Since we have assumed loans have fixed payments, the payment in each period, when the principal is L and the interest rate is r, is:

$$\pi(L,r) \equiv L\frac{r(1+r)^T}{(1+r)^T-1} \tag{2}$$

Define $\phi(r)$ as the fraction of the principal that is paid off in each period, if the interest rate is r. That is,

$$\phi(r) \equiv \frac{\pi(L,r)}{L} = \frac{r(1+r)^T}{(1+r)^T-1} \tag{3}$$

Wealth. Let w_t represent the borrower's baseline wealth in period t. We assume, for simplicity, that the borrower is completely liquidity-constrained, so in the absence of saving instruments, she will consume w_t in each period t. If she borrows L, which corresponds to payment $\pi(L,r)$ in each period in the future, her consumption is thus:

$$c_0 = w_0 + L,\ c_t = w_t - \pi(L,r) \tag{4}$$

For periods $t > 0$, we adopt a linear approximation for $u(c_t)$:

$$u(c_t) \approx u(w_t) + u'(w_t)(c_t - w_t) = u(w_t) - u'(w_t)\,\pi(L,r) \tag{5}$$

This also implies that $u'(w_t)$, the marginal utility of consumption in period t, is fixed and exogenous. Essentially, we ignore concavity of the utility function in future periods.[8]

Combining (1), (4), and (5), we can thus write the borrower's optimization problem as:

$$V(r) = \max_L u_0(w_0 + L) + \sum_{t=1}^{T} \beta^t (1-\delta)^t \left[u(w_t) - u'(w_t)\,\pi(L,r)\right] + \sum_{t=1}^{T} (1-\delta)^{t-1}\,\delta \sum_{\tilde{t}=t}^{T} \beta^{\tilde{t}}\, u(c_D) \tag{6}$$

Let $L^*(r)$ denote the optimal loan size when the interest rate is r; that is,

$$L^*(r) = \arg\max_L u_0(w_0 + L) + \sum_{t=1}^{T} \beta^t (1-\delta)^t \left[u(w_t) - u'(w_t)\,\pi(L,r)\right] + \sum_{t=1}^{T} (1-\delta)^{t-1}\,\delta \sum_{\tilde{t}=t}^{T} \beta^{\tilde{t}}\, u(c_D) \tag{7}$$

Next, we characterize borrower surplus for an individual borrower.

Claim 1. *We have:*

$$V(r) = \bar{V} + \underbrace{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]}_{Utility\ weight} \underbrace{\left[\int_r^{\rho} L^*(\hat{r})\,\frac{d\phi}{d\hat{r}}\,d\hat{r}\right]}_{Payment\ change} \tag{8}$$

where ρ is the interest rate which causes the borrower to stop borrowing, and $\bar{V}$ is the borrower's utility from borrowing nothing and simply consuming wealth:

$$\bar{V} = u_0(w_0) + \sum_{t=1}^{T} \beta^t u(w_t) \tag{9}$$

[8] Effectively, we think of the borrowers' problem as smoothing a large expenditure in a single period $t = 0$, over a large number of future periods where the borrower has close to linear utility. The borrower's utility in the initial period is concave, so she has an incentive to borrow to increase her consumption in the first period, until the point where her marginal utility of consumption in the first period is equal to the sum of discounted marginal utilities in future periods, multiplied by the probabilities of reaching these future periods, multiplied by the payment fractions $\phi(r)$.

In difference terms,

$$V(r) - V(\tilde{r}) = \underbrace{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]}_{Utility\ weight} \underbrace{\left[\int_r^{\tilde{r}} L^*(\hat{r}) \frac{d\phi(\hat{r})}{d\hat{r}} d\hat{r}\right]}_{Payment\ change}$$

Claim 1 is useful because it justifies taking a "money metric" utility approach to borrower welfare. Expression (8) can be interpreted as follows. The marginal utility of a borrower for receiving an extra dollar, in each period $t = 1 \dots T$ if she does not default, is:

$$\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)$$

This is the sum of marginal utilities $u'(w_t)$, multiplied by discount rates β^t and the probability $(1-\delta)^t$ that the borrower reached period t without defaulting. Expression (8) then says that, if a borrower borrows at rate r, her utility is equal to what she would get if she did not borrow at all, $\bar{V}$, and then received a monetary payment in each period of the loan, of the amount:

$$\int_r^{\rho} L^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r}$$

This expression is essentially the borrower surplus triangle, for a single borrower.

For more intuition, consider a borrower who has $L^*(r) = K$; that is, the borrower has totally inelastic loan quantity for $r < \rho$, and borrows the same amount at any interest rate below ρ. Expression (8) then reduces to:

$$V(r) = \bar{V} + \underbrace{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]}_{Utility\ weight} \underbrace{\left[\int_r^{\rho} K \frac{d\phi}{dr} dr\right]}_{Payment\ change}$$

We can alternatively write this as:

$$V(r) = \bar{V} + \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t) [\pi(L,\rho) - \pi(L,r)]$$

The intuition behind this is, if the interest rate is ρ, so the borrower pays $\pi(L,\rho)$ in each period

$t = 1 \ldots T$, the borrower is just indifferent between borrowing and not borrowing; thus, her utility is $\bar{V}$. If the interest rate decreases to $r < \rho$, the borrower's payment in each period decreases to $\pi(L, r)$. Hence, the borrower's utility from borrowing must be equal to her utility from not receiving a loan, plus her utility from receiving a payment of:

$$\pi(L, \rho) - \pi(L, r)$$

in each period $t = 1 \ldots T$. Expression (8) generalizes this to the case where the borrower can have elastic loan size.[9]

2.2.2 Aggregation

Next, we show how to aggregate welfare across borrowers i within a given group j, of the same risk-type. These groups are defined by risk-type, so within the group j borrowers have the same default rate δ and hence are offered the same interest rate r and contracts with loan term T. Borrowers may differ in their utility functions $u_{i,0}(\cdot)$ for consumption in the first period; thus, some agents may have higher demand for borrowing than others. We assume there is a finite number n_j of potential borrowers in group j, indexed by i and for notational convenience, we will suppress the j subscripts in the remainder of this Subsection.

We aggregate utility across agents by assigning equal utility weight to a dollar in extra payments in each period, in every state of the world where the borrower does not default, for each borrower. This generalizes the classic idea of money-metric utility to our setting. Formally, we state this assumption as follows.

Assumption 1. *We will normalize agents' marginal utilities $u'\left(w_{i,t}\right)$, such that:*

$$\sum_{t=1}^{T} \beta^t (1-\delta)^t u'\left(w_{i,t}\right) = 1 \ \forall i \tag{10}$$

[9] Our approach is related to the certainty equivalent approach of Einav et al. (2010). The difference is, in our setting, payments are only made in state of the world without default. Since there is no component of loan contracts which is paid with full certainty, in order to apply a revealed-preference approach to consumer welfare analysis, we thus have to measure welfare in units of payments in non-default periods. Nevertheless, note that non-default period payments are just the unit of accounting. Default states are accounted for via the break-even rate (ρ).

We then define borrower surplus when the interest rate is r, $CS(r)$, as the sum of all agents' value functions:

$$CS(r) \equiv \sum_i V_i(r)$$

Claim 2. *Total borrower surplus satisfies:*

$$CS(r) - CS(\tilde{r}) = \int_r^{\tilde{r}} \Lambda(\hat{r}) \frac{d\phi(\hat{r})}{d\hat{r}} d\hat{r} \tag{11}$$

where $\Lambda(r)$ is aggregate loan volume at rate r:

$$\Lambda(r) = \sum_{i=1}^{n} L_i(r) \tag{12}$$

Next, we characterize lender surplus. Let r_{fair} be the fair interest rate for borrowers, at which lenders make zero profits.[10] r_{fair} is equal across borrowers in the group, since we have assumed borrowers in the group have equal default rates. Lender surplus from providing credit to borrowers at any other rate r, simply as the difference between $\phi(r)$ and marginal costs $\phi\left(r_{fair}\right)$, multiplied by the loan quantity $\Lambda(\phi(r))$:

$$PS = \Lambda(\phi(r))\left(\phi(r) - \phi\left(r_{fair}\right)\right) \tag{13}$$

It is important to note that lender surplus and borrower surplus have the same units: one unit of lender surplus is equal to one extra dollar in each period, paid in states of the world in which the borrower does not default. Adding borrower and lender surplus, total surplus is thus:

$$TS = \sum_{i=1}^{n} \left[\bar{V}_i + \int_{r_i}^{\rho} L_i^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r} + L_i^*(r_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right) \right] \tag{14}$$

These definitions of borrower, lender, and total surplus hinge crucially on the normalization in Assumption 1. This normalizes agents' marginal utilities, so that an extra dollar in each non-default period of the world has equal welfare weight for all borrowers and lenders. This is the

[10] Note that, if lenders make zero profits, prices to borrowers reflect all costs that lenders face for providing credit to borrowers: funding costs, losses upon default, non-payment of interest upon default, logistical costs, and other costs. We do not need to take a stance on the exact breakdown of these costs. In Section 4.1 we extend the framework to a setting with imperfect competition.

classic concept of partial-equilibrium money-metric surplus mapped to our setting: borrowers' welfare weights are calculated in terms of dollar willingness-to-pay. This concept is the basis for the methodology in Einav, Finkelstein and Cullen (2010) to study health insurance markets, and DeFusco, Tang and Yannelis (2022) to study borrower lending markets.

As we discussed towards the end of Subsection 2.1 above, the standard justification for this is that any allocation which maximizes the sum of money-metric utility across agents Pareto dominates any other allocation rules, with appropriately chosen transfers across agents. This property generalizes to our setting: in Appendix A.3, we show that social welfare as defined in (14) is maximized by setting $r_i = r_{fair}$, and also that for any rates $r_i \neq r_{fair}$, it is possible to change rates to r_{fair}, and construct a set of transfers from a social planner paid to borrowers and lenders in non-default states of the world, such that all borrowers and lenders are weakly better off. The normalization in Assumption 1 can be thought of as measuring how much welfare is lost relative to a first-best world in which the social planner can implement the Pareto dominant outcome of $r_i = r_{fair}$, and then use lump-sum transfers to redistribute surplus across agents.

Of course, for various logistical and political-economy reasons, regulators rarely implement such lump-sum transfers in practice. Regulators who are concerned about the distributional consequences of data policy may thus be willing to deviate from the Pareto-dominant optimum towards outcomes with better distributional properties; our core empirical finding will be that the welfare cost of such deviations in the setting of bankruptcy flag removal, compared to the amount of surplus redistributed between consumer groups. We use the idea of partial-equilibrium money-metric surplus, reflected in Assumption 1, as a benchmark for a regulator to measure how poorly imperfect surplus transfer tools perform, relative to the theoretical first-best outcome of maximizing money-metric surplus and redistributing using non-distortive lump-sum transfers.[11]

[11] In the empirical application in Section 3, we evaluate separately the effects of bankruptcy flag removal on total money-metric welfare, and the division of surplus. If we were willing to take an explicit stance on the relative welfare weights assigned to regular borrowers and previously bankrupt borrowers, we would not need to calculate these numbers separately: we could simply calculate the effects of bankruptcy removal on the weighted sum of welfare across borrowers. However, these conclusions would be sensitive to the particular welfare weights chosen. By separately calculating the magnitude of the transfer and the change in money-metric utility, our conclusions are robust to different weights: regulators could in principle decide on how much money-metric surplus loss they are willing to tolerate for every dollar of surplus transferred across groups, and interpret our results accordingly.

2.2.3 Estimation

Next, we bring our framework to data with multiple groups of borrowers and different default rates. As in the stylized example in Subsection 2.1, suppose we observe demand for multiple groups of borrowers with different default rates, indexed by j. Each group of borrowers faces two different price points: r_{pool}, which is the interest rate borrowers face when data is not available, so their pricing is pooled with other borrowers with different default rates; and $r_{j,fair}$, which is the interest rate in competitive markets when data is available and borrowers are not pooled. $r_{j,fair}$ thus reflects the social marginal cost of lending to group j borrowers. We have assumed expected default rates are constant within borrower groups j, so $r_{j,fair}$ is also constant for all borrowers within each group j borrowers. We assume loan demand is linear in the payment $\phi(r)$; that is, loan demand is:[12]

$$\Lambda(\phi(r)) = a - b\phi(r)$$

When there are multiple groups of borrowers with different default rates, we can simply calculate borrower, lender, and total surplus separately for each group using expressions (11), (13), and (14). For each group j, Appendix A.3 shows that welfare is maximized by setting rates equal to $r_{j,fair}$, and any other interest rate are Pareto dominated with transfers.[13] The following claim characterizes lender profits, borrower surplus, and total welfare.

Claim 3. *Suppose data becomes available to distinguish these groups, and the interest rates facing these borrowers changes from the pooled interest rate r_{pool} to the group-specific interest rates $r_{j,fair}$,*

[12] Depending on the empirical variation, in some applications nonlinear demand curves could be estimated and deadweight loss can be computed using numerical integration. The assumption that demand is linear greatly reduces the amount of data required to estimate our model. If demand were nonlinear, the dark gray welfare regions in Figure 1 may have different sizes. A lower bound is that the dark gray regions would both be zero, if demand is very convex for the low-cost group and concave for the high-cost group. An upper bound is that the regions are double the size of the triangles, if demand is concave for the low-cost group and convex for the high-cost group. Thus, the total change in welfare from making data available is at most twice the effect we find under the assumption of linearity, and may be as low as zero. However, the assumption of linearity has a small effect on the imputed transfers from making data available, since the size of the light gray regions in Figure 1 are insensitive to the convexity of the demand curve.

[13] Aggregating welfare across borrower groups with different default rates is somewhat subtle, however, because welfare for group j is measured in terms of non-default periods for group j, and the expected number of non-default periods a loan will last for differs across borrowers with different default rates. In our empirical application in Section 3, we show an approach to adjust for this by computing the expected number of non-default periods per groups.

where $r_{j,fair}$ is the marginal cost of providing credit to group j borrowers. Lenders' profit from borrower group j increases by the net amount:

$$\Lambda\left(\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{j,fair}\right)-\phi\left(r_{pool}\right)\right) \tag{15}$$

Borrower surplus for group j increases by:

$$\left(\phi\left(r_{pool}\right)-\phi\left(r_{i,fair}\right)\right)\left(\frac{\Lambda\left(\phi\left(r_{i,fair}\right)\right)+\Lambda\left(\phi\left(r_{pool}\right)\right)}{2}\right) \tag{16}$$

Total social welfare increases by:

$$\frac{1}{2}\left(\phi\left(r_{j,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\Lambda\left(\phi\left(r_{pool}\right)\right)-\Lambda\left(\phi\left(r_{j,fair}\right)\right)\right) \tag{17}$$

Using the expressions in Claim 3, we can quantify the intuitions in Figure 1. In the two-group example, when data is made available, allowing lenders to charge different prices for high- and low-cost borrowers, this induces a transfer of surplus across the two groups. When data is made available, prices fall for low-cost borrowers and rise for high-cost borrowers. Expression (16) allows us to quantify the transfer across borrower groups, and (17) gives the change in social surplus.[14]

Claim 3 also implies that the data requirements for implementing our model empirically are very low. These quantities can be measured in any setting where new data is introduced into a lending market, and the econometrician observes pre- and post-data availability prices, $\phi\left(r_{j,fair}\right)$ and $\phi\left(r_{pool}\right)$, and quantities, $\Lambda\left(\phi\left(r_{j,fair}\right)\right)$ and $\Lambda\left(\phi\left(r_{pool}\right)\right)-\Lambda\left(\phi\left(r_{j,fair}\right)\right)$, for each group of consumers; these quantities are likely to be observed in most settings where data is introduced or removed from a market. The main benefit of our simplifying assumptions is that they result in welfare expressions which are very simple to bring to data, so welfare analysis is relatively straightforward in settings where our assumptions are likely to approximately hold. However, in Section 4, we show that our framework can easily be extended to

[14]Note that our empirical analysis will focus on the change in borrower surplus for each group, (16), and social welfare, (17); lenders' profits in (15) will not appear in our analysis. This is because we will assume lenders are competitive and make zero profits, so when pricing is pooled, their profits on the low-cost group of borrowers are exactly offset by their losses on the high-cost group: by assumption, lenders' profit changes, (15), will sum to zero across groups.

empirically accommodate a number of forces we have disregarded in the baseline model (e.g., imperfect competition and adverse selection). We show that the qualitative conclusions of the model largely continue to hold, and the econometrician simply needs to observe more data to account for these forces.

2.2.4 Discussion of Model Features

Our framework allows borrowers to be harmed when default occurs. In the baseline model, we think of borrowers as receiving some consumption amount c_D if they default; this can be thought of as modeling, for example, the repossession of valuable assets, impacts on borrowers' credit scores, wage garnishment, and other ways that borrowers are harmed by default. We model c_D as invariant over the course of the loan in the baseline model, but this can easily be relaxed. Intuitively, losses from default do not affect our borrower surplus accounting, because anticipated losses are incorporated into ρ, the max rate at which borrowers are willing to borrow. If c_D decreases, so default harms borrowers more, then ρ will decrease: borrowers will require a lower rate to borrow in the first place, but the result of Claim 2 still holds. This is particularly useful, since it is generally difficult to measure precisely the effects of default on borrowers' welfare.[15]

Our framework is also robust to different sources of costs lenders may face for making loans. There are many possible sources of lenders' costs: lenders may lose money when borrowers default, if collateral values are not high enough to cover the outstanding debts; collateral liquidation may be costly; and lenders face funding and operational costs for originating loans. We do not need to take a stance on the source and structure of these costs, because they are all accounted for in $\phi\left(r_{j,fair}\right)$, the interest rate set by competitive lenders for group j consumers when data is available. Lenders' profits on group j when data is not available are measured simply as the gap between the actual rate $\phi\left(r_j\right)$ and the zero-profit rate $\phi\left(r_{j,fair}\right)$. Since we do not need to directly measure costs, our framework can be applied to both collateralized and

[15]Note that this argument does not generalize easily to settings where adverse selection is a significant concern, since borrowers' average default rates may vary as the interest rate changes. In such settings, borrower surplus could be calculated using an envelope-formula approach: the marginal effect of lowering the interest rate slightly on borrower surplus depends on the probability of default at that interest rate. We discuss how our framework could be extended to accommodate adverse selection briefly in Subsection 4.2. These concerns do not seem first-order in our empirical application, since we find empirically that interest rate changes do not have a statistically significant effect on default rates.

uncollateralized debt contracts; we do not need to take a stance on the structure of lenders' costs upon default, as these are all included in the competitive rate $\phi\left(r_{j,fair}\right)$.

Our framework also allows loan demand to be affected both by intensive and extensive margins. Both effects may be relevant in practice: data changes affect both the fraction of borrowers that borrow, and the size of loans. When bringing the model to data, we ignore loan rejections. Rejections are equivalent to lenders offering infinite prices. In principle, some prospective borrowers may be rejected from all lenders, and are thus unable to borrow at any price. Our methodology cannot capture welfare for these hypothetical borrowers, as we cannot estimate their willingness-to-pay. In many settings, this is unlikely to be a concern as some lenders specialize in riskier borrowers, and the vast majority of borrowers can get a loan offer at some price. In other words, we assume that borrowers can get a loan offer from some lender at some finite price. This is particularly the case in our empirical application to the auto loan market.

We briefly discuss how our welfare criterion relates to Dávila and Schaab (2021), who construct a general framework for welfare comparisons with heterogeneous agents. Dávila and Schaab decompose welfare effects of policies into four components: aggregate efficiency, risk-sharing, intertemporal sharing, and redistribution. The main component of the decomposition which is relevant in our setting is the intertemporal sharing component: the marginal rate of substitution of borrowers between period-zero and future period consumption generally differs from the rate at which lenders can transform period-zero consumption into future period consumption, so there are gains from pricing credit in a way that allows borrowers to smooth consumption over time.

Though it likely plays a smaller role, the redistribution channel of the Dávila and Schaab decomposition is also nonzero in our setting. To see this, consider an allocation which is not socially optimal, and consider data policies which change prices and allocations marginally. Under the normalization in Assumption 1, the planner has no value for transferring a unit of consumption between agents in non-default future periods. However, on the margin, consumers' marginal rate of substitution between period-zero consumption and future period consumption will generally differ; hence, there are in general welfare gains from transferring units of consumption between borrowers in period zero. Thus, transferring a unit of consumption

between borrowers in all periods will generally change social welfare, which corresponds to the redistribution component of the Dávila and Schaab decomposition. The aggregate efficiency component is not relevant in our setting, since we study policies which transfer consumption across borrowers and lenders, but do not change total consumption. Since we only study transfers in non-default periods, the risk-sharing component is also not relevant in our setting.

3 Empirical Application

3.1 Bankruptcy Flag Removals

We apply the methodology discussed in the previous Section on data commonly used to assess consumer credit. In this setting, bankruptcy flags must be removed from credit records after ten years under the Fair Credit Reporting Act (FCRA). There are two main types of consumer bankruptcy in the U.S., Chapter 7 (liquidation) and Chapter 13 (reorganization). Chapter 7 bankruptcy flags are typically removed 117 to 118 months after filing, while Chapter 13 bankruptcy flags are typically removed seven years after filing.

From the perspective of policymakers, bankruptcy flag removals represent a tradeoff between maximizing efficiency and redistributing surplus to previously bankrupt borrowers. As bankruptcy flags are a negative attribute in many credit scoring models, individuals are likely to receive a sharp discontinuous increase in their credit scores when flags are deleted. This transfers surplus to previously bankrupt borrowers, allowing them to attain interest rates similar to comparable consumers who never underwent bankruptcy. However, since lenders have less information on previously bankrupt borrowers, they set prices that differ from marginal costs, decreasing aggregate money-metric welfare. Using our methodology, we quantify the effects of flag removal on money-metric welfare as well as the distribution of surplus, to quantify the tradeoff facing policymakers in this setting.

To estimate the effect of bankruptcy flag removals on outcomes, we estimate variants of the following specification,

$$y_{it} = \gamma_c + \gamma_t + \delta^y \mathbb{1}[FlagRemoved] + \beta X_{it} + \varepsilon_{it} \tag{18}$$

where y_{it} are outcomes for individual i in month t, γ_c are cohort fixed effects, γ_t are calendar period fixed effects, X_{it} are individual controls and ε_{it} is an error term which we assume is uncorrelated with $\mathbb{1}[FlagRemoved]$, conditional on observables. We cluster standard errors at the level of the month in which the bankruptcy flag is removed. $\mathbb{1}[FlagRemoved]$ is an indicator of whether bankruptcy flags have been removed. The main coefficients of interest are the δ^y terms, which identify the difference in the outcome y_{it} following the removal of information.

We explore four outcomes, VantageScore credit scores, interest rates, loan amounts and charge-offs. We observe credit scores and loan quantities in all time periods. We only observe interest rates and charge-offs conditional on contracting. Exploring whether a loan becomes delinquent conditional on contracting is consistent with the theory presented in Section 2, as a loan cannot become delinquent and be charged off if there is no contract. Given that we observe interest rates conditional on contracting, we additionally assume that these reflect the lowest available offer rates for a contract.[16]

The coefficient δ^y captures the difference in the outcome y_{it} under the assumption that nothing else changes other than the removal of information on prior bankruptcies. More precisely, we assume that flag removal is orthogonal to the error term ε_{it}. A potential concern is that due to individuals having flags removed at different points in time, estimates for δ^y may be biased by individuals leaving the control group and treatment effect heterogeneity over time (Goodman-Bacon, 2021; Barrios, 2021). To address this concern, we first provide sharp graphical evidence of breaks when flags are removed, and then implement modern dynamic difference-in-difference estimators in Appendix E.

To provide graphical evidence that the observed effects are indeed driven by flag removals, we further estimate an event-study regression to evaluate the identifying assumption using the following variant of equation (18),

$$y_{it} = \gamma_c + \gamma_{ts} + \sum_{t=-6}^{6} \delta_t \{e_{it} = t\} + \beta X_{it} + \varepsilon_{it} \tag{19}$$

[16] We focus on the case where the bankruptcy decision is already determined, and focus on allocative effects. Appendix B.1 discusses the case where flag removal impacts the filing decision.

where γ_c are cohort-month and γ_{ts} are year-month by score bucket fixed effects.[17]

We plot the coefficients δ_t, along with a 95% confidence interval. The coefficients capture the difference in an outcome in each month before and after flag removal. We exclude the relative time dummy for period -1 as well as relative time dummy for period -6 due to collinearity arising from the age-period-cohort problem common in similar specifications.[18]

3.2 Data

To implement the analysis, we use the Booth TransUnion (TU) Consumer Credit Panel.[19] The data is an anonymized 10% sample of all TU consumer credit records from 2009 to 2020. We restrict the sample to the 2009-2018 period to allow at least two years for delinquency realizations after account openings. The sample is a panel. Individuals in the initial sample have their information updated annually, and each year a new 10% sample of first time borrowers is added. At a monthly frequency, the data contain basic information about borrowers and loans, including the original balance, the current balance, VantageScore credit scores, scheduled payments, the maturity of the loan, geography and importantly bankruptcy flags.

Our main outcomes of interest relating to the welfare framework are credit scores, interest rates, new loan balances and charge-offs. We do not directly observe interest rates, but back these out from scheduled payments using the amortization formula.[20] To avoid selection concerns, we predict interest rates for all individual-month observations. We predict interest rates using a 3rd order polynomial of current and up to 9 months lagged scores, time and cohort dummies.[21] We measure balances as the sum of balances of new auto loan accounts opened by an individual in a given month, and as zero when no account is opened by the individual in a given month. Charge-offs are measured as whether a loan becomes charged-off within 2

[17] We measure score buckets by sorting individuals into one of 20 score buckets in the month before flag removal and hold the sorting constant throughout the 13 months observed.

[18] We ensure that our results are not dependent on this design choice implementing a stacked difference in difference comparing individuals with flag removal to individuals with bankruptcy flag removal 12 to 17 months later in Appendix E.

[19] These data, along with similar credit panel data, are described in more detail in Keys et al. (2020) and Yannelis and Zhang (2021).

[20] Specifically, we let the monthly payment $A = \frac{P \times i}{1-(1+i)^{-n}}$, where P is the principal, n is the maturity, and i is the interest rate. We use a root-solving algorithm to solve for i. Note that we use scheduled, and not actual payments to construct interest rates.

[21] Appendix D shows the effect of flag removal on observed interest rates with fundamentally similar conclusions.

years of account opening. We collapse the data to the borrower-month level, and restrict to individuals who ever had a Chapter 7 or Chapter 13 bankruptcy flag and observations within 6 months around flag removal.

Table 1: Summary Statistics

This table displays basic summary statistics for the main analysis variables: the mean, median, and standard deviation. The first three columns show the statistics for the full sample, the next triplet for individuals pre-flag removal, and the final three columns show summary statistics post-flag removal. Source: TransUnion.

	Full			Pre			Post		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Credit Score	630.67	638.00	83.76	620.63	631.00	76.68	639.28	643.00	88.47
Interest Rate	8.78	7.66	4.31	9.02	7.96	4.30	8.57	7.43	4.30
Quantity Opened Cond. on Opening	20783	19484	10773	20332	19031	10495	21143	19842	10976
Quantity Opened	326.38	0.00	2915.35	306.50	0.00	2792.55	343.41	0.00	3016.52
If Charged-off	0.02	0.00	0.15	0.02	0.00	0.15	0.02	0.00	0.15

Table 1 presents summary statistics for our main analysis variables. There are 865,499 individuals in our final sample over 13 months totaling 11,251,487 individual-month observations. In 1.57 percent of individual-month observations at least one auto-debt account is opened, and the average opening balance conditional on opening is $20,783. Hence, the average monthly account opening is $326.38. The average rate is 8.78%. The average credit score in our sample is 631. Table 1 also shows summary statistics pre and post flag removal. Following flag removal, mean credit scores increase, interest rates decrease and borrowing amounts increase. Of course, these changes may be due to both bankruptcy flag removals, and secular time trends, which motivates our empirical strategy.

3.3 Results

3.3.1 Effects of Flag Removal

We begin with graphical evidence showing point estimates of equation (19). The figures show specifications including cohort and score bucket by year-month fixed effects. Figure 3 shows estimates of the coefficients δ_t, where the outcomes are credit scores, interest rates, and loan quantities. The top panel shows the VantageScore credit score, the middle panel shows interest rates at origination, and the bottom panel shows loan quantities. Consistent with prior work, we see a very sharp increase of almost 20 credit score points following the flag removal. This

translates into a reduction in borrowing costs. Interest rates show a clear drop following the flag removal. The change in scores and corresponding drop in rates seen in Figure 3 gives us the variation needed to estimate welfare losses. Consistent with the drop in interest rates, we see a sharp rise in loan volumes. There is an approximately $20 increase in auto loan openings.

We next quantify the visual results in a regression framework. Table 2 presents variants of equation (18). Column (1) includes a linear time trend. Column (2) adds time period fixed effects, while column (3) adds in cohort fixed effects, based on the bankruptcy filing date. Column (4) includes both cohort and year-month fixed effects, and finally column (5) includes cohort and score bucket by time period fixed effects. We measure score buckets by sorting individuals into one of 20 score buckets in the month before flag removal and hold the sorting constant throughout the 13 months observed. In the top panel, the outcome is the VantageScore credit score, in the middle panel the outcome is interest rates, while in the bottom panel it is loan volumes.

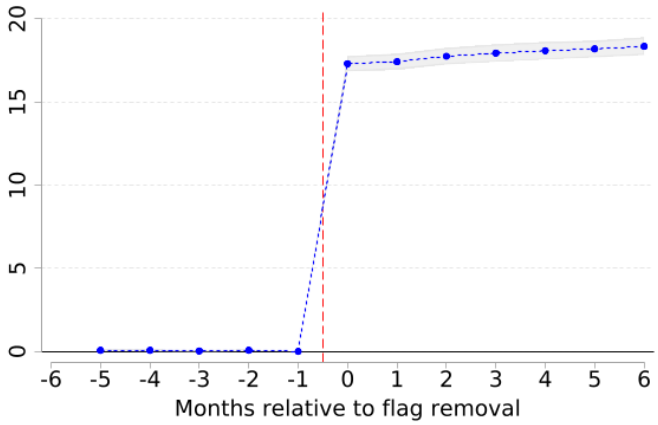
The results are broadly in line with the graphical evidence. The top panel of Table 2 indicates that bankruptcy flag removals lead to a 17.1 to 17.2 point increase in credit scores, or an approximate 2.76% increase in credit scores. The middle panel indicates that this is associated with a 21 to 23 basis point decrease in interest rates, or a 2.4-2.5% decrease in interest rates. The bottom panel shows that average new auto loan balances increase by $17.7 to $18.4, or 5.8-6%. This translates to an absolute price elasticity of approximately 2.3 to 2.5, depending on the specification. In the majority of and in our preferred specification, the effect is significant at the 1% level. In Columns (1) and (3), the effect size for loan quantities is significant at the 5% level.

3.3.2 Welfare Estimates

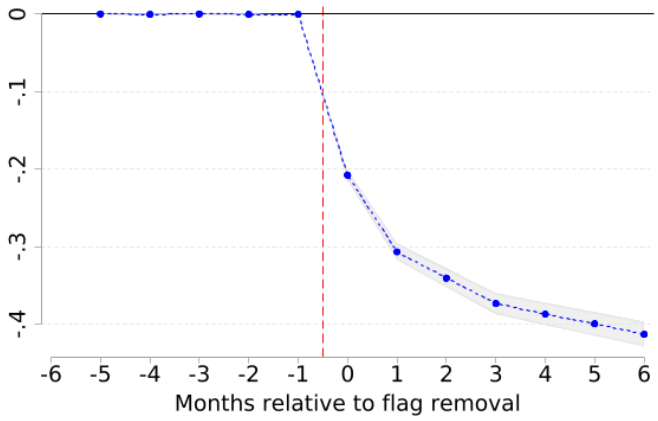
An appealing feature of our approach is that we can directly map regression coefficients into model parameters connecting our theoretical framework and its empirical application. Using our estimates, we can answer the following counterfactual: how does surplus for each type of borrower, and total social welfare, compare to a counterfactual in which bankruptcy flags were never removed from borrowers' credit records? We will estimate welfare by assuming that, previously bankrupt borrowers in counterfactual would attain the same outcomes as previously

Figure 3: Credit Scores, Interest Rates, and Loan Balances

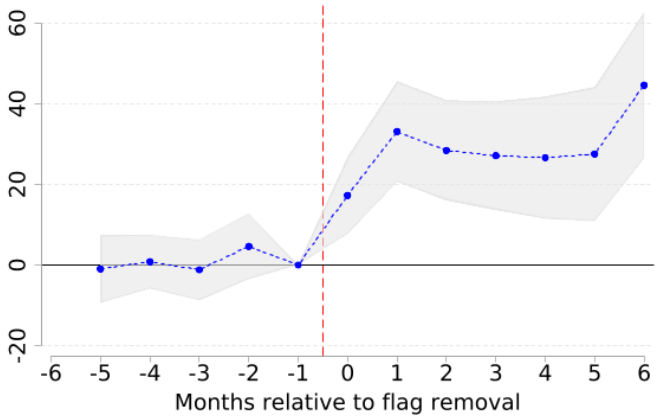
Panel A: Credit Scores



Panel B: Interest Rates



Panel C: Loan Volumes



This figure shows estimates of the coefficients δ_t from the following specification $y_{it} = \gamma_c + \gamma_{ts} + \sum_{t=-6}^{6} \delta_t\{e_{it} = t\} + \beta X_{it} + \varepsilon_{it}$, along with a 95% confidence interval. In Panel A, the outcome y_{it} is credit scores, while in Panel B it is interest rates. In Panel C the outcome is loan volumes. γ_c are cohort fixed effects, and γ_{ts} are time period by score bucket fixed effects. Standard errors are clustered at the cohort level. Source: TransUnion.

Table 2: Credit Scores, Interest Rates, and Loan Volumes

This table shows estimates of the coefficients δ^y from the following specification $y_{it} = \gamma_c + \gamma_t + \delta^y \mathbb{1}[FlagRemoved] + \beta X_{it} + \varepsilon_{it}$. In the top panel, the outcome y_{it} is the Vantage Score, in the middle panel the outcome is interest rates, while in the bottom panel it is loan volumes. Interest rates are predicted with a polynomial of current and past credit scores, period, and cohort fixed effects. γ_c are cohort fixed effects, and γ_t are time period fixed effects. Standard errors are clustered at the cohort level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Source: TransUnion.

	(1)	(2)	(3)	(4)	(5)
Panel A: Credit Scores					
Post	17.216***	17.147***	17.216***	17.125***	17.118***
	(0.298)	(0.258)	(0.298)	(0.239)	(0.237)
Constant	620.075***	621.165***	620.075***	621.448***	621.452***
	(0.488)	(0.497)	(0.188)	(0.129)	(0.128)
Observations	11,251,487	11,251,487	11,251,487	11,251,487	11,251,487
Adjusted R^2	0.012	0.017	0.017	0.018	0.813
Panel B: Interest Rates					
Post	-0.218***	-0.226***	-0.218***	-0.226***	-0.226***
	(0.018)	(0.006)	(0.018)	(0.004)	(0.004)
Constant	9.109***	8.979***	9.109***	8.898***	8.898***
	(0.096)	(0.024)	(0.018)	(0.002)	(0.002)
Observations	11,251,487	11,251,487	11,251,487	11,251,487	11,251,487
Adjusted R^2	0.003	0.044	0.044	0.046	0.878
Panel C: Loan Volumes					
Post	18.344**	17.926***	18.344**	17.979***	17.793***
	(6.270)	(4.049)	(6.270)	(4.017)	(4.010)
Constant	299.363***	310.897***	299.363***	316.696***	316.796***
	(7.473)	(2.901)	(3.098)	(2.163)	(2.159)
Observations	11,251,487	11,251,487	11,251,487	11,251,487	11,251,487
Adjusted R^2	0.000	0.001	0.001	0.001	0.002
Linear Time Trend	Yes	Yes	Yes	No	No
Year-month FE	No	Yes	No	Yes	No
Cohort FE	No	No	Yes	Yes	Yes
Year-month by Score Bucket FE	No	No	No	No	Yes
Clustered SE	Cohort	Cohort	Cohort	Cohort	Cohort

bankrupt borrowers in our data just before bankruptcy flags are removed. They would face persistently higher prices, reflecting their higher default risk, rather than being pooled with never-bankrupt borrowers. On the other hand, never-bankrupt borrowers become separated from previously bankrupt borrowers, so prices for never-bankrupt borrowers would decrease.

Figure 4 illustrates our results intuitively. The horizontal red lines show the break-even payments lending to high $\phi(r_{H,fair})$ and low cost $\phi(r_{L,fair})$ types, respectively. When bankruptcy flag are removed for high cost types, the payment drops to the pooling payment $\phi(r_{pool})$. The increase in borrower surplus for previously bankrupt individuals is illustrated by the light gray trapezoid. Because willingness to pay is below marginal costs for some borrowers, this results in a social welfare decrease from credit overprovision to high cost types, illustrated by the dark gray triangle. When high-cost types are pooled with low-cost, never-bankrupt borrowers, prices are also somewhat above costs for never-bankrupt borrowers, as the right panel shows. This leads to a loss in borrower surplus for the low-cost borrowers, as well as a loss in social welfare, illustrated by the dark gray triangle in the right plot. In a world where bankruptcy flags were never removed, total social welfare would increase by the size of the two dark gray triangles. The surplus of never-bankrupt borrowers would increase, by the size of the light gray rectangle plus the dark gray triangle in the right panel. The surplus of previously bankrupt borrowers would decrease, by the size of the light gray trapezoid in the left panel. Since we assumed lenders are competitive, lenders would make zero profits in both cases. Thus, in our framework, flag removal is an imperfectly efficient way to transfer surplus from never-bankrupt borrowers to previously bankrupt borrowers. We proceed to describe how we calculate welfare and borrower surplus changes formally; the quantitative results are summarized in Table 3.[22]

Let r_{flag} be the interest rate facing previously bankrupt borrowers with the bankruptcy flag, just prior to flag removal; this corresponds to $r_{H,fair}$ in the theoretical model, the fair interest rate for type-H high-cost, previously bankrupt borrowers. Let r_{pool} be the rate facing previously bankrupt borrowers just after flag removal, when they are pooled with never-bankrupt borrowers. When flags are kept on credit reports forever, previously bankrupt borrowers simply see an increase in prices from r_{pool} to r_{flag}. We set the pre-flag-removal interest rate to $r_{flag} = 9.02\%$,

[22] To map the model to the data, we measure time in months, since auto loans are paid monthly. We will measure loan volume as the total dollar volume of new auto loans opened each month; as in the model, this captures both the extensive and intensive margins of loan volume adjustment.

the average interest rate for borrowers before their flags are removed in the data. We set the post-flag-removal rate r_{pool} to $9.02\% + \delta^{InterestRate}$. The pre-flag-removal payment fraction, $\phi\left(r_{flag}\right)$, is 2.077% of the original loan balance per month, and the post-flag-removal payment, $\phi\left(r_{pool}\right)$, is 2.066%.

For loan quantities, we observe that roughly 18.1% of individuals open a loan each year, and the average loan size conditional on opening a loan is \$20,332. Thus, we set the annual average loan volume per borrower to $\Lambda\left(\phi\left(r_{flag}\right)\right) = \$3,678$. We set the post-flag-removal loan quantity to:

$$\Lambda(\phi(r_{pool})) = \Lambda(\phi(r_{flag})) + 12\delta^{LoanVolume}$$

where $\delta^{LoanVolume}$ is the coefficient on loan quantities from our differences-in-differences specification; we multiply $\delta^{LoanVolume}$ by 12 because we run the loan volume regression at a monthly level. Quantitatively, loan quantities increase from $\Lambda(\phi(r_{flag})) = \$3,678$ to $\Lambda(\phi(r_{pool})) = \$3,891.5$ per borrower.

With our estimates of $\phi\left(r_{flag}\right), \phi\left(r_{pool}\right), \Lambda\left(\phi\left(r_{flag}\right)\right), \Lambda\left(\phi\left(r_{pool}\right)\right)$ in hand, we can then use expressions (16) and (17) from Claim 3 to calculate the changes in total welfare and borrower surplus from previously bankrupt borrowers. Graphically, the borrower surplus change corresponds to the light gray trapezoid spanned by repayment fractions and loan sizes, in the left panel of Figure 4.

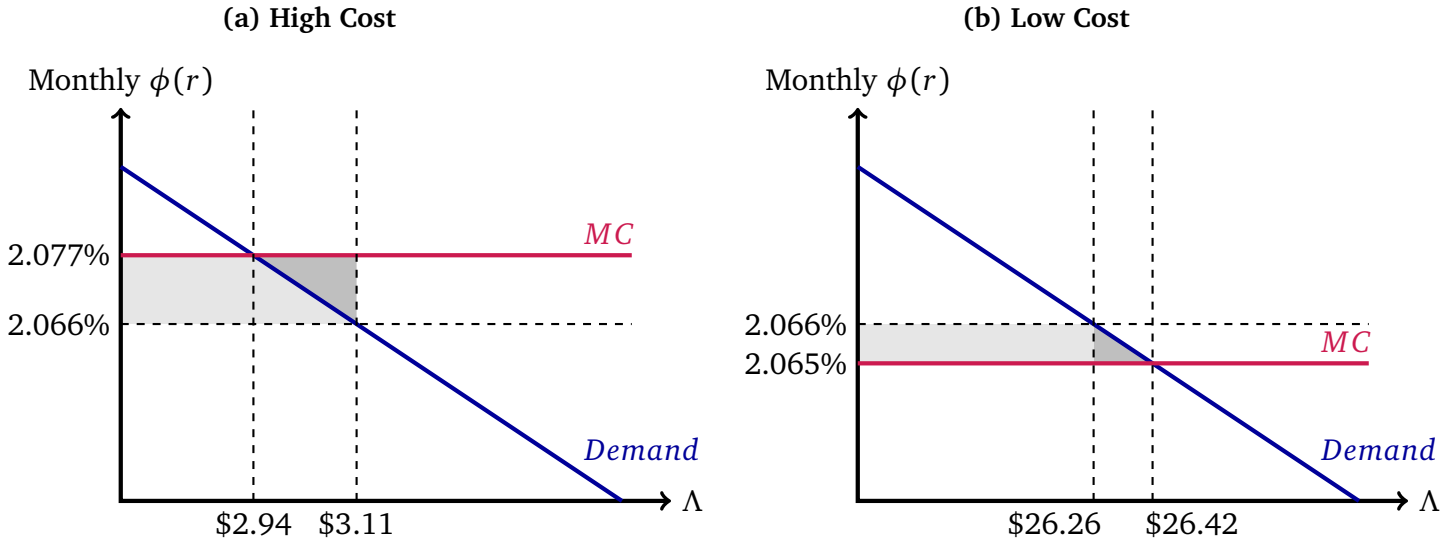
We will report borrower surplus and social welfare in two ways: in terms of dollars per non-default month, and also in total expected dollars transferred to a borrower over the lifetime of a loan, taking into account borrowers' default probability. The expected number of non-default periods that a borrower with default rate δ_H lives through, over the course of a T-period loan, is:

$$\psi_H \equiv \sum_{t=1}^{T}(1-\delta_H)^t = (1-\delta_H)\left(\frac{1-(1-\delta_H)^T}{\delta_H}\right) \tag{20}$$

Plugging in 0.15% for the default rate δ_H, we get ψ_H equal to 57.29.[23] The expected changes in borrower surplus and total welfare, in units of expected dollars over the lifetime of a loan,

[23]For all auto loans ever opened by individuals who are ever bankrupt, we compute the ratio of loans that is charged-off within two years of loan opening as 3.6%. This implies a monthly default probability of $0.15\% = \delta_m = 1-(1-Pr(\text{charged off in first two years}))^{\frac{1}{24}}$. The monthly default rate corresponds to approximately 91% of loans reaching 5-year maturity without being charged-off.

Figure 4: Empirical Welfare Estimates



This figure illustrates the changes in borrower surplus and efficiency induced by bankruptcy flag removal. In each panel, the y-axis, $\phi(r)$, represents the nominal fraction of the loan amount repaid each month, which we calculate as the monthly payment divided by the principal balance. For example, a $\phi(r)$ value of 2.077% means that the borrower pays 2.077% of the principal amount each month, over the course of a five-year loan. The x-axis shows the total loan amount per year, Λ, in billion USD for high and low cost borrowers, respectively. For example, $2.94 billion represents the loan amount that 800,000 high cost borrowers borrow each year prior to bankruptcy flag removal. In panel (a), the light gray trapezoid illustrates the borrower surplus redistributed towards previously bankrupt individuals. The dark gray triangle shows the inefficiency arising from over credit provision to previously bankrupt individuals. In panel (b), the light gray and dark gray area shows the borrower surplus taken from low cost individuals to redistribute towards high cost individuals. The dark gray triangle illustrates the inefficiency arising due to under credit provision to low cost individuals.

are thus equal to the changes per non-default period from Claim 3 multiplied by ψ_H.[24]

We find that the average previously bankrupt individual gains borrower surplus equivalent to $0.41, for each eligible borrower each month. Multiplying by ψ_H, previously bankrupt borrowers gain borrower surplus equal to roughly $23.75 in expectation over the lifetime of a 5-year loan. Restricting attention to the 18.1% of borrowers that get a new auto loan each year, each borrower essentially receives an expected transfer of $131.22 over the lifetime of a 5-year loan. The loss in social welfare from flag removal for previously bankrupt borrowers corresponds to the dark gray triangle in the left panel of Figure 4. We find that due to higher

[24]For example, the social welfare change for previously bankrupt borrowers, over the lifetime of a loan, is given by $\psi_H \frac{1}{2}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\Lambda_L\left(\phi\left(r_{H,fair}\right)\right)-\Lambda_L\left(\phi\left(r_{pool}\right)\right)\right)$.

credit provision to previously bankrupt individuals relative to the efficient benchmark, there is a social welfare loss of \$0.012 per previously bankrupt borrower per month; multiplying by ψ_H, this is \$0.67 in expectation over a 5-year loan. Restricting attention to the 18.1% of borrowers that get a new auto loan each year, there is a welfare loss of \$3.70 per borrower over a 5-year loan due to inefficiently high credit provision.

Next, we aggregate these welfare estimates across borrowers. Approximately 800,000 individuals have their bankruptcy flags removed each year. Multiplying by average loan size, aggregate loan volume is approximately \$2.94 billion at the pre-flag-removal rate r_{flag}, and increases to approximately \$3.11 billion at the lower rate r_{pool}. The borrower surplus increase for previously bankrupt borrowers, which is illustrated as the light gray trapezoid in the left panel of Figure 4, is \$0.33 million per month, or \$19 million in expectation over a five-year loan term. The loss in social welfare at r_{pool}, which is illustrated as the dark gray triangle in the left panel of Figure 4, is \$9,356 per month, or \$536,005 in expectation over a five-year loan term.

Welfare effects for never-bankrupt borrowers. Flag removal policies also have a general equilibrium effect on the prices faced by never-bankrupt borrowers, since they affect whether pricing is pooled with previously bankrupt borrowers. In our application, we can calculate the counterfactual effect on prices faced by never-bankrupt borrowers using lenders' zero-profit condition, and use this to calculate the welfare effects on never-bankrupt borrowers.[25] Since there are many more never-bankrupt borrowers than bankrupt borrowers, the welfare effects for the never-bankrupt group will be small. This is because the change in prices induced by the policy will be much smaller for the never-bankrupt borrowers, whereas the change in loan quantities will be approximately comparable to the effect for the previously bankrupt borrowers. The social welfare change is half the product of the price and quantity changes, hence is much smaller for the never-bankrupt borrowers.

Let H (high cost) denote previously bankrupt borrowers, and L (low cost) denote never-bankrupt borrowers. Let $r_{H,fair}$ and $r_{L,fair}$ represent the interest rates for H and L respectively in competitive markets when bankruptcy flags are never removed, so data is available to distin-

[25] In many cases, data acquisition or removal can be used to estimate price and quantity effects for multiple groups. In our particular context, we only observe flag removal for previously bankrupt borrowers, so much infer prices and quantities for never-bankrupt borrowers.

guish the two groups. These interest rates are also equal to the social cost of providing credit to H and L respectively. We will have $r_{H,fair} > r_{L,fair}$, since previously bankrupt borrowers tend to be more costly to lenders (which is why their rates drop when flags are removed). $r_{H,fair}$ is equal to r_{flag}, the rate faced by previously bankrupt borrowers when flags are present on their credit reports. We also observe r_{pool} in the data, which is the price that previously bankrupt borrowers face when flags are removed and they are pooled with type-L borrowers. Hence, the econometric problem we face is to estimate $r_{L,fair}$, the socially efficient price of credit for the never-bankrupt group. We can calculate $r_{L,fair}$ based on lenders' zero-profit condition, and the relative sizes of the H and L-groups. If we further assume that the demand elasticities in groups H and L are the same, we can then calculate the welfare change for the L-group of never-bankrupt borrowers.

As we did with the H-group, we will report L-group borrower surplus and total welfare in terms of total expected dollars over the lifetime of a loan, which is equal to dollars per non-default period, multiplied by the expected number of non-default periods for an L-type borrower:

$$\psi_L \equiv \sum_{t=1}^{T} (1-\delta_L)^t = (1-\delta_L)\left(\frac{1-(1-\delta_L)^T}{\delta_L}\right) \tag{21}$$

After multiplying by ψ_L, the units of H-group and L-group welfare and surplus are both in dollars, and can be compared consistently.[26] The following claim characterizes the interest rate $r_{L,fair}$, and the welfare change for the L-group.

Claim 4. *When bankruptcy flags are present on credit reports, the price $r_{L,fair}$ that low-cost borrowers face will satisfy:*

$$\psi_H \Lambda_H(\phi(r_H))\left(\phi(r_H) - \phi\left(r_{H,fair}\right)\right) = \psi_L \Lambda_L(\phi(r_L))\left(\phi\left(r_{L,fair}\right) - \phi(r_L)\right) \tag{22}$$

Where,

$$\psi_H \equiv (1-\delta_H)\left(\frac{1-(1-\delta_H)^T}{\delta_H}\right),\ \psi_L = (1-\delta_L)\left(\frac{1-(1-\delta_L)^T}{\delta_L}\right) \tag{23}$$

Suppose the elasticities of demand in groups H and L are the same, that is, the slopes of demand

[26] Note that we cannot directly add welfare per non-default period across H- and L-group borrowers without adjusting by ψ_L, because the default rate differs across the H- and L- groups.

b_L and b_H satisfy:

$$\frac{b_L}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)} = \frac{b_H}{\Lambda_H\left(\phi\left(r_{pool}\right)\right)} \tag{24}$$

Then, the gain in social welfare from keeping bankruptcy flags on credit reports for the L-group (never bankrupt borrowers), in expected dollars over the term of a loan, is:

$$\frac{1}{2}\frac{\psi_H^2}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\left[b_H\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\right] \tag{25}$$

Expression (25) is a factor $\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ times the social welfare change, in expected dollars over the term of a loan, for the H-group. The increase in borrower surplus for the L-group from keeping bankruptcy flags on credit reports is equal to the loss in borrower surplus from the H-group, plus the increase in social welfare across both groups.

Claim 4 states that the change in social welfare for the L-group in (25) is a factor $\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ times the change in welfare for the H-group, where $\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ is the ratio of total loan volume across all borrowers in the H and L groups, and $\frac{\psi_H}{\psi_L}$ is an adjustment term which takes into account the differences in default rates between the two groups. In the data, only 10.6% of borrowers are ever bankrupt; thus the ratio $\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ is low. The ratio $\frac{\psi_H}{\psi_L}$ is equal to 0.982. As a result, the majority of the social welfare effect comes from the H-group. The intuition behind this result is that, from (22), the price change for the L-group is a factor $\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ times smaller than the price change for the H-group. This price change induces a quantity change in the L-group of :

$$\frac{\psi_H}{\psi_L}b_H\left(\phi\left(r\right)-\phi\left(r_{H,fair}\right)\right)$$

which is approximately the quantity change in the H-group. Thus, the quantity changes in the two groups are approximately the same, but the price change is a factor $\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ smaller for the L-group. From (17) of Claim 3, the social welfare change induced by the price change is equal to half the product of the price and quantity changes. Multiplied by the expected number of non-default periods, ψ_L, the absolute welfare change is thus a factor $\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ lower for the L-group.

Using Claim 3, we can calculate the effects of flag removal on social welfare. Total surplus for previously bankrupt borrowers increases by roughly $19 million per annual cohort. Next, we compute the borrower surplus loss for low-cost, never-bankrupt individuals from pooling with previously bankrupt borrowers. The difference between borrower surplus taken from low cost individuals and borrower surplus redistributed to high cost individuals is the efficiency cost of the redistribution. From (22), the zero profit condition implies a spread between the pooling and break-even payment for the low cost group of:

$$\frac{\psi_H \Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)$$

This increase in prices for never-bankrupt individuals implies a loss in borrower surplus for never-bankrupt borrowers, equal to the sum of the light gray rectangle and dark gray triangle in the right panel of Figure 4. Since we know the payment change for high cost borrowers from our regression estimates, we only need to compute relative loan quantities for high and low cost types to determine the area of the light gray rectangle and dark gray triangle in the right panel of Figure 4. Approximately 10.6% of individuals ever go bankrupt. Assuming that high and low cost individuals demand the same loan amount at the same price, we obtain relative loan quantities to be

$$\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}=\frac{0.106}{1-0.106}\approx 0.1186.$$

As we observe the high cost pooling quantity in the data, we can, therefore, also compute the low cost pooling quantity:

$$\Lambda_L\left(\phi\left(r_{pool}\right)\right)=\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{0.1186}$$

We show this quantity scaled by the number of people with flag removals in the right panel of Figure 4. To compute the default rate adjustment terms, ψ_L and ψ_H, note that monthly default rates are 0.15% among previously bankrupt customers, and 0.09% among never-bankrupt customers. Thus, we have:

$$\psi_H = 57.29,\ \psi_L = 58.34$$

Knowing the implied price change, the relative group sizes and assuming equal demand

elasticities, we can compute the borrower surplus taken from low-cost, never-bankrupt individuals; this works out to a loss of $2.9 for each never-bankrupt individual, in expectation over the course of a 5-year loan. In the aggregate, this implies that approximately $19.6 million in borrower surplus is taken from never-bankrupt individuals per annual-cohort of flag removals. Note that the borrower surplus taken from never-bankrupt individuals somewhat exceeds the borrower surplus redistributed to previously bankrupt individuals; the difference is the social welfare loss from flag removal.

To compute the total social welfare change from flag removal, we can add the changes in borrower surplus for the high- and low-cost groups. Our estimates imply an expected welfare loss of $0.75 per eligible individual with flags removed over the course of a 5-year loan. Multiplying by the total number of borrowers affected by flag removals, we have a total welfare change of approximately $598,385 in the U.S. per year.[27]

In summary, our estimates imply that flag removal transfers roughly $19 million per year from never-bankrupt borrowers to previously bankrupt borrowers, at the cost of destroying $598,385 in social surplus. Thus, for each dollar of surplus transferred to previously bankrupt borrowers, 3.15 cents of social surplus is destroyed. This quantifies the intuition from Subsection 2.1, showing that flag removals are a fairly low-cost way to redistribute surplus to affected borrowers: the social deadweight losses from credit misallocation in this setting are small.[28]

[27] Note also that we assumed that the elasticities of demand for the two groups were equal in (24) for simplicity of exposition. If instead we assumed that the L-group elasticity was k_L times larger than the H-group elasticity, then (25) would become $\frac{1}{2}k_L \frac{\psi_H^2}{\psi_L} \frac{\Lambda_H(\phi(r_{pool}))}{\Lambda_L(\phi(r_{pool}))} \left(\phi\left(r_{H,fair}\right) - \phi\left(r_{pool}\right)\right)\left[b_H\left(\phi\left(r_{pool}\right) - \phi\left(r_{H,fair}\right)\right)\right]$. In order for the L-group welfare change to be non-negligible, the factor k_L would have to be approximately equal to $\frac{\psi_L}{\psi_H} \frac{\Lambda_L(\phi(r))}{\Lambda_H(\phi(r))}$. Since we find that the elasticity of demand for the H-group is around 2.32, the L-group would need to have an unreasonably high elasticity, of approximately 19.89, for the L-group welfare change to be equal to that of the H-group.

[28] A natural question is how our conclusions are affected by the informativeness of the data that is removed. In Appendix B.2, we show that the efficiency ratio is worse when the signal is more informative: that is, if data is more informative about default rates, then removing data will have a larger negative impact on social welfare, for each dollar of surplus transferred between groups. Quantitatively, however, we show that data removal would remain a low-cost way to transfer surplus between groups, costing less than 21 cents of deadweight loss for each dollar transferred, even for signals that induce price changes up to eight times as large as the price changes induced by bankruptcy flag removals.

Table 3: Welfare Estimates

This table summarizes our main estimates implied by the specifications of Table 2. Panel A shows average interest rates in the six months before flag removal (r_{flag}), the interest rate effect of flag removal ($r_{pool} - r_{flag}$), and the effect of flag removal on the fraction of the principal repaid each month in a standardized five-year loan ($\phi(r_{pool}) - \phi(r_{flag})$). Panel B shows average loan quantities in the six months before flag removal and the quantity effect of flag removal. Panel C summarizes surplus changes implied by the estimates in Table 2..The first row shows the average change in consumer surplus for individuals with flag removal for the average five-year loan. It is the sum of monthly non-default period surpluses. The number of non-default periods is derived from the probability of loans to individuals who ever have a bankruptcy flag to be charged off within two years of loan opening. The second row shows the social surplus change in expected dollars over the term of a loan, for each individual whose flag is removed. The third row provides the efficiency change per dollar redistributed to bankrupt individuals by removing the bankruptcy flag. It is computed dividing the second row by the first row. Source: TransUnion.

	(1)	(2)	(3)	(4)	(5)
Panel A: Prices					
Pre-flag-removal loan interest rate (%)	9.02%	9.02%	9.02%	9.02%	9.02%
Flag removal-induced change in interest rate (%)	-0.218%	-0.226%	-0.218%	-0.226%	-0.226%
Change in monthly repayment as share of loan principal (%)	-0.011%	-0.011%	-0.011%	-0.011%	-0.011%
Panel B: Quantities					
Pre-flag-removal loan quantity (Average $ per borrower per year)	$3,678.00	$3,678.00	$3,678.00	$3,678.00	$3,678.00
Flag removal-induced change in loan quantity (Average $ per borrower per year)	$220.13	$215.11	$220.13	$215.75	$213.52
Panel C: Average Surplus Changes					
Average consumer surplus redistributed to individuals with flag removal over 5 years ($ per eligible borrower with flag removal)	$22.93	$23.76	$22.93	$23.76	$23.75
Change in social surplus per individual over 5 years ($ per eligible borrower with flag removal)	-$0.74	-$0.75	-$0.74	-$0.76	-$0.75
Welfare change per dollar redistributed to bankrupt individuals	-0.0324	-0.0317	-0.0324	-0.0318	-0.0315

4 Extensions

In the baseline model, we assume that markets are perfectly competitive, and that changes in price do not affect the selection of borrowers. In this section, we relax each assumption, and discuss how this would affect our results.

4.1 Imperfect Competition

We can extend our framework to imperfect competition, under which lenders are able to charge a markup above marginal cost. In general, if data availability changes the markups lenders can charge, outcomes are known to be complex (He et al., 2020; Huang, 2022). By analogy to the classic literature on third-degree price discrimination, it is not in general possible to say whether data will increase or decrease markups and thus social welfare (Schmalensee, 1981; Varian, 1985, 1989). Nonetheless, while markups make our theoretical conclusions less sharp, we show that markups can be accommodated in the empirical application of our methodology, with some additional data. Then, we show that, under a certain set of parameter restrictions – when the demand elasticities in the two consumer groups are equal, and when markups pre- and post-data availability are the same – the existence of markups in fact does not affect welfare gains from data availability. While this set of assumptions may not always be satisfied exactly in reality, this result suggests that the existence of markups per se do not dramatically affect our results. The two groups would need to have quite different demand elasticities, or markups would have to change significantly due to data availability, for our results to be substantially inaccurate.

As in the main model, let H denote previously bankrupt borrowers, and L denote never-bankrupt borrowers. Let $r_{H,fair}$ and $r_{L,fair}$ represent the interest rates for H and L respectively if markets were fully competitive, which are also equal to the social cost of providing credit to H and L respectively. We will have $r_{H,fair} > r_{L,fair}$, since previously bankrupt customers tend to be more costly to lenders (which is why their rates drop when flags are removed). As in the baseline model, we assume demand in both groups is linear, with possibly different slopes and intercepts:

$$\Lambda_L\left(\phi\left(r_L\right)\right) = a_L - b_L\phi\left(r_L\right) \tag{26}$$

$$\Lambda_H\left(\phi\left(r_H\right)\right) = a_H - b_H\phi\left(r_H\right) \tag{27}$$

In contrast to the main text, we assume that lenders may charge markups over marginal cost, both before and after data is made available. Rather than take a stance on the particular theoretical model generating markups, we will simply take markups as exogenous, and express welfare in terms of markups over marginal costs before and after data is made available. Let m_H, m_L be the markups charged over the competitive prices $\phi\left(r_{H,fair}\right), \phi\left(r_{L,fair}\right)$ when data is available to distinguish the two groups, and m_{pool} be the markup when data is not available.

When lenders charge markups, interest rates for each group when data are available, $r_{H,data}$ and $r_{L,data}$, will be higher than lenders' break-even interest rates $r_{H,fair}$ and $r_{L,fair}$. We will write these as:

$$\phi\left(r_{H,data}\right) = \phi\left(r_{H,fair}\right) + m_H,\ \phi\left(r_{L,data}\right) = \phi\left(r_{L,fair}\right) + m_L \tag{28}$$

When data is not available, the average cost across both groups is:

$$\frac{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{L,fair}\right)\right) + \psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{H,fair}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right) + \psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)} \tag{29}$$

That is, (29) is a weighted average of the cost of serving L and H type borrowers, with weights equal to loan volumes Λ_H, Λ_L multiplied by the expected number of non-default periods, ψ_H, ψ_L, defined in (23) of Claim 4. If lenders set a markup m_{pool} above average costs, the price that borrowers face without data is then:

$$\phi\left(r_{pool}\right) = \frac{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{L,fair}\right)\right) + \psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{H,fair}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right) + \psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)} + m_{pool} \tag{30}$$

The following claim characterizes the welfare effects of data availability in this setting.

Claim 5. *The change in total welfare when data is made available, in expected dollars over the*

term of a loan, is:

$$\Delta Welfare = \underbrace{\psi_H \frac{b_H}{2}\left(\left(m_{pool} - s_L\Delta\right)^2 - m_H^2\right)}_{High\ group} + \underbrace{\psi_L \frac{b_L}{2}\left(\left(m_{pool} + (1-s_L)\Delta\right)^2 - m_L^2\right)}_{Low\ group} \tag{31}$$

Where:

$$s_L \equiv \frac{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right)}{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right) + \psi_H \Lambda_H\left(\phi\left(r_{pool}\right)\right)} \tag{32}$$

is the share of loans given to low-cost borrowers, at the pooled price r_{pool}*,*

$$\Delta \equiv \phi\left(r_{H,fair}\right) - \phi\left(r_{L,fair}\right) \tag{33}$$

is the difference in costs between the two groups, and ψ_H, ψ_L*, are the expected number of non-default periods per group, defined in (23) of Claim 4.*

Expression (31) is the most general expression for the change in welfare when lenders set markups above marginal costs. In the fully general case to say anything about how data availability impacts welfare we must estimate markups m_H, m_L, m_{pool} in addition to prices, quantities, and the terms ψ_L, ψ_H.

However, if markups are constant and demand elasticities are identical across groups, then equation (31) collapses to our earlier result, and we can estimate welfare changes using price and quantity data alone. To see this, first, suppose that markups are constant, across groups, and before and after data is available:

$$m_{pool} = m_H = m_L = m \tag{34}$$

This assumption is likely to approximately hold in our empirical setting. We calculate state-level Herfindahl-Hirschman indices (HHI) in our empirical setting for previously bankrupt borrowers, and never-bankrupt borrowers with similar credit scores. The HHI in a market is defined as the sum of squared market shares ($\sum_{l=1}^{N} s_l^2$), where l indexes lenders. In most models of imperfect competition, markups depend on measures of market concentration, so two markets for similar products which have similar HHI values are likely to have similar markups. We find that the state-level Herfindahl index is 0.0376 for previously bankrupt borrowers, and

0.0330 for never-bankrupt borrowers with similar credit scores. The HHIs in both cases are low and fairly similar, lending support to the assumption that markups are similar across groups in our empirical setting.

Given (34), the welfare change in (31) then simplifies to:

$$\Delta Welfare = \psi_H \frac{b_H}{2}\left(s_L^2\Delta^2 - 2ms_L\Delta\right) + \psi_L \frac{b_L}{2}\left((1-s_L)^2\Delta^2 + 2m(1-s_L)\Delta\right) \tag{35}$$

For additional intuition, note that we can write (35) as:

$$\Delta Welfare = \underbrace{\psi_H \frac{b_H}{2}s_L^2\Delta^2 + \psi_L \frac{b_L}{2}(1-s_L)^2\Delta^2}_{A} \underbrace{-\psi_H b_H m s_L\Delta + \psi_L b_L m(1-s_L)\Delta}_{B} \tag{36}$$

Suppose we set markups to zero in Expression (36). The change in welfare is then term A in (36). This term thus represents the welfare gain from data availability in competitive markets. Term A is equivalent to the sum of expression (17) of Claim 3 for the low and high cost groups.[29]

Term B in (36) thus captures how markups change welfare gains, relative to the competitive case. To understand term B, first we consider a special case of the result, where the elasticities of demand in the two groups, around the pooled-pricing rate, are the same. That is, assume that the slopes of demand are proportional to the size of each group:

$$\frac{b_H}{\Lambda_H\left(\phi\left(r_{pool}\right)\right)} = \frac{b_L}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)} \tag{37}$$

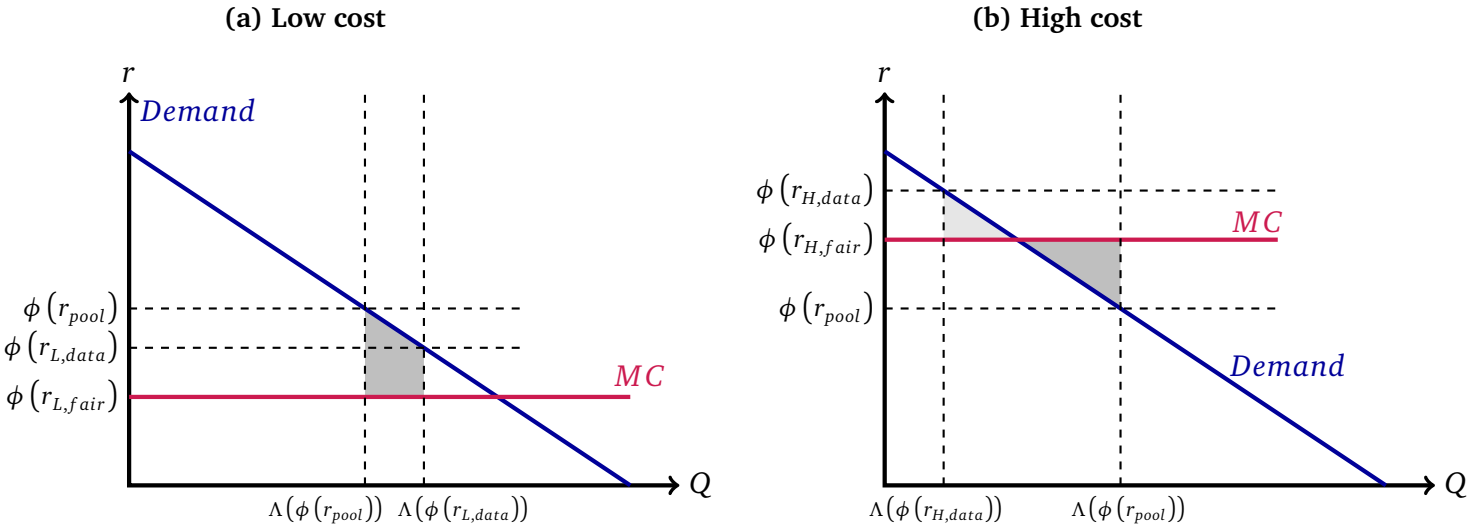
If (37) holds, given the definition of s_L in (32), we have:

$$\frac{\psi_H b_H}{\psi_L b_L} = \frac{1-s_L}{s_L} \tag{38}$$

Now, under (38), term B in (36) then becomes $m\psi_L\Delta\left(-b_L(1-s_L) + b_L(1-s_L)\right) = 0$. Thus, when the elasticities of demand in the two groups are equal, and markups pre- and post-data availability are equal, then the welfare change with markups is exactly the same as if markets

[29] To see this, note that when markets are competitive and data becomes available, prices for the high-cost group increase by $s_L\Delta$, so quantities decrease by $b_H s_L\Delta$. Taking the product of the price change and the quantity change and dividing by 2, according to (17), we get $\frac{b_H}{2}s_L^2\Delta^2$. The same calculation for the low-cost group gives $\frac{b_L}{2}(1-s_L)^2\Delta^2$.

Figure 5: Price Discrimination with Imperfect Competition



This figure illustrates how third-degree price discrimination affects welfare in credit markets, when there is imperfect competition and prices may be higher than costs. Suppose there are two groups of prospective borrowers, low cost (panel a) and high cost (panel b). The red lines show the cost of serving each group, and the blue lines show borrowers' demand curve. Lenders are initially unable to distinguish between these borrowers, so set price $\phi(r_{pool})$. After lenders are able to distinguish the two groups of borrowers, suppose they set $\phi(r_{L,data})$ for the low-cost group (panel a) and $\phi(r_{H,data})$ for the high-cost group (panel b). The dark gray shaded area in panel (a) shows the welfare gain for the low cost group, where prices decrease. The dark gray (light gray) shaded triangle in panel (b) shows welfare gains (losses) for the high cost group. The total welfare effect on the high cost group is the difference between the size of the dark gray and the light gray triangles.

were competitive.

The intuition behind this result is illustrated in Figure 5, which graphically depicts the welfare effects of data availability when there are markups. When there are markups, the right panel of Figure 5 shows that the welfare gains from *raising* prices a given amount for H-type borrowers are smaller, since prices are already closer to their marginal costs. This is reflected by the negative $-b_H\psi_H ms_L\Delta$ term in (36). However, the left panel shows that the welfare gains from *lowering* prices for L-type borrowers are larger, since prices are further above marginal costs. This is captured by the positive $b_L\psi_L m(1-s_L)\Delta$ term in (36). When markups and demand elasticities are the same across groups, these two effects exactly offset each other, so welfare gains in the case with imperfect competition are exactly the same as in the competitive case; releasing data will generally increase social welfare.

A more basic intuition for the result that data will generally increase welfare, even with markups, is the following. Suppose prices for the H- and L-groups are the same, though prices may be much higher than costs. Then the willingness-to-pay of the marginal borrower in group H and group L are the same. Suppose we remove a small number of marginal H-type borrowers from the borrowing pool, and add an equal number of marginal L-type borrowers, so total loan amount across the two groups is unchanged. Since the marginal WTP is the same, total borrower utility across the two groups is unchanged. However, reallocating from H-types to L-types decreases the average social cost of serving these borrowers. Thus, social welfare must increase. This argument holds regardless of whether markets are competitive or not.

We can think of the general case in terms of its deviations from the special case of constant markups and elasticities across the two groups. First, suppose we hold markups fixed, but relax the elasticity assumption in (37). The sign of term B in (36) depends on the relative elasticities in the two groups. When we have:

$$\frac{b_H}{\Lambda_H\left(\phi\left(r_{pool}\right)\right)} > \frac{b_L}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$$

so the elasticity in the high group is greater (smaller) than the elasticity in the low group, then the welfare gain from making data available is smaller (greater) than the competitive case. The intuition is that, when the high group has higher demand elasticity, the decreased welfare gains from raising prices for H-group borrowers tend to dominate, and vice versa.

Second, suppose we allow markups to vary before and after data availability. Note that (31) is strictly decreasing in m_H and m_L, the size of post-data-availability markups. The intuition is simply that higher average markups are worse for social welfare. Thus, if data availability tends to increase (decrease) the level of overall markups, this tends to decrease (increase) social welfare.[30]

In summary, our results imply that, if data availability does not affect markups substantially, and if the demand elasticities in the two groups are similar, data availability tends to increase welfare even when there is market power, through a similar mechanism of reallocat-

[30]It is known in the theoretical literature that third-degree price discrimination has ambiguous effects on the effect of overall markups: with different demand functions in the two groups, essentially any pattern of markup increases or decreases is possible (Bergemann et al., 2015).

ing towards lower-cost borrowers. If there were sufficient data available, one could quantify the most general expression for welfare changes, (31), by measuring pre- and post-change prices, quantities, ψ_L and ψ_H, and markups for all borrower groups.

Next, we apply the results from this extension to imperfect competition and empirically show how the efficiency ratio - the welfare cost per USD redistributed- depends on market power in credit markets. The thought experiment considers the redistributional and welfare consequences if the observed data was generated in a market with market power. Following the derivations in section 4.1 and equation (36), we decompose the effects of data removal into the effect in the absence of market power and the consequences of data removal due to market power. Table 4 shows that the efficiency ratio is in the range of a few cents per USD redistributed. We compute the HHI implied mark-up as 34 basis points (=9.02% $*$ 0.0376). For robustness, we also consider mark-ups twice the size of HHI implied mark-ups. As argued above, Table 4 illustrates that under similar mark-ups across markets and with similar demand elasticity across high and low cost group, the consequences of data removal are independent of the mark-up (compare columns 2 and 4). When demand elasticities vary across groups, market power can either improve ($\frac{b_L}{\Lambda_L}/\frac{b_H}{\Lambda_H} = 0.5$) or deteriorate ($\frac{b_L}{\Lambda_L}/\frac{b_H}{\Lambda_H} = 2$) the efficiency ratio by a few cents. In cases with market power, data removal can even slightly increase efficiency in some cases.

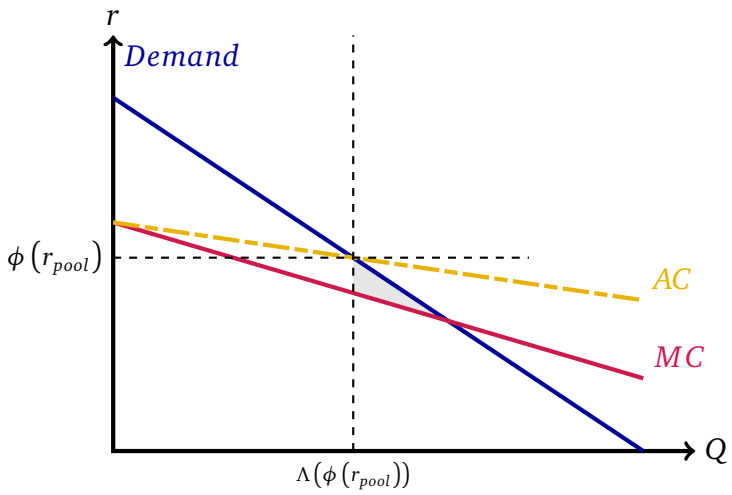
4.2 Adverse Selection

In the main text, we assume that there is no selection: we assume costs depend on borrowers' types, but are not correlated with borrowers' willingness to pay. In this section, we relax this assumption and allow prices to be correlated with costs. In a case with adverse selection and competitive markets, the top panel of Figure 6 depicts market outcomes in this case. There is a deadweight loss triangle (i.e., the light gray shaded region in the figure) because markets reach the point where average costs are equal to the marginal borrower's willingness to pay. However, at this point, marginal costs are *below* willingness to pay.

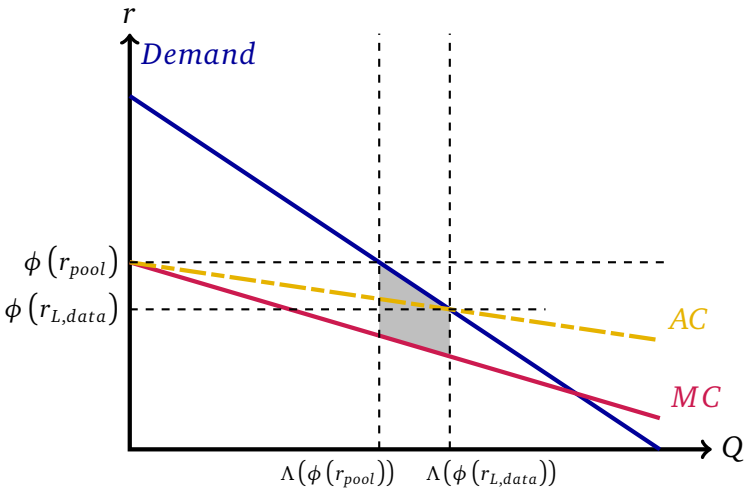
Suppose that data becomes available, but that there is also adverse selection in each of the two submarkets. For the low-cost group, prices decrease. When data is not available, prices are too high for these borrowers, relative to the socially efficient point, for two reasons. First, they

Figure 6: Price Discrimination with Adverse Selection

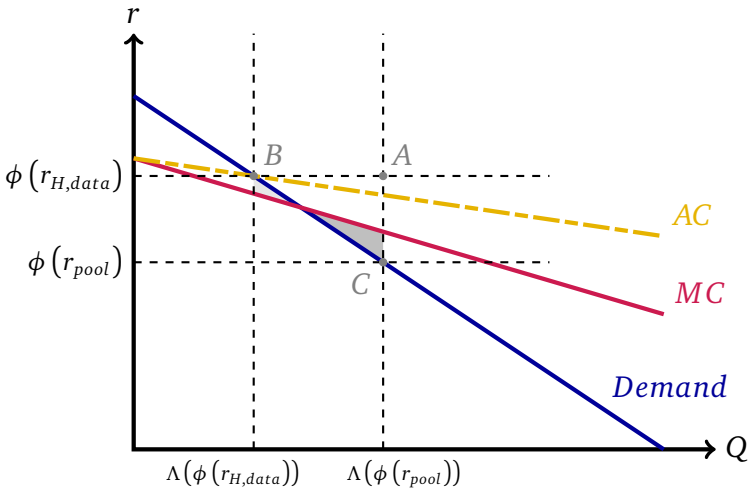
(a) Adverse selection



(b) Low cost



(c) High cost



This figure illustrates how third-degree price discrimination affects welfare in credit markets in the presence of adverse selection or moral hazard. With these frictions, costs and prices vary. Panel (a) illustrates that, under adverse selection, prices in competitive equilibrium will be equal to average costs. The light gray triangle shows the welfare loss, relative to the constrained optimum of setting prices equal to average costs. If data becomes available on high-cost and low-cost borrowers as illustrated in panels (b) and (c), prices will be set equal to average costs separately for each group in competitive equilibrium. Prices will tend to fall for the low-cost group as shown in panel (b) and rise for the high-cost group as shown in panel (c). After lenders are able to distinguish the two groups of borrowers, suppose they set $\phi\left(r_{L,data}\right)$ for the low-cost group and $\phi\left(r_{H,data}\right)$ for the high-cost group. The dark gray shaded area in panel (b) shows the welfare gain for the low cost group, where prices decrease. The dark gray (light gray) shaded triangle in panel (c) shows welfare gains (losses) for the high cost group.

Table 4: Efficiency Ratio of Data Removal with Market Power

This table summarizes the Efficiency Ratio (Welfare Loss per USD redistributed) of removing data under various elasticity and mark-up assumptions. The first three columns summarize the consequences of removing data when the observed data was generated in a market with a 34basis point mark-up. The mark-up is calibrated to the HHI. Columns four to six double the HHI implied mark-up. The mark-up is assumed to be similar for high cost, low cost, and pooling market. Low Cost to High Cost Elasticity shows varying assumptions about the low cost group demand elasticity in relation to the high cost group. The efficiency ratio with market power includes losses due to both data removal and credit rationing - it combines the effects of A and B in equation 36. The efficiency ratio without market power only shows the effect of A in equation 36, which is the effect of data removal in the absence of market power. The efficiency ratio due to market power only shows the effect of B in equation 36, which are the additional losses (gains) from data removal due to prevailing market power.

	34bp mark-up			68bp mark-up		
Low Cost to High Cost Elasticity	0.5	1	1.5	0.5	1	1.5
Efficiency Ratio						
With Market Power	\$0.01	-\$0.03	-\$0.08	\$0.06	-\$0.03	-\$0.12
Without Market Power	-\$0.03	-\$0.03	-\$0.03	-\$0.03	-\$0.03	-\$0.03
Due to Market Power	\$0.04	\$0.00	-\$0.04	\$0.09	\$0.00	-\$0.09

are pooled with the high-cost types. Second, there is adverse selection. When firms have data on these borrowers, they lower prices to the point where average cost is equal to marginal WTP; this is the point $\Lambda(\phi(r_{L,data}))$. This increases consumer surplus and social welfare, though not to the socially optimal point, as distortions from adverse selection remain. The bottom-left panel of Figure 6 illustrates the welfare gain as the dark gray shaded region.

For the high-cost group, prices increase. When data is not available, prices may be too high or too low for these borrowers relative to the socially efficient price, as there are two counteracting forces. Adverse selection tends to cause prices to be too high relative to the social optimum. However, pooling with the low-type borrowers tends to make the prices too low.

The bottom-right panel of Figure 6 illustrates a case where the price without data is below the social optimum. After data is available, firms will set prices where the average cost curve crosses the demand curve. This is always above the socially optimal point, where the marginal cost curve crosses the demand curve. As a result, the effects on consumer welfare are ambiguous: there is a welfare gain as prices increase to the social optimum, represented by the dark gray triangle, but there is a welfare loss from prices increasing further, represented by the light gray shaded triangle.

This framework also lets us analyze how our estimates of changes in total surplus are likely biased by our assumption that there is no adverse selection. Fixing prices $\phi\left(r_{pool}\right)$ and $\phi\left(r_{H,data}\right)$, and loan amounts $\Lambda\left(\phi\left(r_{H,data}\right)\right)-\Lambda\left(\phi\left(r_{pool}\right)\right)$, the calculated welfare gain from making data available is always larger if we assume there is no adverse selection, so the marginal cost curve is flat and equal to $\phi\left(r_{H,data}\right)$. To see this, note that the welfare gain that we calculate in the main text, (17) of Claim 3, corresponds to the triangular area enclosed by points A, B, and C in the bottom-right panel of Figure 6. In contrast, the welfare gain under adverse selection is the dark gray area, which is weakly smaller than the ABC triangle, minus the light gray area.[31]

Thus, when there is adverse selection, the actual welfare gains from data availability for the H-group must be even smaller than we find in the main text, whereas the change in consumer surplus is identical to expression (16). Conversely, the total welfare effects of *removing* data such as bankruptcy flags are *smaller* if there is adverse selection. Thus, our conclusion that flag removals are a quantitatively efficient way to redistribute surplus would not change significantly if there was adverse selection.[32]

If data were available on lenders' costs before and after lenders are able to use new data for pricing, the triangles in Figure 6 could be quantified to calculate the welfare gains from making data available. DeFusco, Tang and Yannelis (2022) demonstrate this in the case of a single set of borrowers; to quantify the effects of data availability, the methodology in DeFusco, Tang and Yannelis (2022) could be applied to the high-cost and low-cost borrower groups separately. In our empirical application, however, adverse selection does not appear to be present in our setting. We conduct a test similar in spirit to Chiappori and Salanie (2000). Asymmetric information would suggest a positive correlation between rates and charge-offs, as adverse selection would mean that riskier borrowers select higher interest loans, and moral hazard would mean that higher rates induce borrowers into default. This test is show in Figure C.1, we cannot reject that the price changes induced by bankruptcy flag removal have no effect on

[31] If adverse selection is sufficiently severe, so r_{pool} is higher than the efficient price, it is possible that the dark gray area is empty, and the light gray area is a trapezoid; it is then the case that data availability lowers aggregate welfare.

[32] We disregard the low-cost group in this discussion; based on arguments in Subsection 3.3.2, the welfare effects for the low-cost group will tend to be much smaller, since the group is larger and thus the change in prices is smaller.

default rates. This is confirmed in Table C.1, which shows similar results to those in Table 2, replacing the outcomes with charge-offs. The table shows that flag removal is associated with an insignificant increase in charge-offs.[33]

4.3 Application to Other Estimates of Data Removal

While in our main analysis we focus on auto lending, in this section, we apply the developed cost curve approach to estimate the consequences of data acquisition/removal to the bankruptcy flag removal in Gross, Notowidigdo and Wang (2020) (GNW 2020). We find similar estimates for the welfare loss per USD redistributed to previously bankrupt individuals using the regression estimates of GNW 2020. This application serves two purposes (1) we show that the finding of low efficiency losses relative to the surplus redistributed is robust (2) we demonstrate the easy applicability of the methodology to estimates by other researchers. However, it has the disadvantage that due to a lack of price data, aggregate surplus and welfare losses cannot be computed. Additionally, the mortgage and other credit markets analyzed in this section have many institutional features such as government guarantees (Bachas et al., 2021), which are not modeled but can affect welfare, so we treat these estimates with some caution.

Table 5 summarizes the welfare cost per USD redistributed for various credit markets based on the quantity estimates of GNW 2020. While price data is necessary to compute aggregate surplus and welfare consequences, the efficiency ratio can be computed by exploiting changes in data availability and quantity data only. Appendix section B.3 derives the efficiency ratio as a function of loan quantities and shows details of the computation. Panel A shows the efficiency ratio using account openings and balances on accounts. Intuitively, the 'High Cost' row shows the ratio of the dark gray triangle relative to the light gray trapezoid, as illustrated in Panel B of Figure 1. The combined estimate, also taking the inefficiency arising from under-credit provision to low-cost consumers into account, is shown in the 'High and Low Cost' row.

[33] We do not mean to argue that adverse selection and moral hazard are not concerns in auto lending settings. Some studies have found that adverse selection and moral hazard are important particularly in the period prior to the Great Financial Crisis, for example Adams et al. (2009). Other more recent papers such as Argyle et al. (2020) find no effects, which is consistent with auto loans being highly collateralized. In our setting we do not find any effects of flag removal in particular on default rates, which is suggestive that there is no asymmetric information.

Table 5: Welfare cost per USD redistributed using GNW 2020

This table summarizes the efficiency ratio (welfare loss per USD redistributed) for various credit markets. The estimates are based on the quantity estimates of GNW 2020 Appendix Table A3. Panel A summarizes the efficiency ratio at a 6 month horizon using quantity estimates based on new account openings and balances. Panel B repeats the exercise using new accounts and their principal. Pre-removal quantities are computed using the number of open accounts and their balance (principal). Post removal quantities are computed using pre-removal quantities plus the effect of flag removals on the number of accounts and the size of the balance (new principal). To compute welfare losses for the never-bankrupt group, we made the following assumptions: Elasticities for high and low cost borrowers are similar; Relative group size is determined by the share of individuals ever having a bankruptcy flag; Relative number of repayment periods (ψ_H/ψ_L) is assumed to be the same as before. Source of regression estimates: Gross, Notowidigdo and Wang (2020).

	All	Mortgage	Auto	Other
Panel A: New Accounts and Balance Effect				
High Cost	-0.0261	-0.0063	-0.0096	-0.0317
High and Low Cost	-0.0292	-0.0070	-0.0107	-0.0354
Panel B: New Accounts and Principal Effect				
High Cost	-0.0296	-0.0068	-0.0091	-0.0385
High and Low Cost	-0.0331	-0.0076	-0.0101	-0.0430

Intuitively, it shows the area of the dark gray triangles in Panels A and B of Figure 1, relative to the light gray trapezoid in Panel B of Figure 1. The estimates for the efficiency ratio across all considered consumer credit markets is 2.61 cents for the high-cost only and 2.92 cents, also considering the inefficiency for the low-cost group. These estimates are very close to the Table 3 estimates of around 3 cents. For mortgage markets, the efficiency cost varies from 0.63 to 0.7 cents per dollar redistributed, depending on the group considered. For auto loans, estimates are around 1 cent per dollar redistributed. In other consumer credit markets, which includes credit cards and student loans, the estimate varies from 3.17 cents considering only the high-cost inefficiency to 3.54 cents, also considering under-credit provision for low-cost individuals. Panel B repeats the exercise using account openings and the principals on new accounts as quantity estimates instead. However, the efficiency costs estimated are very close to those estimated in Panel A. In summary, across various credit markets and using estimates from other researchers, we show that the cost for redistribution via data removal is a few cents per dollar redistributed.

5 Concluding Remarks

This paper presents a new framework for studying the role of data acquisition in consumer credit markets. We show that prices and borrowing changes resulting from new data are sufficient statistics for welfare analysis. Using administrative data, we apply a novel methodology to the removal of bankruptcy flags and examine the effects of such removals in auto lending. We find that the welfare losses from removing bankruptcy flags are small relative to the welfare transfers. Thus, flag removal is an efficient way to transfer surplus to previously bankrupt consumers.

While we present a specific application associated with the welfare benefits of data acquisition, the method is broadly applicable in financial markets. Future work could study data acquisition in other lending markets, and explore other contexts in which data acquisition leads to consumer benefits. For example, certain types of data acquisition may lead to large welfare gains relative to welfare transfers. Further, the welfare gains associated with data acquisition may be small in certain settings, suggesting that the privacy or equity gains from making data unavailable may outweigh the direct benefits of using the data to screen borrowers.

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar, "Too Much Data: Prices and Inefficiencies in Data Markets," Technical Report, National Bureau of Economic Research 2019.

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 2009, *99* (1), 49–84.

Argyle, Bronson, Benjamin Iverson, Taylor D Nadauld, and Christopher Palmer, "Personal Bankruptcy and the Accumulation of Shadow Debt," Technical Report, National Bureau of Economic Research 2021.

_ , Sasha Indarte, Ben Iverson, and Christopher Palmer, "Explaining Racial Disparities in Personal Bankruptcy Outcomes," 2022. Working Paper.

_ , Taylor D Nadauld, and Christopher Palmer, "Real Effects of Search Frictions in Consumer Credit Markets," Technical Report, National Bureau of Economic Research 2020.

Bachas, Natalie, Olivia S Kim, and Constantine Yannelis, "Loan Guarantees and Credit Supply," *Journal of Financial Economics*, 2021, *139* (3), 872–894.

Barrios, John Manuel, "Staggeringly Problematic: A Primer on Staggered DiD for Accounting Researchers," *Working Paper*, 2021.

Begenau, Juliane, Maryam Farboodi, and Laura Veldkamp, "Big Data in Finance and the Growth of Large Firms," *Journal of Monetary Economics*, 2018, *97*, 71–87.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris, "The Limits of Price Discrimination," *American Economic Review*, 2015, *105* (3), 921–57.

Blattner, Laura, Jacob Hartwig, and Scott Nelson, "Information Design in Consumer Credit Markets," Technical Report, Working paper 2022.

Bos, Marieke, Emily Breza, and Andres Liberman, "The Labor Market Effects of Credit Market Information," *The Review of Financial Studies*, 2018, *31* (6), 2005–2037.

Calomiris, CW and L Neal, "History of Financial Globalization, Overview," *Handbook of Key Global Financial Markets, Institutions and Infrastructure*, 2013, *1*, 3–14.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-wage Jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chatterjee, Satyajit, Dean Corbae, Kyle P Dempsey, and José-Víctor Ríos-Rull, "A Quantitative Theory of the Credit Score," Technical Report, National Bureau of Economic Research 2020.

Chen, Yongmin and Marius Schwartz, "Differential Pricing when Costs Differ: A Welfare Analysis," *The RAND Journal of Economics*, 2015, *46* (2), 442–460.

Chiappori, Pierre-André and Bernard Salanie, "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 2000, *108* (1), 56–78.

Dávila, Eduardo, "Using Elasticities to Derive Optimal Bankruptcy Exemptions," *The Review of Economic Studies*, 2020, *87* (2), 870–913.

_ **and Andreas Schaab**, "Welfare Assessments with Heterogeneous Individuals," 2021.

_ **and Ansgar Walther**, "Corrective Regulation with Imperfect Instruments," Technical Report, National Bureau of Economic Research 2021.

DeFusco, Anthony A, Huan Tang, and Constantine Yannelis, "Measuring the Welfare Cost of Asymmetric Information in Consumer Credit Markets," *Journal of Financial Economics*, 2022, *146* (3), 821–840.

Dobbie, Will, Paul Goldsmith-Pinkham, and Crystal S Yang, "Consumer Bankruptcy and Financial HHealth," *Review of Economics and Statistics*, 2017, *99* (5), 853–869.

_ , _ , **Neale Mahoney, and Jae Song**, "Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports," *The Journal of Finance*, 2020, *75* (5), 2377–2419.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen, "Estimating Welfare in Insurance Markets Using Variation in Prices," *Quarterly Journal of Economics*, 2010, *125* (3), 877–921.

Elul, Ronel and Piero Gottardi, "Bankruptcy: Is it enough to forgive or must we also forget?," *American Economic Journal: Microeconomics*, 2015, *7* (4), 294–338.

European Data Protection Supervisor, "Opinion on the Proposal for a Directive on Consumer Credits," 2021.

Farboodi, Maryam and Laura Veldkamp, "Long-run Growth of Financial Data Technology," *American Economic Review*, 2020, *110* (8), 2485–2523.

_ , **Roxana Mihet, Thomas Philippon, and Laura Veldkamp**, "Big Data and Firm Dynamics," in "AEA papers and proceedings," Vol. 109 2019, pp. 38–42.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, "Predictably Unequal? The Effects of Machine Learning on Credit Markets," *The Journal of Finance*, 2022, *77* (1), 5–47.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021.

Gross, Tal, Matthew J Notowidigdo, and Jialan Wang, "The Marginal Propensity to Consume Over the Business Cycle," *American Economic Journal: Macroeconomics*, 2020, *12* (2), 351–84.

_ , **Raymond Kluender, Feng Liu, Matthew J Notowidigdo, and Jialan Wang**, "The Economic Consequences of Bankruptcy Reform," *American Economic Review*, 2021, *111* (7), 2309–41.

Harberger, Arnold C., "The Measurement of Waste," *American Economic Review*, 1964, *54* (3), 58–76.

He, Zhiguo, Jing Huang, and Jidong Zhou, "Open Banking: Credit Market Competition When Borrowers Own the Data," Technical Report, National Bureau of Economic Research 2020.

Herkenhoff, Kyle, "The Impact of Consumer Credit Access on Unemployment," *The Review of Economic Studies*, 2019, *86* (6), 2605–2642.

_ , Gordon M Phillips, and Ethan Cohen-Cole, "The Impact of Consumer Credit Access on Self-employment and Entrepreneurship," *Journal of Financial Economics*, 2021, *141* (1), 345–371.

Huang, Jing, "Fintech Expansion," Technical Report, Working Paper 2022.

Indarte, Sasha, "Moral Hazard versus Liquidity in Household Bankruptcy," 2021. Working Paper.

Jehle, G.A. and P.J. Reny, *Advanced Microeconomic Theory* The Addison-Wesley series in economics, Financial Times/Prentice Hall, 2011.

Jones, Charles I and Christopher Tonetti, "Nonrivalry and the Economics of Data," *American Economic Review*, 2020, *110* (9), 2819–58.

Keys, Benjamin J, Neale Mahoney, and Hanbin Yang, "What Determines Consumer Financial Distress? Place-and Person-Based Factors," Technical Report, National Bureau of Economic Research 2020.

Kovbasyuk, Sergey and Giancarlo Spagnolo, "Memory and Markets," *Available at SSRN 2756540*, 2021.

Liberman, Andres, Christopher A. Neilson, Luis Opazo, and Seth Zimmerman, "The Equilibrium Effects of Information Deletion: Evidence from Consumer Credit Markets," 2019. Working Paper.

Liu, Zhuang, Michael Sockin, and Wei Xiong, "Data Privacy and Algorithmic Inequality," Technical Report, National Bureau of Economic Research 2023.

Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta, "Mortgage Modification and Strategic Behavior: Evidence from a Legal Settlement with Countrywide," *American Economic Review*, 2014, *104* (9), 2830–57.

Musto, David K, "What Happens When Information Leaves a Market? Evidence from Post-bankruptcy Consumers," *The Journal of Business*, 2004, *77* (4), 725–748.

Nelson, Scott, "Private Information and Price Regulation in the US Credit Card Market," *Unpublished Working Paper*, 2018.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Padilla, A Jorge and Marco Pagano, "Sharing default information as a borrower discipline device," *European Economic Review*, 2000, *44* (10), 1951–1980.

Pagano, Marco and Tullio Jappelli, "Information Sharing in Credit Markets," *The Journal of Finance*, 1993, *48* (5), 1693–1718.

Schmalensee, Richard, "Output and Welfare Implications of Monopolistic Third-degree Price Discrimination," *The American Economic Review*, 1981, *71* (1), 242–247.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Tang, Huan, "The Value of Privacy: Evidence from Online Borrowers," 2019.

Varian, Hal R, "Price Discrimination and Social Welfare," *The American Economic Review*, 1985, *75* (4), 870–875.

_ , "Price Discrimination," *Handbook of Industrial Organization*, 1989, *1*, 597–654.

Vercammen, James A, "Credit Bureau Policy and Sustainable Reputation Effects in Credit Markets," *Economica*, 1995, pp. 461–478.

Yannelis, Constantine, "Strategic Default on Student Loans," Technical Report, Working paper 2016.

_ **and Anthony Lee Zhang**, "Competition and Selection in Credit Markets," Technical Report, National Bureau of Economic Research 2021.

Appendix

A Proofs

A.1 Proof of Claim 1

First, we will rewrite the borrower's optimization problem (6), using $\phi(r)$ from expression (3):

$$V(r) = \max_{L} u_0(w_0 + L) + \sum_{t=1}^{T} \beta^t (1-\delta)^t \left[u(w_t) - u'(w_t) L\phi(r)\right] + \sum_{t=1}^{T} (1-\delta)^{t-1} \delta \sum_{\tilde{t}=t}^{T} \beta^{\tilde{t}} u(c_D) \quad (39)$$

Rearranging,

$$V(r) = \max_{L} u_0(w_0 + L) - L\phi(r) \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t) + \sum_{t=1}^{T} (1-\delta)^{t-1} \delta \sum_{\tilde{t}=t}^{T} \beta^{\tilde{t}} u(c_D) + \sum_{t=1}^{T} \beta^t (1-\delta)^t u(w_t) \quad (40)$$

To prove Claim 1, we will first prove Claims 6 and 7.

Claim 6. *There is some maximum interest rate, ρ, above which the borrower optimally borrows nothing. Thus,*

$$V(r) = \bar{V} \ \forall r \geq \rho$$

Proof. First, in (40) note that $u_0(\cdot)$ is concave, the term:

$$\sum_{t=1}^{T} \beta^t (1-\delta)^t \left[u(w_t) - u'(w_t) L\phi(r)\right]$$

is linear in L, and all other terms do not depend on L. Thus, (40) is concave in L, so the first-order condition with respect to L is sufficient for optimality. Differentiating (40) and rearranging, the borrower's FOC for optimal loan size is:

$$u_0'(w_0 + L) = \phi(r) \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)$$

whenever $L^*(r) > 0$. If

$$u_0'(w_0) < \phi(r) \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t) \tag{41}$$

then the optimal loan size is 0. Now, from (3), $\phi(\cdot)$ increases towards ∞ as r increases towards ∞; thus, for r sufficiently large, the RHS of (41) becomes unbounded. Thus, there always exists some r such that the inequality (41) holds, and the borrower's optimal choice is to borrow nothing, $L^*(r) = 0$. □

Claim 7. *We have, for any r:*

$$\frac{dV(r)}{dr} = -L^*(r) \frac{d\phi(r)}{dr} \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t) \tag{42}$$

Proof. Differentiating (40) using the envelope theorem, we have:

$$\frac{dV(r)}{dr} =$$

$$\frac{\partial}{\partial r} \left[u_0(w_0 + L) - \underbrace{L\phi(r) \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)}_{Rate-dependent\ term} + \sum_{t=1}^{T} (1-\delta)^{t-1} \delta \sum_{\tilde{t}=t}^{T} \beta^t u(c_D) + \sum_{t=1}^{T} \beta^t (1-\delta)^t u(w_t) \right]$$

Note that only the term with an underbrace depends on r. Hence, we have:

$$\frac{dV(r)}{dr} = -L^*(r) \frac{d\phi(r)}{dr} \sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)$$

This proves Claim 7. In particular, note that Claim 7 holds even for $r > \rho$, since $L^*(r) = 0$, and thus from (42), $\frac{dV}{dr} = 0$. □

The intuition behind Claim 7 is as follows. The effect of a small change in r is simply to make the borrower pay slightly more, $L^*(r) \frac{d\phi}{dr}$, in each future period assuming she does not default. Thus, multiplying by the sum of marginal utilities in these periods, multiplied by the probability of reaching these periods, we get $\frac{dV}{dr}$.

Now, to prove Claim 1, note that from (8), for $r = \rho$, the borrower does not borrow, $L^*(r) = 0$, hence $V(r) = \bar{V}$, hence (8) holds. For $r < \rho$, (8) follows from integrating the derivative $\frac{dV}{dr}$, from (42) of Claim 7.

A.2 Proof of Claim 2

Applying (8), we have:

$$CS(r) = \sum_{i=1}^{N} V_i(r) = \sum_{i=1}^{N} \bar{V}_i + \left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t) \int_r^{\rho} L^*(r) \frac{d\phi}{dr} dr \right]$$

Hence, we have:

$$CS(r) - CS(\tilde{r}) = \sum_{i=1}^{N} \left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t) \int_r^{\tilde{r}} L^*(r) \frac{d\phi}{dr} dr \right]$$

Given the normalization in (10) of assumption 1, we have:

$$CS(r) - CS(\tilde{r}) = \sum_{i=1}^{N} \int_r^{\tilde{r}} L^*(r) \frac{d\phi}{dr} dr$$

Switching the integral and the sum,

$$CS(r) - CS(\tilde{r}) = \int_r^{\tilde{r}} \sum_{i=1}^{N} L^*(r) \frac{d\phi}{dr} dr$$

Using aggregate loan volume, defined in (12), we have:

$$CS(r) - CS(\tilde{r}) = \int_r^{\tilde{r}} \Lambda(r) \frac{d\phi}{dr} dr$$

which is (11).

The intuition here is simply that, using the normalization in Assumption 1, we can aggregate (8) of Claim 1 across borrowers. The change in borrower surplus, when rates change from $\tilde{r}$ to r, is the borrower surplus triangle $\int_r^{\tilde{r}} \Lambda(r) \frac{d\phi}{dr} dr$, adjusted by a factor $\frac{d\phi}{dr}$, which says how loan payments change as r changes.

A.3 Welfare Maximization and Pareto Efficiency

Consider allocations in which borrower i faces rate r_i. As defined in Subsection 2.2.2 of the main text, aggregate welfare across borrowers is:

$$CS = \sum_{i=1}^{n} V(r_i) = \sum_{i=1}^{n} \left[\bar{V}_i + \int_{r_i}^{\rho} L_i^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r} \right] \tag{43}$$

Lender surplus, summing across all borrowers i, is:

$$PS = \sum_{i=1}^{n} L_i^*(r_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right) \tag{44}$$

Total surplus, defined as the sum of (43) and (44), can thus be written as:

$$TS = \sum_{i=1}^{n}\left[\bar{V}_i + \int_{r_i}^{\rho} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r} + L_i^*(r_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right)\right] \tag{45}$$

We now prove two claims. Claim 8 shows that social welfare, defined as in (45), is maximized by setting $r_i = r_{fair}$ for all borrowers i. Claim 9 then shows that allocations which do not maximize social welfare defined this way are Pareto dominated by an allocation setting rates to $r_i = r_{fair}$, and implementing some transfers to borrowers and lenders in non-default periods.

A.3.1 Social welfare maximization

Claim 8. *Social welfare for group i, defined as the sum of (43) and (44), is maximized by setting $r_i = r_{fair}$ for all agents i.*

Proof. The maximization problem is:

$$\max_{r_i \ldots r_n} \sum_{i=1}^{n}\left[\bar{V}_i + \int_{r_i}^{\rho} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r} + L_i^*(r_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right)\right] \tag{46}$$

Note that this problem is separable across borrowers: for borrower i, we solve:

$$\max_{r_i}\left[\bar{V}_i + \int_{r_i}^{\rho} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r} + L_i^*(r_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right)\right]$$

For borrower i, total welfare at r_{fair} is:

$$\bar{V}_i + \int_{r_{fair}}^{\rho} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r} \tag{47}$$

Now, consider any other higher price $r_{i,high} > r_{fair}$. The difference between borrowers' surplus at $r_{i,high}$ and r_{fair} is:

$$\left(\bar{V}_i + \int_{r_{fair}}^{\rho} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r}\right) - \left(\bar{V}_i + \int_{r_{i,high}}^{\rho} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r}\right) = \int_{r_{fair}}^{r_{i,high}} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r} \tag{48}$$

Lenders make zero profits at r_{fair}, and at $r_{i,high}$ their profits are:

$$L_i^*\left(r_{i,high}\right)\left(\phi\left(r_{i,high}\right)-\phi\left(r_{fair}\right)\right) \tag{49}$$

Now, note that we can write borrowers' problem, (6) in the main text, as:

$$V(r)=\max_L u_0(w_0+L)+ \\ \sum_{t=1}^{T}\beta^t(1-\delta)^t\left[u(w_t)-u'(w_t)\phi(r)L\right]+\sum_{t=1}^{T}(1-\delta)^{t-1}\delta\sum_{\tilde{t}=t}^{T}\beta^t u(c_D) \tag{50}$$

Expression (50) shows that $V(r)$ has increasing differences in r and $-L$. Thus, by Topkis' theorem, the optimal loan choice $L^*(r)$ is weakly decreasing in r. Thus, we have:

$$\int_{r_{fair}}^{r_{i,high}} L_i^*(\hat{r})\frac{d\phi}{d\hat{r}}d\hat{r} \geq \int_{r_{fair}}^{r_{i,high}} L_i^*\left(r_{i,high}\right)\frac{d\phi}{d\hat{r}}d\hat{r} = L_i^*\left(r_{i,high}\right)\left(\phi\left(r_{i,high}\right)-\phi\left(r_{fair}\right)\right)$$

This shows that borrowers' losses from increasing rates to $r_{i,high}$, in (48), are always larger than lenders' profit increases, in (49). Analogously, for any $r_{i,low} < r_{fair}$, borrowers' losses outweigh lenders' profit increases. This shows that any r_i which is different from r_{fair} is weakly dominated by r_{fair}, proving that setting $r_i = r_{fair}$ maximizes welfare for group i. Thus, to maximize (46), it is optimal to set $r_i = r_{fair}$ for all groups. □

A.3.2 Pareto optimality

Next, we will show that any interest rates which do not maximize social welfare, as defined in (45), are Pareto dominated with some set of transfers. First, we define the set of transfers we consider. Suppose the social planner imposes taxes or subsidies on lenders and borrowers, which are constant for each borrower or lenders within every non-default period. Let t_i^C denote the transfer in each non-default period paid to borrower i, and let t_i^P denote the transfer paid to the lender to borrower i, in each period when borrower i does not default. Adding the transfers for borrower i, and the lender lending to borrower i, the planner's net expenditure in each non-default period for borrower i is $t_i^C + t_i^P$. Hence, the planner breaks even or makes a surplus in each non-default period if and only if $t_i^C + t_i^P \leq 0$; that is, if net subsidies paid to borrowers and lenders sum to weakly less than 0.[34]

[34]Note that, with this class of transfers, the planner receives a net amount $-\left(t_i^C + t_i^P\right)$ in each non-default period, and receives nothing when borrowers default. Hence, the planner has no risk of having to pay any net amount: if $t_i^C + t_i^P \leq 0$, the planner breaks even or makes a surplus in each non-default period for each borrower, and makes no payments upon default, so there is no state of the world where the planner ever makes a net payment to borrowers and lenders.

Since we have assumed transfers are paid in periods after the first, and we have adopted the linear approximation of borrower utility in (5), transfers do not affect borrowers' optimal loan demand $L^*(r)$. Under a transfer t_i^C and an interest rate r, i's utility is:

$$V\left(r, t_i^C\right) = \bar{V} + \left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]\left[t_i^C + \int_r^{\rho} L^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r}\right] \tag{51}$$

Expression (51) is just borrower utility in (43), plus the size of the transfer t_i^C multiplied by the utility weight $\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]$. For lenders, note that lenders' per-period profit with transfers t_i^P is:

$$\Pi\left(r, t_i^P\right) = \underbrace{L_i^*(\tilde{r}_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right)}_{A} + t_i^P \tag{52}$$

That is, lenders simply receive their per-period profits from lending to a borrower i, which is term A in (52), plus the transfer t_i^P.

Claim 9. *Consider any interest rates $\tilde{r}_i \neq r_{fair}$, which do not maximize social welfare as defined in (45). These interest rates are Pareto dominated with transfers: it is possible to change rates to $r_i = r_{fair}$, and implement transfers in every non-default period across borrowers and lenders, which make all borrowers and lenders weakly better off.*

Proof. We will first demonstrate that the claim holds for a single borrower and lender, and then generalize. Suppose there is a single lender lending only to borrower i. The utility of borrower i at rate r_i is:

$$V_i(r) = \bar{V} + \underbrace{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]}_{Utility\ weight}\underbrace{\left[\int_{r_i}^{\rho} L_i^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r}\right]}_{Payment\ change}$$

The lender's profit from serving borrower i at rate r_i is:

$$\Pi_i(r) = L_i^*(r_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right)$$

Now, consider an alternative interest rate $\tilde{r}_i \neq r_{i,fair}$ for borrower i. Define:

$$V(\tilde{r}_i) = \bar{V} + \underbrace{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]}_{Utility\ weight}\underbrace{\left[\int_{\tilde{r}_i}^{\rho} L^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r}\right]}_{Payment\ change}$$

$$\Pi_i(\tilde{r}_i) = L_i^*(\tilde{r}_i)\left(\phi(\tilde{r}_i) - \phi\left(r_{fair}\right)\right)$$

We will show that it is possible to change prices to r_{fair}, and construct transfers such that the

borrower and lender are both just as well off as they are under $\tilde{r}_i$, while making the planner a budget surplus. First, let the transfer to borrower i be:

$$t_{C,i} = \frac{V(\tilde{r}_i) - V\left(r_{fair}\right)}{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]} \tag{53}$$

□

Plugging in to (51), borrower i's utility under r_{fair} is thus:

$$V_i\left(r_{fair}, t_{C,i}\right) = \bar{V} + \left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]\left[\int_{r_{fair}}^{\rho} L^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r} + \frac{V(\tilde{r}_i) - V(r)}{\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)}\right]$$

$$= \bar{V} + \left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]\left[\int_{\tilde{r}_i}^{\rho} L^*(\hat{r}) \frac{d\phi}{d\hat{r}} d\hat{r}\right] = V(\tilde{r}_i)$$

Hence, borrower i is indifferent between receiving rate $\tilde{r}_i$ with no transfers, and rate r_{fair} with a transfer of t_i^C.

Similarly, let us pay the lender a net transfer in each period of:

$$t_{P,i} = \Pi(\tilde{r}_i) - \Pi(r_i) \tag{54}$$

From (52), the lender's total per period profit, if they lent at $r_{i,fair}$ and receive transfers of $t_{P,i}$, is thus:

$$L_i^*(\tilde{r}_i)\left(\phi(r_i) - \phi\left(r_{fair}\right)\right) + \Pi(\tilde{r}_i) - \Pi(r_i) = \Pi(\tilde{r}_i)$$

Hence, the lender is indifferent between getting $\tilde{r}_i$ with no transfers, and r_{fair} with transfers $t_{P,i}$.

Next, we need to show that the social planner at least breaks even. Adding the transfers, we have:

$$\begin{aligned} t_{C,i} + t_{P,i} &= \left(V(\tilde{r}_i) - V\left(r_{fair}\right)\right) + \left(\Pi(\tilde{r}_i) - \Pi\left(r_{fair}\right)\right) \\ &= \underbrace{\left(V(\tilde{r}_i) + \Pi(\tilde{r}_i)\right)}_{A} - \underbrace{\left(V\left(r_{fair}\right) + \Pi\left(r_{fair}\right)\right)}_{B} \end{aligned} \tag{55}$$

Term A and B in (55) are respectively just social welfare for group i, defined using the normalizations in Assumption 1, evaluated at interest rates $\tilde{r}_i$ and r_i respectively. We showed in Claim 8 that social welfare, defined thus, is maximized at r_{fair}; thus, term B is always greater than term A. As a result, (55) is weakly negative: the total cost to the social planner of implementing the transfers in (53) and (54) is negative, so the planner collects a surplus. Thus, beginning from rates $\tilde{r}_i \neq r_{fair}$, the social planner can change the rate to $r_i = r_{fair}$, and implement the

transfers in (53) and (54); these transfers always leave borrower i and her lender equally well-off. If welfare is strictly higher under r_{fair} than $\tilde{r}_i$, then the planner makes a strictly positive budget surplus; this surplus can be redistributed to borrower i and her lender, by increasing both t_i^C and t_i^P by some small amount, creating a strict Pareto improvement.

We have thus shown that, for any individual borrower i, given any rate $\tilde{r}_i \neq r_{fair}$, outcomes are Pareto dominated by an allocation in which the rate is r_{fair} and the social planner implements the transfers in (53) and (54). When there are many borrowers i, in borrower groups j with possibly different default rates δ_i and break-even interest rates r_{fair}, assuming each borrower is served by a separate lender, the result follows by applying the transfers in (53) and (54) within each group. Note in particular that the optimality result does not even require implementing transfers across groups of consumers with different default rates.

When a single lender may lend to multiple borrowers, the result still holds: to show this, note that we can think of an aggregated lender as consisting of multiple subsidiaries, each of which lends to a single borrower. Claim 9 shows that all subsidiaries of a given lender have equal profits under rate r_{fair} with the transfers in (53) and (54), compared to $\tilde{r}_i$. Profits of the lender as a whole are the sum of profits of the lender's subsidiaries; hence, the lender as a whole has equal profits under rate r_{fair} with transfers, compared to $\tilde{r}_i$.

A.4 Proof of Claim 3

Lender profits. Since we have assumed $r_{j,fair}$ is equal to the marginal cost of providing credit, at rate $r_{j,fair}$, and payment $\phi\left(r_{j,fair}\right)$, lenders break even on borrowers. At rate r_{pool}, lenders' profit, per non-default period and per dollar lent, is:

$$\phi\left(r_{pool}\right) - \phi\left(r_{j,fair}\right)$$

Thus, lenders' total profit at rate r_{pool} is:

$$\Lambda\left(\phi\left(r_{pool}\right)\right)\left[\phi\left(r_{pool}\right) - \phi\left(r_{j,fair}\right)\right] \tag{56}$$

The change in profits on group j, when moving from rate r_{pool} to $r_{j,fair}$, is thus the negative of (56).

Borrower welfare. From Claim 2, we have:

$$CS\left(r\right) - CS\left(\tilde{r}\right) = \int_r^{\tilde{r}} \Lambda\left(\phi\left(\hat{r}\right)\right) \frac{d\phi\left(\hat{r}\right)}{d\hat{r}} d\hat{r}$$

Changing variables to ϕ, we can alternatively write this as:

$$CS(r) - CS(\tilde{r}) = \int_{\phi(r)}^{\phi(\tilde{r})} \Lambda(\phi)\, d\phi \tag{57}$$

Since $\Lambda(\phi)$ is linear, (57) is equal to (16). Summing (15) and (16) and rearranging, we get (17).

A.5 Proof of Claim 4

If we removed bankruptcy flags on all H borrowers, they would be indistinguishable from L borrowers. Lenders' net profits from group L at interest rate r, in each non-default period, are:

$$\Lambda_L(\phi(r))\left(\phi(r) - \phi\left(r_{L,fair}\right)\right)$$

This is the loan volume to L, multiplied by the difference between the price $\phi(r)$ and the fair price $\phi\left(r_{L,fair}\right)$. When flags are removed, so the two groups cannot be distinguished, lenders will set a price such that they break even on average across both groups, so their losses on the type-H group are made up for by gains on the type-L group. To write lenders' zero-profit condition, we must also account for the fact that default rates differ across groups. Let δ_H and δ_L respectively be the default rates in group H and L. The total profits that lenders attain from lending $\Lambda_L(\phi(r_L))$ to group L and $\Lambda_H(\phi(r_H))$ to group H, summed across all time periods, is:

$$\sum_{t=1}^{T}(1-\delta_H)^t\left[\Lambda_H(\phi(r_H))\left(\phi(r_H) - \phi\left(r_{H,fair}\right)\right)\right] + (1-\delta_L)^t\left[\Lambda_L(\phi(r_L))\left(\phi(r_L) - \phi\left(r_{L,fair}\right)\right)\right] \tag{58}$$

In words, profits from the H-group is the sum of profits per non-default period, multiplied by the geometric sum $\sum_{t=1}^{T}(1-\delta_H)^t$, which is the expected number of non-default periods among type H borrowers, and likewise for the L-group. Setting profits in (58) to zero, and applying the geometric series formula for $\sum_{t=1}^{T}(1-\delta_H)^t$, we have:[35]

$$\psi_H \Lambda_H(\phi(r_H))\left(\phi(r_H) - \phi\left(r_{H,fair}\right)\right) = \psi_L \Lambda_L(\phi(r_L))\left(\phi\left(r_{L,fair}\right) - \phi(r_L)\right) \tag{59}$$

Where,

$$\psi_H \equiv (1-\delta_H)\left(\frac{1-(1-\delta_H)^T}{\delta_H}\right),\ \psi_L = (1-\delta_L)\left(\frac{1-(1-\delta_L)^T}{\delta_L}\right) \tag{60}$$

The terms ψ_H and ψ_L capture the expected number of non-default periods for groups H

[35] Note that we do not need to account for recovery rates in (59), because these are included in the break-even interest rates $r_{H,fair}$ and $r_{L,fair}$.

and L respectively. Rearranging, we have:

$$\phi\left(r_{pool}\right)-\phi\left(r_{L,fair}\right)=\frac{\psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right) \tag{61}$$

This is (22) of Claim 4. In words, (61) states that the price increase for L borrowers is the product of the ratio of total loan volumes, $\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}$, the ratio of expected non-default periods $\frac{\psi_H}{\psi_L}$, and the price change to the H-group, $\left(\phi\left(r\right)-\phi\left(r_{H,fair}\right)\right)$. The ratio $\frac{\psi_H}{\psi_L}$ will tend to be close to 1 when default rates are similar across groups; when the L-group has lower default rate than the H group, this ratio will tend to be small. When the L-group is much larger than the H-group, $\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}$ will tend to be small, so the price change for the L-group will tend to be small.

Using (61), we can calculate the change in borrower surplus and social welfare for the L-group. From (17) of Claim 3, the change in social welfare for the L-group from removing data is:

$$\frac{1}{2}\left(\phi\left(r_{L,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\Lambda_L\left(\phi\left(r_{L,fair}\right)\right)-\Lambda_L\left(\phi\left(r_{pool}\right)\right)\right)$$

Substituting using (61), this is:

$$\frac{1}{2}\frac{\psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\left(\Lambda_L\left(\phi\left(r_{L,fair}\right)\right)-\Lambda_L\left(\phi\left(r_{pool}\right)\right)\right) \tag{62}$$

Now, writing (62) using the demand slope b_L, this becomes:

$$\frac{1}{2}\frac{\psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\left(b_L\left(\phi\left(r_{pool}\right)-\phi\left(r_{L,fair}\right)\right)\right)$$

Substituting again using (61), this is:

$$=\frac{1}{2}\frac{\psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\times \left(b_L\left(\frac{\psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\right)\right) \tag{63}$$

Next, we can use the assumption in (24) that the elasticities of demand in the two groups are the same, so that $b_L=b_H\frac{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}$. Substituting into (63), we get:

$$= \frac{1}{2}\frac{\psi_H \Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\left(b_H\left(\frac{\psi_H}{\psi_L}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\right)\right)$$

This simplifies to:

$$= \frac{1}{2}\left(\frac{\psi_H}{\psi_L}\right)^2\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\left[b_H\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\right] \tag{64}$$

Expression (64) is the welfare change for L-group borrowers, in dollars per non-default period. To convert this into expected dollars over the lifetime of a loan, we multiply by ψ_L, to get:

$$\frac{1}{2}\frac{\psi_H^2}{\psi_L}\frac{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\left[b_H\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right)\right] \tag{65}$$

This is (25) of Claim 4.

Now, from (17) of Claim 3, and substituting for the Λ_H terms using (27), the change in social welfare for the H-group, in dollars per non-default period, is:

$$\frac{1}{2}b_H\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right) \tag{66}$$

Multiplying by ψ_H, the change in welfare in expected dollars over the lifetime of a loan is:

$$\psi_H\frac{1}{2}b_H\left(\phi\left(r_{H,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\phi\left(r_{pool}\right)-\phi\left(r_{H,fair}\right)\right) \tag{67}$$

Comparing (67) and (65), the change in welfare for the L-group is a factor $\frac{\psi_H\Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L\Lambda_L\left(\phi\left(r_{pool}\right)\right)}$ times the change in welfare for the H-group.

The change in borrower surplus for the L-group can be calculated by plugging expressions for $\phi\left(r_{L,fair}\right)-\phi\left(r_{pool}\right)$, $\Lambda_L\left(\phi\left(r_{pool}\right)\right)$, b_L, and $\Lambda_L\left(\phi\left(r_{L,fair}\right)\right)$ into (16) of Claim 3. The change in borrower surplus is:

$$\left(\phi\left(r_{L,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\frac{\Lambda_L\left(\phi\left(r_{pool}\right)\right)+\Lambda_L\left(\phi\left(r_{L,fair}\right)\right)}{2}\right)$$

We can rearrange this to:

$$\left(\phi\left(r_{L,fair}\right)-\phi\left(r_{pool}\right)\right)\left(\Lambda_L\left(\phi\left(r_{pool}\right)\right)+\frac{\Lambda_L\left(\phi\left(r_{L,fair}\right)\right)-\Lambda_L\left(\phi\left(r_{pool}\right)\right)}{2}\right)$$

$$= \left(\phi\left(r_{L,fair}\right) - \phi\left(r_{pool}\right)\right)\left(\Lambda_L\left(\phi\left(r_{pool}\right)\right) + \frac{b_L}{2}\left(\phi\left(r_{pool}\right) - \phi\left(r_{L,fair}\right)\right)\right)$$

Plugging in $b_L = b_H \frac{\Lambda_L(\phi(r_{pool}))}{\Lambda_H(\phi(r_{pool}))}$, and plugging in for $\phi\left(r_{pool}\right) - \phi\left(r_{L,fair}\right)$ using (61), we get that the borrower surplus change is:

$$\frac{\psi_H \Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{pool}\right) - \phi\left(r_{H,fair}\right)\right) \times \left(\Lambda_L\left(\phi\left(r_{pool}\right)\right) + \frac{b_H}{2}\frac{\Lambda_L\left(\phi\left(r_{pool}\right)\right)}{\Lambda_H\left(\phi\left(r_{pool}\right)\right)}\left(\frac{\psi_H \Lambda_H\left(\phi\left(r_{pool}\right)\right)}{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right)}\left(\phi\left(r_{H,fair}\right) - \phi\left(r_{pool}\right)\right)\right)\right) \tag{68}$$

A.6 Proof of Claim 5

Repeating the definitions in (32) and (33), we have:

$$s_L \equiv \frac{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right)}{\psi_L \Lambda_L\left(\phi\left(r_{pool}\right)\right) + \psi_H \Lambda_H\left(\phi\left(r_{pool}\right)\right)} \tag{69}$$

$$\Delta \equiv \phi\left(r_{H,fair}\right) - \phi\left(r_{L,fair}\right) \tag{70}$$

Using (69) and (70), we can write (30) as:

$$\phi\left(r_{pool}\right) = \phi\left(r_{L,fair}\right) + (1 - s_L)\Delta + m_{pool} \tag{71}$$

Or,

$$\phi\left(r_{pool}\right) = \phi\left(r_{H,fair}\right) - s_L \Delta + m_{pool} \tag{72}$$

To calculate welfare, we will calculate the welfare of each group, when data is available and is not, relative to the fully efficient case, using the result of Claim 3. For the low group, when data is available, the social welfare loss relative to the fully efficient case can be obtained using (17):

$$\frac{1}{2}\left(\phi\left(r_L\right) - \phi\left(r_{L,fair}\right)\right)\left(\Lambda_L\left(\phi\left(r_{L,fair}\right)\right) - \Lambda_L\left(\phi\left(r_L\right)\right)\right) \tag{73}$$

$$= \frac{1}{2}\left(\phi\left(r_{L,fair}\right) + m_L - \phi\left(r_{L,fair}\right)\right)\left(\left(a_L - b_L \phi\left(r_{L,fair}\right)\right) - \left(a_L - b_L\left(\phi\left(r_{L,fair}\right) + m_L\right)\right)\right) \tag{74}$$

$$Loss_{L,data} = \frac{b_L}{2} m_L^2 \tag{75}$$

Intuitively, there is a welfare loss from markups, which depends on the size of the markup m_L,

and the slope of demand b_L. Similarly, for the high group, the welfare loss from markups is:

$$Loss_{H,data} = \frac{b_H}{2} m_H^2 \tag{76}$$

Using these expressions, we calculate the welfare loss, relative to the fully efficient benchmark, in the case of pooled pricing. Note that we can write (71) as:

$$\phi\left(r_{pool}\right) = \phi\left(r_{L,fair}\right) + (1 - s_L)\,\Delta + m_{pool}$$

Plugging this into (73), expanding and simplifying, the welfare loss for the L group in the no-data case is:

$$Loss_{L,nodata} = \frac{b_L}{2}\left(m_{pool} + (1 - s_L)\,\Delta\right)^2 \tag{77}$$

And for the H group, we have:

$$Loss_{H,nodata} = \frac{b_H}{2}\left(m_{pool} - s_L \Delta\right)^2 \tag{78}$$

Hence, the welfare change when going from the pooled case to the case with data is

$$\frac{b_L}{2}\left(\left(m_{pool} + (1 - s_L)\Delta\right)^2 - m_L^2\right) \tag{79}$$

for the L group, and

$$\frac{b_H}{2}\left(\left(m_{pool} - s_L \Delta\right)^2 - m_H^2\right) \tag{80}$$

for the H group, in terms of dollars per non-default period. To convert these quantities into expected dollars over the term of a loan, we will multiply each by the expected number of non-default periods, ψ_L and ψ_H, defined in (23) of Claim 4. The total welfare change from data availability is thus:

$$\Delta Welfare = \psi_H Loss_{H,nodata} + \psi_L Loss_{L,nodata} - \psi_H Loss_{H,data} - \psi_L Loss_{L,data}$$

$$\Delta Welfare = \underbrace{\psi_H \frac{b_H}{2}\left(\left(m_{pool} - s_L \Delta\right)^2 - m_H^2\right)}_{High\ group} + \underbrace{\psi_L \frac{b_L}{2}\left(\left(m_{pool} + (1 - s_L)\,\Delta\right)^2 - m_L^2\right)}_{Low\ group} \tag{81}$$

This is (31).

B Additional Theoretical Results

B.1 Strategic Bankruptcy and Incentive Effects of Flag Removals

Bankruptcy flags on credit reports provide information to the market about borrowers' default risks, increasing the efficiency of credit allocation. While only 39.6% of bankrupt individuals know the duration of bankruptcy flags on their credit file,[36] bankruptcy flags may still have an incentive effect: borrowers who declare bankruptcy face higher interest rates in the future, creating a disincentive to declare bankruptcy. If policymakers force credit reporting agencies to remove the bankruptcy flag from the credit report, then this would also affect the bankruptcy incentives facing borrowers, which in turn increases bankruptcy rates. If bankruptcies decrease social welfare on the margin, the incentive effect is important to account for in a full welfare accounting of the effects of flag removals. While the main focus of the paper is on the allocative effects of flag removal, in this appendix, we construct a simple model to illustrate how to evaluate the effects of consumer bankruptcy on social welfare, when flag removal affects the bankruptcy filing decision. Using the model, we then calculate the incentive effects for borrowers of the flag removal, and show that these incentive effects can be much larger than the allocative effects.

B.1.1 Model of Bankruptcy Decisions and Welfare

We consider a two-stage game. The second stage, which we call the "downstream" market, is identical to the model in the main text. Previously bankrupt consumers face some cost r of getting credit, which may be affected by policies such as bankruptcy flag removal. We add a first stage, in which borrowers have some heterogeneous cost $c \sim F(\cdot)$ of declaring bankruptcy. Costs may differ because consumers have different subjective valuations of bankruptcy. In the first stage, consumers can choose whether to declare bankruptcy, or not declare bankruptcy and receive value V_{NB}. There are three kinds of agents: borrowers, "downstream" lenders who lend to consumers in the second stage, and "upstream" lenders who have outstanding loans to consumers at the point where they can declare bankruptcy. We will separately characterize the surplus of each kind of agent, then add these terms to analyze social welfare.

First, we analyze the downstream market. For simplicity, suppose all previously bankrupt consumers have the same default rate, and thus the cost of serving previously bankrupt consumers is some constant $\phi\left(r_{fair}\right)$, where r_{fair} is the break-even interest rate for these borrowers. We focus on high-cost borrowers; by arguments analogous to the main text, the welfare effects for never-bankrupt borrowers will be small, since the never-bankrupt group is much

[36] See Table 6 of Gross et al. (2020)

larger than the previously bankrupt group. As in (7) in the main text, let $L^*(r;c)$ denote the loan demand of type c of the prospective borrower, at interest rate r. We maintain assumption 1, normalizing the marginal value of a dollar in each period to 1. Now, the total amount of loans made at price r is:

$$\int_{c\leq\bar{c}(r)} L^*(r;c)\,dF(c)$$

where $\bar{c}(r)$ reflects the fact that the opt-out condition depends on r.

Let V_{NB} represent the value from not declaring bankruptcy; for simplicity, assume this is a constant for all consumers, though this can be relaxed without affecting the results. In the second stage, years after a consumer has declared bankruptcy, she faces some price r for loans. To calculate surplus in lending markets, note that the surplus of a consumer with type c is:

$$V_B(c) = \underbrace{\int_{\hat{r}=r}^{\bar{r}(c)} L^*(\hat{r};c)\,d\phi(\hat{r})}_{Downstream} - c \tag{82}$$

where $\bar{r}(c)$ is the maximum rate at which a consumer of type c borrows positive amounts. Expression (82) is just (8) from Claim 1 in the main text, normalizing the utility weight equal to 1 as in Assumption 1. Taking into account the fixed cost c of bankruptcy, a consumer with cost c has a value of declaring bankruptcy:

$$\int_{\hat{r}=r}^{\bar{r}(c)} L^*(\hat{r};c)\,d\phi(\hat{r}) - c \tag{83}$$

A consumer with bankruptcy cost c optimally declares bankruptcy if (83) is greater than the value of not declaring bankruptcy, V_{NB}. We can thus define a function $\bar{c}(r)$ as the marginal bankrupt consumer, given the downstream rate r:

$$\bar{c}(r) = \left\{ c:\ V_{NB} = \int_{\hat{r}=r}^{\bar{r}(c)} L^*(\hat{r};c)\,d\phi(\hat{r}) - c \right\} \tag{84}$$

All consumers with $c \leq \bar{c}(r)$ declare bankruptcy, and all consumers with $c > \bar{c}(r)$ do not. From (84), $\bar{c}(r)$ is always decreasing in r: the higher the rate post-bankruptcy, the lower consumer surplus in the post-bankruptcy market, and thus the less types c will declare bankruptcy. The function $\bar{c}(r)$ thus captures the elasticity of the bankruptcy decision to the post-bankruptcy interest rate r.

Consumer surplus. Integrating over all consumers with different bankruptcy costs c, con-

sumers' surplus is thus:

$$CS = \underbrace{\int_{c>\bar{c}(r)} V_{NB} dF(c)}_{No\ bankruptcy} + \underbrace{\int_{c\leq\bar{c}(r)} \left[\int_{\hat{r}=r}^{\bar{r}(c)} L^*(\hat{r};c)\, d\phi(\hat{r}) - c \right] dF(c)}_{Bankruptcy} \tag{85}$$

We wish to characterize how consumer surplus changes as we shift r, the interest rate facing previously bankrupt consumers in lending markets. Differentiating (85) with respect to r, we have:

$$\frac{\partial CS}{\partial r} = \underbrace{-\int_{c\leq\bar{c}(r)} L^*(r;c)\, \phi'(r)\, dF(c)}_{Term\ 1} + \underbrace{\bar{c}'(r) f(\bar{c}) \left(\int_{\hat{r}=r}^{\bar{r}(c)} L^*(\hat{r};\bar{c}(r))\, d\phi(\hat{r}) - \bar{c}(r) - V_{NB} \right)}_{Term\ 2} \tag{86}$$

Now, from the definition of $\bar{c}(r)$ in (84), the rightmost piece of term 2 is 0; thus, we have:

$$\frac{\partial CS}{\partial r} = -\int_{c\leq\bar{c}(r)} L^*(r;c)\, \phi'(r)\, dF(c) \tag{87}$$

In words, (87) states that the derivative of total consumer surplus with respect to r is the standard envelope formula: it is the change in payments, $\phi'(r)$, multiplied by loan size $L^*(r;c)$, integrated over all consumers. Changing r also changes the set of consumers that declare bankruptcy. However, the marginal consumers are indifferent between declaring bankruptcy and not doing so, hence there is no first-order welfare effect of moving these consumers into or out of bankruptcy.

Downstream lender profits. As in (13), profits of downstream lending firms, who lend to previously bankrupt consumers, are simply demand minus costs:

$$\Pi_D = \int_{c\leq\bar{c}(r)} L^*(r;c) \left(\phi(r) - \phi\left(r_{fair}\right) \right) dF(c) \tag{88}$$

Differentiating (88) with respect to r, we have:

$$\frac{\partial \Pi_D}{\partial r} = \underbrace{\int_{c \leq \bar{c}(r)} L^*(r;c)\,\phi'(r)\,dF(c)}_{Term\ 1} + \underbrace{\int_{c \leq \bar{c}(r)} \frac{\partial L^*(r;c)}{\partial r}\left(\phi(r) - \phi\left(r_{fair}\right)\right) dF(c)}_{Term\ 2} + \underbrace{\bar{c}'(r) f\left(\bar{c}(r)\right) L^*(r;c)\left(\phi(r) - \phi\left(r_{fair}\right)\right)}_{Term\ 3} \quad (89)$$

In (89), term 1, which is exactly the negative of $\frac{\partial CS}{\partial r}$ in (87), reflects the fact that, when rates increase, welfare is transferred from borrowers to downstream lenders. Term 2 is the marginal change in the deadweight loss triangle, as r increases: it is the height of the deadweight loss triangle, $\left(\phi(r) - \phi\left(r_{fair}\right)\right)$, multiplied by the change in loan amount, $\frac{\partial L^*(r;c)}{\partial r}$. Both these terms are also present in the baseline model, where there is no bankruptcy margin. Term 3 is novel to the setting where bankruptcy decisions are elastic. When $r \neq r_{fair}$ in downstream markets, the marginal consumer's decision to declare bankruptcy imposes an externality on downstream lenders, of size:

$$L^*(r;c)\left(\phi(r) - \phi\left(r_{fair}\right)\right) \quad (90)$$

For example, if downstream lenders lose money on previously bankrupt consumers, so $\phi(r) < \phi\left(r_{fair}\right)$, then the marginal consumer who declares bankruptcy imposes a negative externality on lenders. The size of the effect depends on the size of the negative externality, (90), multiplied by the measure of marginal consumers, $\bar{c}'(r) f\left(\bar{c}(r)\right)$

Upstream lender profits. Suppose that type c consumers, at the time that they declare bankruptcy, have some outstanding debt $D(c)$ with upstream lenders. Suppose that their decision to declare bankruptcy causes lenders to lose a fraction ψ of the debt. Upstream firms' welfare, as a function of r, is thus:

$$\Pi_U = \int_{c \leq \bar{c}(r)} -\psi D(c)\,dF(c) \quad (91)$$

That is, upstream firms lose $\psi D(c)$ on all consumer types that default, $c \leq \bar{c}(r)$. Differentiating with respect to r, we have:

$$\frac{\partial \Pi_U}{\partial r} = -\bar{c}'(r) f\left(\bar{c}(r)\right) \psi D\left(\bar{c}(r)\right) \quad (92)$$

In words, decreasing r slightly causes a measure $-\bar{c}'(r) f\left(\bar{c}(r)\right)$ of marginal consumers with type $\bar{c}(r)$ to declare bankruptcy (note that $\bar{c}'(r)$ is negative). This decreases upstream firms'

profits by the losses on their loans for these consumers, which is $\psi D\left(\bar{c}\left(r\right)\right)$.

Now, total social welfare is just the sum of the welfare of consumers, upstream producer profits, and downstream producer profits. To find the effect of a small change in r on total welfare, we sum (87), (89), and (92). Consumer surplus (87) cancels with term 1 in downstream firms' profits (89), so we get:

$$\frac{\partial CS}{\partial r}+\frac{\partial \Pi_D}{\partial r}+\frac{\partial \Pi_U}{\partial r}=\underbrace{\int_{c\leq\bar{c}(r)}\frac{\partial L^*(r;c)}{\partial r}\left(\phi\left(r\right)-\phi\left(r_{fair}\right)\right)dF\left(c\right)}_{Deadweight\ Loss}$$
$$+\underbrace{\bar{c}'\left(r\right)f\left(\bar{c}\left(r\right)\right)L^*\left(r;c\right)\left(\phi\left(r\right)-\phi\left(r_{fair}\right)\right)}_{Downstream\ externality}-\underbrace{\bar{c}'\left(r\right)f\left(\bar{c}\left(r\right)\right)\psi D\left(\bar{c}\left(r\right)\right)}_{Upstream\ externality} \quad (93)$$

In (93), the first term is the marginal change in the size of the deadweight loss triangle, which is analogous to the depiction in Figure 1. When there is no incentive effect of bankruptcy, so $\bar{c}'\left(r\right)=0$ and bankruptcy decisions are perfectly inelastic, then the change in welfare is simply the change in the deadweight loss triangle. When bankruptcy is elastic, there are two additional terms: the externality on downstream firms, which is term 3 in (89), and the externality on upstream firms, which is (92). The upstream externality term will always be positive (that is, decreasing rates will tend to lower welfare), since bankruptcies create negative externalities on upstream lenders. When $\phi\left(r\right)<\phi\left(r_{fair}\right)$, so prices are lower than marginal costs for previously bankrupt consumers, and the downstream externality term is also positive, so decreasing rates will tend to lower social welfare. Thus, if bankruptcy decisions are sensitive to rates in downstream markets, there are two additional forces causing lower rates to tend to decrease social welfare.

B.1.2 Estimate of Welfare Costs with Elastic Bankruptcy

Next, we do a back-of-the-envelope calculation of how large the incentive effects of bankruptcy on welfare could be in the data. There are a variety of estimates in the literature on how strategic borrowers are in their decisions to default on loans and declare bankruptcy. Yannelis (2016) provides evidence for strategic default on student loans, showing that introducing bankruptcy protection for student loans would increase loan default by 18%, and increasing garnishable income by $10,000 would lead to a 15% decrease in defaults. Mayer et al. (2014) argue that a legal settlement offering modifications to delinquent borrowers increased delinquency rates by 10%. Argyle et al. (2021) find that borrowers with increased cash flows tend to delay filing for bankruptcy.

It is difficult to extrapolate the effect of a particular policy, i.e., bankruptcy flag removal, on

strategic bankruptcies. The lower bound of the incentive effect in the literature—that borrowers are completely non-strategic in their default and bankruptcy decisions—would imply that there is no incentive effect of flag removal on bankruptcy. To gauge the size of the incentive effect, we do a quick back-of-the-envelope calculation: suppose that flag removal, relative to keeping bankruptcy flags on borrowers' credit records indefinitely, would increase bankruptcy filing rates by 1%. With approximately 800,000 bankruptcy filings annually, this implies 8,000 additional bankruptcies.

According to our model, we must estimate two numbers. For upstream lenders, we must estimate the effect of bankruptcy filing on lenders' losses. Since our main analysis focuses on auto loans, we also consider losses to auto lenders. At the time of bankruptcy, the average consumer has $19,865 in auto loan debt.[37] In most states, borrowers lose their cars in bankruptcy; thus, we will assume a loss rate of 40%.[38]. Hence, the loss per loan is $7,946 (=0.4*19,865) or $3,575.7 per consumer.[39] With approximately 800,000 filers per year, the aggregate loss to lenders would be $28.6 million (=0.01*800,000*3,575).

For downstream lenders, we must estimate their losses on each customer, multiplied by the set of marginal consumers. From our baseline estimates in Table 2, the average loss per customer of lenders in downstream markets is $24.42 over a five year loan term or $19.5 million for 800,000 bankruptcy filers every year. Multiplying by the 1% increase in bankruptcy filings, this would create an additional loss of $0.195 million to social surplus.

Adding the effects on upstream and downstream lenders, we get a total effect of $28.8 million. This quantity is large relative to the allocative welfare quantities we calculated in the main text. Accounting for a 1% increase in bankruptcies due to incentive effects, flag removal transfers $19 million to previously bankrupt consumers, at the cost of $29.4 million (=$28.8 million+$0.6 million) in social welfare.

Why are the welfare effects through the incentive channel so large, relative to the allocative effects? Intuitively, we showed in the main text that, because competitive lending markets lead to efficient credit allocations, the removal of small amounts of data has only a second-order effect on the allocative efficiency of lending. This does not apply to the effects of data removal on bankruptcy incentives. Borrowers do not internalize the costs to lenders of their bankruptcy decisions, so they default more than the socially optimal level. Data removal can affect borrowers' incentives to declare bankruptcy, and this generally has a first-order effect on social welfare. From a policy perspective, however, note that increased incentives from flag removal could be offset by increasing the cost of bankruptcy, e.g., through more stringent

[37] See Experian's auto loan debt study.

[38] See American Banker.

[39] Table 1 of Dobbie et al. (2017) indicates that approximately 45% of consumers have an auto loan when filing for bankruptcy.

repayment plans or lower thresholds for asset protection in liquidation.

B.2 Varying Signal Informativeness

Different kinds of data may be differentially informative about customers' default rates, and thus the costs of lending to these customers. In this appendix, we show that, when data is more informative about default rates, the social welfare losses from data removal tend to be larger relative to the surplus transfers, so data removal is less efficient as a tool for transferring surplus. However, plugging in our demand elasticity estimate from the main text, we show that flag removal would remain a quantitatively efficient way to transfer surplus to previously bankrupt individuals, even if bankruptcy flags were much more informative about default rates than we find in our analysis. As in the main text, we assume demand to be linear in the payment $\phi(r)$:

$$\Lambda_H(\phi(r_H)) = a_H - b_H\phi(r_H) \tag{94}$$

$$\Lambda_L(\phi(r_L)) = a_L - b_L\phi(r_L) \tag{95}$$

Let the data removal-induced change in interest rates be:

$$t = \phi(r) - \phi\left(r_{H,fair}\right) \tag{96}$$

Claim 3 shows that consumer surplus changes by

$$\left(\phi\left(r_{H,fair}\right) - \phi(r)\right)\left(\frac{\Lambda_H(\phi(r)) + \Lambda_H\left(\phi\left(r_{H,fair}\right)\right)}{2}\right) \tag{97}$$

as data is removed. High cost types gain as they are charged the lower pooling price. However, data removal also induces an inefficiency due to over credit provision to high cost types. Claim 3 also shows that this efficiency consequence of the data removal can be expressed as:

$$\frac{1}{2}\left(\phi(r) - \phi\left(r_{H,fair}\right)\right)\left(\Lambda_H(\phi(r)) - \Lambda_H\left(\phi\left(r_{H,fair}\right)\right)\right) \tag{98}$$

In addition, data removal increases prices for low-cost consumers leading to under credit provision for low-cost consumers relative to the efficient benchmark. The inefficiency consequence of the data removal due to under credit provision for low-cost consumers can be expressed as:

$$\frac{1}{2}(\phi(r_{L,fair}) - \phi(r))(\Lambda_L\left(\phi\left(r_{L,fair}\right)\right) - \Lambda_L(\phi(r))) \tag{99}$$

To account for differential default rates across high and low cost groups, we multiply each of those quantities, by the expected number of non-default periods in each group. Dividing efficiency changes by redistributive consequences, we obtain the efficiency ratio, that is, the welfare cost per dollar redistributed to high-cost individuals:

$$\text{Efficiency Ratio} = \frac{\psi_H \frac{1}{2}\left(\phi(r)-\phi\left(r_{H,fair}\right)\right)\left(\Lambda(\phi(r))-\Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)+\psi_L \frac{1}{2}\left(\phi\left(r_{L,fair}\right)-\phi(r)\right)\left(\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda(\phi(r))\right)}{\psi_H\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\left(\frac{\Lambda(\phi(r))+\Lambda\left(\phi\left(r_{H,fair}\right)\right)}{2}\right)} \tag{100}$$

Plugging linear demand (94) and (95) into the efficiency ratio (100), exploiting the zero-profit condition of the competitive equilibrium, and writing the pooling price as sum of the fair price and price distortion, we obtain:

$$\text{Efficiency Ratio} = \frac{1}{2}\left(\epsilon_H + \frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\frac{\psi_H}{\psi_L}\epsilon_L\right)\frac{t}{\frac{\Lambda_H\left(\phi\left(r_{H,fair}\right)\right)}{\Lambda_H(\phi(r))} - \frac{1}{2}\epsilon_H t} \tag{101}$$

where, ϵ_H and ϵ_L are respectively the demand elasticities in the H and L groups at r, that is:

$$\epsilon_H \equiv \frac{b_H}{\Lambda_H(\phi(r))}, \epsilon_L \equiv \frac{b_L}{\Lambda_L(\phi(r))}$$

Taking the derivative with respect to the data induced price distortion:

$$\frac{\partial \text{Efficiency Ratio}}{\partial t} = \frac{1}{2}\left(\epsilon_H + \frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\frac{\psi_H}{\psi_L}\epsilon_L\right)\frac{\frac{\Lambda_H\left(\phi\left(r_{H,fair}\right)\right)}{\Lambda_H(\phi(r))}}{\left(\frac{\Lambda_H\left(\phi\left(r_{H,fair}\right)\right)}{\Lambda_H(\phi(r))} - \frac{1}{2}\epsilon_H t\right)^2} \tag{102}$$

When t is small, expression (102) is approximately constant. Thus, the efficiency ratio will increase roughly linearly in the price change t induced by flag removal, for relatively small values of t.[40] Intuitively, this is because the social welfare loss is a triangle, which is quadratic in the size of the deviation of prices from their efficient level, whereas the transfer is a trapezoid. The ratio of the two is therefore larger when data is more informative, leading to larger price changes upon its removal.

We can bring this theory to our data, to evaluate how efficient flag removals would be as a tool for transferring surplus, in a counterfactual scenario where bankruptcy flags were more informative about default rates, so their removal decreased interest rates more. Under our baseline estimates, flag removal decreases interest rates by 0.226%. We consider counterfactual scenarios in which flag removal induces a rate change two, four, eight, and sixteen

[40] Technically, the Efficiency Ratio exhibits a convexity in the price distortion t as can also be numerically seen for larger price changes in our numerical example (see the fifth row of Table B.1). However, this convexity only attenuates the linear decline in the Efficiency Ratio for larger price decreases, further enhancing our result that allocative efficiency costs are small relative to distributional consequences of data removal.

Table B.1: Welfare cost by price effect size

This table summarizes consumer surplus changes, welfare consequences, and efficiency ratios for the average five-year loan, under counterfactual scenarios in which bankruptcy flag removal is increasingly informative about costs. We consider scenarios in which flag removal induces a rate change equal to our baseline estimate of 0.226%, and then two, four, eight, and sixteen times higher. "Multiple of Effect" shows the x fold of the true price change. "Induced Rate Change" is the counterfactual interest rate variation induced by the flag removal (in percentage points). The true effect size is shown in the first row. Counterfactual price changes are depicted in rows two to five. "Consumer Surplus Redistribution" is the $ change in consumer surplus per individual, and "Welfare Change" depicts the change in social welfare per individual. Both variables are in units of expected dollars per individual, over the course of a five-year loan. "Efficiency Ratio" is the ratio of the welfare change to consumer surplus redistributed; that is, the dollars of social surplus lost, per dollar redistributed to bankrupt individuals.

Multiple of Effect	Induced Rate Change	Consumer Surplus Redistribution	Welfare Change	Efficiency Ratio
1x	-0.226%	23.75	-0.75	-0.0315
2x	-0.452%	48.76	-2.98	-0.0612
4x	-0.904%	102.51	-11.86	-0.1156
8x	-1.808%	224.56	-46.82	-0.2085
16x	-3.616%	524.04	-182.52	-0.3483

times larger. We can then plug these changes into our expressions for welfare changes, surplus transferred, and the efficiency ratio, holding fixed the demand elasticity at our estimate in the main text, and evaluate how efficiency ratios would change if bankruptcy flags were more informative about default rates.

The results of this exercise are shown in Table B.1. As expected, consumer surplus changes increase approximately linearly with the induced change in interest rates. Welfare changes vary approximately quadratically with price distortions. Thus, the efficiency ratio changes approximately linearly with the price distortion. As we double and quadruple the price effect of bankruptcy flag removals from 22.6bps reductions to 45.2bps and 90.4bps reductions, we can see that the efficiency cost per dollar redistributed to previously bankrupt individuals doubles and approximately quadruples from 3 cents to 6 cents and 12 cents, respectively. Thus, even when flag removals are fairly strong signals of default rates – 4 or 8 times more informative than we find in the main text – flag removal remains a relatively low-cost way to redistribute surplus, costing less than $0.21 in social surplus per dollar transferred between groups. However, the efficiency ratio begins to decline for larger rate changes.

B.3 Efficiency Ratios with Quantity Data Only

B.3.1 If the high cost group is small

The baseline model only requires price variation and quantity data to estimate the efficiency and redistributional consequences of data availability. However, researchers cannot always observe price data. In this section, we show that under our baseline assumptions a key quantity of our model - the efficiency loss per USD redistributed - can be computed with quantity data

and variation in data availability only.

Similar to our baseline case, we assume competitive credit provision, two cost types (high and low), and demand linear in the repayment:

$$\Lambda_L(\phi(r_L)) = a_L - b_L\phi(r_L)$$

$$\Lambda_H(\phi(r_H)) = a_H - b_H\phi(r_H)$$

For simplicity, let's assume that the high cost group (H) is very large in comparison to the low cost group (L). We will later relax this assumption.

Under those conditions, consumer surplus redistributed to the high cost group is equivalent to expression 16 in the main text

$$(\phi(r_{H,fair}) - \phi(r))(\frac{\Lambda(\phi(r) + \Lambda(\phi(r_{H,fair})}{2})$$

and the efficiency loss due to over creditprovision to the high cost type is:

$$\frac{1}{2}\left(\phi\left(r_{H,fair}\right) - \phi(r)\right)\left(\Lambda(\phi(r)) - \Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)$$

Hence, the efficiency loss per USD redistributed is:

$$\frac{\frac{1}{2}\left(\phi\left(r_{H,fair}\right) - \phi(r)\right)\left(\Lambda(\phi(r)) - \Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)}{(\phi(r_{H,fair}) - \phi(r))(\frac{\Lambda(\phi(r)) + \Lambda(\phi(r_{H,fair}))}{2})}$$

The price change cancels and the efficiency ratio can be expressed as:

$$\text{Efficiency ratio}_{\text{High Cost}} = \frac{\Lambda(\phi(r)) - \Lambda\left(\phi\left(r_{H,fair}\right)\right)}{\Lambda(\phi(r)) + \Lambda(\phi(r_{H,fair}))} \tag{103}$$

Hence, the efficiency ratio is simply the change in loan quantity due to data removal relative to the sum of pre and post data removal loan quantities.

B.3.2 For any group size

In the previous section, we have neglected the inefficiency arising from increased prices to the low cost group. Because we assumed the low cost group to be very large, the price increase is very small, and the resulting inefficiency is negligible. However, relative group sizes may not always be disproportionate. This subsection allows for flexible high and low cost type group sizes and derives the efficiency ratio as function of quantities only. To compute the efficiency

ratio under those conditions three observations are useful.

First, in competitive markets, the gain from lending to low cost types must equal the loss from lending to high cost types. We, therefore, have that

$$\psi_L \Lambda_L(\phi(r))\left(\phi(r)-\phi\left(r_{L,fair}\right)\right)+\psi_H \Lambda_H(\phi(r))\left(\phi(r)-\phi\left(r_{H,fair}\right)\right)=0$$

$$\psi_L \Lambda_L(\phi(r))\left(\phi(r)-\phi\left(r_{L,fair}\right)\right)=-\psi_H \Lambda_H(\phi(r))\left(\phi(r)-\phi\left(r_{H,fair}\right)\right)$$

Hence, we can express the price increase for the low cost types as:

$$\phi(r)-\phi\left(r_{L,fair}\right)=\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\left(\phi\left(r_{H,fair}\right)-\phi(r)\right) \tag{104}$$

Second, note that we can write the price change underlying the quantity changes for the high cost types as a function of those quantity changes and the demand elasticity:

$$\Lambda_H(\phi(r_H))-\Lambda_H(\phi(r))=a_H-b_H\phi(r_H)-a_H+b_H\phi(r)=b_H(\phi(r)-\phi(r_H))$$

Divide by the pooling quantity for high cost types:

$$\frac{\Lambda_H(\phi(r_H))-\Lambda_H(\phi(r))}{\Lambda_H(\phi(r))}=\frac{b_H}{\Lambda_H(\phi(r))}(\phi(r)-\phi(r_H))=\epsilon_H*(\phi(r)-\phi(r_H))$$

Rearrange for the quantity implied price change:

$$\frac{\Lambda_H(\phi(r_H))-\Lambda_H(\phi(r))}{\Lambda_H(\phi(r))}*\frac{1}{\epsilon_H}=\phi(r)-\phi(r_H)$$

Third, we can write the quantity change for the low cost types as a function of high cost type quantities and price elasticities:

$$\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda(\phi(r))=a_L-b_L\phi(r_L)-a_L+b_L\phi(r)=b_L(\phi(r)-\phi(r_L))$$

Plug in for the loss implied price change

$$\begin{aligned}\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda(\phi(r)) &= b_L\left(\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\right)\\ &= \frac{b_L}{\Lambda_L(\phi(r))}\frac{\psi_H}{\psi_L}\Lambda_H(\phi(r))\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\\ &= \epsilon_L\frac{\psi_H}{\psi_L}\Lambda_H(\phi(r))\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\end{aligned}$$

Use the price change calculated for the high type as a function of elasticity:

$$\begin{aligned}\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda\left(\phi\left(r\right)\right) &= \epsilon_L\frac{\psi_H}{\psi_L}\Lambda_H\left(\phi\left(r\right)\right)\left(\phi\left(r_{H,fair}\right)-\phi\left(r\right)\right)\\ &= -\epsilon_L\frac{\psi_H}{\psi_L}\Lambda_H\left(\phi\left(r\right)\right)\left(\phi\left(r\right)-\phi\left(r_{H,fair}\right)\right)\\ &= -\epsilon_L\frac{\psi_H}{\psi_L}\Lambda_H\left(\phi\left(r\right)\right)\frac{\Lambda_H\left(\phi\left(r_H\right)\right)-\Lambda_H\left(\phi\left(r\right)\right)}{\Lambda_H\left(\phi\left(r\right)\right)}*\frac{1}{\epsilon_H}\\ &= \epsilon_L\frac{\psi_H}{\psi_L}\Lambda_H\left(\phi\left(r\right)\right)\frac{\Lambda_H\left(\phi\left(r\right)\right)-\Lambda_H\left(\phi\left(r_H\right)\right)}{\Lambda_H\left(\phi\left(r\right)\right)}*\frac{1}{\epsilon_H}\end{aligned}$$

With those three expressions:

$$\phi\left(r\right)-\phi\left(r_{L,fair}\right)=\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r\right)\right)}{\Lambda_L\left(\phi\left(r\right)\right)}\left(\phi\left(r_{H,fair}\right)-\phi\left(r\right)\right)$$

$$\frac{\Lambda_H\left(\phi\left(r_H\right)\right)-\Lambda_H\left(\phi\left(r\right)\right)}{\Lambda_H\left(\phi\left(r_H\right)\right)}*\frac{1}{\epsilon_H}=\phi\left(r\right)-\phi\left(r_H\right)$$

$$\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda\left(\phi\left(r\right)\right)=\frac{\epsilon_L}{\epsilon_H}\frac{\psi_H}{\psi_L}\frac{\Lambda_H\left(\phi\left(r\right)\right)}{\Lambda_H\left(\phi\left(r\right)\right)}(\Lambda_H\left(\phi\left(r\right)\right)-\Lambda_H\left(\phi\left(r_H\right)\right))$$

We can rewrite the efficiency ratio as a function of quantities only.

$$\text{Efficiency ratio}=\frac{\psi_H\frac{1}{2}\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\left(\Lambda(\phi(r))-\Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)+\psi_L\frac{1}{2}\left(\phi(r)-\phi\left(r_{L,fair}\right)\right)\left(\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda(\phi(r))\right)}{\psi_H(\phi(r_{H,fair})-\phi(r))(\frac{\Lambda(\phi(r))+\Lambda(\phi(r_{H,fair}))}{2})}$$

Replace the low cost price change:

$$=\frac{\psi_H\frac{1}{2}\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\left(\Lambda(\phi(r))-\Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)+\psi_L\frac{1}{2}\left(\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\left(\phi\left(r_{H,fair}\right)-\phi(r)\right)\right)\left(\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda(\phi(r))\right)}{\psi_H(\phi(r_{H,fair})-\phi(r))(\frac{\Lambda(\phi(r))+\Lambda(\phi(r_{H,fair}))}{2})}$$

Cancel the price change:

$$=\frac{\psi_H\left(\Lambda(\phi\left(r\right))-\Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)+\psi_L\left(\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\right)\left(\Lambda\left(\phi\left(r_{L,fair}\right)\right)-\Lambda\left(\phi\left(r\right)\right)\right)}{\psi_H(\Lambda(\phi(r))+\Lambda(\phi(r_{H,fair})))}$$

Plugging in the low cost type quantity change:

$$=\frac{\psi_H\left(\Lambda(\phi\left(r\right))-\Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)+\psi_L\left(\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\right)\left(\frac{\epsilon_L}{\epsilon_H}\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_H(\phi(r))}(\Lambda_H\left(\phi\left(r\right)\right)-\Lambda_H\left(\phi\left(r_H\right)\right))\right)}{\psi_H(\Lambda(\phi(r))+\Lambda(\phi(r_{H,fair})))}$$

$$=\frac{\left(\Lambda(\phi\left(r\right))-\Lambda\left(\phi\left(r_{H,fair}\right)\right)\right)+\frac{\psi_L}{\psi_H}\left(\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}\right)\left(\frac{\epsilon_L}{\epsilon_H}\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_H(\phi(r))}(\Lambda_H\left(\phi\left(r\right)\right)-\Lambda_H\left(\phi\left(r_H\right)\right))\right)}{(\Lambda(\phi(r))+\Lambda(\phi(r_{H,fair})))}$$

and rearranging:

$$\text{Efficiency ratio}_{\text{High and Low Cost}} = (1 + \frac{\epsilon_L}{\epsilon_H}\frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))})\underbrace{\frac{\Lambda(\phi(r)) - \Lambda\left(\phi\left(r_{H,fair}\right)\right)}{\Lambda(\phi(r)) + \Lambda(\phi(r_{H,fair}))}}_{\text{Efficiency ratio}_{\text{High Cost}}}$$

The last expression shows that if high and low cost demand elasticities are the same, we can write the efficiency ratio as a function of quantities only (assuming $\epsilon_L = \epsilon_H$).

$$\text{Efficiency ratio}_{\text{High and Low Cost}} = (1 + \frac{\psi_H}{\psi_L}\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))})\underbrace{\frac{\Lambda(\phi(r)) - \Lambda\left(\phi\left(r_{H,fair}\right)\right)}{\Lambda(\phi(r)) + \Lambda(\phi(r_{H,fair}))}}_{\text{Efficiency ratio}_{\text{High Cost}}} \tag{105}$$

Furthermore, we can bound the efficiency ratio with quantity data only if we have some theory-guided intuition for the maximum low type demand elasticity and the minimum demand elasticity for the high type. Lastly, note that as the ratio $\frac{\Lambda_H(\phi(r))}{\Lambda_L(\phi(r))}$ is getting small, we arrive at the expression of the previous section.

B.3.3 Detailed computations for GNW 2020

In this section, we outline the computation of the efficiency ratios displayed in Table 5 of the main text using the expressions above. The computations are based on the effects of flag bankruptcy flag removals on loan quantities displayed in Appendix Table A3 of GNW 2020. As equation 103 shows, the efficiency ratio (welfare cost per USD redistributed) is a function of loan quantities with and without data. For Panel A of Table 5, we compute the pre-removal loan quantities as the product of the pre-removal average number of new accounts and the pre-removal average balance on new accounts. For all credit markets combined, this means the average new balance was \$341,920 (=4.789*71,397). The new quantities without bankruptcy flag are computed using the pre-means and the effects of bankruptcy flag removal on new account openings and balances on new accounts. The effects are scaled to reflect the 15 credit score point increase.[41] Hence, loan quantities without bankruptcy flag are \$360,259 (=(4.789+0.204)*(71,397+755.75)). Plugging those numbers into equation 103 yields our estimate for the efficiency ratio in all credit markets for the high-cost group as shown in Table 5 . Equation 105 extends this framework to account for the inefficiency arising from under-credit provision to low-cost consumers in the pooling equilibrium. Due to the large difference in group sizes - the number of never-bankrupt individuals far exceeds the number of bankrupt

[41] Table A3 of GNW 2020 shows effects per 10-point increase in credit scores while the average effect of bankruptcy flag removal on credit scores is 15.455 according to the GNW 2020 Table 1 estimates.

individuals - the additional inefficiency is small. The additional parameters required are the ratio of demand elasticity, the ratio of repayment periods for high and low-cost individuals and the ratio of loan quantities across the two groups. We use the share of bankrupt to never bankrupt individuals as estimate for the quantity fraction ($\frac{0.106}{1-0.106}$). We further assume similar demand elasticities and the same ratio of repayment periods as computed in the main text. Due to the differential group sizes, the efficiency ratio is not sensitive to these assumptions. Plugging everything into equation 105, we obtain the efficiency ratio estimate of 2.92 cents shown in Table 5. We then repeat the same analysis separately for mortgage, auto, and other credit markets. For computations in Panel B, we use estimates based on new principal amounts instead of new balances. However, this has little effect on the estimates of the efficiency ratio.

C Flag Removals and Charge-Offs

In this section, we examine the effect of flag removals on charge-offs. Figure C.1 is comparable to Figure 3 in the main text. However, it replaces the outcome variable with a dummy variable equal to one if a loan gets charged-off within two years of loan opening and zero if the opened loan does not get charged-off within two years of opening. In the graphical evidence, we cannot reject the null of no effect of flag removal on charge-offs. As charge-offs do not appear to decrease when prices decrease, we do not find evidence of adverse selection. We also quantify the visual result in a regression framework and present the results in Table C.1. The regression evidence is in line with the graphical evidence.

Figure C.1: Charge-offs

This figure shows estimates of the coefficients δ_t from the following specification $y_{it} = \gamma_c + \gamma_{ts} + \sum_{t=-6}^{6} \delta_t \{e_{it} = t\} + \beta X_{it} + \varepsilon_{it}$, along with a 95% confidence interval. The outcome y_{it} is charge-offs. γ_c are cohort fixed effects, and γ_{ts} are time period by score bucket fixed effects. Standard errors are clustered at the cohort level. Source: TransUnion.

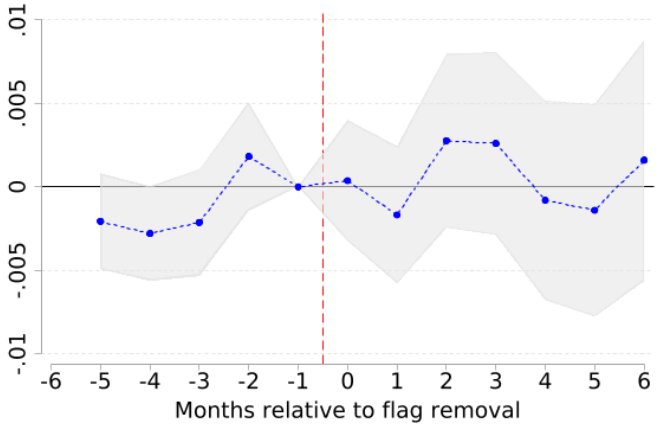


Table C.1: Charge-offs Around Flag Removal

This table shows estimates of the coefficients δ_y from the following specification $y_{it} = \gamma_c + \gamma_t + \delta^y \mathbb{1}[FlagRemoved] + \beta X_{it} + \varepsilon_{it}$. The outcome y_{it} is charge-offs. γ_c are cohort fixed effects, and γ_t are time period fixed effects. Standard errors are clustered at the cohort level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Source: TransUnion.

	(1)	(2)	(3)	(4)	(5)
Post	0.001	0.001	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.024***	0.024***	0.024***	0.023***	0.023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	176690	176690	176690	176690	176686
Adjusted R^2	0.000	0.001	0.001	0.001	0.030
Linear Time Trend	Yes	Yes	Yes	No	No
Year-month FE	No	Yes	No	Yes	No
Cohort FE	No	No	Yes	Yes	Yes
Year-month by Score Bucket FE	No	No	No	No	Yes
Clustered SE	Cohort	Cohort	Cohort	Cohort	Cohort

D Flag Removals and Observed Interest Rates

In the main text, we show the effect of flag removals on interest rates. This section shows that the effect of bankruptcy flag removals is qualitatively and quantitatively similar using observed interest rates only. Table D.1 shows variations of the main specification with observed interest rates as the outcome variable. Note that the effect size is of comparable magnitude to the effect on predicted interest rates in Table 2 in the main text. One potential concern with our estimates of the effect of flag removal on interest rates is that we only observe interest rates for loan offers that were actually taken up by customers, whereas our model is about offered rates. This could introduce downwards bias in our estimates of effects on interest rates, if customers with different characteristics are offered different rates, but customers who receive higher rate offers are less likely to accept. However, if unobserved heterogeneity between customers affects loan rates in this manner, then controlling for *observable* heterogeneity between customers should also affect our coefficient estimates (Oster, 2019). Our estimates of effects on interest rates in columns 1 to 5 of Table D.1 are very stable, suggesting that this bias is not likely to be quantitatively important.

The graphical evidence in Figure D.1 confirms the findings of Table D.1. The removal of bankruptcy flags reduces interest rates by approximately 20bps.

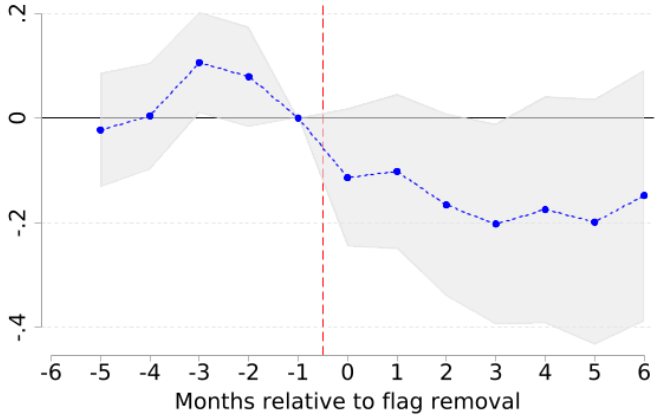
Table D.1: Interest Rates Around Flag Removal

This table shows estimates of the coefficients δ_y from the following specification $y_{it} = \gamma_c + \gamma_t + \delta^y \mathbb{1}[FlagRemoved] + \beta X_{it} + \varepsilon_{it}$. The outcome y_{it} is observed interest rates. γ_c are cohort fixed effects, and γ_t are time period fixed effects. Standard errors are clustered at the cohort level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Source: TransUnion.

	(1)	(2)	(3)	(4)	(5)
Post	-0.192***	-0.202***	-0.183***	-0.198***	-0.169***
	(0.066)	(0.060)	(0.064)	(0.059)	(0.050)
Constant	8.160***	8.131***	8.176***	7.799***	7.782***
	(0.053)	(0.033)	(0.029)	(0.033)	(0.028)
Observations	176690	176690	176690	176690	176686
Adjusted R^2	0.003	0.009	0.009	0.010	0.328
Linear Time Trend	Yes	Yes	Yes	No	No
Year-month FE	No	Yes	No	Yes	No
Cohort FE	No	No	Yes	Yes	Yes
Year-month by Score Bucket FE	No	No	No	No	Yes
Clustered SE	Cohort	Cohort	Cohort	Cohort	Cohort

Figure D.1: Interest rates

This figure shows estimates of the coefficients δ_t from the following specification $y_{it} = \gamma_c + \gamma_{ts} + \sum_{t=-6}^{6} \delta_t\{e_{it} = t\} + \beta X_{it} + \varepsilon_{it}$, along with a 95% confidence interval. The outcome y_{it} is observed interest rates. γ_c are cohort fixed effects, and γ_{ts} are time period by score bucket fixed effects. Standard errors are clustered at the cohort level. Source: TransUnion.



E Stacked Dynamic Difference in Difference Estimation

Our main estimator is a two-way fixed effect estimator with heterogeneous treatment timing. If fully saturated and under homogeneous treatment effects, our estimator provides an unbiased estimate of a treatment effect. In the presence of heterogeneous treatment effects, the estimator may suffer from negatively weighting contrasts, and leads may reflect lags (Sun and Abraham (2021)). To ensure that the choice of the estimator does not drive our results, we follow best practices in Barrios (2021) and Cengiz et al. (2019) (Appendix D) in implementing a Stacked Difference in Differences Estimation. The results remain qualitatively unchanged.

In particular, we implement the stacked difference in differences as follows: For each treated cohort, which is defined by the month in which the individuals have their bankruptcy flags removed, we construct a separate control group. The control group consists of individuals with their bankruptcy flags removed 12 to 17 months after the treated cohort. We then restrict the dataset of treated and control individuals to the six months surrounding the flag removal of the treated cohort. That is, for individuals who have their bankruptcy flag removed in July of 2009, we compare credit scores, interest rates, and auto loan quantities during 2009 to outcomes for individuals who will have their bankruptcy flags removed from July to December 2010. The identifying assumption is that the change in outcomes for individuals with flag removal in July 2009 would have been the same as the change in outcomes for the control cohorts, in the absence of the flag removal for the treated cohort (parallel trends assumption). We repeat this dataset construction for all treated cohorts from July 2009 to June 2017 and stack the separate datasets together. We call each dataset a group and run variants of the following regression:

$$y_{itg} = \gamma_{cg} + \gamma_{tsg} + \sum_{t=-6}^{6} \delta_t \{e_{itg} = t\} + \varepsilon_{itg} \tag{106}$$

e_{itg} indicates time relative to the treatment of the treated cohort. We plot the coefficients δ_t, along with a 95% confidence interval. The coefficients capture the difference in an outcome in each month before and after flag removal relative to the months prior to flag removal.[42] We include cohort-month fixed effects, as well as year-month by score bucket fixed effects. We allow those to differ by the respective dataset. Standards errors are clustered at the cohort-month by group level. Figure E.1 plots estimation results and validates our findings from the main specifications.

We further validate the graphical evidence by implementing a regression framework and showing the results in Table E.1. To be computationally able to account for group-specific linear

[42] We exclude the relative time dummy for period -1.

time trends, we run the stacked regression at the month-cohort month-score bucket-group level and weight by the number of observations. Following standard practice, standard errors are clustered at the cohort-month by group level. When not prohibited by multi-collinearity, the specifications are chosen to match the main specifications in Table 2. Overall, the results in Table E.1 confirm the visual results in Figure E.1. To ensure that the estimates from the stacked specification do not substantially change our welfare computations, we repeat the exercise illustrated in Table 3 and replace the regression estimates with the estimates from the stacked specifications. The resulting welfare estimates are shown in Table E.2. We find an efficiency ratio of approximately 0.059 (=1.83/31.04): for each dollar transferred, 5.9 cents of social welfare is lost, not changing our conclusions from the main text.

To address concerns that one particularly influential group drives the results, we also aggregate treatment and control outcomes for each generated dataset and, subsequently, average relative time means across datasets. Hence, each generated dataset has the same weight in the plotted mean scores, quantities, and interest rates. Figure E.2 illustrates that mean score, interest, and quantity outcomes move in parallel in the pre-period. Besides, the estimated treatment effects appear to be driven by trend breaks in the treated group at the time of flag removal. While it is strictly speaking not a necessary condition for identification, we find this observation comforting. To further address treatment effect heterogeneity, we sort and plot the treat $\times$ post coefficients for each of the stacked datasets in Figure E.3. The majority of point estimates are in line with our overall conclusions.

Table E.1: Credit Scores, Interest Rates, and Loan Volumes

This table shows estimates of the coefficients δ_y from the following specification $y_{itg} = \gamma_{cg} + \gamma_{tsg} + \delta^y \mathbb{1}[FlagRemoved] + \beta X_{itg} + \varepsilon_{itg}$. In the top panel, the outcome y_{itg} is the Vantage Score, in the middle panel the outcome is observed interest rates, while in the bottom panel it is loan volumes. γ_{cg} are cohort fixed effects, and γ_{tg} are time period fixed effects that can vary by group. Standard errors are clustered at the cohort by group level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Source: TransUnion.

	(1)	(2)	(3)	(4)	(5)
Panel A: Credit Scores					
Post=1 × Treat=1	16.916***	17.186***	17.198***	17.186***	17.407***
	(0.381)	(0.208)	(0.219)	(0.208)	(0.220)
Observations	174,720	174,720	174,720	174,720	174,720
Panel B: Interest Rates					
Post=1 × Treat=1	-0.269**	-0.303***	-0.302***	-0.303***	-0.288***
	(0.085)	(0.006)	(0.009)	(0.006)	(0.007)
Observations	174,720	174,720	174,720	174,720	174,720
Panel C: Loan Volumes					
Post=1 × Treat=1	28.547***	34.309***	33.715***	34.309***	34.193***
	(7.744)	(2.489)	(3.130)	(2.489)	(2.656)
Observations	174,720	174,720	174,720	174,720	174,720
Group Specific Linear Trend	Yes	No	Yes	No	No
Year-month by Group FE	No	Yes	No	Yes	No
Cohort by Group FE	No	No	Yes	Yes	Yes
Year-month by Score Bucket by Group FE	No	No	No	No	Yes
Clustered SE Cohort by Group	Yes	Yes	Yes	Yes	Yes

Table E.2: Summarizing Estimates Implied by Stacked Specifications

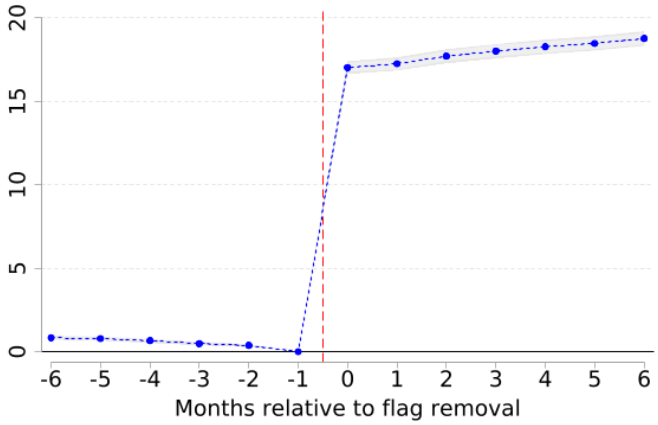
This table is comparable to Table 3 in the main text. We replace the regression coefficients obtained from the main specifications in Table 2 with coefficients of the stacked specifications in Table E.1. This table then summarizes our estimates implied by the stacked specifications of Table E.1. Panel A shows average interest rates in the six months before flag removal (r_{flag}), the interest rate effect of flag removal ($r_{pool} - r_{flag}$), and the effect of flag removal on the fraction of the principal repaid each month in a standardized five-year loan ($\phi(r_{pool}) - \phi(r_{flag})$). Panel B shows average loan quantities in the six months before flag removal and the quantity effect of flag removal. Panel C shows the market demand elasticity implied by our estimates, the inverse demand slope in terms of the interest rate ($\frac{r_{pool}-r_{flag}}{\Lambda_{pool}-\Lambda_{flag}}$), and the inverse demand slope in terms of the repayment fraction ($\frac{\phi(r_{pool})-\phi(r_{flag})}{\Lambda_{pool}-\Lambda_{flag}}$). Panel D summarizes surplus changes implied by the estimates in Table 2. The first row shows the average change in consumer surplus for individuals with flag removal for the average five-year loan. It is the sum of monthly non-default period surpluses. The number of non-default periods is derived from the probability of loans to individuals who ever have a bankruptcy flag to be charged off within two years of loan opening. The second row shows the aggregate change in consumer surplus for individuals with bankruptcy flags when flags are removed for 800,000 individuals. The third row of Panel D shows the implied consumer surplus loss for never-bankrupt individuals for the average 5 year loan scaled by the number of flag removals. It is the sum of non-default period surpluses. The number of non-default periods is derived from the probability of loans to individuals who never have a bankruptcy flag to be charged off within two years of loan opening. The fourth row scales the implied consumer surplus loss for never-bankrupt individuals over 5 years by the occurrence of never-bankrupt individuals and is, consequently, showing the average burden carried by individuals in the never-bankrupt group. The fifth row calculates the consumer surplus loss for never-bankrupt people when 800,000 bankruptcy flags are removed. The sixth row shows the social surplus change over five years scaled by the number of people with flag removal. It is the sum of first and third row. The sixth row shows the total change in social surplus when flags are removed for 800,000 individuals. It is the sum of the second and fifth row. The seventh row provides the efficiency change per dollar redistributed to bankrupt individuals by removing the bankruptcy flag. Source: TransUnion.

	(1)	(2)	(3)	(4)	(5)
Panel A: Prices					
Pre-flag-removal loan interest rate (%)	9.02%	9.02%	9.02%	9.02%	9.02%
Flag removal-induced change in interest rate (%)	-0.269%	-0.303%	-0.302%	-0.303%	-0.288%
Change in monthly payments (%)	-0.013%	-0.015%	-0.015%	-0.015%	-0.014%
Panel B: Quantities					
Pre-flag-removal loan quantity (Average $ per borrower per year)	$3,678.00	$3,678.00	$3,678.00	$3,678.00	$3,678.00
Flag removal-induced change in loan quantity (Average $ per borrower per year)	$342.56	$411.71	$404.58	$411.71	$410.32
Panel C: Elasticity and Slope					
Market Demand Elasticity	-3.12	-3.33	-3.29	-3.33	-3.49
Inverse demand slope (Interest rate % per $100)	-0.0785	-0.0736	-0.0746	-0.0736	-0.0702
Inverse demand slope (Repayment fraction % per $100)	-0.0038	-0.0036	-0.0036	-0.0036	-0.0034
Panel D: Surplus Changes					
Average consumer surplus redistributed to individuals with flag removal over 5 years ($ per eligible borrower with flag removal)	$28.74	$32.66	$32.52	$32.66	$31.04
Total consumer surplus redistributed to individuals with flag removal over 5 years ($)	$22,995,245	$26,128,165	$26,018,217	$26,128,165	$24,832,830
Average consumer surplus taken from never-bankrupt individuals over 5 years ($ per eligible borrower with flag removal)	-$30.17	-$34.59	-$34.42	-$34.59	-$32.87
Average consumer surplus taken from never-bankrupt individuals over 5 years ($ per eligible never bankrupt borrower)	-$3.58	-$4.10	-$4.08	-$4.10	-$3.90
Total consumer surplus taken from never-bankrupt individuals over 5 years ($)	-$24,137,612	-$27,674,278	-$27,532,558	-$27,674,278	-$26,297,587
Change in social surplus per individual over 5 years ($ per eligible borrower with flag removal)	-$1.50	-$1.43	-$1.93	-$1.93	-$1.83
Total change in social surplus over 5 years ($)	-$1,142,368	-$1,546,113	-$1,514,341	-$1,546,113	-$1,464,757
Welfare change per dollar redistributed to bankrupt individuals	-0.0497	-0.0592	-0.0582	-0.0592	-0.0590

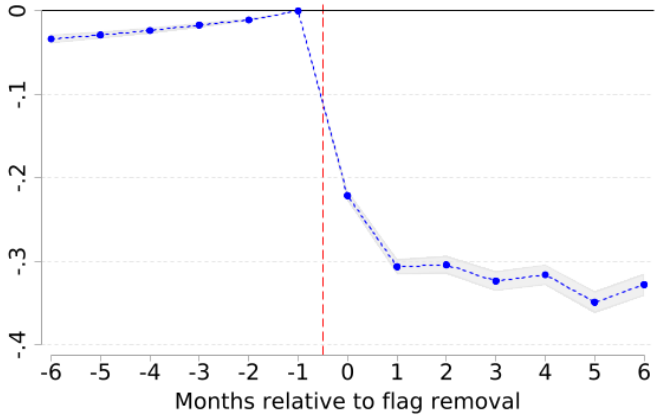
Figure E.1: Credit Scores, Interest Rates, and Loan Balances

This figure shows estimates of the coefficients δ_t from the following specification $y_{itg} = \gamma_{cg} + \gamma_{tsg} + \sum_{t=-6}^{6} \delta_t \{e_{itg} = t\} + \varepsilon_{itg}$, along with a 95% confidence interval. In the first panel, the outcome y_{itg} is credit scores, while in the second panel it is interest rates. In the third panel, the outcome is loan balances. γ_{cg} are cohort by group fixed effects, and γ_{tsg} are time period by score bucket by group fixed effects. Standard errors are clustered at the cohort by group level. Source: TransUnion.

Panel A: Credit Scores



Panel B: Interest Rates



Panel C: Loan Balances

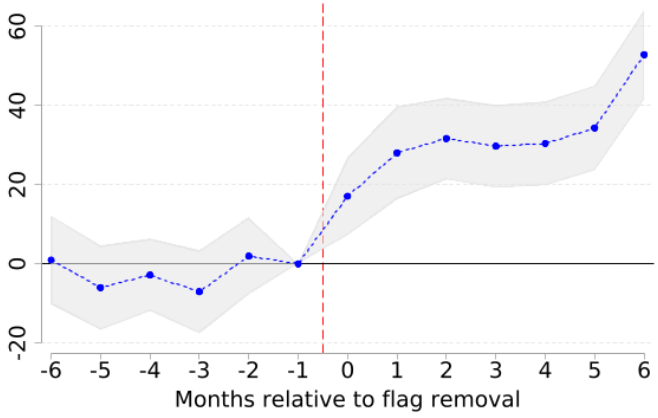


Figure E.2: Mean Outcomes: Credit Scores, Interest Rates, and Loan Volumes

This figure shows average treatment and control outcomes in relative time. We aggregate treatment and control outcomes for each generated dataset and, subsequently, average relative time means across datasets. Each generated dataset has the same weight in the plotted mean scores, predicted interest rates, and quantities. Source: TransUnion.

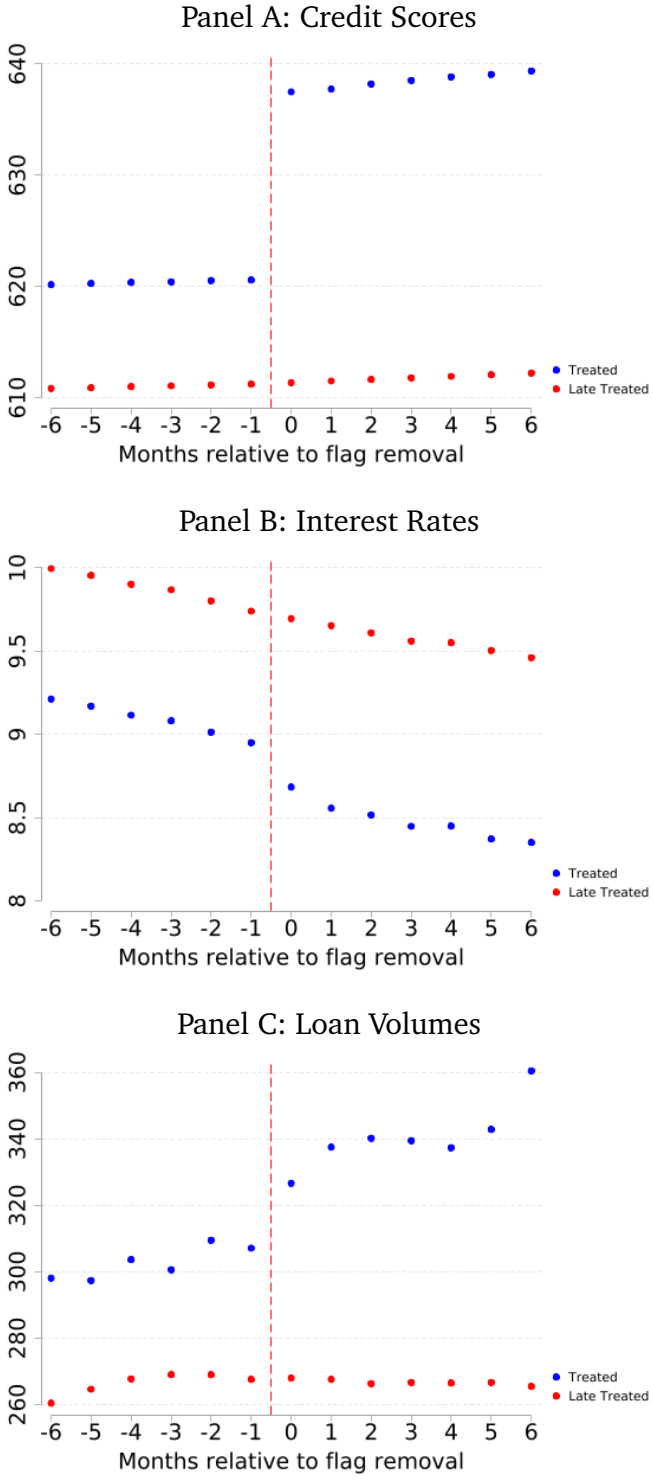
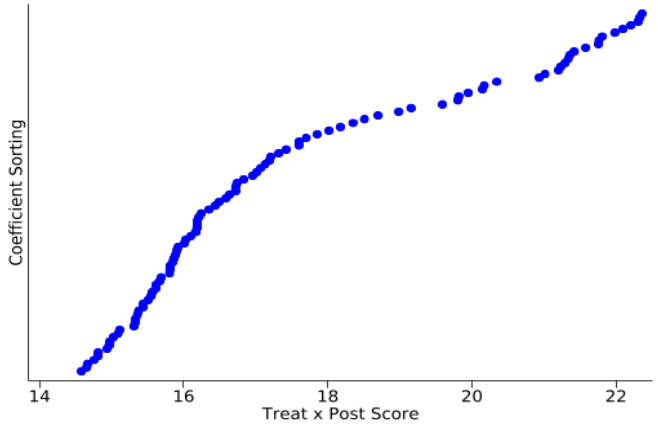


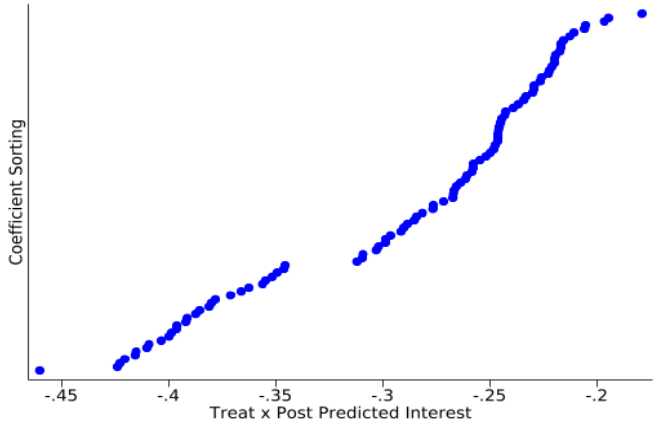
Figure E.3: Individual Events: Credit Scores, Interest Rates, and Loan Volumes

This figure shows estimates of the coefficients δ_t from the following specification $y_{itg} = \gamma_{cg} + \gamma_{tsg} + \delta_t \text{Treat} \times \text{Post} + \varepsilon_{itg}$. Point estimates across Panels do not correspond to each other as the coefficient sorting is Panel specific. Source: TransUnion.

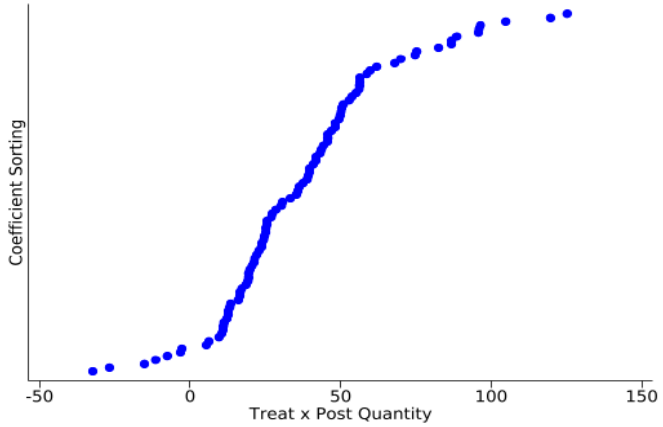
Panel A: Credit scores



Panel B: Interest rates



Panel C: Loan volumes



F Variable Definitions

Table F.1: Variable Description

This table denotes the construction of the main analysis variables. The source for all variables is TransUnion.

Variable	Description
Credit Score	VantageScore 3.0
Quantity Opened	Sum of balances on new auto accounts opened by an individual in a given month; zero when no account opened by the individual in the given month
Quantity Opened Cond. on Opening	Sum of balances on new auto accounts opened by an individual in a given month conditional on an opening being reported
Auto Interest Rate	Credit amount weighted interest of auto accounts at opening. Missing when no auto account opened by individual in a given month
Charged-off	1 if one of the auto loans opened by an individual in a given month is charged-off within the 2 years after opening and zero otherwise
Score Bucket	One of 20 score buckets assigned in the month before flag removal and held constant throughout