Disaster Flags: Credit Reporting Relief from Natural Disasters

Benedict Guttman-Kenney*

12th February, 2025

Abstract

I study the use of "disaster flags" that are applied to consumer credit reports and aim to provide relief to consumers exposed to natural disasters. Between 2015 and 2024, 68 million consumers had a disaster flag on their US credit report. Consumers with pre-disaster financial distress experience the largest, but temporary, VantageScore credit score increases from flags, however, their credit access does not improve. A counterfactual policy that automatically masks all defaults in credit reports for all consumers exposed to natural disasters, masks 2% to 33% of all defaults at the cost of reducing predictive performance by 0.1% to 0.5%.

*Rice University, Jones Graduate School of Business. Email: benedictgk@rice.edu. I am grateful for feedback from my curriculum paper advisors, Matt Notowidigdo & Neale Mahoney, and I also thank Aditya Chaudhry, Andrés Shahidinejad, Bob Hunt, Constantine Yannelis, Eric Budish, Evan White, Karthik Srinivasan, Jack Mountjoy, John Heilbron, Jonah Kaplan, Judith Ricks, Kelly Cochran, Larry Santucci, Lauren Lambie-Hanson, Lucy Msall, Melody Harvey, Michael Varley, Pascal Noel, Pauline Mourot, Rob Marquez, Sally Parker, Scott Nelson, Stephanie Moulton, Thomas Covert, Walter Zhang, Chicago Booth microeconomics and finance groups, industry participants, the CFPB, the Federal Reserve Bank of Philadelphia's Consumer Finance Institute, FinRegLab, International Conference on Credit Risk Evaluation, and American Finance Association (AFA) Annual Meeting, and APPAM conference participants. I am grateful to support from the NBER's PhD Dissertation Fellowship on Consumer Financial Management funded by the Institute of Consumer Money Management, Bradley Fellowship, Katherine Dusak Miller PhD Fellowship, and the Sanford J. Grossman Fellowship in Honor of Arnold Zellner. I am grateful to the University of Chicago Booth School of Business's Kilts Center for Marketing (especially Art Middlebrooks and Heather McGuire) for supporting my research. This paper is an extension of one chapter of Guttman-Kenney (2024)'s PhD Thesis. The results in this paper were 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 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.

1 Introduction

The United States is increasingly affected by more numerous and more economically damaging natural disasters such as flooding, hurricanes, tornadoes, and wildfires.[1] This climate change poses many challenges to financial markets (e.g., Giglio et al., 2021). Climate change means that some consumers are increasingly exposed to disasters that can cause or exacerbate financial distress. Consumer financial distress in the form of missing one or more payments ("defaults") on credit products during a natural disaster can have longer-term adverse effects, including reduced access to credit. This is because information on defaults from the last seven years remains on a consumer's credit report, and therefore can have persistent impacts on a consumer's credit score and access to credit.

In this paper, I consider the role of masking defaults in credit reports to provide relief to consumers exposed to natural disasters. Masking defaults may mean trading-off heterogeneous impacts. It may improve credit access for masked defaulters, who appear less risky to lenders, but reduce credit access for non-defaulters, who instead appear riskier having been pooled with masked defaulters. Whether to mask the defaults that occur during natural disasters ("disaster defaults") depends on how informative such information is in predicting future defaults. If disaster defaults are highly predictive, then masking this information is expected to be costly to lenders and reduce the market efficiency of lending. Whereas if disaster defaults offer limited predictive value, then a social planner may consider it worthwhile to mask this information, as it would indicate that defaults arising during natural disasters are not informatively revealing a consumer's risk type.

I research this topic by first documenting and evaluating the existing voluntary system of natural disaster credit reporting relief developed by the market. Second, I consider a counterfactual government policy that could be implemented, which would automatically, temporarily mask all defaults in credit reports for all consumers that reside in areas affected by natural disasters. I show how lenders currently respond to natural disasters

[1] Between 1980 to 2010, there were only two years, 1998 and 2008, with at least ten weather/climate disasters each resulting in damages of one billion-dollars or more. In contrast, every year from 2011 to 2024, except 2014, has experienced at least ten weather/climate disasters where each caused at least one billion dollars in damages. There were a record-breaking 28 and 27 billion-dollar weather/climate disasters in 2023 and 2024. Source: National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disaster. Billion-dollar disasters are inflation-adjusted.

by applying “disaster flags” to their customers’ credit reports. Disaster flags are designed to provide relief to consumers exposed to natural disasters, with the aim of helping to protect their credit access. Lenders voluntarily apply disaster flags. These flags temporarily mask negative information (i.e., defaults) about a consumer in the calculation of their VantageScore credit score. The only other study on this topic is a short report by the Consumer Financial Protection Bureau (Banko-Ferran and Ricks, 2018) that concludes that “more analysis is needed to better understand whether and how the furnishing of information on natural disasters affects consumer credit”. My paper addresses this research gap.

I use a 10% representative sample of monthly US consumer credit reporting data from 2000 to 2024 to document five new facts on disaster flag’s use. First, there is growth in the use of disaster flags on credit reports over time. 68.1 million consumers had a disaster flag on their credit report between 2015 and 2024. This is 6.9 times the number of consumers who became bankrupt in the same period. Disaster flags were rarely used until Hurricane Katrina in 2005, and use increased tenfold in 2017 with Hurricanes Harvey and Irma. Second, there is broad geographic coverage of disaster flags. Disaster flags are most commonly used in the South East coastal areas prone to hurricanes, however, they are increasingly used across the US, most broadly during 2020 and 2021 in response to the COVID-19 pandemic and other disasters. Third, the majority of disaster flags only remain on a credit tradeline for a few months. 88% of disaster flags are present for six months or less. Fourth, the majority of accounts with disaster flags do not also have deferments reported, except for student loans. Fifth, disaster flags are typically only applied to a subset of consumers’ credit accounts. Only 11% of consumers with at least one disaster flag have disaster flags on all their credit reporting accounts.

I examine the information value of disaster flags to understand how costly it is for lenders to apply disaster flags. I find that consumers with disaster flags are a selected sample of the population. They are, on average, more indebted than unflagged consumers in the same geographic areas. I construct credit scoring models to evaluate the predictive value of defaults that also have disaster flags, “flagged defaults”, compared to “unflagged defaults”, defaults that do not also have disaster flags. I find that flagged defaults are slightly riskier signals than unflagged defaults. Although a model predicting a new credit default that masks flagged defaults as an input performs worse than a model without such masking, the difference between the two appears economically small. It therefore appears that lenders incur a seemingly small cost to apply disaster flags to mask defaults.

I use event study and difference-in-differences methodologies to evaluate the benefits of disaster flags to consumers. My results are consistent across these approaches. Disaster flags mask a subset of a consumer's defaults, which initially leads to an economically small 2.8 points increase in the average VantageScore credit score, relative to consumers exposed to the same disaster. These average effect sizes are smaller than those found in research on the effects of removing bankruptcy flags from US consumer credit reports (e.g., Dobbie et al., 2020; Gross et al., 2020; Jansen et al., 2024). There is heterogeneity in the effects of disaster flags on credit scores. For the subgroups of consumers with subprime credit scores or with any defaults in the twelve months before the disaster flag was applied, the initial effect of disaster flags on credit scores is larger, 10 to 15 points, however, it dissipates within twelve months. Even among consumers who experience the largest temporary increases in their credit scores, I do not find evidence that flags improve consumer credit access, measured by new account openings. Instead, the number of new account openings decline. Flags may not improve credit access, partially because the credit score increases that I find are temporary, dissipating within twelve months. Also, although disaster flags temporarily affect VantageScore credit scores, that are observed in my data, disaster flags do not affect FICO credit scores, that are unobserved in my data, and therefore credit decisions taken using FICO scores would be unaffected.

In the last part of the paper, I consider a counterfactual feasible policy that automatically masks all defaults in credit reports for consumers when they are exposed to natural disasters. Such place-based policies may be motivated by redistributive objectives (e.g., Gaubert et al., 2024), given the large and persistent geographic inequalities in financial distress across the US (e.g., Keys et al., 2022). I merge my credit reporting data with public government data on the timing and location of natural disasters. I examine a variety of ways to mask information based on different thresholds of masking defaults within three, six, and twelve months from a disaster, and whether such information is masked only temporarily during a disaster, or permanently afterwards. I find that masking all defaults of consumers exposed to natural disasters would mask 2% to 33% of all defaults of consumers with US credit reports. I estimate that masking all disaster defaults would reduce the ability to predict a new credit default, as measured by AUC, by 0.1% to 0.5%, with more generous policies leading to larger predictive losses. These losses can be benchmarked relative to a model where all defaults (i.e., irrespective of whether during a disaster or not) are masked, which would reduce AUC by 1.0%. My results quantify the trade-off that policymakers face, showing that offering more generous relief masking defaults for a longer period of time increasingly reduces predictive performance, and the costs depend upon how defaults that initially occur during disaster are recorded in credit reports after disasters.

A growing literature studies the effects of disasters, and government assistance, on household finances (e.g., Gallagher and Hartley, 2017; Deryugina et al., 2018; Farrell and Greig, 2018; Bleemer and van der Klaauw, 2019; Billings et al., 2022; Gallagher et al., 2023; Begley et al., 2024; Collier et al., 2024a; Cookson et al., 2024). The first contribution of my paper to this literature is to understand a little-known but widely used form of relief: "disaster flags". My second contribution is to provide quantitative evidence to inform recent public policy discussions (e.g., Banko-Ferran and Ricks, 2018; National Consumer Law Center, 2019; Urban Institute, 2019; FinRegLab, 2020) on the potential for an alternative policy that automatically provides credit reporting relief from natural disasters.

Prior household finance literature has studied the effects of changes in credit contract terms, such as reductions in principal or monthly payments, to alleviate consumer financial distress (e.g., Agarwal et al., 2017; Dobbie and Song, 2020; Ganong and Noel, 2020; Cherry et al., 2021; Chava et al., 2023; Goodman and Zhu, 2023; Aydin, 2024; Di Maggio et al., 2024; Dinerstein et al., 2024; Katz, 2024; Kim et al., 2024). My contribution is to add evidence on a new policy tool to this prior literature. Masking defaults during natural disasters is a form of relief that does not change the contract terms and instead only changes how information on a consumer's defaults appear in credit reports.

My research informs a literature on the economics of credit information, reviewed in Gibbs et al. (2024). Prior literature has studied the effects of removing bankruptcies (e.g., Musto, 2004; Dobbie et al., 2020; Gross et al., 2020; Herkenhoff et al., 2021; Jansen et al., 2024), and historical defaults (e.g., Bos et al., 2018; Liberman et al., 2020; Blattner et al., 2022). Research has studied how hospital admissions and insurance coverage affect medical debt in collections (Dobkin et al., 2018; Kluender et al., 2021; Batty et al., 2022), the credit risk value of information on medical debt in collections (Brevoort and Kambara, 2015), and the implications of removing this information (Kluender et al., 2024). Information on medical debt is increasingly removed from credit reports and credit scores, partially driven by fairness concerns that such information, relative to other debts, arguably may arise more from bad luck rather than being an informative negative signal of a consumer's type. A somewhat similar concern may apply to disaster defaults. I contribute to this literature by introducing a new source of variation, disaster flags, and evaluating the predictive value of flagged and disaster defaults compared to non-disaster defaults. Considering the implications of masking information in credit reports is more broadly important in the wake of the COVID-19 pandemic, which resulted in widespread, but untested, laws preventing lenders from updating adverse information in US credit reports (Cherry et al., 2021), with similar policies implemented in other countries. Understanding the implications of different information being shared is therefore increasingly important to study (e.g., Guttman-Kenney and Shahidinejad, 2025).

Finally, my study also contributes to the social insurance literature by studying a voluntary form of social insurance. One of the main roles of public policymaking is to provide social insurance (e.g., Chetty and Finkelstein, 2013): Providing insurance against adverse shocks such as being unemployed or suffering poor health or distress from a natural disaster. Previous research has studied the connections between social insurance and household debt (e.g., Hsu et al., 2018; Bornstein and Indarte, 2023; Braxton et al., 2024), and Deryugina (2017) shows that the fiscal costs of social insurance payments (e.g., unemployment insurance, public medical payments) significantly outweigh the costs of direct disaster aid. My contribution to this literature is showing a case of voluntary social insurance, where disaster flags "tag" (e.g., Akerlof, 1978) a group of consumers affected by natural disasters, which is an interesting case given how there is less use of tagging across domains than theory recommends (e.g., Weinzierl, 2012).

The paper proceeds as follows. Section 2 provides the institutional background, a motivating framework, and explains the data I use. Section 3 documents five new facts about how disaster flags are used. Section 4 shows the characteristics of consumers with disaster flags and examines the information value of masking defaults. Section 5 uses event study and difference-in-differences methodologies to evaluate the benefits of disaster flags for consumers. In Section 6, I examine the loss of information from automatically masking all defaults during natural disasters. Finally, Section 7 concludes.

2 Institutional Background and Data

Section 2.1 provides institutional details on disaster flags. Section 2.2 provides a motivating framework for considering the masking information in credit reports, such as through disaster flags. Section 2.3 describes the data used in this paper.

2.1 What Are Disaster Flags?

Lenders can apply a "disaster flag" to their customer's credit report to show that they have been affected by natural or declared disasters. These flags appear as a comment code "AW" added to an individual tradeline account.[2] Disaster flags are intended to provide credit reporting relief to consumers by protecting credit access following exposure to natural disasters such as hurricanes, forest fires, and COVID-19.

There are no governmental or regulatory requirements for lenders to use disaster flags, nor is there explicit guidance on whether or how to do so. The industry organization that governs information sharing, the Consumer Data Industry Association, is not prescriptive in its guidance on lenders' use of disaster flags. Lenders have complete discretion over whether to apply disaster flags and, if so, which consumers and tradelines to apply them (e.g., all or a subset in an area exposed to a natural disaster) and how many months to keep flags on a consumer's credit report for. Disaster flags are a separate field from the reporting of defaults in credit reports. Discussions with industry participants indicate that some lenders sometimes do not report new defaults during natural disasters, however, it is unclear how frequently such non-reported defaults are as they are, by definition, there are unobserved in credit reporting data..[3] Disaster flags may be applied instead of or in addition to changes in contract terms (e.g., deferring payments or offering forbearance)

[2] Credit Reporting Resource Guide FAQ 58 explains how these are recorded in credit reports with a comment code "AW" added to the tradeline. In TransUnion data, the comment (remark) code is technically named "AND" instead of "AW". https://www.fico.com/sites/default/files/upload_files/FAQ.pdf

[3] Although not the focus of this study, such non-reporting of defaults may help to explain why the average effects of natural disasters on defaults, measured by those observed in credit reporting data, found in prior literature have been described as "modest" (Gallagher and Hartley, 2017). In 2018, Fannie Mae and Freddie Mac introduced the requirement that mortgage servicers temporarily not update the default status in credit reports for their mortgages if they default when affected by natural disasters, this was removed in 2020 following the CARES Act taking effect.

that may also be recorded in credit reports.[4]

Disaster flags mask negative information only on the flagged tradeline in the calculation of VantageScore credit scores.[5] Flags only mask information when the flag is currently present on a tradeline. Once a flag is removed, the previously masked information is revealed. Disaster flags do not factor into the calculation of FICO credit scores, i.e., they do not mask negative information, and FICO states that "the reporting of special comment code AW alone will not affect a consumer's FICO Score".[6] Manual underwriters that review a consumer's credit report can observe disaster flags and consider them in their credit decisions. There are potential parallels between the application of a disaster flag and the removal of a bankruptcy flag 7 to 10 years after bankruptcy, studied in many prior papers (e.g., Musto, 2004; Dobbie et al., 2020; Gross et al., 2020; Jansen et al., 2024). Both disaster flag application and bankruptcy flag removal mask information in credit reports, resulting in the pooling of consumers with different credit risks. Therefore, one may consider "disaster flags" as a type of temporary low-cost bankruptcy providing a non-governmental form of social insurance for consumers affected by natural disasters.

2.2 Motivating Framework

I use a stylized framework of credit scoring to motivate this paper. The basis of most lending decisions in the US, and many other developed countries, are credit scores, derived from credit reporting data, that predict the likelihood of future default (i.e., a missed payment). A credit applicant's credit score determines whether their application is accepted and, if so, what contractual terms, such as the interest rate and amount of credit, are offered. A higher credit score represents a lower probability of default, i.e., lower credit risk. Equation 1 shows a simple example, where a credit score predicts at time t, an outcome,

[4] Credit Reporting Resource Guide FAQ 44 and 45 explain how these "accommodations" are recorded in credit reports by setting payments due equal to zero or adding codes to show that the payment is deferred or the agreement is in forbearance. The Consumer Data Industry Association defines a deferred payment as "A loan arrangement in which the borrower is allowed to start making payments at some specified time in the future." and forbearance as: "A period during repayment in which a borrower is permitted to temporarily postpone making regular monthly payments. The debt is not forgiven, but regular payments are suspended until a later time...The consumer may be making reduced payments, interest-only payments or no payments." https://www.cdiaonline.org/wp-content/uploads/2020/03/CDIA-NEWS_Coronavirus-The-Credit-Bureaus-Response_3.15.2020.pdf

[5] https://cdn2.hubspot.net/hubfs/431136/Forbearance%20Hub/VantageScore%20Code%20AW%20Update%20and%20FAQs%20(003).pdf

[6] https://www.fico.com/en/covid-19-credit-reporting-impact-US/

$Y_{i,t+1}$, the likelihood that the consumer i will default on a credit agreement one period in the future. For simplicity, here I assume that the consumer only has one credit agreement. The credit score has some generic function $f(.)$, historically this is typically a logistic, and I have partitioned the predictive inputs into a single binary default component, $d_{i,t}$, which takes a value of one if defaults and zero otherwise, and a vector of all other non-default inputs, $X'_{i,t}$, such as product holdings, balances, credit card utilization. These inputs that are measured in consumer credit reporting data (see Gibbs et al., 2024 for more general information on credit reporting data and credit scoring).

$$Pr\Big(Y_{i,t+1} = 1\Big) = f\Big(X'_{i,t}\beta_1 + \theta_1 d_{i,t}\Big) \tag{1}$$

In such credit scoring models, past defaults are a strong predictor of future defaults with $\theta_1 > 0$. Consumers with past defaults have lower credit scores, resulting in lower access to credit and higher interest rates. As credit scores are predictive models, the relationships between inputs and the outcome are not causal. Credit scoring models regard the predictive value of a default as being homogeneous irrespective of the underlying cause or heterogeneity by socioeconomic characteristics, despite it masking variation that may improve prediction. This is due to a mixture of a lack of data and legal constraints that limit the predictive accuracy of credit scoring models. For example, lenders have limited visibility of life events, such as income shocks, and, for equity reasons, cannot discriminate on the basis of protected characteristics, such as gender and race. One source of heterogeneity that is observable in the data and is not a protected characteristic is whether defaults differ with natural disasters (e.g., wildfires, floods, hurricanes). Does the ability to predict future defaults vary depending on whether a default occurs during a natural disaster or not? Equation 2 allows for this by including an interaction term between the binary default term, $d_{i,t}$, and a binary variable, $N_{g,t}$, which takes a value of one if the consumer resides in a geographical area g at time t where and when that area is exposed to a natural disaster, and if not, is zero.

$$Pr\Big(Y_{i,t+1} = 1\Big) = f\Big(X'_{i,t}\beta_2 + \theta_2 d_{i,t} + \pi(d_{i,t} \times N_{g,t})\Big) \tag{2}$$

The value of the π parameter in Equation 2 is informative of the marginal predictive value of defaults during natural disasters, "disaster defaults", compared to non-disaster defaults. It may be that $\pi < 0$, meaning that disaster defaults are lower risk than non-disaster defaults. This could be due to disasters being exogenous shocks to households, disasters disrupting communications making it difficult for households to make payments on time, and households being better able to recover due to Federal assistance that may only arrive with a lag so be unable to prevent the original default. In contrast, it may be that $\pi > 0$, meaning that disaster defaults may be higher risk than non-disaster defaults, possibly due to disasters causing longer-term damage to household resilience, or revealing riskier types. If $\pi = 0$, and the predictive performance does not improve, then differentiating disaster defaults from non-disaster defaults may not be informative to improve credit risk prediction. If disaster defaults are uninformative noise, then predictive performance may even be improved by masking such information.

Given this framework, I can quantify how costly it would be for the credit industry to mask disaster defaults in credit reporting data by adapting Equation 1 to Equation 3, where disaster defaults are masked to be recorded as not in default. Comparison of the predictive performance of these two models can be informative. If the difference in predictive performance is small, the credit industry may voluntarily agree to mask disaster defaults. However, if the difference in predictive performance is large, lenders would be reluctant to voluntarily mask such information, and then the government has to decide the merits based on its social welfare function. By masking default information in credit reporting data, consumers of different risks are pooled together with the same credit scores, and, therefore, it may help to preserve the credit access of those affected by natural disasters. Although my study examines natural disaster defaults, this framework could be applied to evaluate other characteristics with richer data merged in, e.g., masking defaults linked to life events such as divorce, income shocks, or expenditure shocks.

$$Pr\Big(Y_{i,t+1} = 1\Big) = f\Big(X'_{i,t}\beta_3 + \theta_3\tilde{d}_{i,t}\Big), \text{ where } \tilde{d}_{i,t} \begin{cases} 0 \text{ if } N_{g,t} = 1 \\ d_{i,t} \text{ otherwise} \end{cases} \tag{3}$$

2.3 Data

2.3.1 Consumer Credit Reporting Data

This research uses a large, anonymized, representative sample of US consumer credit reporting data: The University of Chicago Booth School of Business TransUnion Consumer Credit Panel (BTCCP). The BTCCP is provided by TransUnion to the University of Chicago Booth School of Business (TransUnion, 2024). The data is a 10% sample of consumers with a TransUnion credit report in July 2000 supplemented with 10% of new entrants added each month to ensure the sample remains representative. The data is at the individual tradeline account level, i.e., showing each mainstream credit account held by a consumer, at the monthly frequency from July 2000 to December 2024. Each month of data is an archive that recreates the consumer's credit report as it would have appeared at that point-in-time and as lenders would take credit decisions on. Individual tradelines and consumers are tracked over time with anonymized identifiers. In addition to the tradeline data, each month of data also includes the consumer's VantageScore 3.0 credit score and other consumer-level attributes. From January 2009, the data contain more detailed data, so my research focuses on this period. See Gibbs et al., 2024 for a broader review of consumer credit reporting data.

For each consumer, I observe the state, zip code, and the census block group of their primary address each month. Census block groups are units of geography that typically contain 600 to 3,000 consumers and are more granular than census tracts. I keep observations for consumers in the US, with a birth date and restrict to where the birth year is after 1920 and before 2007, and when a consumer has tradeline data, at any point 2000 to 2024, to remove low-quality, fragmented credit records (e.g., "consumers" with only inquiries), following Gibbs et al. (2024).

Importantly, for my study, I observe whether a disaster flag was applied for each tradeline, each month. This monthly tradeline-level view is crucial. Disaster flags would not be visible in credit reporting variables that are aggregated to the consumer level. Furthermore, quarterly or annual tradeline-level data would not observe disaster flags applied intra-quarter unless such flags were still present on a tradeline at the end of a quarter.

2.3.2 Natural Disasters Data

When a major disaster occurs, it is declared as such by the US President under the Stafford Act. I use public government data on these declarations provided by the Federal Emergency Management Agency (FEMA)'s Disaster Declarations Summaries.[7] I restrict analysis to natural disasters e.g., flooding, hurricanes, wildfires, severe storms, tornadoes. This excludes chemical, toxic substances, terrorist, or other disaster events. The data report the timing and location of all federally declared disasters. These events are generally at the county-level, however, there are five cases that are statewide. The data is merged with the BTCCP by county, state, and date, using the public HUD USPS zip code crosswalk files for linking zip codes to counties.

3 Disaster Flag Facts

I use my data to document five new facts describing the use of disaster flags in the US.

FACT 1. Growth in the use of disaster flags.

68.1 million consumers in the US had a disaster flag on their credit report between January 2015 and December 2024, and 72.9 million consumers across my entire dataset going back to July 2000. These statistics are calculated as a disaster flag on at least one open tradeline in their consumer credit report for at least one month.[8] This is a substantial number of consumers. To provide a benchmark, this is 6.9 times the number of consumers who became bankrupt in the US between 2015 and 2024.[9] The large number of consumers with disaster flags in their credit reports makes this an important practice to understand. Disaster flags are applied across all mainstream credit types (auto loans, credit cards, mortgages, and student loans), and across all lender types (banks, non-bank finance companies, and credit unions), see Internet Appendix Table A1 and Figures A1 and A2.

[7] https://www.fema.gov/openfema-data-page/disaster-declarations-summaries-v2

[8] If closed tradelines are included, this is 70.6 million consumers (2015 to 2024). If only open tradelines with positive balances are included, this is 63.1 million consumers (2015 to 2024).

[9] I estimate 9.2 million bankruptcies with filing dates between January 2015 and December 2024 based on chapter 7 or chapter 13 filings, dismissals, or discharges observed.

Figure 1 Panel A shows that the use of disaster flags has increased greatly with time. Disaster flags were very rarely used until Hurricane Katrina in 2005. There are spikes in the use of disaster flags in 2017 that are mainly driven by Hurricanes Harvey and Irma, in 2020 due to COVID-19 and other disasters, and in 2024 primarily due to Hurricanes Milton and Helene. The growth is so large in 2017 that I separately present the period up to July 2017 in Figure 1 Panel B with a scale that is ten times smaller than Panel A. The growth over time is consistent with Banko-Ferran and Ricks (2018) that examined Hurricane Harvey, and found that very few tradelines in Texas already had disaster flags in the months just before the hurricane.

FACT 2. Broad geographic usage of flags.

The panels of Figure 2 display the fraction of consumers in US counties with a credit report who had a disaster flag for each year 2015 to 2024, selecting the month each year that has the highest number of consumers with disaster flags. This shows that disaster flags are most commonly used in the South East coastal areas, which are more prone to hurricanes. However, over time there are increasing pockets of usage elsewhere in the country. For example, we see areas of the North West affected by wildfires, and in 2018 for Maine following severe storms. Figure 2 Panel F shows that disaster flags appear in credit reports across the country in response to COVID-19 and other natural disasters in 2020, and Panel G shows that they are often present in 2021. Coverage is broad based across counties, however, there is noticeable regional variation in the intensity of usage. Figure 2 Panel J shows that in November 2024, we see that the use of disaster flags follows a path corresponding to the impacts of Hurricanes Milton and Helene.

FACT 3. The majority of flags only remain on a tradeline for a few months.

Figure 3 Panel A shows how long disaster flags remain on a tradeline, in months since the flag was first applied. I observe that disaster flags typically only remain on a credit tradeline for up to three months and rarely more than six months. 31% of tradelines with disaster flags are only flagged for one month, 48% for three months or less, 88% for six months or less, and 92% for twelve months or less. Figure 3 Panel B shows that the flags on auto loans are likely to remain on those tradelines slightly longer than for credit cards, mortgages, or student loans. This short duration limits the potential relief that disaster flags can provide to consumers, as the disruption that consumers may experience from disasters may last more than a few months. These results are broadly similar across lender types and over time, except for an increase in duration for flags applied during COVID-19, as shown in Internet Appendix Figures A4 and A5.

FACT 4. The majority of accounts with flags do not also have deferments reported.

Figure 4 shows that disaster flags are typically applied to tradelines without deferments reported, except for student loans. Deferments are measured by either a deferment being listed on the tradeline or when a tradeline has a positive balance but has zero payments due. Between 2009 and 2024, 17% of all tradelines, excluding student loans, that have disaster flags are also deferred at the same time. During the pre-COVID-19 period, from 2009 to 2019, only 6% of tradelines with disaster flags are also deferred. Between 2009 and 2024, 81% of student loans with disaster flags are also deferred. Deferments have become more common on flagged accounts since the onset of COVID-19 in 2020 when federally-mandated payment deferments occurred more broadly, and visibly recently for student loans. Disaggregating by lender and credit type shows that since 2020 mortgages are more likely to be deferred than auto loans or credit cards, and banks are more likely to defer loans than non-bank finance lenders (Internet Appendix Figure A3).

FACT 5. Flags are usually only applied to a subset of a consumer's accounts.

Among consumers with flags, typically only a third of their tradeline accounts on their credit report have disaster flags, with 34% between 2009 and 2024, and also 34% as of December 2024. Figure 5 shows the intensive margin of the use of flags by the number of tradelines held: the fraction of a flagged consumer's tradelines flagged in Panel A, and the share of flagged consumers with all tradelines flagged in Panel B.

Disaster flags are only attached to the individual tradeline accounts to which they are applied. This means that a consumer's entire portfolio only has disaster flags on it if all lenders add disaster flags to all a consumer's tradelines. This is a rare event. Across 2009 to 2024, only 11% of consumers, with at least one disaster flag on one tradeline, have disaster flags on *all* of their open tradelines, and only 9% in December 2024. Figure 5 Panel A shows the fraction of tradelines flagged decreases with the number of tradelines held. Figure 5 Panel B shows that it is extremely rare, below 2%, for consumers with three or more tradelines to have flags on all of their tradelines. This indicates that frictions exist in the use of disaster flags, consistent with Kim et al. (2024) who show intermediation frictions in COVID-19 mortgage forbearance being applied. Although we do observe a slight trend of increasing intensity of use over time, see Internet Appendix Figure A6. As only a small subset of a consumer's tradelines are typically flagged, this limits the potential relief that disaster flags can provide to consumers. This is because only negative information on flagged accounts is masked in their VantageScore calculation, and therefore even if a consumer has a disaster flag on one account, negative information on that same consumer's other unflagged accounts still impacts their credit score.

4 Information Costs of Disaster Flags Masking Defaults

Having documented how disaster flags are used, I now evaluate the information on credit risk that they contain. Section 4.1 describes the selection of consumers with disaster flags, and I build predictive models to quantify the information value contained in flagged defaults, with the methodology in Section 4.2 and the results in Section 4.3.

4.1 Describing Selection

What are the characteristics of consumers with disaster flags? I examine this in Table 1 that compares (1) consumers with disaster flags, based on the time when they are first flagged, to (2) consumers who never have disaster flags but are in the same geographical region (a combination of census block group and zip code) at the same time as those that do, and (3) consumers who never have disaster flags and are in other geographical regions and/or time periods without disaster flags. Consumers with disaster flags are a selected sample of the population. Consumers with disaster flags are, on average, more indebted, with more tradelines, more defaults, and higher balances. This selection holds both when comparing flagged consumers with unflagged consumers in the same geographical region and also when comparing with unflagged consumers across the US. Credit scores of flagged consumers are slightly higher than those of unflagged consumers, although this varies over time, for example, cohorts of consumers flagged in 2017 typically have higher credit scores, whereas cohorts flagged in 2020 typically have lower credit scores, as shown in Internet Appendix Figure A7.

4.2 Predictive Methodology

I apply my motivating framework from Section 2.2 to my data to evaluate the information costs of disaster flags that mask defaults in credit reports. I use a representative sample containing 23.3 million consumers who have tradeline data, a non-missing VantageScore, and are in the US as of October 2017. I train models on 70% of the data, 16.3 million consumers, and test model performance out-of-sample on the remaining 30%, 6.99 million consumers. This time period is chosen to ensure that there is a sufficiently large sample of consumers with disaster flags in my data, and to ensure my inputs and my outcome are not affected by the non-reporting of defaults that occurred during COVID-19. My outcome, Y_{t+24}, is a binary outcome for whether a consumer has any *new* default in the 24 months after October 2017, measured as 90 or more days past due. A new default is one where a tradeline was not in default in October 2017 but is in default at any point between November 2017 and October 2019. This outcome is positive for 16.9% of my sample.

I build a series of models to evaluate the value of information for predicting default. I calculate my models in two steps. First, I construct a credit score ($S_{i,t}$) using only the vector of non-default variables as inputs. Second, I use this credit score and evaluate how varying information on defaults as inputs to improves predictive performance.

In the first step, I want to capture predictive information in credit reports except for information on defaults. I therefore build a credit score without default information using 171 non-default variables as predictors. This uses 138 consumer-level attributes available in my dataset that do not use defaults information (and this does *not* include VantageScore credit score). I construct 33 additional consumer-level variables. 31 variables are created by aggregating tradeline-level data to the consumer-level. These are the number of accounts, outstanding balances, number of closed accounts, and number of open accounts with a positive balance for each of: all accounts, non-mortgage accounts, auto loan accounts, credit card accounts, mortgage accounts, student loan, and unsecured personal loan accounts. I also construct from tradeline-data the total value of credit card limits and their utilization rate. From the headers file, I construct the number of months with a credit report, and finally from the public records file, I construct an indicator for any bankruptcy. To avoid overfitting outliers and dropping missing observations, I winsorize all these variables at their 99th percentile of non-zero, non-missing values, with missing values imputed at zeros (except for variables of the format measuring the number of months since an event, which I instead impute at the variable's 99th percentile). In this first step, I use a XGBoost machine learning method to predict default.[10] This is a methodology that has been previously shown to be highly predictive of default in consumer credit applications (e.g., Blattner and Nelson, 2024; Fuster et al., 2022), and is also used in industry (FinRegLab et al., 2022). This model without default information predicts defaults well, with an area under the receiver operating characteristic curve (AUC) of 0.8902 (Table 4), where an AUC of 0.5 would mean there is no information in the classification and an AUC of 1 would be classifying cases perfectly.

In the second step, I use either a logistic regression approach, enabling me to interpret coefficients on default variables, or a XGBoost machine learning methodology to better assess the ability of defaults to improve credit risk prediction. I construct a vector of default variables, $D'_{i,t} = [d^{12}_{i,t}, d^{24}_{i,t}, d^{36}_{i,t}, d^{84}_{i,t}]'$, for each consumer i, at time t, where $d^k_{i,t}$ denotes the number of accounts in default over the last k months. These are calculated from tradeline-level data and aggregated to the consumer-level. $X_{i,t}$ denotes the non-default variables included in this prediction. In the logistic regression, I include variables for the number of accounts 30 or more days past due in the last 12, 24, 36, and 84 months, along with my credit score variable created in the first step. In the machine learning approach, I include variables for the number of accounts 30 or more days past due in the last 6, 12, 24, 36, and 84 months, and the number of accounts 90 or more days past due in the last

[10]I use as hyperparameters a maximum depth of 10, learning rate of 0.3, using 80% of data for each tree, and 80% of features for each tree, using 1,000 rounds, and stop if there has been no improvement after 50 rounds.

6, 12, 24, 36, and 84 months, as well my credit score variable created in the first step, and the 171 non-default variables. I do not include information beyond 84 months because 84 months is the maximum duration for which defaults remain in credit reports under existing US law (Gibbs et al., 2024).

My baseline predictive model is shown in Equation 4. This is a traditional credit score that includes historical default information. My second model is shown in Equation 5. This adds a vector of "flagged default" variables, $F'_{i,t} = [f^{12}_{i,t}, f^{24}_{i,t}, f^{36}_{i,t}, f^{84}_{i,t}]'$, over the same time horizons as $D_{i,t}$, and are constructed from interactions between defaults and disaster flags. For constructing $F'_{i,t}$, each consumer's tradeline default status each month is assigned a value of one only if it is both in default and also has a disaster flag that month, and zero otherwise. I aggregate this to the consumer level. I then construct the consumer level variable for the number of accounts that are both in default and with disaster flags, e.g., $f^{12}_{i,t}$ denotes the number of accounts flagged defaults in the last twelve months. Comparing the marginal effects on the π coefficients when estimating Equation 5 using a logistic regression informs of the informativeness of this flagged default variable. My third model, "masked flagged defaults", shown in Equation 6, adjusts the input data to mask all defaults that also have disaster flags. The vector $\tilde{D}'_{i,t} = [\tilde{d}^{12}_{i,t}, \tilde{d}^{24}_{i,t}, \tilde{d}^{36}_{i,t}, \tilde{d}^{84}_{i,t}]'$ reclassifies such "flagged defaults" as not in default. Comparing the predictive performance, measured by AUC, of this third model with the baseline model shows the information costs of masking flagged defaults.

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_1 + D'_{i,t}\theta_1\Big) \tag{4}$$

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_2 + D'_{i,t}\theta_2 + F'_{i,t}\pi\Big) \tag{5}$$

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_3 + \tilde{D}'_{i,t}\theta_3\Big) \tag{6}$$

4.3 Predictive Results

Figure 6 shows the average marginal effects of an account in default, the θ_2 in black, and of a flagged account in default, the π in orange, coefficients from the logistic regression in

Equation 5, with estimates shown in Internet Appendix Table A2. The average marginal effects show that $\theta_2 > 0$, which means that an extra account in default in the past increases the risk of a consumer defaulting in the future. For flagged defaults in the last twelve months, the average marginal effects of $\pi > 0$, and are much greater than those for θ_2. There is some evidence that flagged defaults are economically and statistically significantly higher risk. The average marginal effects of an additional account in default in the last 84 months is 0.001266 (s.e. 0.000063) whereas the average marginal effects of an additional flagged account in default are substantially larger at 0.019841 (s.e. 0.003856), also in the last 84 months. The average marginal effects of an additional flagged account in default in the last 12, 24, and 36 months are all insignificant from zero, so the evidence is noisy. Masking flagged defaults reduces the number of defaults in the last 84 months by 0.03% (Table 3) with a 0.02% reduction in the number of consumers with any defaults in the last 84 months (Internet Appendix Table A7).

Table 4 compares the out-of-sample predictive performance from XGBoost models measured by AUC from the baseline credit risk model, 0.890236, to one masking flagged defaults, 0.890233. Masking flagged defaults therefore reduces predictive performance by a trivial amount 0.0003% (Table 4), whereas including flagged defaults as separate predictors increases predictive performance by 0.0751% (Table 2). Consistent conclusions are reached when examining a variety of alternative measures to evaluate predictive performance, accuracy, balanced accuracy, sensitivity, specificity, shown in Internet Appendix Tables A8 and A9. Such an economically small information cost of flags masking defaults can help to understand why lenders voluntarily temporarily apply disaster flags.

5 Consumer Benefits of Disaster Flags

What are the benefits to consumers of having disaster flags on their credit report? I study this using an event study design with the methodology explained in Section 5.1, and then show descriptive results in Section 5.2. I then use a difference-in-differences methodology, explained in Section 5.3, to study the effects of disaster flags on credit access, with results shown in Section 5.4.

5.1 Event Study Methodology

My event study methodology exploits the timing of disaster flags being applied being quasi-random as a function of the timing and geography of natural disasters. This descriptive methodology evaluates changes in a consumer's finances relative to their pre-disaster trend. I take the first time a consumer has a disaster flag applied to their credit report. I exclude consumers where the first time they have a disaster flag applied occurs only for a student loan, since these commonly contemporaneously have payments deferred. I keep cohorts of consumers with their first flags applied between January 2010 and December 2018 to ensure that I observe sufficient pre- and post-periods of each cohort without being affected by COVID-19 disruptions or contemporaneous deferrals that more commonly appear in more recent cohorts. I retain consumers with open tradelines with positive balances and credit scores observed twelve months before first being flagged as a group of active consumers. I also restrict to consumers who are aged 18 to 65. This produces a dataset of 2.8 million consumers representative of 28 million consumers. For all of these consumers, I construct a balanced panel of 25 months showing twelve months pre- and post-flags being first applied. This time window is driven by my earlier descriptive evidence that shows that disaster flags only remain on credit reports for a short period of time, and therefore any impacts of these flags are expected to be observed within twelve months. To assist with interpreting the event studies, I add linear time trends. Linear time trends are calculated from OLS regressions on data $t-12$ to $t-1$. Such linear trends may be a reasonable counterfactual over a short time horizon for how consumers' credit scores would have evolved without a disaster flag, see Dobbie et al. (2020) and Gross et al. (2020) for examples using similar approaches when studying the effects of bankruptcy flag removals.

I show event studies across flagged consumers and also by two sources of heterogeneity that measure pre-disaster financial distress. The first heterogeneous measure is whether a consumer's credit score twelve months before first being flagged was low-score "subprime" (300 to 600) or high-score "non-subprime" (601 to 850). 14.8% of the consumers in my sample are subprime. The second heterogeneous measure is whether a consumer has any defaults (30+ days past due) on open tradelines with positive outstanding balances on their credit report twelve months before being flagged. 6.3% of the consumers in my sample had any defaults by this measure. These two measures are highly correlated; 79% of consumers with any defaults have subprime credit scores, and 34% of consumers with subprime credit scores have any defaults. This heterogeneity is motivated by prior research showing that the effects of natural disasters on consumers' finances vary by pre-disaster financial distress (e.g., Billings et al., 2022). The second measure is also motivated by the institutional details of flags, where one would expect the potential gains from using disaster flags to be largest for those who already have defaults that could be masked by flags.

5.2 Event Study Results

Defaults. I examine the mechanism through which disaster flags can affect credit scores and credit access: masking defaults. Figure 7 shows the prevalence of defaults in credit reports of flagged consumers. The black line on Figure 7 Panel A shows the fraction of consumers with any defaults, before masking by the flags. This trend increases slightly over event time. The orange line masks defaults that occur on the tradeline months where the flags also appear. Flag masking immediately reduces the fraction of consumers with any defaults appearing in their credit report by 1.5 percentage points, however, defaults do not go to zero but remain at 5.7 percent, which is a 21 percent decline in defaults. These consumers have defaults on their other tradelines without flags, and these remain unmasked. The two default series quickly converge within twelve months, showing that any potential benefit of flags masking defaults is temporary.

These small average results are largely driven by consumers experiencing pre-disaster financial distress. Panels B and C of Figure 7 repeat this exercise for subsamples of the data by pre-disaster financial distress where $t-1$ values of defaults are normalized to zero. This shows that the temporary changes in defaults are concentrated among the minority of consumers who experience pre-disaster financial distress: subprime credit scores or those with any defaults. Flags initially mask defaults for 6.4 percent of pre-disaster subprime consumers, a 24 percent decline in defaults, and 8.2 percent of consumers with any pre-disaster defaults, a 23 percent decline in defaults. Masking of defaults is still only temporary for these groups. There is no discernible difference in defaults before or after masking among consumers without pre-disaster financial distress. This evidence indicates that any positive effects of flags on credit scores and credit access would be expected to be concentrated among consumers experiencing pre-disaster financial distress and only occur with a few months of the flag being first applied.

Credit Scores. How do credit scores change after disaster flags are applied? Figure 8 Panel A shows an average increase in credit scores of 2.9 points in the month the flag was applied, and of 3.0 points after twelve months. This is an increase of 0.4 percent relative to the $t-1$ baseline mean of 705.6 points and, after twelve months is as is predicted by a linear pre-trend. This average change in credit scores is too small to generate economically meaningful differences in credit access. This change is economically small relative to the average increase of approximately 15 points from removing bankruptcy flags in credit reports (e.g., Dobbie et al., 2020; Gross et al., 2020; Jansen et al., 2024). One might not expect the effects of disaster flags to be as large as the effect of removing bankruptcy flags for a couple of reasons. First, bankruptcy is one of the most negative signals about a consumer's creditworthiness, affecting highly financially distressed consumers who have high potential gains from these negative signals being removed to pool them with more creditworthy consumers. Second, the average positive effect of disaster flags is expected to be a dilution of a larger positive effect given that only a small subset of consumers have defaults, whereas all bankrupt consumers have bankruptcy flags (before their removal).

Figure 8 Panels B and C show how the most financially distressed consumers—with subprime credit scores or any defaults—receive the largest increases in their credit scores from disaster flags, with results for all credit score segments shown in Internet Appendix Figure A9. These panels normalize the credit scores for each subgroup relative to their baseline mean at $t-1$. Financially distressed consumers experience increases of 11 points for subprime consumers, relative to the baseline mean of 565, and 14 points for consumers with defaults, relative to the baseline mean of 576. Such estimates are similar in magnitude to those of bankruptcy flag removal, which, in turn, had real effects. The increases in credit scores from disaster flags appear to be short-lived, in contrast to the persistent effects after bankruptcy flag removal shown in prior literature (e.g., Dobbie et al., 2020; Gross et al., 2020; Jansen et al., 2024). Within twelve months credit scores become *lower* than that predicted by a linear pre-trend. Credit scores of consumers without pre-disaster financial distress show little meaningful changes: increasing by 2.0 points, compared to 0.7 predicted by a linear pre-trend, on a baseline of 737 points for the No Defaults group, and increasing by 1.5 points, compared to a 1 point decrease predicted by a linear pre-trend, by on a baseline of 730 points for the Non-Subprime group.

Credit Access. I do not find that disaster flags improve credit access for consumers who experience pre-disaster financial distress. This is despite these consumers receiving the largest boosts to their credit scores. I examine the extensive margin of credit access using the number of new credit card account openings. New account openings often have lags of several months before they are recorded in a consumer's credit report (Gibbs et al., 2024). To address this, I use the variable recording an account's opening date to create a new time series of the number of new credit card account openings, rather than the credit report archive date, and record zeros for months where no new credit card accounts are opened.

Figure 9 Panel A shows that the number of new credit card account openings decreases over time and there is no sign of improvement following the application of disaster flags. If anything, there are slight *decreases* in credit access with the number of new credit card account openings falling below their linear pre-trend. Figure 9 Panels B and C show similar conclusions for consumers experiencing pre-disaster financial distress. There is no sign of improved credit access, and there appears to be a slight reduction relative to a linear pre-trend. These results are robust to alternative measures of credit access: the total number of account openings (across auto loans, credit cards, mortgages, and unsecured personal loans), shown in Internet Appendix Figure A10, and the value of new credit card limits, shown in Internet Appendix Figure A11.

5.3 Difference-in-Differences Methodology

I find consistent results to my event study when I now use a difference-in-differences design to estimate the causal effects of disaster flags relative to a matched sample of unflagged consumers in the same geographic area, who therefore are also exposed to the disaster. I estimate the causal effects of adding disaster flags to a credit report on consumers using a stacked difference-in-differences empirical design used in prior work (Cengiz et al., 2019; Deshpande and Li, 2019; Jansen et al., 2024; Cherry, 2024). Dube et al. (2024) show that a stacked approach is equivalent to using a local projections estimator, and it corrects for the potential bias of negative weighting that arises in designs with staggered, heterogeneous, or dynamic treatments that is shown in Baker et al. (2022). This stacked difference-in-differences approach differences out the contemporaneous effects of the natural disaster to leave only the effects of the disaster flag. This methodology exploits the timing of disaster flags being applied as a quasi-random function of natural disasters.

I keep consumers who first received a disaster flag between January 2010 and December 2018. This stacked difference-in-differences empirical design stacks data from each flag event study and, for each event, constructs a clean control group of "unflagged" consumers that have never been flagged between July 2000 and December 2019. The control group is constructed using variables calculated twelve months prior to the date the flagged group is first flagged. The clean control of unflagged consumers is in the same combination of geographic area—the same census block group $\times$ zip code—, credit score group, any defaults, and any mortgage debt, to the flagged consumer, and also aged 18 to 65. I drop cases where either all or no consumers in that combination are flagged. Within the combinations where flagged consumers with potential controls exist, I choose potential controls by the nearest neighbor in Euclidian distance by standardized values of credit score, credit card limit, number of trades, and outstanding balances, and keeping cases where there are close matches, using a distance less than or equal to one. This leaves 57% of consumers in the event study dataset.

In this dataset, each flagged consumer is matched with an unflagged consumer to produce a total dataset of 3.2 million consumers representative of 32 million consumers. For each of these consumers, I take twelve months of observations before and twelve months after the flagged event to create a balanced panel of observations: 25 months per consumer stacked into a single dataset. This ensures that my results are not driven by compositional changes. I then aggregate the data to the group-cohort-event-time level. Group, g, groups consumers by the calendar year-month when they are first flagged, and their matched control. Cohort, c, is whether or not a consumer is flagged. Event time, t, ranges from -12 to $+12$.

I estimate the regression shown in Equation 7. This regression includes fixed effects for each group-by-event-time ($\gamma_{g,t}$) and for each cohort-group ($\gamma_{g,c}$). The term $FLAG_c$ is an indicator that takes a value of one for flagged cohorts, and a value of zero if a consumer is in the unflagged control cohorts. Standard errors are clustered at the group-level. In my regressions, I weight data based on the number of consumers, and doing so gives equal weight to the trends for the controls as the treatments to avoid the issues discussed in Wing et al. (2024).

$$Y_{g,c,t} = \sum_{\tau \neq -1} \delta_\tau \Big(FLAG_c \times D_t^\tau \Big) + \gamma_{g,c} + \gamma_{g,t} + \varepsilon_{g,c,t} \tag{7}$$

The parameters of interest are δ_τ, which are the interaction on the event time dummies, D_t^τ, and the indicator for the flagged group, $FLAG_c$. Under the assumption of common trends, δ_τ estimates the effect of disaster flags, among those selected in and with suitable controls, on outcomes, $Y_{g,c,t}$, after τ months.

5.4 Difference-in-Differences Results

Credit Scores. My difference-in-differences results for all consumers, and by subsamples of pre-disaster financial distress, are shown in Table 5 for the effects in the month of flagging and in Table 6 for the effects after twelve months.[11] In line with my event study results, my difference-in-differences results show that disaster flags cause a 2.82 points significant average increase to Vantagescore credit score in the month they are applied, with a standard error of 0.22 points, relative to a baseline mean of 704, (Table 5) that dissipate to being insignificant from zero within six months (Internet Appendix Figure A12). After twelve months, the effects are not significantly different from zero, with an estimate of -0.36 and a standard error of 0.35 as shown in Table 6.

The average temporary positive effect of credit scores is driven by consumers experiencing pre-disaster financial distress. 4.4% of the consumers in my sample, 0.14 million consumers, experience any pre-disaster defaults at $t-12$. Figure 10 Panel A shows that at $t = 0$, consumers with any pre-disaster defaults experience a significant average increase in credit score of 14.93 points with a standard error of 1.59 points, 2.7% relative to the baseline mean of 547. As credit scores are non-linear predictors of default risk, a 15 points increase for a consumer with a score of 550 is expected to be more valuable than a similarly sized increase for consumers with higher scores. Consumers without pre-disaster defaults also experience a significant increase of 2.27 points, with a standard error of 0.17 points, which is 0.3% of the baseline mean of 711. The effects for both groups dissipate

[11] Internet Appendix Tables A4 and A5 show additional results by all pre-disaster credit score groups.

within twelve months, making them insignificant from zero. After twelve months, the estimated effects for those with any pre-disaster defaults are -0.29, with a standard error of 0.83, and for those without pre-disaster defaults is -0.36, with a standard error of 0.33. Internet Appendix Figure A14 Panel A shows the results by all credit score groups. The positive effects are concentrated among the 16.2% consumers with subprime credit scores. For these subprime consumers, the effects peak at $t = 0$ at 10.18 points with a standard error of 0.92 points, an increase of 1.8% relative to baseline mean credit score of 560. These increases among subprime consumers are short-lived and turn significantly negative within twelve months.

Credit Access. I find no effects of flags increasing credit access. I find no positive effect of flags increasing credit access even for the consumers who received the largest temporary boosts to their credit scores: those with pre-disaster defaults. There is no significant effect on whether a consumer opened any new credit card account each month. Figure 10 Panel B shows credit access declines for both consumers with and without pre-disaster defaults. After twelve months the estimated effects for those with any pre-disaster defaults are slightly negative -0.0036, with a standard error of 0.0022, shown in Table 6, which is a decrease of 10% relative to the baseline mean of 0.0346. As also shown in Table 6, the effect for those without pre-disaster defaults is also significantly negative -0.0175, with a standard error of 0.0021, which is a decrease of 27% relative to the baseline mean of 0.0646. These results are not specific to new credit cards. Examining the number of new accounts opened across credit types shows the same pattern of results, see Internet Appendix Figure A15.

There does not appear to be an effect on the intensive margin. Figure 10 Panel B shows no positive effect on the value of new credit card limits. On this margin, consumers with pre-disaster defaults do not experience a significant difference after 12 months, with an estimate of \$6.3 and a standard error of \$9.3, as shown in Table 6, relative to their $t-1$ mean of \$64.6. Whereas consumers without pre-disaster defaults experience a significant decrease of 24%, with an estimate of $-$\$108.1 and a standard error of \$15.8, also shown in Table 6, relative to their $t-1$ mean of \$450.4. These results on credit access are consistent with segmenting by credit score (Internet Appendix Figures A13 and A14) and when averaging across all consumers (Internet Appendix Figure A12).

Our findings provide a useful example of how credit score increases do not necessarily translate into improved credit access, a point raised in Gibbs et al. (2024). Disaster flags may not improve credit access because credit score increases are not for a long enough period of time for consumers to realize the potential benefits. Furthermore, although disaster flags temporarily affect VantageScore credit scores, that are observed in my data, disaster flags do not affect FICO credit scores, that are unobserved in my data (and in other credit reporting datasets used by researchers Gibbs et al., 2024), and therefore credit decisions taken using FICO scores would be unaffected.[12] This motivates considering a counterfactual system to more effectively preserve the credit access of consumers affected by disasters.

6 Loss of Information From Masking All Disaster Defaults

6.1 Methodology

The existing voluntary regime of disaster flags appears to have limited costs to lenders but does not appear to benefit consumer credit access. What are the feasible alternatives? I quantify the loss of information from a counterfactual government policy that automatically requires the masking of all defaults during disasters ("disaster defaults"). This counterfactual are designed to apply equitably to all consumers subject to a disaster: removing selection. Doing so pools consumers affected by disasters with those unaffected by disasters.[13] Applying such an automatic flagging approach removes frictions for lenders and consumers. Such a counterfactual policy has been proposed by consumer organizations (e.g., National Consumer Law Center, 2019; Urban Institute, 2019; FinRegLab, 2020), however, the costs of removing such information from lenders have not been estimated.

[12] Credit cards are a domain where any positive effects of disaster flags on credit access are most likely to be visible, as VantageScore is sometimes used for credit cards, whereas FICO is predominantly used for mortgages.

[13] Practically implementing this would mean disaster defaults are masked *before* they appear on credit reports so that disaster defaults cannot be observed by lenders in their credit decisions. This could be required of firms furnishing data to credit reporting agencies or could be required of credit reporting agencies before they release data to be used by lenders. Consumer credit reporting data is an important source of information for regulators, and therefore there is potentially value for regulators observing disaster defaults, as enabled by flagging disaster defaults rather than non-reporting disaster defaults, even if these are masked from lenders, to evaluate the impacts of disasters on consumers and lenders, and to monitor such a policy's effectiveness.

If such a counterfactual policy was required by law, it would affect the underlying credit reporting data which all credit scores (e.g., FICO, VantageScore) and manual underwriters rely on and, therefore, would be expected to have downstream impacts on consumer credit access.[14] Although I do not estimate supply responses, previous research in Cortés and Strahan (2017) that studies supply responses to natural disasters may be indicative of lenders being willing to meet increased local credit demand, while Blickle et al. (2022) shows that disasters increase loan demand, which offsets losses and actually boosts profits at larger banks. A larger predictive loss from masking disaster defaults indicates that lenders would be expected to be more likely to restrict credit supply, whereas lenders may be expected to be more able to absorb a small predictive loss.

I evaluate this policy using an analogous approach to Section 4.2, with the same dataset and logistic and XGBoost methods, with the only difference being that instead of studying defaults with disaster flags, I now vary the inclusion of information on all defaults that occur in areas and time periods exposed to natural disasters. I examine the marginal effects of a vector of "disaster defaults", $N'^{j}_{i,t} = [n^{12,j}_{i,t}, n^{24,j}_{i,t}, n^{36,j}_{i,t}, n^{84,j}_{i,t}]'$, in the regression model specified in Equation 8. This regression uses the same vectors of default, $D'_{i,t} = [d^{12}_{i,t}, d^{24}_{i,t}, d^{36}_{i,t}, d^{84}_{i,t}]'$, and non-default variables, $X'_{i,t}$, as the earlier Equation 4. Disaster defaults are constructed from the interactions between defaults and natural disasters, over the same time horizons as $D_{i,t}$ as previously studied in Section 4.2. For constructing $N'^{j}_{i,t}$, I first construct a binary variable taking a value of one only if a consumer, i, has any defaults in an archive month and the consumer resides in a county, g, where a FEMA natural disaster was declared in the j months to that archive month, and zero otherwise. Then from this I can construct the consumer level variables in vector $N'^{j}_{i,t}$ for the number of disaster default accounts. I vary j to be three, six, and twelve months to examine how sensitive the results are to different policy thresholds for defining disaster defaults. This means that, for example, $n^{12,03}_{i,t}$ denotes the number of accounts that are both in default and exposed to a natural disaster in the last twelve months, where disaster defaults are classified as such if they occur within three months of a FEMA event. The coefficients on $N'^{j}_{i,t}$ in a logistic model inform about whether disaster defaults are different from other defaults at being informative of a consumer's credit risk.

[14]Collier et al. (2024b) and Collier et al. (2024a) show the benefits of emergency liquidity to households and lenders, and Collier and Ellis (2024) estimates consumer demand

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_4 + D'_{i,t}\theta_4 + N'^{j}_{i,t}\phi\Big) \tag{8}$$

Choosing a j threshold of six months would ensure that any defaults during that six-month period only affect subsequent credit access if they are still present after six months. This limits the ability of temporary adverse shocks to propagate and have long-term impacts. Three to twelve months are studied as alternative thresholds. Such thresholds provide time for consumers to be able to apply for and receive federal social insurance and disaster aid, as well as to contact their creditors to adjust payments, if required. These are also short enough durations to limit the potential moral hazard of strategically defaulting consumers, which may be more of a concern with a twelve-month threshold and less of a concern with a three-month threshold, and to prevent consumers in high risk locations repeatedly exposed to disasters having their defaults perennially masked.[15]

Analogous to my earlier masking of flagged defaults in Equation 6, masking disaster defaults takes the form shown in Equation 9, where the term $\tilde{D}'^{j}_{i,t}$ masks disaster defaults, based on the policy threshold j, which masks defaults based on whether exposed to a disaster in the one of three, six, or twelve months from a FEMA event. I examine how masking disaster defaults affects predictive performance using my XGBoost machine learning approach and varying the masking of disaster defaults. I benchmark the models' performance with masked defaults to the baseline model with all defaults, shown in Equation 4, and also to a model without defaults information, displayed in Equation 10.

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_5 + \tilde{D}'^{j}_{i,t}\theta_5\Big) \tag{9}$$

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_6\Big) \tag{10}$$

[15]Recent empirical evidence potentially indicates that moral hazard may be less of a concern driving defaults than life events (e.g., Ganong and Noel, 2023; Low, 2023), and non-financial concerns also motivate consumers to repay their debts (e.g., Guiso et al., 2013; Bursztyn et al., 2019; Martínez-Marquina and Shi, 2024), along with non-credit reporting financial incentives to repay debt on time, such as the increased borrowing costs, and the risk that late or non-payment reduces a consumer's ability to borrow from that lender in the future when they may need it.

6.2 Results

Figure 6 Panel B shows the average marginal effects of an additional trade in default from my logistic regressions, with estimates also shown in Internet Appendix Table A2. The marginal effects of disaster defaults are generally lower than those on all defaults. The disaster defaults terms are generally insignificant from zero, although they are sometimes significantly negative, most notably the number of disaster defaults in the last 36 months, when using any of three, six, or twelve months as j thresholds for capturing defaults around disasters. Consistent with this is the uplift in predictive performance, measured by AUC from the XGBoost machine learning versions of these models, is 0.08% as shown in Table 2. I interpret this as recent disaster defaults appear to have limited information of a consumer's future risk of default beyond non-disaster defaults.

What happens when I mask disaster defaults? Rows three to five of Table 3 shows that using j thresholds of three to six to twelve months for masking disaster defaults reduces the mean number of accounts with defaults in the last 84 months by 2.38% to 5.16% to 14.03% respectively, and reduces the number of consumers with any defaults in the last 84 months by 1.5% to 3.1% to 8.1% (Internet Appendix Table A7). These are economically large removals of data from credit reports.

Figure 11 shows the ROC curves for masking disaster defaults ("Masked Disaster Defaults") under different duration thresholds. It is difficult to distinguish the different lines, though there is a clear gap to the no defaults model shown in orange, showing how some defaults information is clearly valuable for prediction. Table 4 shows how much predictive performance, measured by AUC, declines as more information is masked. The baseline model has an AUC of 0.890236, which decreases to 0.889206, a 0.12% decrease on the baseline, with a three month policy threshold for masking disaster defaults. It decreases to 0.888825 (-0.16%) with a six month threshold, and to 0.887044 (-0.36%) with a twelve month threshold. These effects of masking disaster defaults can be benchmarked against a counterfactual policy masking all defaults, i.e., disaster and non-disaster defaults, that significantly reduces predictive performance with AUC declining to 0.880902, a 1.05% on baseline. My pattern of results are consistent when examining a variety of alternative measures to evaluate predictive performance (accuracy, balanced accuracy, sensitivity, specificity) as shown in Internet Appendix Table A9.

These reductions in predictive performance may appear small relative to the amount of information masked, and therefore policymakers may consider automatically masking disaster defaults to be a proportionate policy to help provide a temporary source of credit reporting relief for consumers from natural disasters, however, it would worsen lenders' credit risk assessments. One way to help benchmark the size of these AUC changes is to compare them to the gains of moving from logistic models to machine learning models. Fuster et al. (2022) show that the ability to predict mortgage default, measured by AUC, increases from $0.8486 - 0.8537$, in logistic models, to 0.8602 using a random forest model, which are improvements of 0.8% to 1.3%. While Blattner and Nelson (2024) find that moving from a commercial credit score to XGBoost algorithm increases AUC predicting non-mortgage default from 0.835 to 0.879, an improvement of 5.0%.

An alternative way to implement a policy masking disaster defaults could be to adapt the approach used in the in the US by the Coronavirus Aid, Relief, and Economic Security (CARES) Act of 2020, only masking defaults if the default status worsens during the disaster. This "CARES Masked Disaster Defaults" approach means that accounts that were already in default pre-disaster remain in default, whereas in the "Masked Disaster Defaults" approach they were temporarily masked. Rows six to eight of Table 3 (and Internet Appendix Table A7) show that this masks slightly less information, reducing mean defaults (relative to our all defaults baseline) by 2.33%, 4.88%, and 11.40% under three, six, and twelve month thresholds, than our earlier "Masked Disaster Defaults" approach. The changes in AUC predictive performance are similar, -0.12%, -0.16%, -0.32% based on three, six, and twelve month thresholds, as shown in rows six to eight of Table 4 (ROC curves and alternative performance measures are in Internet Appendix Figure A16 and Table A9).

However, the above approaches only temporarily mask defaults, masking defaults if they are present within j periods of a FEMA disaster. However, defaults that arise during a disaster may persist beyond j periods in my historical data, but may not be under a counterfactual policy. I examine how "permanently" masking defaults that arise during natural disasters would affect prediction in my "PCARES Masked Disaster Defaults" approach. In this approach, I take each account's default status in the month preceding a disaster, and mask the defaults on the accounts if their account status worsens during the disaster threshold, and, unlike the CARES approach, I keep masking the default status on account months after the disaster.[16] I leave default statuses of accounts that are opened after a disaster threshold unaffected, and the default status of accounts whose status did not worsen during a disaster, and, of course, the default status of consumers in areas unaffected by disasters remains unchanged. A wider threshold means that more new delinquencies would be masked, and their pre-disaster default status may be earlier due to the consumer experiencing repeated disasters.

This "PCARES Masked Disaster Defaults" approach may be considered an upper bound on the amount of information that a policy masking disaster defaults may mask. Rows 9 to 11 of Table 3 show that this masks 10.3%, 18.7%, and 32.6% of defaults, depending on whether three, six, or twelve month thresholds are used, and 4%, 8%, and 17% fewer consumers have any defaults in the last seven years (Internet Appendix Table A7). Rows 9 to 11 of Table A9 show that the predictive performance worsens by 0.22% to 0.32% to 0.51% as three to six to twelve months of information is removed (ROC curves and alternative performance measures in Internet Appendix Figure A17 and Table A9). I quantify the trade-off that policymakers face, masking defaults for a longer period of time increasingly reduces predictive performance, and the costs depend upon how defaults that initially occur during disasters are recorded in credit reports after the disasters.

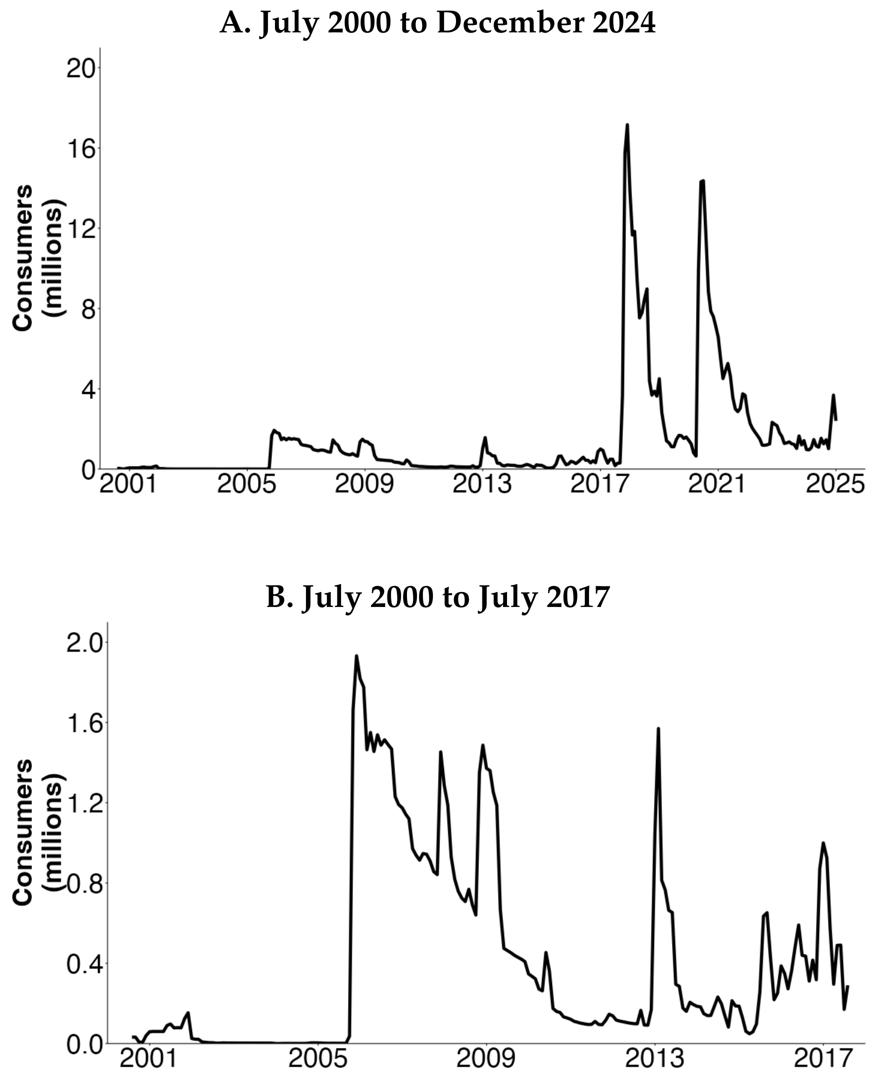
7 Conclusions

This research provides new facts documenting the widespread use of "disaster flags" on US credit reports. Disaster flags are intended to provide relief to consumers affected by natural disasters. I show that these flags are widely used, however, they do not increase credit access. I consider the trade-off of a counterfactual policy that automatically masks all defaults of consumers exposed to natural disasters. Doing so removes 2% to 33% of all defaults from credit reports, at the cost of reducing predictive performance by 0.1% to 0.5%. More generous policies trade-off greater predictive losses. My quantification of predictive losses can help inform policy discussions on the design of credit information markets and how to alleviate financial distress from natural disasters.

[16] For example, an account not in default at $t = -1$, and in default at both $t = 0$ and $t = +1$, where $t = 0$ is the time of a disaster. A one-month CARES policy, and also a "Masked Disaster" policy, would mask defaults at $t = 0$ but not $t + 1$, whereas the PCARES policy would mask defaults at both $t = 0$ and $t = +1$.

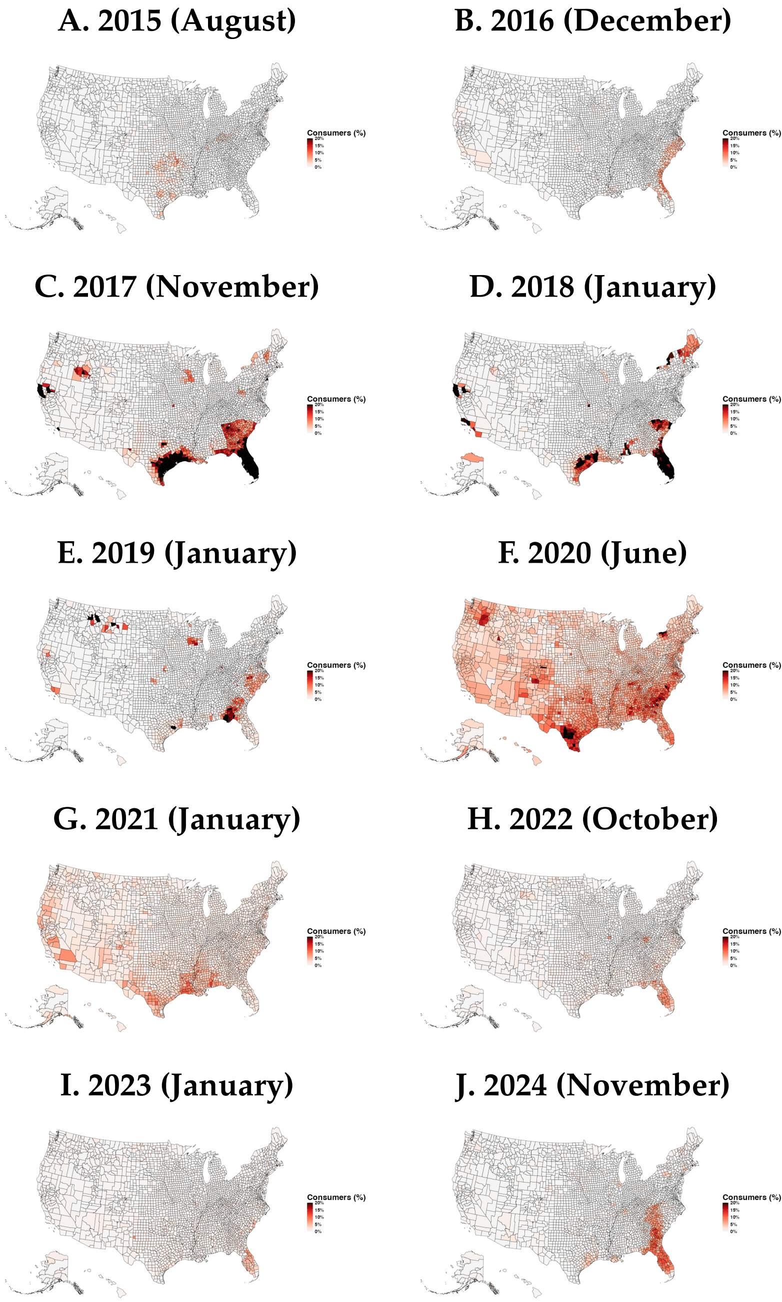
8 Figures and Tables

Figure 1: Consumers with Any Credit Report Disaster Flag



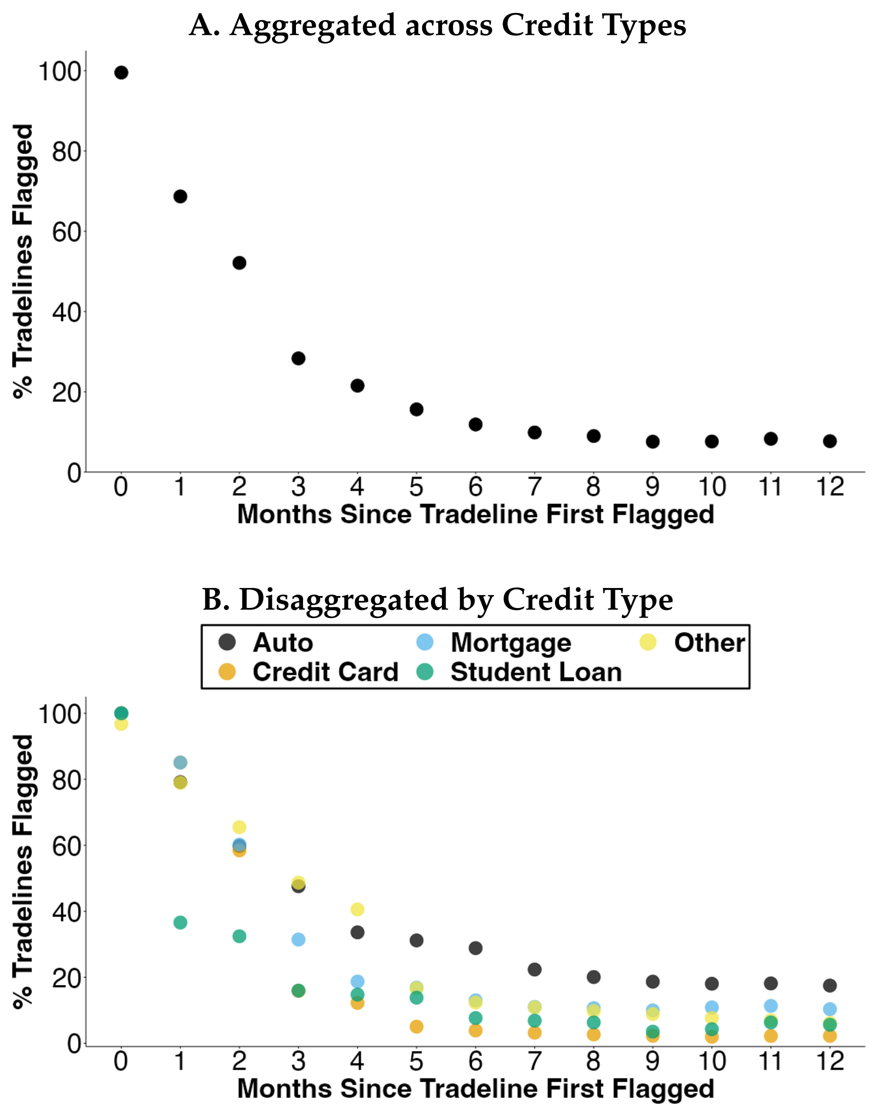
Notes: BTCCP data. Consumers with a credit report disaster flag on at least one open tradeline in their credit report. The number of consumers is extrapolated to population estimates from the BTCCP data's 10% sample of consumers.

Figure 2: Consumers (%) in a County with Any Disaster Flag



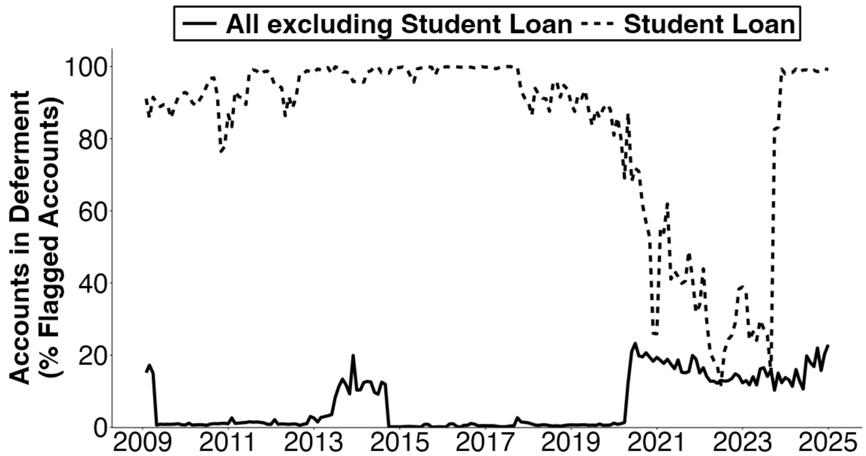
Notes: BTCCP data. Each panel shows the percentage of consumers in a county that have any disaster flag. The denominator in this calculation is the number of consumers with an open tradeline with a positive balance on their credit report in a county that month. The numerator in this calculation is the subset of these consumers with a credit report disaster flag on at least one of these tradelines that month. The values in each county are top-coded at 20% to ease presentation. The months shown are the month with the highest number of consumers with disaster flags in each of the years from 2015 to 2024.

Figure 3: Persistence of Disaster Flags on Credit Report Tradelines



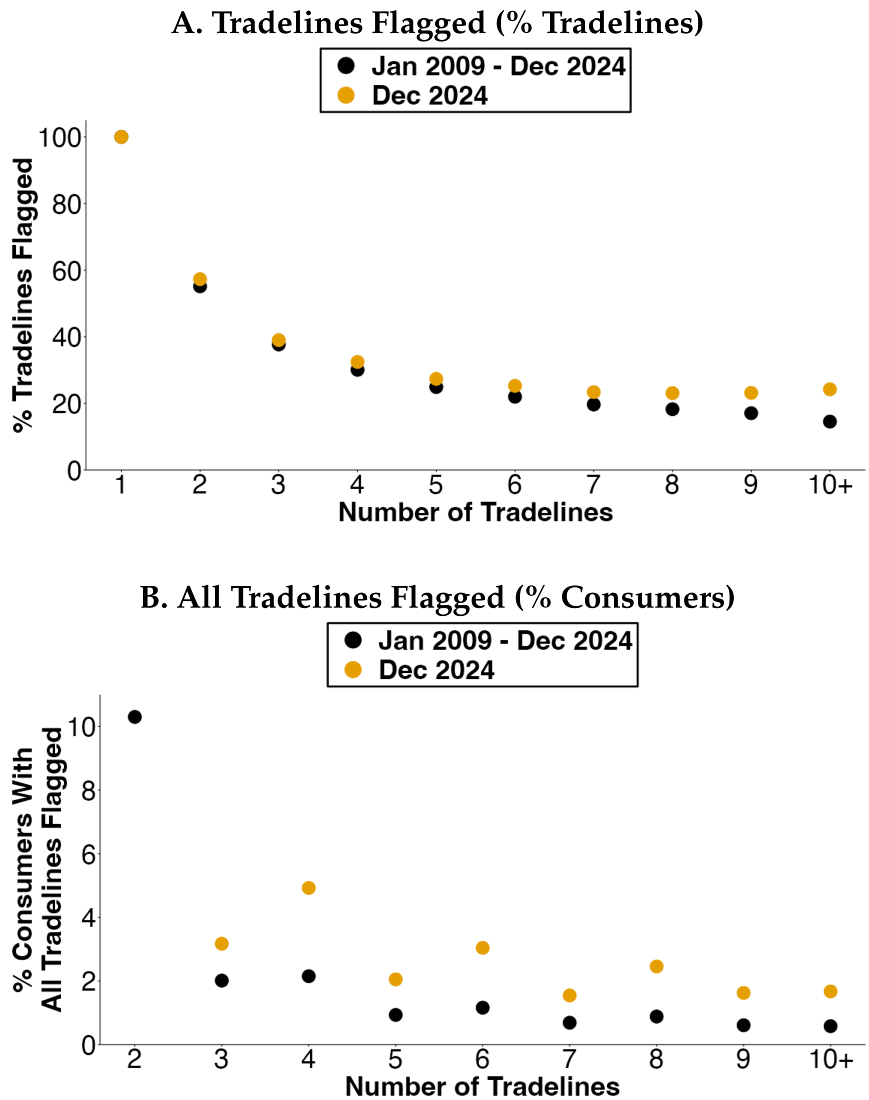
Notes: BTCCP data. These figures take open credit report tradelines with positive balance that first have a disaster flag added between January 2010 and December 2023. These figures plot the fraction of these tradelines that still have disaster flags present 1 to 12 months later. Panel A shows this aggregated across tradelines of different credit types. Panel B disaggregates tradelines by their credit types, where the 'other' category contains retail cards and unsecured loans.

Figure 4: Tradelines with Credit Report Disaster Flags and Deferments Reported



Notes: BTCCP data. This figure shows the fraction of accounts that have a disaster flag that also have a deferment reported. The denominator in this calculation is the number of open tradelines with a positive balance in their credit report with a credit report disaster flag. The numerator in this calculation is the subset of these accounts that also have deferments reported. Deferments are tradelines where deferments are listed on the account or the tradeline has a positive balances but zero payments due. The solid line excludes student loans from both the numerator and denominator of this calculation. The dashed line shows this calculation for student loans.

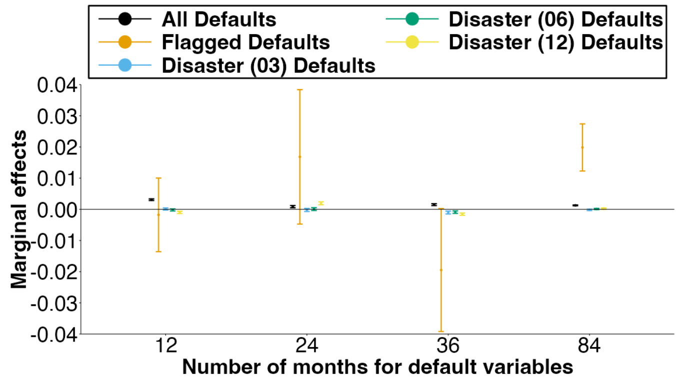
Figure 5: Disaster Flags by Total Number of Consumer Tradelines



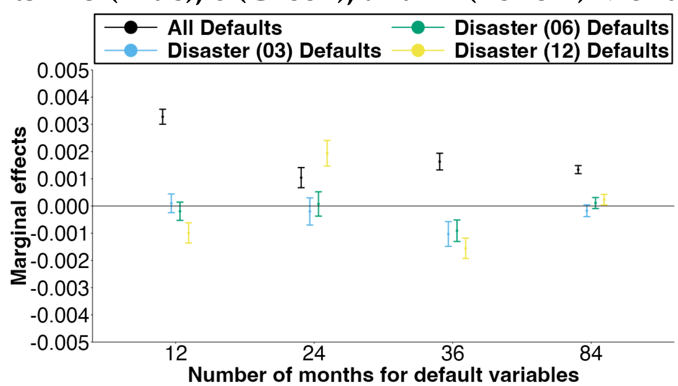
Notes: BTCCP data. Both Panels A and B restrict to consumer months where the consumer has a credit report disaster flag on at least one open tradeline with a positive balance in their credit report. Panel A shows the fraction of these consumers' tradelines that have a credit report disaster flag. Panel B shows the fraction of consumers where all of their tradelines have credit report disaster flags reported. The x axes on both panels plot the number of open trades with a positive balance a consumer has on their credit report. Statistics shown combining observations January 2009 to December 2024 (black) and also for December 2024 only (orange).

Figure 6: Average Marginal Effects of Past Defaults Predicting Any New Default

A. All Defaults (Black), Flagged Defaults (Orange),
Disaster Defaults in 3 (Blue), 6 (Green), and 12 (Yellow) Months of a Disaster

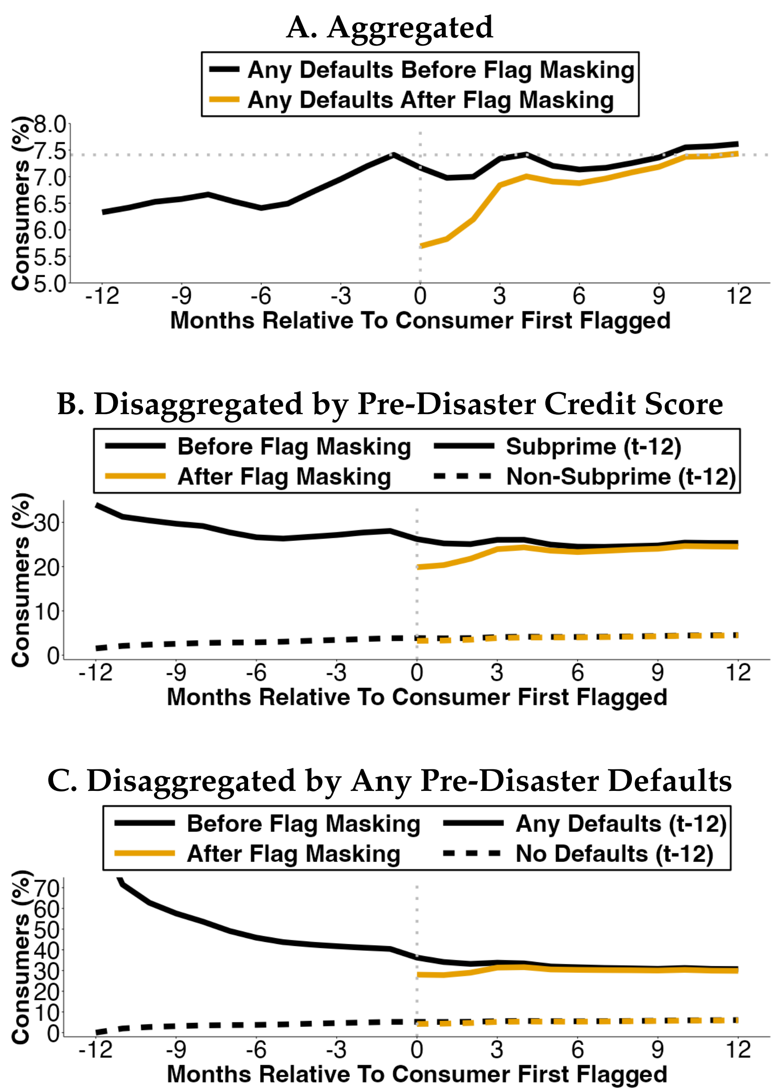


B. All Defaults (Black),
Disaster Defaults in 3 (Blue), 6 (Green), and 12 (Yellow) Months of a Disaster



Notes: BTCCP data. The X axes show the predictive variables for defaults in the last 12, 24, 36, and 84 months respectively. The Y axes shows the average marginal effects from coefficients from logistic regressions predicting any new defaults in the next 24 months. The dots show the average marginal effects and the lines show their 95% confidence intervals. The black dots in Panel A are the marginal effects from the θ_2 *coefficients on default terms* ($D'_{i,t}$) *from Equation 5 in Panel A (and the* θ_4 *coefficients on default terms* ($D'_{i,t}$) *from Equation 5 in Panel B), and the orange dots in Panel A are the marginal effects from the* π *coefficients on flagged default terms* ($F'_{i,t}$) *from Equation 5, and in both panels the blue, green, and yellow dots are the marginal effects from the* ϕ *coefficients on disaster default terms* ($N'_{i,t}$) *from Equation 8, using different definitions of disaster defaults: blue captures defaults within 3 months of a disaster, green within 6 months, and yellow within 12 months. Marginal effects are calculated using data on 6.99 million consumers (the 30% testing sample) to October 2017. The default terms used as predictors measure the number of accounts that a consumer is 30 or more days past due over varying durations. The logistic regressions use a control variable of a credit score that is built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs.*

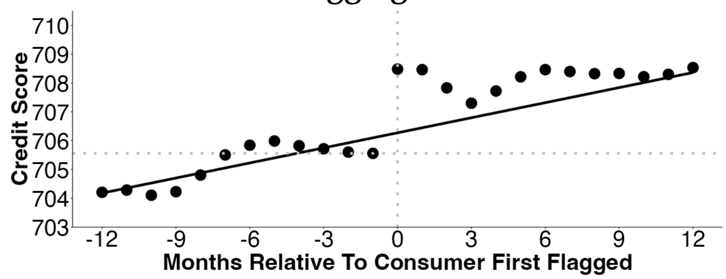
Figure 7: Event Study of Consumers (%) with Any Defaults on Credit Reports before (Black) and After (Orange) Masking by Disaster Flags



Notes: BTCCP data. Unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. The X axes show the number of months since a consumer first has a disaster flag. The Y axes show the fraction of consumers with any defaults, measured as 30 or more days past due, before (black) and after (orange) tradeline months where defaults masked by flags. Panel A shows this for flagged consumers and Panels B and C disaggregates this by measures of pre-disaster financial distress. Panel B splits by whether a consumer has subprime credit score (300 - 600) twelve months prior to first being flagged ($t-12$)*. In Panel B, 14.8% of consumers have subprime credit scores twelve months prior to first being flagged. Panel C splits by whether a consumer has any defaults twelve months prior to first being flagged* ($t-12$)*. In Panel C, 6.3% of consumers have any defaults twelve months prior to first being flagged.*

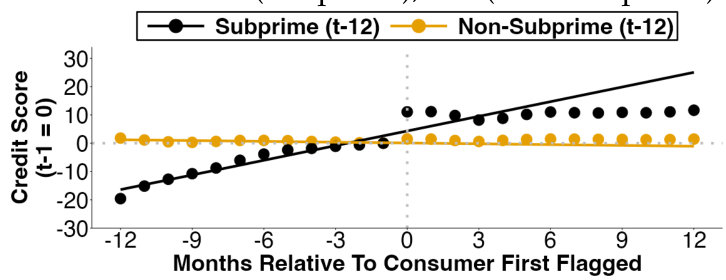
Figure 8: Event Study of VantageScore Credit Scores Relative to Linear Pre-Trends

A. Aggregated



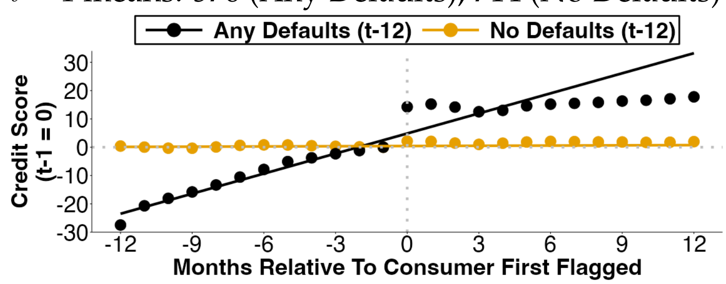
B. Disaggregated by Pre-Disaster Credit Score

$t-1$ means: 565 (Subprime), 730 (Non-Subprime)



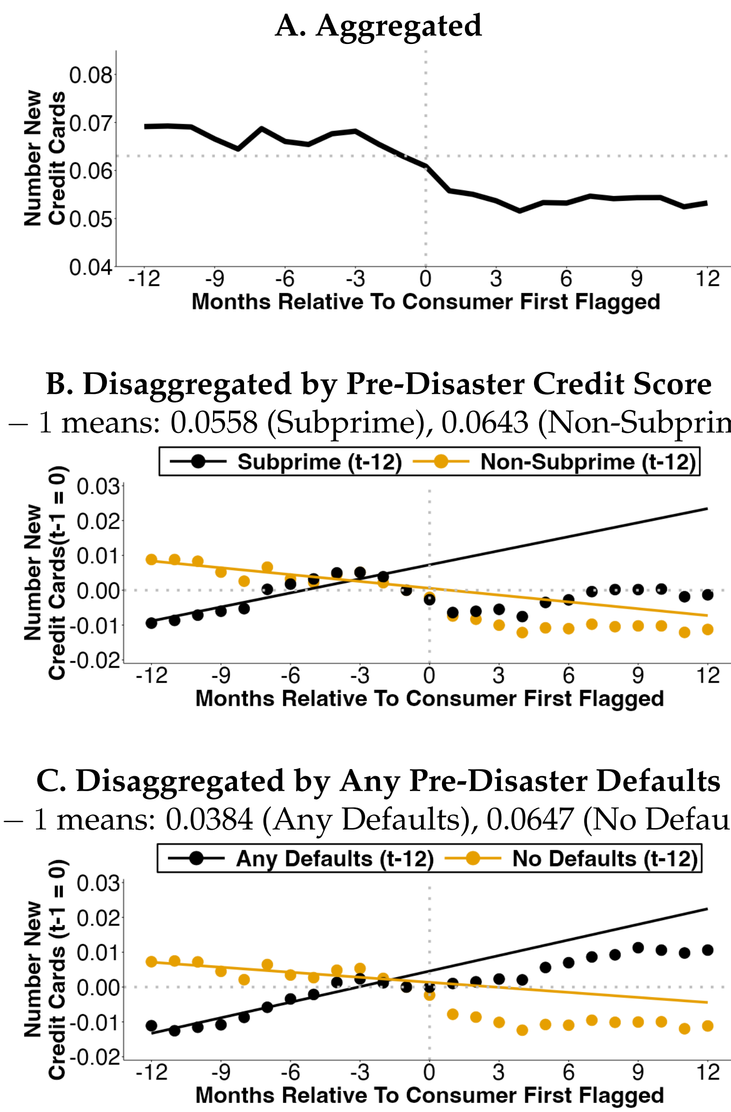
C. Disaggregated by Any Pre-Disaster Defaults

$t-1$ means: 576 (Any Defaults), 714 (No Defaults)



Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. The solid lines are linear pre-trends from OLS regressions on data $t-12$ to $t-1$. The X axes show the number of months since a consumer is first flagged. The Y axes show VantageScore credit scores. Panel A shows this for flagged consumers and Panels B and C disaggregates this by measures of pre-disaster financial distress. Panel B splits by whether a consumer has subprime credit score (300 - 600) twelve months prior to first being flagged ($t-12$). In Panel B, 14.8% of consumers have subprime credit scores twelve months prior to first being flagged. Panel C splits by whether a consumer has any defaults twelve months prior to first being flagged ($t-12$). In Panel C, 6.3% of consumers have any defaults twelve months prior to first being flagged. Panels B and C normalize credit scores for each subgroup to $t-1=0$.

Figure 9: Event Study of New Credit Card Account Openings Relative to Linear Pre-Trend

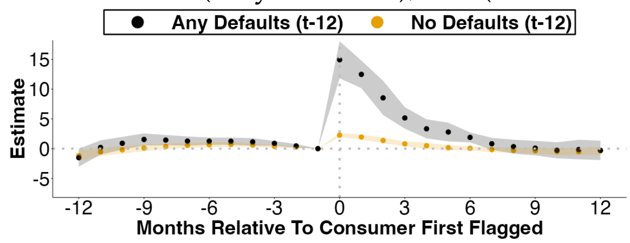


Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied between January 2010 and December 2018. The solid lines are linear pre-trends from OLS regressions on data $t - 12$ to $t - 1$. The X axes show the number of months since a consumer is first flagged. The Y axes show the new credit card account openings. Panel A shows this for flagged consumers and Panels B and C disaggregates this by measures of pre-disaster financial distress. Panel B splits by whether a consumer has subprime credit score (300 - 600) twelve months prior to first being flagged ($t - 12$). In Panel B, 14.8% of consumers have subprime credit scores twelve months prior to first being flagged. Panel C splits by whether a consumer has any defaults twelve months prior to first being flagged ($t - 12$). In Panel C, 6.3% of consumers have any defaults twelve months prior to first being flagged. Panels B and C normalize new account openings for each subgroup to $t - 1 = 0$.

Figure 10: Difference-in-Differences Estimates of Heterogeneous Treatment Effects of Disaster Flags on Credit Scores and Credit Openings, By Any Pre-Disaster Defaults

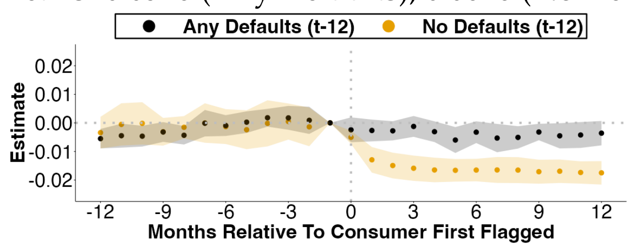
A. Credit Score

$t-1$ means: 547 (Any Defaults), 711 (No Defaults)



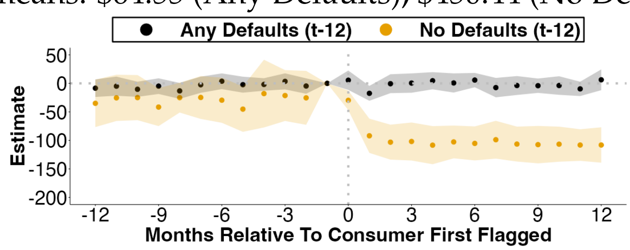
B. Number New Credit Cards

$t-1$ means: 0.0346 (Any Defaults), 0.0646 (No Defaults)



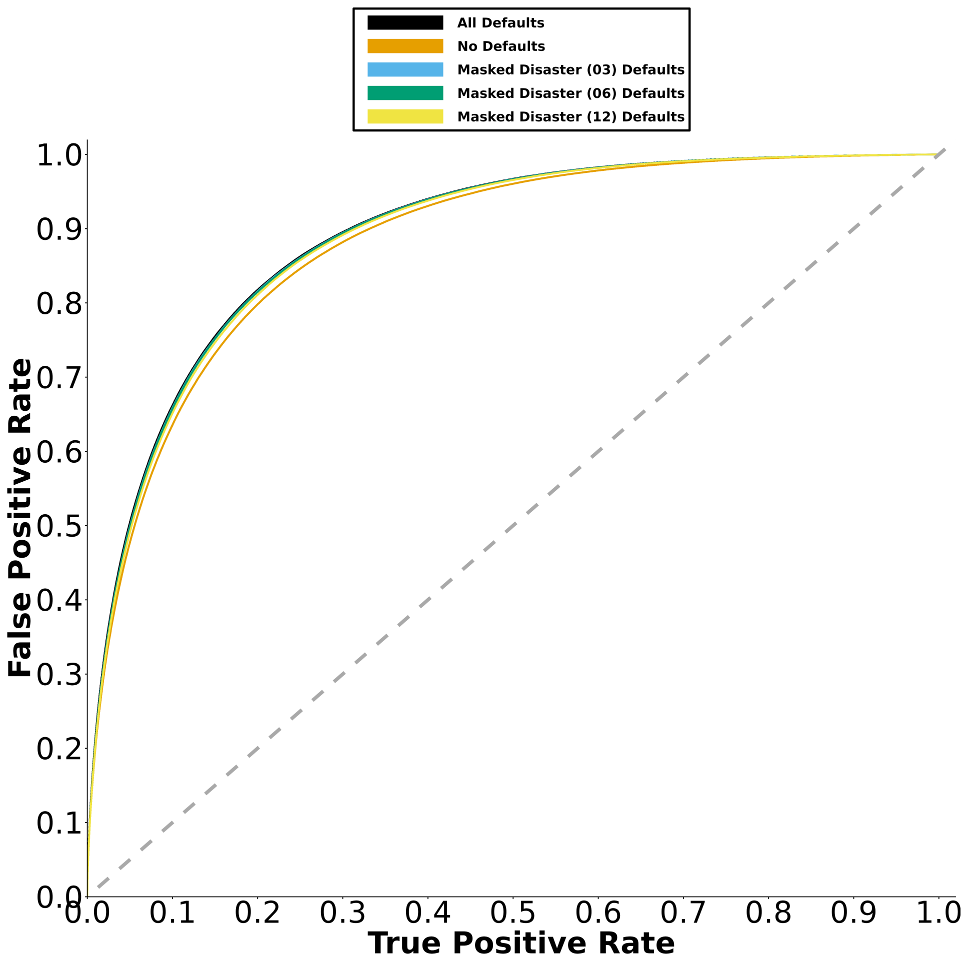
C. Value New Credit Card Limits

$t-1$ means: \$64.55 (Any Defaults), \$450.44 (No Defaults)



Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 7's δ_τ. δ_τ are the coefficients on the interaction between event time indicators after τ periods and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each panel shows results for different outcomes: A. credit score (VantageScore 3.0), B. the number of new credit cards opened that month, and C. the value of new credit card limits opened that month. Within each panel, different colors denote separate regressions for consumers based on whether they had any pre-disaster defaults at event time $t-12$. The black estimates use data for the 4.4% (0.14 million) subsample of consumers who experience any pre-disaster defaults ($t-12$), and the orange estimates use data for the 95.6% (3.06 million) subsample without any pre-disaster defaults.

Figure 11: Receiver Operating Characteristic Curves (ROC) Showing the Predictive Performance of Models Masking Default Information. Baseline Model Includes Information on Defaults (Black), Compared to Models Masking Defaults over Three (Blue), Six (Green), and Twelve (Yellow) Months of a Disaster, and a Model Including No Defaults Information (Orange)



Notes: BTCCP data. Receiver Operating Characteristic Curves (ROC) from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. The black line is the baseline model that includes defaults information the prediction. The orange line respectively show performance from a model without defaults information. The blue, green, and yellow lines show performance from models with defaults information where defaults that occur within three (blue), six (green), and twelve (yellow) months of a disaster are masked. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables. True Positive Rate is defined as the fraction of predicted new defaulters out of actual new defaulters. False Positive Rate is defined as the fraction of predicted new defaulters out of actual no new defaulters.

Table 1: Summarizing Consumers with (1) Disaster Flags Compared to (2) Unflagged in Same Census Block Group $\times$ Zip code (CBGZIP) and (3) Unflagged in US

	(1) Flagged	(2) Unflagged in CBGZIP	(3) Unflagged in US
Consumer Months (mn)	6.9	252.9	4067.2
Credit Score	679.7	676.3	676.3
Age	48	50	51
Accounts (#)	7.1	3.9	3.8
Any 30+ Defaults (%)	9	5	6
30+ Defaults (#)	0.17	0.08	0.09
Any Balance (%)	95	71	70
Any Auto (%)	51	28	26
Any Credit Card (%)	75	56	54
Any Mortgage (%)	41	23	25
Any Non Mortgage (%)	93	69	68
Total Balances	124168	60068	59624
Auto Balances	10645	5333	4610
Credit Card Balances	5795	3148	3002
Mortgage Balances	95433	46286	47477
Non Mortgage Balances	28735	13782	12147
Credit Card Limits	23525	15294	14539

Notes: BTCCP data. Table summarizes data for consumers using characteristics twelve months prior to date of disaster flag. Column (1) "Flagged" shows characteristics of consumers with any disaster flags based on the first time a flag is applied. Flagged consumers are those who are first flagged between January 2010 and December 2024. Column (2) "Unflagged in CBGZIP" shows consumers who never have disaster flags and are in the same census block group $\times$ zip code where any other consumers had disaster flags at the same time. Column (3) "Unflagged in US" shows consumers who never have disaster flags and are in a census block group $\times$ zip code where other consumers did not have disaster flags at the same time.

Table 2: Predictive Performance of Models Varying Inclusion of Flagged Defaults and Disaster Defaults

Model	AUC	Change from Baseline
1. All Defaults	0.890236	
2. Flagged Defaults	0.890904	0.0751%
3. Disaster (03) Defaults	0.890927	0.0777%
4. Disaster (06) Defaults	0.890905	0.0751%
5. Disaster (12) Defaults	0.890905	0.0751%

Notes: BTCCP data. Area under the receiver operating characteristic curve (AUC) from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Row 1 is the baseline model that includes defaults information the prediction. Row 2 is performance of a model that includes inputs for defaults that have disaster flags. Rows 3, 4, and 5 show performance from models that include inputs for defaults within three, six, and twelve months of a disaster. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables.

Table 3: Quantity of Defaults Information Masked Across Models

Model	Mean Defaults	Change from Baseline
1. All Defaults	2.243	
2. Masked Flagged Defaults	2.242	-0.033%
3. Masked Disaster (03) Defaults	2.190	-2.375%
4. Masked Disaster (06) Defaults	2.127	-5.162%
5. Masked Disaster (12) Defaults	1.928	-14.026%
6. CARES Masked Disaster (03) Defaults	2.191	-2.327%
7. CARES Masked Disaster (06) Defaults	2.134	-4.875%
8. CARES Masked Disaster (12) Defaults	1.987	-11.396%
9. PCARES Masked Disaster (03) Defaults	2.011	-10.332%
10. PCARES Masked Disaster (06) Defaults	1.824	-18.663%
11. PCARES Masked Disaster (12) Defaults	1.511	-32.637%
12. No Defaults	0.000	-100.000%

Notes: BTCCP data. Calculated using data on 6.99 million consumers (the 30% testing sample) to October 2017. In this table mean defaults is measured by the total number of accounts that are 30 or more days past due in the last 84 months divided by the total number of consumers. Row 1 is the baseline model that includes all defaults. Row 2 masks defaults that have disaster flags. Rows 3, 4, and 5 masks defaults within three, six, and twelve months of a disaster are masked. Rows 6, 7, and 8 show 'CARES' that temporarily keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. Rows 9, 10, and 11 show 'PCARES' that permanently keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. Row 12 excludes all defaults.

Table 4: Predictive Performance of Models Varying Masking of Defaults

Model	AUC	Change from Baseline
1. All Defaults	0.890236	
2. Masked Flagged Defaults	0.890233	-0.0003%
3. Masked Disaster (03) Defaults	0.889206	-0.1157%
4. Masked Disaster (06) Defaults	0.888825	-0.1585%
5. Masked Disaster (12) Defaults	0.887044	-0.3586%
6. CARES Masked Disaster (03) Defaults	0.889176	-0.1191%
7. CARES Masked Disaster (06) Defaults	0.888824	-0.1586%
8. CARES Masked Disaster (12) Defaults	0.887381	-0.3207%
9. PCARES Masked Disaster (03) Defaults	0.888244	-0.2238%
10. PCARES Masked Disaster (06) Defaults	0.887343	-0.3249%
11. PCARES Masked Disaster (12) Defaults	0.885713	-0.5081%
12. No Defaults	0.880902	-1.0485%

Notes: BTCCP data. Area under the receiver operating characteristic curve (AUC) from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Each row shows performance under different models that vary the use of defaults data used as inputs. Row 1 is the baseline model that includes all defaults. Row 2 masks defaults that have disaster flags. Measures of defaults in rows 3, 4, and 5 masks defaults within three, six, and twelve months of a disaster are masked. 'CARES' measures of defaults in rows 6, 7, and 8 temporarily keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. 'PCARES' measures of defaults in rows 9, 10, and 11 permanently keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. Row 12 does not use any defaults information. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables.

Table 5: Difference-in-Differences Estimates (s.e.) of Effects of Disaster Flags in Month First Applied ($t = 0$)

Sample	(1) Credit Score	(2) Number New Credit Cards	(3) Value New Credit Card Limits
A. All	2.82	-0.0049	-28.2
	(0.22)	(0.0012)	(8.5)
B. Any Defaults	14.93	-0.0024	5.5
	(1.59)	(0.0022)	(8.6)
C. No Defaults	2.27	-0.005	-29.7
	(0.17)	(0.0012)	(8.7)
D. Subprime	10.18	-0.0071	-11.5
	(0.92)	(0.0017)	(5.1)
E. Non-Subprime	1.4	-0.0045	-31.4
	(0.09)	(0.0012)	(10)

Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Table show estimates of Equation 7's δ_0 with standard errors in parentheses. δ_0 are the coefficients on the interaction between event time indicator, in the month a consumer first has a disaster flag, and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each column shows results from separate regressions with different outcomes. The outcome in column (1) is credit score (VantageScore 3.0). The outcome in column (2) is the number of new credit card opened that month. The outcome in column (3) is the value of new credit card limits opened that month (\$). Row 'A. All' uses data for all 3.8 million consumers. Rows B to E are for subsamples of consumers by their pre-disaster characteristics at $t - 12$. Row 'B. Any Defaults' uses data for the 4.4% (0.14 million) consumers with any pre-disaster defaults. Row 'C. No Defaults' uses data for the 95.6% (3.06 million) consumers without any pre-disaster defaults. Row 'D. Subprime' uses data for the 16.2% (0.52 million) consumers with subprime (300 to 600) pre-disaster credit scores. Row 'E. Non-Subprime' uses data for the 83.8% (2.68 million) consumers with non-subprime (601 to 850) pre-disaster credit scores.

Table 6: Difference-in-Differences Estimates (s.e.) of Effects of Disaster Flags after Twelve Months ($t = 12$)

Sample	(1) Credit Score	(2) Number New Credit Cards	(3) Value New Credit Card Limits
A. All	-0.36	-0.0169	-103.1
	(0.35)	(0.0021)	(15.4)
B. Any Defaults	-0.29	-0.0036	6.3
	(0.83)	(0.0022)	(9.3)
C. No Defaults	-0.36	-0.0175	-108.1
	(0.33)	(0.0021)	(15.8)
D. Subprime	-2.74	-0.016	-27.7
	(1.18)	(0.0031)	(9.6)
E. Non-Subprime	0.11	-0.017	-117.7
	(0.19)	(0.0021)	(16.6)

Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Table show estimates of Equation 7's δ_{12} with standard errors in parentheses. δ_{12} are the coefficients on the interaction between event time indicator, in the twelfth month after a consumer first has a disaster flag, and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each column shows results from separate regressions with different outcomes. The outcome in column (1) is credit score (VantageScore 3.0). The outcome in column (2) is the number of new credit card opened that month. The outcome in column (3) is the value of new credit card limits opened that month ($). Row 'A. All' uses data for all 3.8 million consumers. Rows B to E are for subsamples of consumers by their pre-disaster characteristics at $t - 12$. Row 'B. Any Defaults' uses data for the 4.4% (0.14 million) consumers with any pre-disaster defaults. Row 'C. No Defaults' uses data for the 95.6% (3.06 million) consumers without any pre-disaster defaults. Row 'D. Subprime' uses data for the 16.2% (0.52 million) consumers with subprime (300 to 600) pre-disaster credit scores. Row 'E. Non-Subprime' uses data for the 83.8% (2.68 million) consumers with non-subprime (601 to 850) pre-disaster credit scores.

References

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., Piskorski, T., and Seru, A. (2017). Policy intervention in debt renegotiation: Evidence from the home affordable modification program. *Journal of Political Economy*, 125(3):654–712.

Akerlof, G. A. (1978). The economics of “tagging” as applied to the optimal income tax, welfare programs, and manpower planning. *American Economic Review*, 68(1):8–19.

Aydin, D. (2024). Forbearance vs. interest rates: Experimental tests of liquidity and strategic default triggers. *Working Paper*.

Baker, A. C., Larcker, D. F., and Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2):370–395.

Banko-Ferran, D. and Ricks, J. (2018). Natural disasters and credit reporting. *Consumer Financial Protection Bureau Quarterly Consumer Credit Trends Report*, November 2018.

Batty, M., Gibbs, C., and Ippolito, B. (2022). Health insurance, medical debt, and financial well-being. *Health Economics*, 31(5):689–728.

Begley, T. A., Gurun, U. G., Purnanandam, A., and Weagley, D. (2024). Disaster lending:“fair” prices but “unfair” access. *Management Science*, 70(12):8484–8505.

Billings, S. B., Gallagher, E., and Ricketts, L. (2022). Let the rich be flooded: The distribution of financial aid and distress after Hurricane Harvey. *Journal of Financial Economics*, 146:797–819.

Blattner, L., Hartwig, J., and Nelson, S. (2022). Information design in consumer credit markets. *Working Paper*.

Blattner, L. and Nelson, S. (2024). How costly is noise? data and disparities in consumer credit. *Working Paper*.

Bleemer, Z. and van der Klaauw, W. (2019). Long-run net distributionary effects of federal disaster insurance: The case of Hurricane Katrina. *Journal of Urban Economics*, 110:70–88.

Blickle, K., Hamerling, S. N., and Morgan, D. P. (2022). How bad are weather disasters for banks? *Federal Reserve Bank of New York Staff Report No. 990*.

Bornstein, G. and Indarte, S. (2023). The impact of social insurance on household debt. *Working Paper*.

Bos, M., Breza, E., and Liberman, A. (2018). The labor market effects of credit market information. *Review of Financial Studies*, 31(6):2005–2037.

Braxton, J. C., Herkenhoff, K. F., and Phillips, G. M. (2024). Can the unemployed borrow? implications for public insurance. *Journal of Political Economy*, 132(9):3025–3076.

Brevoort, K. P. and Kambara, M. (2015). Are all collections equal? the case of medical debt. *Journal of Credit Risk*, 11(4).

Bursztyn, L., Fiorin, S., Gottlieb, D., and Kanz, M. (2019). Moral incentives in credit card debt repayment: Evidence from a field experiment. *Journal of Political Economy*, 127(4):1641–1683.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

Chava, S., Tookes, H., and Zhang, Y. (2023). Leaving them hanging: Student loan forbearance, distressed borrowers, and their lenders. *Working Paper*.

Cherry, S. (2024). Regulating credit: The impact of price regulations and lender technologies on financial inclusion. *Working Paper*.

Cherry, S. F., Jiang, E. X., Matvos, G., Piskorski, T., and Seru, A. (2021). Government and private household debt relief during covid-19. *Brookings Papers on Economic Activity*, Fall 2021:141–199.

Chetty, R. and Finkelstein, A. (2013). Social insurance: Connecting theory to data. *Handbook of Public Economics*, 5:111–193.

Collier, B. and Ellis, C. (2024). A demand curve for disaster recovery loans. *Econometrica*, 92(3):713–748.

Collier, B. L., Hartley, D. A., Keys, B. J., and Ng, J. X. (2024a). Credit when you need it. *NBER Working Paper No. 32845*.

Collier, B. L., Howell, S. T., and Rendell, L. (2024b). After the storm: How emergency liquidity helps small businesses following natural disasters. *NBER Working Paper No. 32326*.

Cookson, J. A., Gallagher, E., and Mulder, P. (2024). Money to burn: Crowdfunding wildfire recovery. *Working Paper*.

Cortés, K. R. and Strahan, P. E. (2017). Tracing out capital flows: How financially integrated banks respond to natural disasters. *Journal of Financial Economics*, 125(1):182–199.

Deryugina, T. (2017). The fiscal cost of hurricanes: Disaster aid versus social insurance. *American Economic Journal: Economic Policy*, 9(3):168–98.

Deryugina, T., Kawano, L., and Levitt, S. (2018). The economic impact of Hurricane Katrina on its victims: evidence from individual tax returns. *American Economic Journal: Applied Economics*, 10(2):202–33.

Deshpande, M. and Li, Y. (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–248.

Di Maggio, M., Kalda, A., and Yao, V. (2024). Second chance: Life with less student debt. *Journal of Finance*, Forthcoming.

Dinerstein, M., Yannelis, C., and Chen, C.-T. (2024). Debt moratoria: Evidence from student loan forbearance. *American Economic Review: Insights*, 6(2):196–213.

Dobbie, W., Goldsmith-Pinkham, P., Mahoney, N., and Song, J. (2020). Bad credit, no problem? credit and labor market consequences of bad credit reports. *Journal of Finance*, 75(5):2377–2419.

Dobbie, W. and Song, J. (2020). Targeted debt relief and the origins of financial distress: Experimental evidence from distressed credit card borrowers. *American Economic Review*, 110(4):984–1018.

Dobkin, C., Finkelstein, A., Kluender, R., and Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. *American Economic Review*, 108(2):308–352.

Dube, A., Girardi, D., Jorda, O., and Taylor, A. M. (2024). A local projections approach to difference-in-differences event studies. *NBER Working Paper No. 31184*.

Farrell, D. and Greig, F. (2018). Weathering the storm: The financial impacts of Hurricanes Harvey and Irma on one million households. *Working Paper*.

FinRegLab (2020). Disaster-related credit reporting options. *Research Brief*.

FinRegLab, Blattner, L., and Spiess, J. (2022). Machine learning explainability & fairness: Insights from consumer lending. *FinRegLab Empirical Whitepaper*.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., and Walther, A. (2022). Predictably unequal? the effects of machine learning on credit markets. *Journal of Finance*, 77(1):5–47.

Gallagher, J. and Hartley, D. (2017). Household finance after a natural disaster: The case of Hurricane Katrina. *American Economic Journal: Economic Policy*, 9(3):199–228.

Gallagher, J., Hartley, D., and Rohlin, S. (2023). Weathering an unexpected financial shock: The role of disaster assistance on household finance and business survival. *Journal of the Association of Environmental and Resource Economists*, 10(2):525–567.

Ganong, P. and Noel, P. (2020). Liquidity vs. wealth in household debt obligations: Evidence from housing policy in the great recession. *American Economic Review*, 110(10):3100–3138.

Ganong, P. and Noel, P. (2023). Why do borrowers default on mortgages? *Quarterly Journal of Economics*, 138(2):1001–1065.

Gaubert, C., Kline, P. M., and Yagan, D. (2024). Place-based redistribution. *NBER Working Paper No. 28337*.

Gibbs, C., Guttman-Kenney, B., Lee, D., Nelson, S., van der Klauuw, W., and Wang, J. (2024). Consumer credit reporting data. *Journal of Economic Literature*, Forthcoming.

Giglio, S., Kelly, B., and Stroebel, J. (2021). Climate finance. *Annual Review of Financial Economics*, 13:15–36.

Goodman, L. and Zhu, J. (2023). Single borrowers versus coborrowers in the pandemic: Mortgage forbearance take-up and performance. *Journal of Housing Economics*, 59:101909.

Gross, T., Notowidigdo, M. J., and Wang, J. (2020). The marginal propensity to consume over the business cycle. *American Economic Journal: Macroeconomics*, 12(2):351–84.

Guiso, L., Sapienza, P., and Zingales, L. (2013). The determinants of attitudes toward strategic default on mortgages. *Journal of Finance*, 68(4):1473–1515.

Guttman-Kenney, B. (2024). *Essays on Household Finance*. PhD thesis, The University of Chicago.

Guttman-Kenney, B. and Shahidinejad, A. (2025). Unraveling information sharing in consumer credit markets. *Working Paper*.

Herkenhoff, K., Phillips, G., and Cohen-Cole, E. (2021). The impact of consumer credit access on employment, earnings and entrepreneurship. *Journal of Financial Economics*, 141(1):345–371.

Hsu, J. W., Matsa, D. A., and Melzer, B. T. (2018). Unemployment insurance as a housing market stabilizer. *American Economic Review*, 108(1):49–81.

Jansen, M., Nagel, F., Yannelis, C., and Zhang, A. L. (2024). Data and welfare in credit markets. *NBER Working Paper No. 30235*.

Katz, J. (2024). Saving and consumption responses to student loan forbearance. *Working Paper*.

Keys, B. J., Mahoney, N., and Yang, H. (2022). What determines consumer financial distress? place- and person-based factors. *Review of Financial Studies*, 36:42–69.

Kim, Y. S., Lee, D., Scharlemann, T. C., and Vickery, J. I. (2024). Intermediation frictions in debt relief: evidence from cares act forbearance. *Journal of Financial Economics*, 158:103873.

Kluender, R., Mahoney, N., Wong, F., and Yin, W. (2021). Medical debt in the us, 2009-2020. *JAMA*, 326(3):250–256.

Kluender, R., Mahoney, N., Wong, F., and Yin, W. (2024). The effects of medical debt relief: Evidence from two randomized experiments. *Quarterly Journal of Economics*, Forthcoming.

Liberman, A., Neilson, C., Opazo, L., and Zimmerman, S. (2020). The equilibrium effects of information deletion: Evidence from consumer credit markets. *Working Paper*.

Low, D. (2023). What triggers mortgage default? new evidence from linked administrative and survey data. *Review of Economics and Statistics*, Forthcoming.

Martínez-Marquina, A. and Shi, M. (2024). The opportunity cost of debt aversion. *American Economic Review*, 114(4):1140–1172.

Musto, D. K. (2004). What happens when information leaves a market? evidence from postbankruptcy consumers. *Journal of Business*, 77(4):725–748.

National Consumer Law Center (2019). Letter urging credit bureaus to provide credit reporting relief to consumers affected by natural disaster. *Public Letter, January 18, 2019*.

TransUnion (2024). University of chicago booth transunion consumer credit panel (btccp), 2000 to 2024.

Urban Institute (2019). Insult to injury: Natural disasters and residents' financial health. *Research Report*.

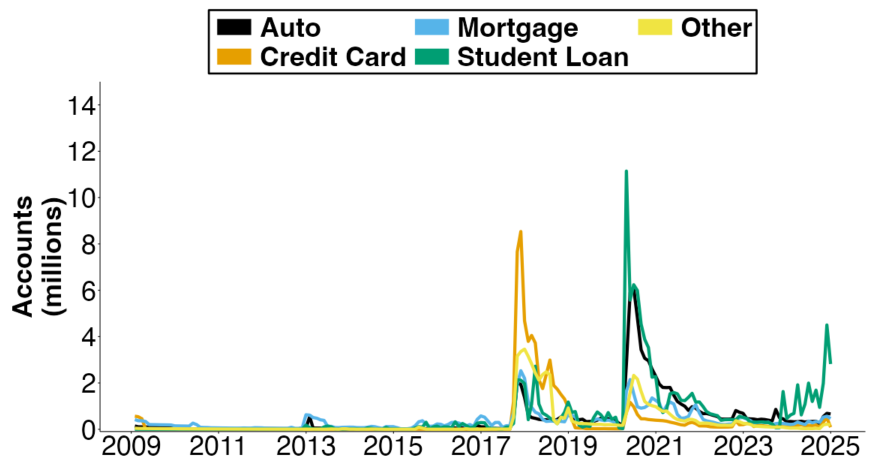
Weinzierl, M. (2012). Why do we redistribute so much but tag so little? the principle of equal sacrifice and optimal taxation. *NBER Working Paper No. 18045*.

Wing, C., Freedman, S. M., and Hollingsworth, A. (2024). Stacked difference-in-differences. *NBER Working Paper No. 32054*.

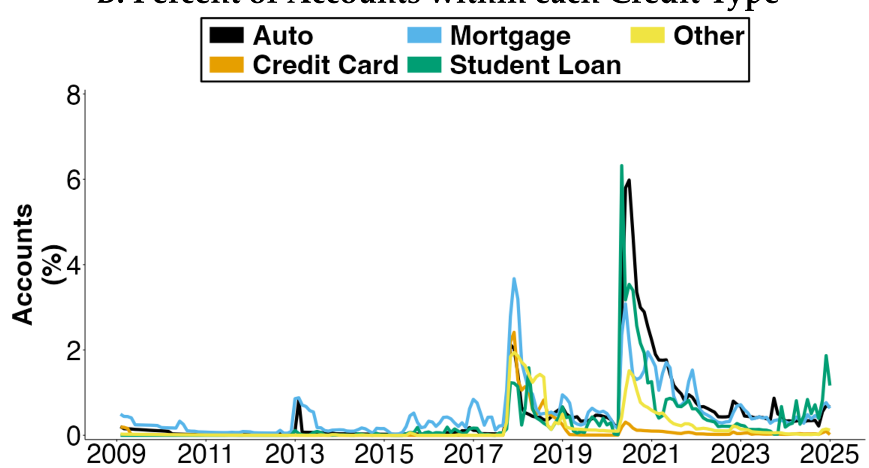
A. Internet Appendix accompanying *"Disaster Flags: Credit Reporting Relief from Natural Disasters"*

Figure A1: Trades with credit report disaster flag by credit type, 2009 - 2024

A. Number of Accounts

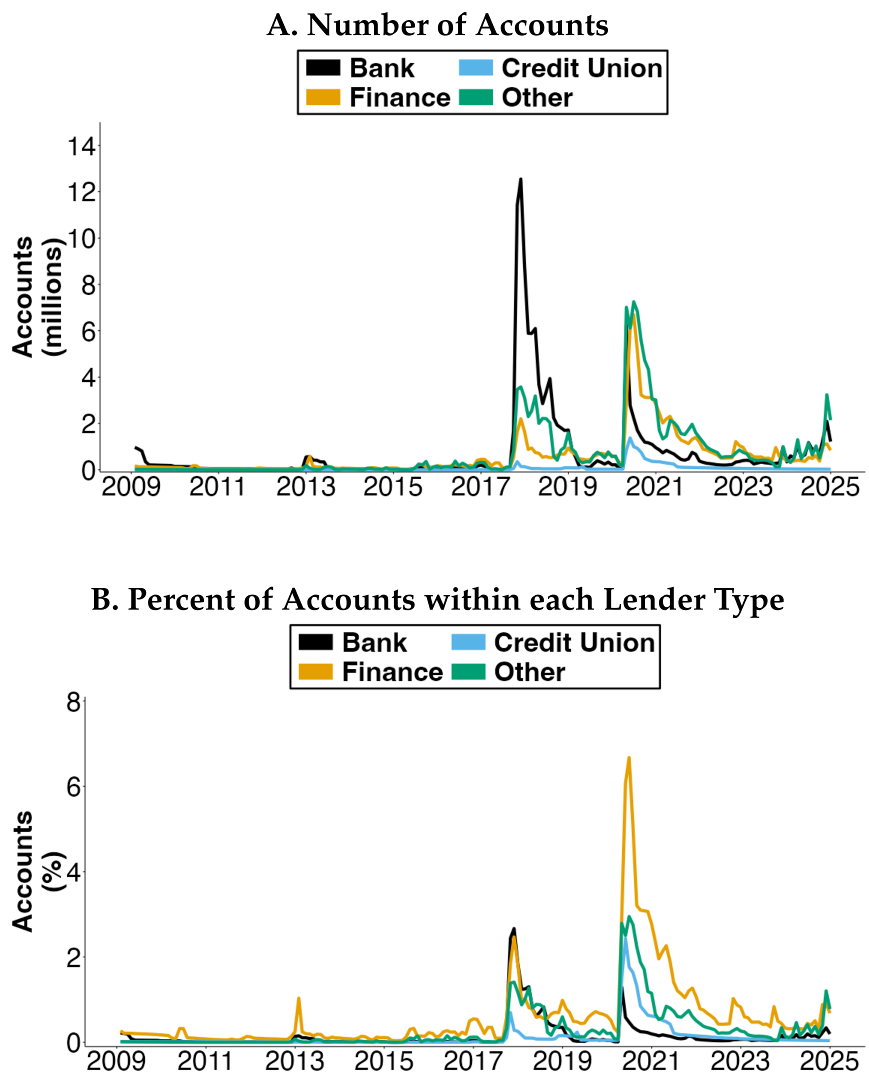


B. Percent of Accounts within each Credit Type



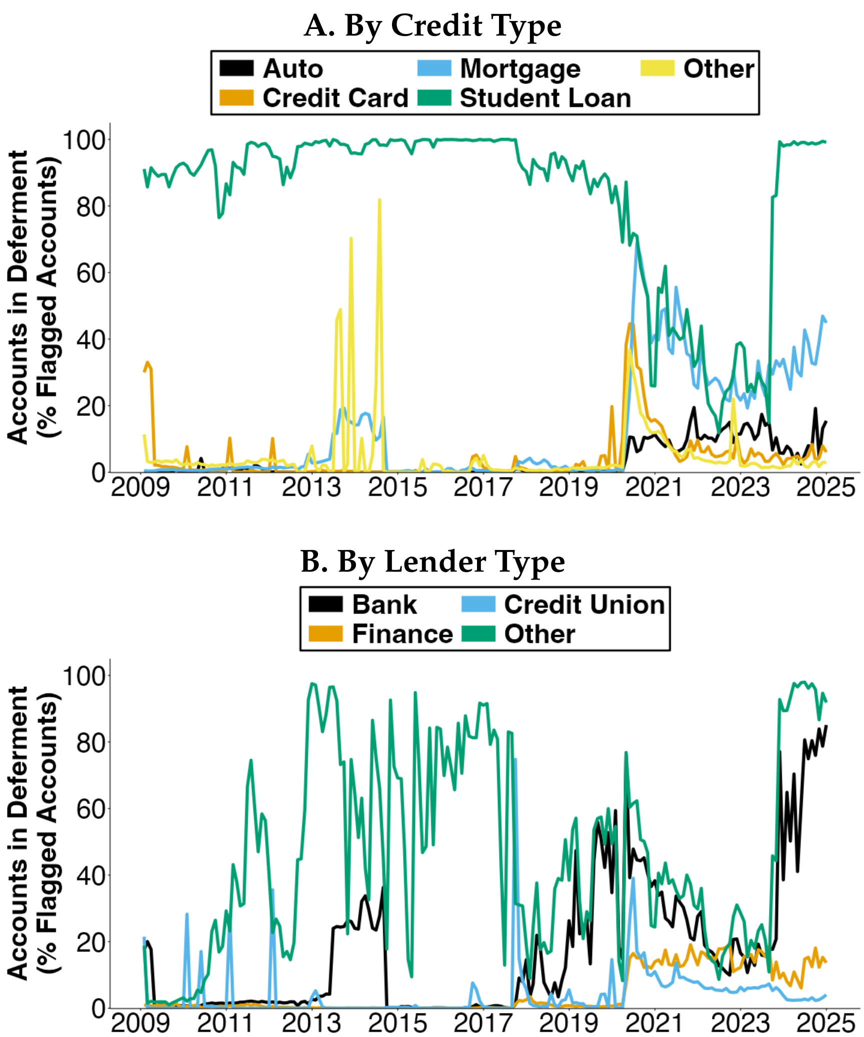
Notes: BTCCP data. Open tradelines with a positive balance in their credit report with a credit report disaster flag. Numbers in Panel A are extrapolated to population estimates from 10% sample. Other contains retail cards and unsecured loans.

Figure A2: Trades with credit report disaster flag by lender type, 2009 - 2024



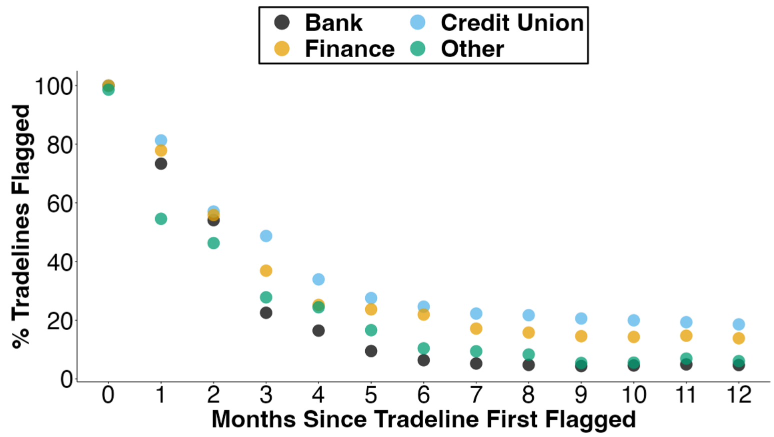
Notes: BTCCP data. Open tradelines with a positive balance in their credit report with a credit report disaster flag. Numbers in Panel A are extrapolated to population estimates from 10% sample.

Figure A3: Trades with credit report disaster flag that also had deferments, 2009 - 2024, (A) by credit type, (B) by lender type



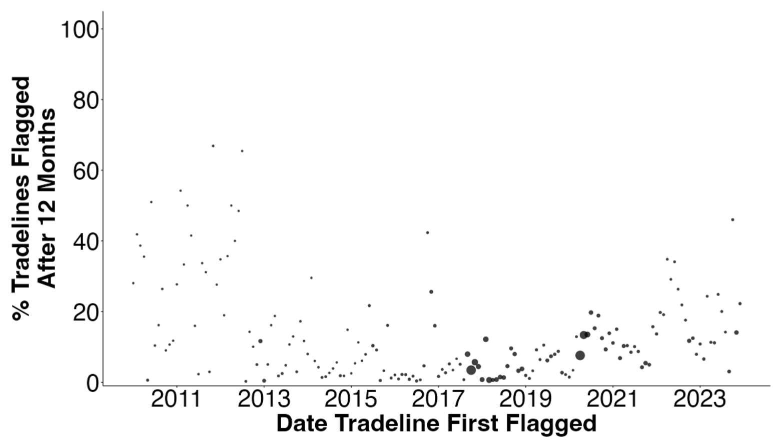
Notes: BTCCP data. Open tradelines with a positive balance in their credit report with a credit report disaster flag. Lines show fractions of flagged tradelines that also have deferments: accounts listed with deferments and tradelines with positive balances but zero payments due.

Figure A4: Duration of disaster flags remaining on a credit report tradeline, by lender type



Notes: BTCCP data. This takes open credit report tradelines with positive balance that first have a disaster flag added between January 2010 and December 2023. Plots the fraction of these with disaster flags still present 1 to 12 months later. Colors are lender types.

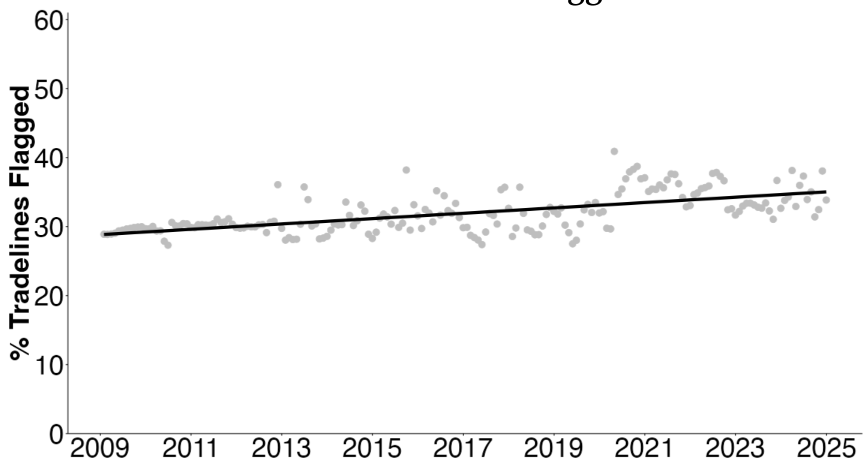
Figure A5: Fraction of disaster flags remaining on a credit report tradeline after 12 months, by cohort



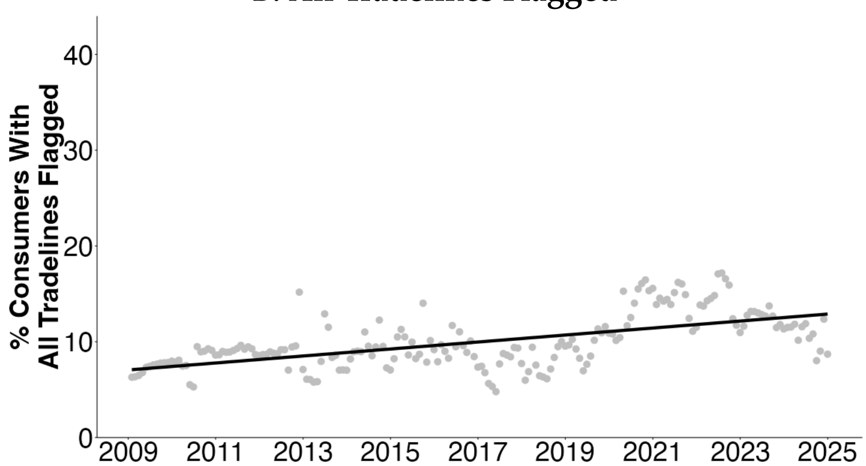
Notes: BTCCP data. This takes open credit report tradelines with positive balance that first have a disaster flag added between January 2010 and December 2023. Plots the fraction of these with disaster flags still present 12 months later for each cohort. X axis is cohort date when disaster flag first added to tradeline. Size of dot is proportional to initial disaster flag cohort size.

Figure A6: Intensive Margin: Among consumers with disaster flags, mean fraction of tradelines flagged (Panel A) and fraction with all tradelines flagged (Panel B), 2009 to 2024

A. Mean Tradelines Flagged

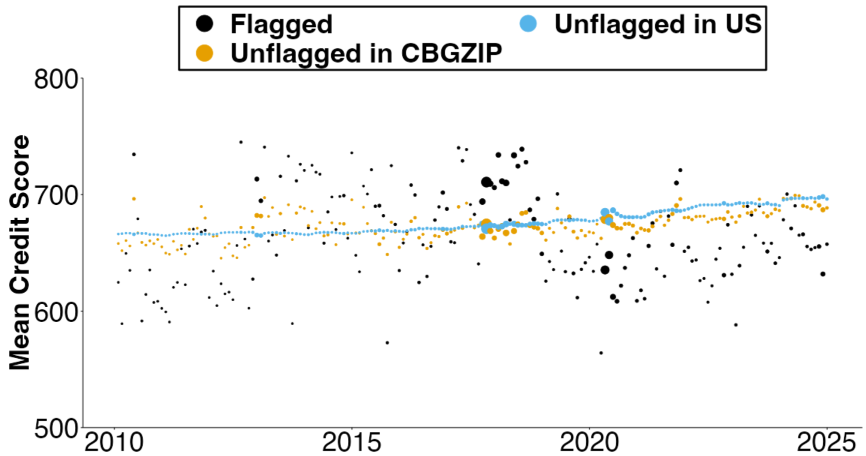


B. All Tradelines Flagged



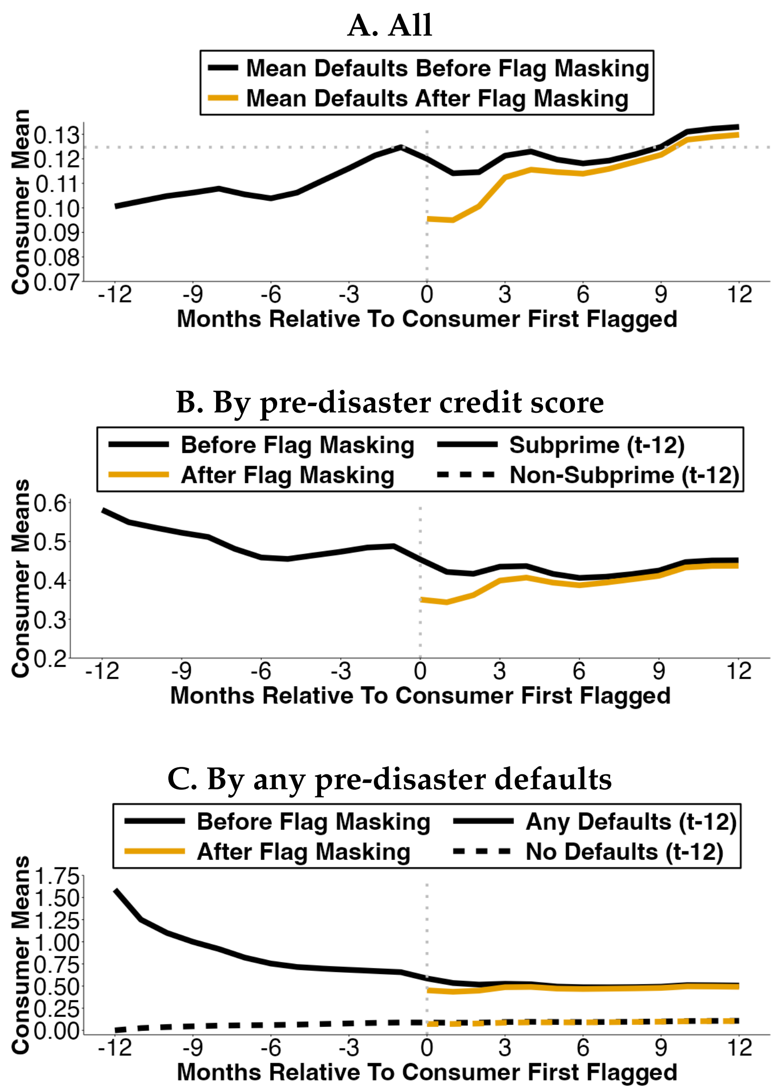
Notes: BTCCP data. Both Panels A and B restrict to consumer months where the consumer has a credit report disaster flag on at least one open tradeline with a positive balance in their credit report. Panel A shows for these consumers, the mean of tradelines with a credit report disaster flag. Panel B shows the fraction of these consumers where all their tradelines have a credit report disaster flag. Linear time trends added in both Panels.

Figure A7: Mean Credit Score of Cohorts of Consumers First Flagged (black) Compared to Unflagged Consumers in the Same Census Block Group × Zip code (CBGZIP, orange), and Unflagged in US (blue)



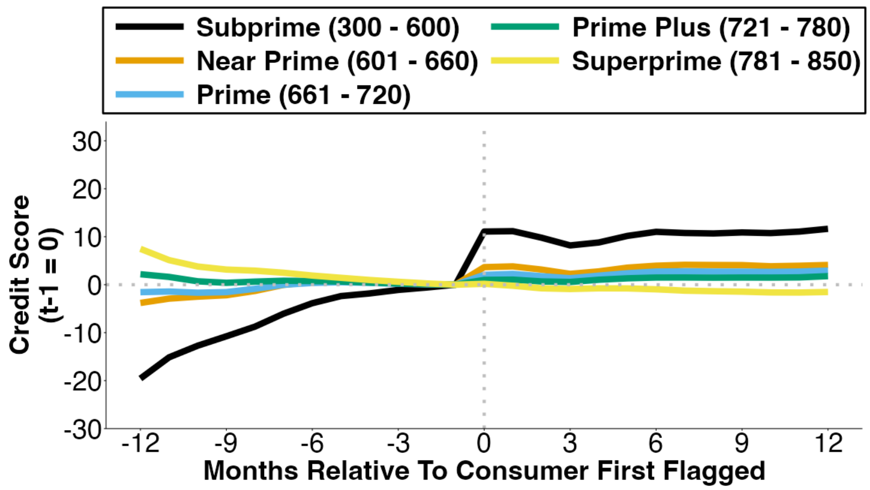
Notes: BTCCP data. Table summarizes data for consumers using credit scores twelve months prior to date of disaster flag. Black dots show "Flagged" groups of consumers with any disaster flags based on the first time a flag is applied. Orange dots show "Unflagged in CBGZIP" groups of consumers who never have disaster flags and are in the same census block group × zip code where any other consumers had disaster flags at the same time. Blue dots show "Unflagged in US" shows consumers who never have disaster flags and are in areas where other consumers did not have disaster flags at the same time. The size of the dots corresponds to the number of consumers in the flagged group at each point-in-time.

Figure A8: Event study of mean number of defaults on credit reports before (black) and after (orange) flag masking for (A) all flagged consumers, (B) by pre-disaster credit score, and (C) by any pre-disaster defaults



Notes: BTCCP data. Lines are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. X axis shows months since consumer first flagged. Y axes are mean number of defaults on consumers' credit reports before (black) and after (orange) tradeline months where defaults are masked by flags. Panel B splits by whether a consumer has subprime credit score (300 - 600) twelve months prior to first being flagged ($t-12$). Panel C splits by whether a consumer has any defaults twelve months prior to first being flagged ($t-12$). Panels B and C normalize each series to $t-1=0$.

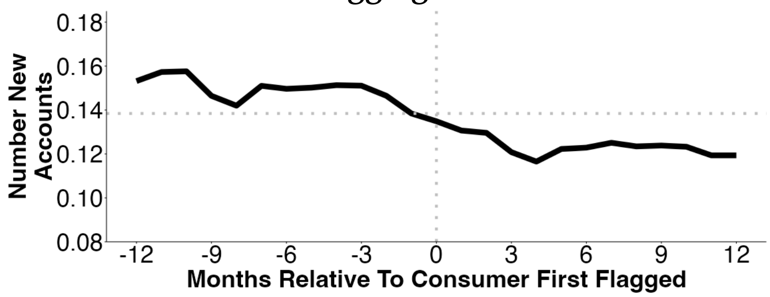
Figure A9: Event study of VantageScore Credit Scores by Pre-Disaster Credit Score ($t-12$)



Notes: BTCCP data. Solid lines are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. X axis shows months since consumer first flagged. Y axis is credit score. Splits by consumer's credit score twelve months prior to first being flagged ($t-12$). Credit score for each subgroup is normalized to $t-1=0$.

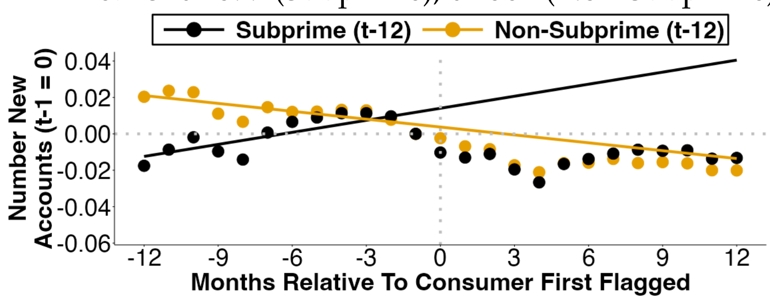
Figure A10: Event Study of Number of New Accounts Relative to Linear Pre-Trend

A. Aggregated



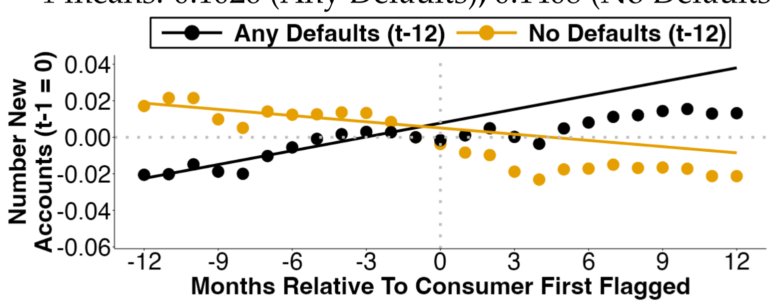
B. Disaggregated by Pre-Disaster Credit Score

$t-1$ means: 0.1399 (Subprime), 0.1382 (Non-Subprime)



C. Disaggregated by Any Pre-Disaster Defaults

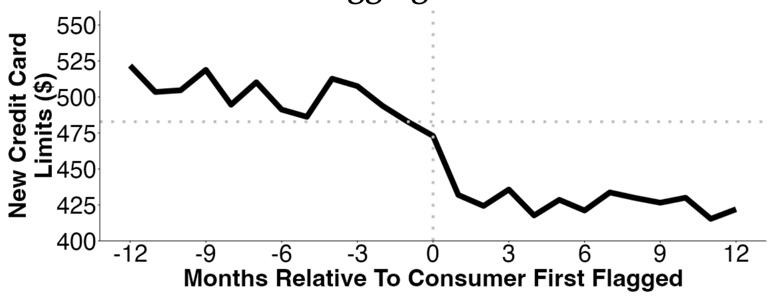
$t-1$ means: 0.1026 (Any Defaults), 0.1408 (No Defaults)



Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. Solid lines are linear time trend from OLS regressions on data $t-12$ to $t-1$. X axis shows months since consumer first flagged. Y axis shows new account openings. Panel B splits by whether a consumer has subprime credit score (300 - 600) twelve months prior to first being flagged ($t-12$). Panel C splits by whether a consumer has any defaults twelve months prior to first being flagged ($t-12$). Panels B and C normalize new account openings for each subgroup to $t-1=0$.

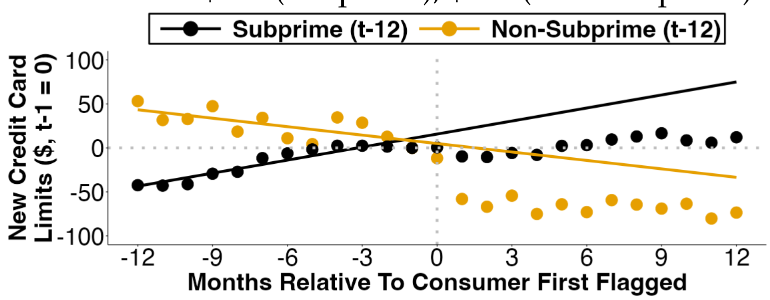
Figure A11: Event Study of Value of New Credit Card Limits Relative to Linear Pre-Trend

A. Aggregated



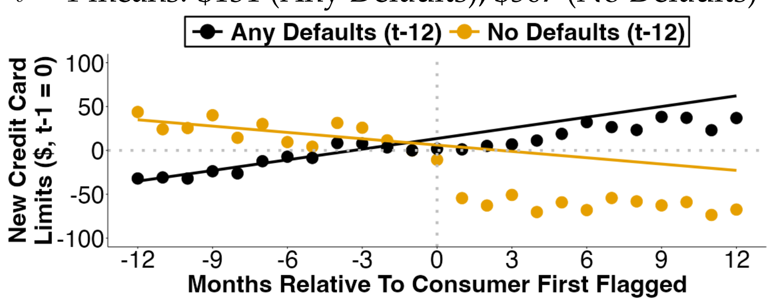
B. Disaggregated by Pre-Disaster Credit Score

$t-1$ means: \$133 (Subprime), \$544 (Non-Subprime)



C. Disaggregated by Any Pre-Disaster Defaults

$t-1$ means: \$131 (Any Defaults), \$507 (No Defaults)

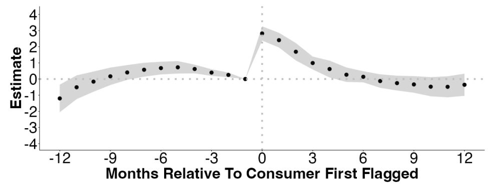


Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. Solid lines are linear time trend from OLS regressions on data $t-12$ to $t-1$. X axis shows months since consumer first flagged. Y axis shows new credit card account openings. Panel B splits by whether a consumer has subprime credit score (300 - 600) twelve months prior to first being flagged ($t-12$). Panel C splits by whether a consumer has any defaults twelve months prior to first being flagged ($t-12$). Panels B and C normalize value new credit card limits for each subgroup to $t-1=0$.

Figure A12: Difference-in-differences estimates of effects on (A) credit score, (B) any new credit card openings, and (C) value of new credit card limits

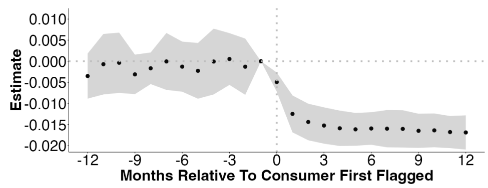
A. Credit score

$t-1$ mean: 704



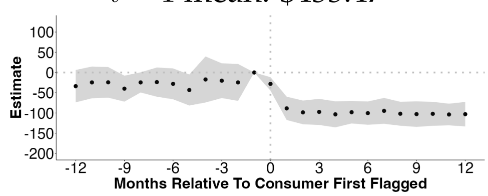
B. Number new credit card openings

$t-1$ mean: 0.0633



C. Value new credit card limits

$t-1$ mean: \$433.47

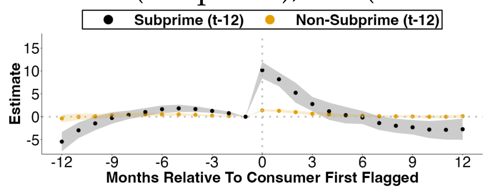


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 7's δ_τ. δ_τ are the coefficients on the interaction between event time indicators after τ periods and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each panel shows results for different outcomes: A. credit score (VantageScore 3.0), B. the number of new credit cards opened that month, and C. the value of new credit card limits opened that month.

Figure A13: Difference-in-differences estimates of effects on (A) credit score, (B) any new credit card openings, and (C) value of new credit card limits, all by pre-disaster credit score

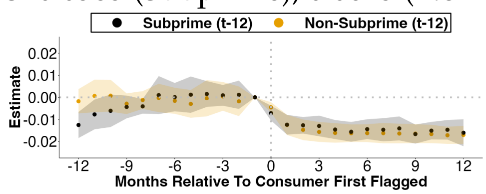
A. Credit score

$t-1$ means: 560 (Subprime), 731 (Non-Subprime)



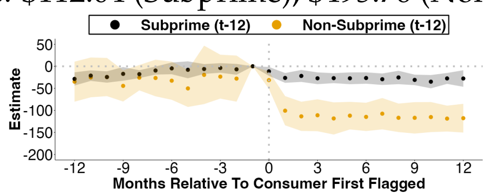
B. Number new credit card openings

$t-1$ means: 0.0583 (Subprime), 0.0643 (Non-Subprime)



C. Value new credit card limits

$t-1$ means: \$112.04 (Subprime), \$495.76 (Non-Subprime)

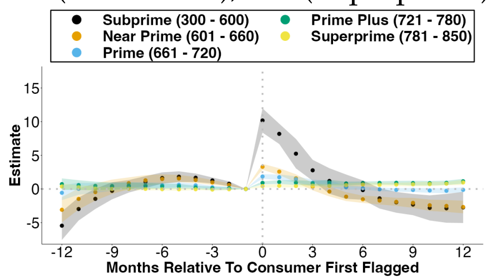


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 7's δ_τ. δ_τ are the coefficients on the interaction between event time indicators after τ periods and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each panel shows results for different outcomes: A. credit score (VantageScore 3.0), B. the number of new credit cards opened that month, and C. the value of new credit card limits opened that month. Within each panel, different colors denote separate regressions for consumers based on their pre-disaster credit score at event time $t-12$. The black estimates use data for the 16.2% (0.52 million) subsample of consumers with subprime (300 to 600) pre-disaster credit scores, and the orange estimates use data for the 83.8% (2.68 million) subsample with non-subprime (601 to 850) pre-disaster credit scores.

Figure A14: Difference-in-differences estimates of effects on (A) credit score, (B) any new credit card openings, and (C) value of new credit card limits, all by pre-disaster credit score

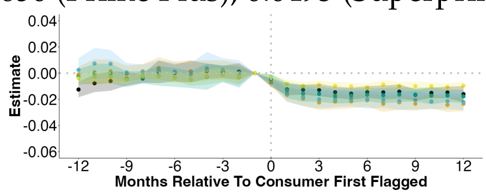
A. Credit score

$t-1$ means: 560 (Subprime), 635 (Near Prime), 689 (Prime)
746 (Prime Plus), 800 (Superprime)



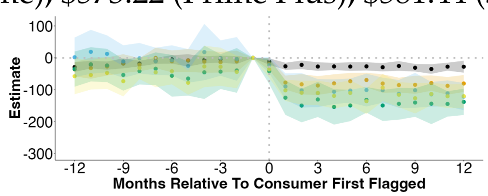
B. Number new credit card openings

$t-1$ means: 0.0583 (Subprime), 0.0764 (Near Prime), 0.0762 (Prime)
0.0650 (Prime Plus), 0.0493 (Superprime)



C. Value new credit card limits

$t-1$ means: \$112.04 (Subprime), \$292.27 (Near Prime),
\$442.16 (Prime), \$573.22 (Prime Plus), \$581.44 (Superprime)

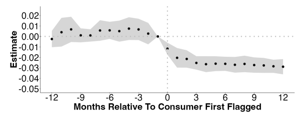


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 7's δ_τ. δ_τ are the coefficients on the interaction between event time indicators after τ periods and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. The outcome in panel A is VantageScore credit score, the outcome in panel B is the number of new credit card account openings each month, and the outcome in panel C is the value of new credit card limits. Each panel shows results for different samples. Within each panel, different colors denote separate regressions for consumers based on their pre-disaster credit scores at event time $t-12$: 16.2% (0.52 million) consumers with subprime (300 to 600) in black, 15.6% (0.50 million) consumers with near prime (601 to 660) in orange, the 17.8% (0.57 million) consumers prime (661 to 720) in blue, the 22.5% (0.72 million) consumers with prime plus (721 to 780) in green, and the 27.9% (0.89 million) consumers with superprime (891 to 850) in yellow.

Figure A15: Difference-in-differences estimates of effects on the number of new account openings across credit types for (A) all consumers, (B) by any pre-disaster defaults, and (C) pre-disaster credit score

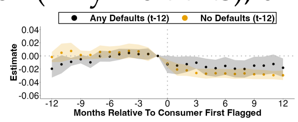
A. All

$t - 1$ mean: 0.1380



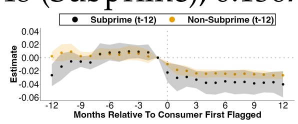
B. By any pre-disaster defaults

$t - 1$ means: 0.0984 (Any Defaults), 0.1398 (No Defaults)



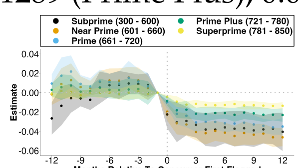
C. By pre-disaster credit score

$t - 1$ means: 0.1448 (Subprime), 0.1367 (Non-Subprime)



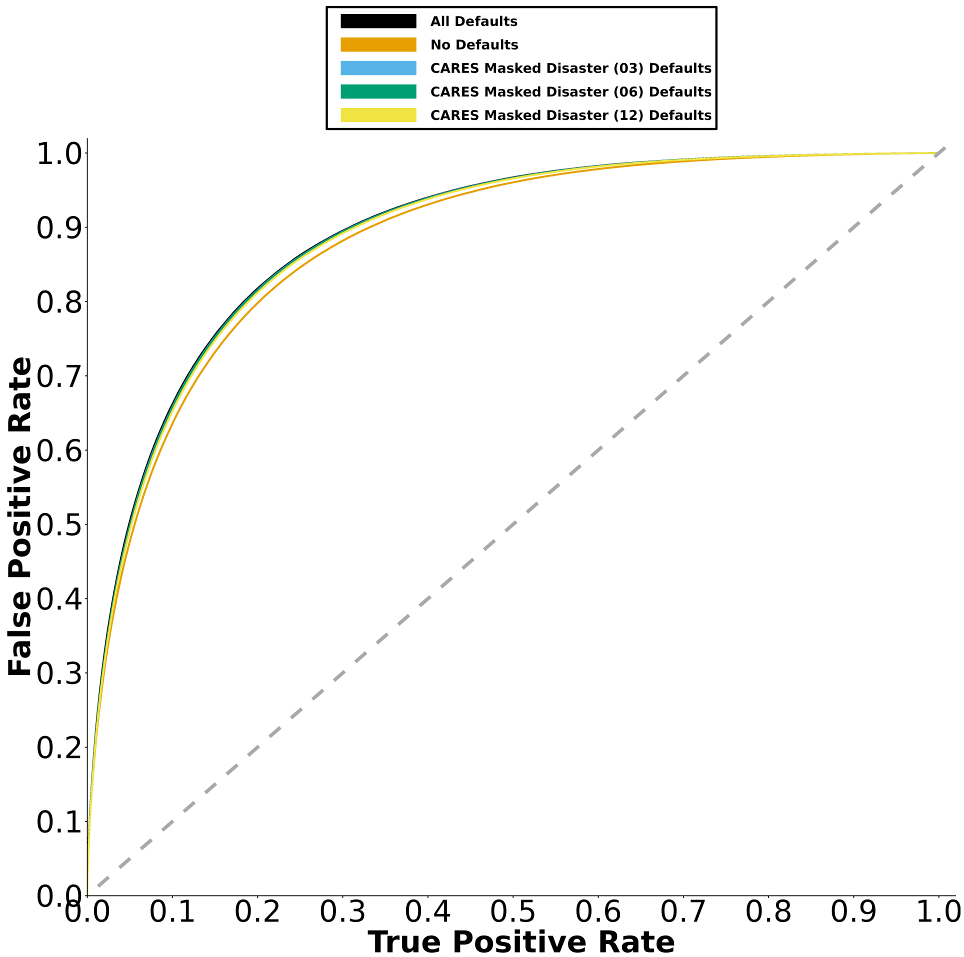
D. By pre-disaster credit score

$t - 1$ means: 0.1448 (Subprime), 0.1868 (Near Prime),
0.1644 (Prime), 0.1289 (Prime Plus), 0.0972 (Superprime)



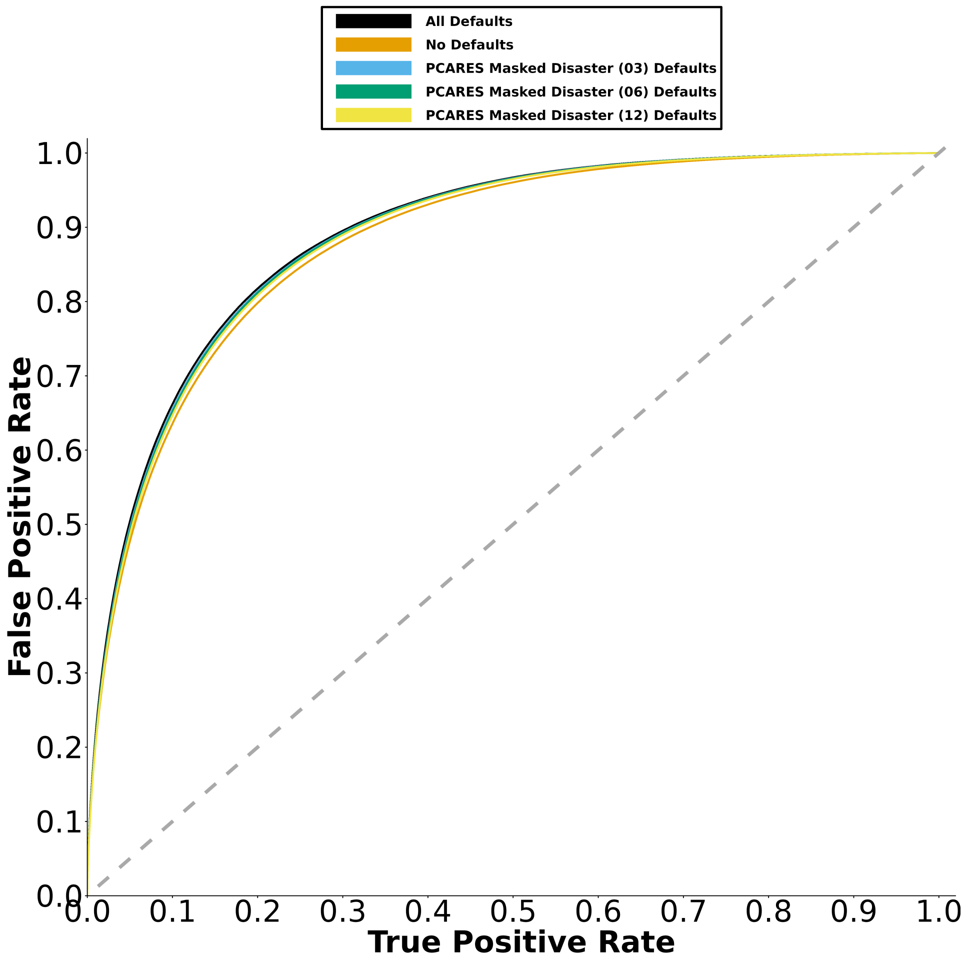
Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 7's δ_τ. δ_τ are the coefficients on the interaction between event time indicators after τ periods and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Outcomes in all panels are the number of new account openings across credit types each month. Each panel shows results for different samples. Within each panel, different colors denote separate regressions for consumers based on their pre-disaster characteristics at event time $t - 12$. Panel A uses data for all 3.8 million consumers. Panel B. the black estimates are for the 4.4% (0.14 million) consumers with any pre-disaster defaults, and the orange estimates for the 95.6% (3.06 million) consumers without any pre-disaster defaults. In Panel C, the black estimates use data for the 16.2% (0.52 million) subsample of consumers with subprime (300 to 600) pre-disaster credit scores, and the orange estimates use data for the 83.8% (2.68 million) subsample with non-subprime (601 to 850) pre-disaster credit scores. Panel D shows results by credit scores: 16.2% consumers with subprime (300 to 600) in black, 15.6% consumers with near prime (601 to 660) in orange, the 17.8% consumers prime (661 to 720) in blue, the 22.5% consumers with prime plus (721 to 780) in green, and the 27.9% consumers with superprime (891 to 850) in yellow.

Figure A16: Receiver Operating Characteristic Curves (ROC) Showing the Predictive Performance of Models Masking Default Information. Baseline Model Includes Defaults (Black), CARES Masking Disaster Defaults over three (Blue), six (Green), and twelve (Yellow) months, and No Defaults (Orange)



Notes: BTCCP data. Receiver Operating Characteristic curves (ROC) from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. The black line is the baseline model that includes defaults information the prediction. The orange line respectively show performance from a model without defaults information. The blue, green, and yellow lines show performance from 'CARES' models with defaults information that mask defaults by temporarily keeping accounts at their pre-disaster delinquency status for three (blue), six (green), and twelve (yellow) months from a disaster starting. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables. True Positive Rate is defined as the fraction of predicted new defaulters out of actual new defaulters. False Positive Rate is defined as the fraction of predicted new defaulters out of actual no new defaulters.

Figure A17: Receiver Operating Characteristic Curves (ROC) Showing the Predictive Performance of Models Masking Default Information. Baseline Model Includes Defaults (Black), PCARES Masking Disaster Defaults over three (Blue), six (Green), and twelve (Yellow) months, and No Defaults (Orange)



Notes: BTCCP data. Receiver Operating Characteristic curves (ROC) from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. The black line is the baseline model that includes defaults information the prediction. The orange line respectively show performance from a model without defaults information. The blue, green, and yellow lines show performance from 'PCARES' models with defaults information that mask defaults by permanently keeping accounts at their pre-disaster delinquency status for three (blue), six (green), and twelve (yellow) months from a disaster starting. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables. True Positive Rate is defined as the fraction of predicted new defaulters out of actual new defaulters. False Positive Rate is defined as the fraction of predicted new defaulters out of actual no new defaulters.

Table A1: Summarizing Tradeline Months with Disaster Flags by Credit Type

Credit Type	(1) 2009 - 2019 (%)	(2) 2009 - 2024 (%)	(3) December 2024 (%)
Auto	12.26	21.94	15.69
Credit Card	31.25	16.87	2.1
Mortgage	20.76	17.1	10.97
Student Loan	15.44	29.42	66.78
Other	20.29	14.68	4.46

Notes: BTCCP data. Columns display each credit type's share of all disaster flagged trade months as of (1) January 2009 to December 2019 (2) January 2009 to December 2024 (3) December 2024. Other contains retail cards and unsecured loans.

Table A2: Average marginal effects of default variables as predictors of any new default. Regressions differentially add default information: (1) ALL is Baseline, (2) FLAGGED includes terms for disaster flagged defaults, (3) DISASTER includes terms for FEMA natural disaster defaults

	(1)	(2)	(3)	(4)	(5)
All Defaults 12	0.003150	0.003096	0.003280	0.003305	0.003434
	(0.000135)	(0.000135)	(0.000141)	(0.000142)	(0.000150)
All Defaults 24	0.000851	0.000859	0.001042	0.000975	0.000424
	(0.000185)	(0.000185)	(0.000188)	(0.000189)	(0.000196)
All Defaults 36	0.001503	0.001496	0.001631	0.001664	0.001953
	(0.000154)	(0.000154)	(0.000156)	(0.000158)	(0.000164)
All Defaults 84	0.001270	0.001266	0.001338	0.001220	0.001141
	(0.000063)	(0.000063)	(0.000076)	(0.000079)	(0.000084)
Flagged Defaults 12		-0.001785			
		(0.006025)			
Flagged Defaults 24		0.016827			
		(0.010995)			
Flagged Defaults 36		-0.019482			
		(0.010060)			
Flagged Defaults 84		0.019841			
		(0.003856)			
Disaster Defaults 12			0.000099	-0.000193	-0.000989
			(0.000175)	(0.000171)	(0.000189)
Disaster Defaults 24			-0.000200	0.000075	0.001939
			(0.000255)	(0.000228)	(0.000240)
Disaster Defaults 36			-0.001029	-0.000908	-0.001553
			(0.000232)	(0.000202)	(0.000190)
Disaster Defaults 84			-0.000173	0.000109	0.000234
			(0.000110)	(0.000103)	(0.000099)

Notes: BTCCP data. Table shows average marginal effects (standard errors in parenthesis) from coefficients from logistic regressions predicting any new defaults in the next 24 months. Each column shows results from a separate regression. Column 1 shows the marginal effects from the θ_1 coefficients on the default terms ($D'_{i,t}$) from Equation 4. Column 2 shows the marginal effects from the θ_2 coefficients on the default terms ($D'_{i,t}$) and the π coefficients on the flagged default terms ($F'_{i,t}$) from Equation 5. Column 3 shows the marginal effects from the θ_3 coefficients on the default terms ($D'_{i,t}$) and also the ϕ coefficients on disaster default terms ($N'_{i,t}$) from Equation 8, when disaster defaults are measured as within 3 months of a disaster. Column 4 changes this definition of disaster defaults to be within 6 months of a disaster, and column 5 changes this definition of disaster defaults to be within 12 months of a disaster. Marginal effects are calculated using data on 6.99 million consumers (the 30% testing sample) to October 2017. The default terms used as predictors measure the number of accounts in a time window that a consumer is 30 or more days past due. The suffixes on predictors 12, 24, 36, 84 respectively denotes the number of accounts delinquent in the last 12, 24, 36, and 84 months. Logistic regressions also use a control variable a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs.

Table A3: Difference-in-Differences Estimates (s.e.) of Effects of Disaster Flags on New Account Openings in Month First Applied ($t = 0$) and after Twelve Months ($t = 12$)

Sample	(1) t=0	(2) t=12
A. All	-0.0117	-0.0289
	(0.0029)	(0.0037)
B. No Defaults	-0.0116	-0.0294
	(0.0028)	(0.0037)
C. Any Defaults	-0.0128	-0.018
	(0.0067)	(0.0072)
D. Non-Subprime	-0.0096	-0.0267
	(0.0021)	(0.0029)
E. Subprime	-0.0224	-0.0404
	(0.0073)	(0.0101)
F. Near Prime	-0.0196	-0.046
	(0.0047)	(0.0047)
G. Prime	-0.0104	-0.0349
	(0.0032)	(0.0044)
H. Prime Plus	-0.0084	-0.023
	(0.0024)	(0.0035)
I. Superprime	-0.0045	-0.0135
	(0.0012)	(0.0019)

Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Table show estimates of Equation 7's δ_τ with standard errors in parentheses. δ_τ are the coefficients on the interaction between event time indicator, relative to the month a consumer first has a disaster flag, and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. The outcome for all regressions in this table is the number of new account openings (across auto loans, credit cards, mortgages, and unsecured personal loans) in a month. Each column shows results at different time horizons. Each row shows results for a separate regression. Row 'A. All' uses data for all 3.8 million consumers. Rows B to I are for subsamples of consumers by their pre-disaster characteristics at $t - 12$. Row 'B. No Defaults' uses data for the 95.6% (3.06 million) consumers without any pre-disaster defaults. Row 'C. Any Defaults' uses data for the 4.4% (0.14 million) consumers with any pre-disaster defaults. Row 'D. Non-Subprime' uses data for the 83.8% (2.68 million) consumers with non-subprime (601 to 850) pre-disaster credit scores. Row 'E. Subprime' uses data for the 16.2% (0.52 million) consumers with subprime (300 to 600) credit scores. Row 'F. Near Prime' uses data for the 15.6% (0.50 million) consumers with near prime (601 to 660) credit scores. Row 'G. Prime' uses data for the 17.8% (0.57 million) consumers prime (661 to 720) credit scores. Row 'H. Prime Plus' uses data for the 22.5% (0.72 million) consumers with prime plus (721 to 780) credit scores. Row 'I. Superprime' uses data for the 27.9% (0.89 million) consumers with superprime (891 to 850) credit scores.

Table A4: Difference-in-Differences Estimates (s.e.) of Effects of Disaster Flags in Month First Applied ($t = 0$) by Pre-Disaster Credit Score

Sample	(1) Credit Score	(2) Number New Credit Cards	(3) Value New Credit Card Limits
A. Subprime	10.18	-0.0071	-11.5
	(0.92)	(0.0017)	(5.1)
B. Near Prime	3.3	-0.0071	-19.9
	(0.22)	(0.0015)	(9.3)
C. Prime	1.86	-0.0049	-21.9
	(0.17)	(0.0018)	(13)
D. Prime Plus	0.91	-0.0045	-41.7
	(0.11)	(0.002)	(20.3)
E. Superprime	0.43	-0.0028	-35.6
	(0.1)	(0.0009)	(13.6)

Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Table show estimates of Equation 7's δ_0 with standard errors in parentheses. δ_0 are the coefficients on the interaction between event time indicator, in the month a consumer first has a disaster flag, and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each column shows results from separate regressions with different outcomes. The outcome in column (1) is credit score (VantageScore 3.0). The outcome in column (2) is the number of new credit card opened that month. The outcome in column (3) is the value of new credit card limits opened that month (\$). Rows A to E are for subsamples of consumers by their pre-disaster credit scores at $t - 12$. Row 'A. Subprime' uses data for the 16.2% (0.52 million) consumers with subprime (300 to 600) credit scores. Row 'B. Near Prime' uses data for the 15.6% (0.50 million) consumers with near prime (601 to 660) credit scores. Row 'C. Prime' uses data for the 17.8% (0.57 million) consumers prime (661 to 720) credit scores. Row 'D. Prime Plus' uses data for the 22.5% (0.72 million) consumers with prime plus (721 to 780) credit scores. Row 'E. Superprime' uses data for the 27.9% (0.89 million) consumers with superprime (891 to 850) credit scores.

Table A5: Difference-in-Differences Estimates (s.e.) of Effects of Disaster Flags After Twelve Months ($t = 12$) by Pre-Disaster Credit Score

Sample	(1) Credit Score	(2) Number New Credit Cards	(3) Value New Credit Card Limits
A. Subprime	-2.74	-0.016	-27.7
	(1.18)	(0.0031)	(9.6)
B. Near Prime	-2.67	-0.0235	-80.2
	(0.48)	(0.0028)	(13.5)
C. Prime	-0.13	-0.0221	-121.3
	(0.19)	(0.0031)	(21.6)
D. Prime Plus	1.15	-0.0178	-137.9
	(0.18)	(0.0026)	(21)
E. Superprime	0.97	-0.0096	-120
	(0.19)	(0.0016)	(19.4)

Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Table show estimates of Equation 7's δ_{12} with standard errors in parentheses. δ_{12} are the coefficients on the interaction between event time indicator, in the twelfth month after a consumer first has a disaster flag, and an indicator for the flagged cohorts. The omitted category is event time $\tau = -1$. Regressions estimated also contain fixed effects for cohort-group and group-by-event-time. 95% confidence intervals shown where standard errors are clustered at the group-level. Each column shows results from separate regressions with different outcomes. The outcome in column (1) is credit score (VantageScore 3.0). The outcome in column (2) is the number of new credit card opened that month. The outcome in column (3) is the value of new credit card limits opened that month ($). Rows A to E are for subsamples of consumers by their pre-disaster credit scores at $t - 12$. Row 'A. Subprime' uses data for the 16.2% (0.52 million) consumers with subprime (300 to 600) credit scores. Row 'B. Near Prime' uses data for the 15.6% (0.50 million) consumers with near prime (601 to 660) credit scores. Row 'C. Prime' uses data for the 17.8% (0.57 million) consumers prime (661 to 720) credit scores. Row 'D. Prime Plus' uses data for the 22.5% (0.72 million) consumers with prime plus (721 to 780) credit scores. Row 'E. Superprime' uses data for the 27.9% (0.89 million) consumers with superprime (891 to 850) credit scores.

Table A6: Means for Outcome (Any New Defaults in Next 24 Months) and Default Variables

	Means
1. Outcome	0.168874
2. All Defaults 12	0.940522
3. All Defaults 24	1.222895
4. All Defaults 36	1.467395
5. All Defaults 84	2.242902
6. Flagged Defaults 12	0.004877
7. Flagged Defaults 24	0.005147
8. Flagged Defaults 36	0.005267
9. Flagged Defaults 84	0.005891
10. Disaster (03) Defaults 12	0.288569
11. Disaster (03) Defaults 24	0.410655
12. Disaster (03) Defaults 36	0.472400
13. Disaster (03) Defaults 84	0.831895
14. Disaster (06) Defaults 12	0.341319
15. Disaster (06) Defaults 24	0.467467
16. Disaster (06) Defaults 36	0.548176
17. Disaster (06) Defaults 84	0.978757
18. Disaster (12) Defaults 12	0.453650
19. Disaster (12) Defaults 24	0.579803
20. Disaster (12) Defaults 36	0.688245
21. Disaster (12) Defaults 84	1.202234

Notes: BTCCP data. Calculated using data on 6.99 million consumers (the 30% testing sample) to October 2017. In this table row 1 'Outcome' is the outcome being predicted, a binary variable for any new defaults, measured by 90 or more days past due that were not delinquent as of October 2017, in the next 24 months. Other rows show means for predictive inputs as measures of defaults. Rows 2 to 5 show the mean number of accounts in default in the last 12, 24, 36, and 84 months respectively. Rows 6 to 9 show the mean number of accounts in default, after masking delinquent account months with disaster flags, in the last 12, 24, 36, and 84 months respectively. Rows 10 to 13 show the mean number of accounts in default, after masking delinquent account months within 3 months of a natural disaster, in the last 12, 24, 36, and 84 months respectively. Rows 14 to 17 do the same but masking delinquent account months within 6 instead of 3 months of a natural disaster. Rows 18 to 21 do the same but masking delinquent account months within 12 instead of 6 months of a natural disaster.

Table A7: Consumers with Any Defaults Across Models Varying Masking of Defaults

Model	Any Defaults	Change from Baseline
1. All Defaults	52.209%	
2. Masked Flagged Defaults	52.199%	-0.020%
3. Masked Disaster (03) Defaults	51.412%	-1.528%
4. Masked Disaster (06) Defaults	50.583%	-3.115%
5. Masked Disaster (12) Defaults	47.975%	-8.110%
6. CARES Masked Disaster (03) Defaults	51.427%	-1.498%
7. CARES Masked Disaster (06) Defaults	50.665%	-2.958%
8. CARES Masked Disaster (12) Defaults	48.710%	-6.702%
9. PCARES Masked Disaster (03) Defaults	50.071%	-4.096%
10. PCARES Masked Disaster (06) Defaults	47.941%	-8.176%
11. PCARES Masked Disaster (12) Defaults	43.429%	-16.818%
12. No Defaults	0.000%	-100.000%

Notes: BTCCP data. Calculated using data on 6.99 million consumers (the 30% testing sample) to October 2017. In this table any defaults is measured by the fraction of consumers that have any accounts that are 30 or more days past due in the last 84 months. Row 1 is the baseline model that includes all defaults. Row 2 masks defaults that have disaster flags. Rows 3, 4, and 5 masks defaults within three, six, and twelve months of a disaster are masked. Rows 6, 7, and 8 show 'CARES' that temporarily keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. Rows 9, 10, and 11 show 'PCARES' that permanently keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. Row 12 excludes all defaults.

Table A8: Additional Measures of Predictive Performance of Models Varying Inclusion of Flagged Defaults and Disaster Defaults

Model	Accuracy	Balanced Accuracy	Precision	Recall	True Positive	True Negative
1. All Defaults	0.875008	0.718415	0.900711	0.954870	0.954870	0.481959
2. Flagged Defaults	0.875506	0.717824	0.900425	0.955923	0.955923	0.479725
3. Disaster (03) Defaults	0.875499	0.718364	0.900633	0.955637	0.955637	0.481091
4. Disaster (06) Defaults	0.875496	0.718135	0.900545	0.955750	0.955750	0.480519
5. Disaster (12) Defaults	0.875489	0.718177	0.900562	0.955717	0.955717	0.480637

Notes: BTCCP data. Receiver operating characteristic curves (ROC) from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Row 1 is the baseline model that includes defaults information the prediction. Row 2 is performance of a model that includes inputs for defaults that have disaster flags. Rows 3, 4, and 5 show performance from models that include inputs for defaults within three, six, and twelve months of a disaster. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables.

Table A9: Additional Measures of Predictive Performance of Models Varying Masking of Defaults

Model	Accuracy	Balanced Accuracy	Precision Precision	Recall Recall	True Positive	True Negative
1. All Defaults	0.875008	0.718415	0.900711	0.954870	0.954870	0.481959
2. Masked Flagged Defaults	0.874940	0.717854	0.900503	0.955054	0.955054	0.480655
3. Masked Disaster (03) Defaults	0.874129	0.715893	0.899845	0.954830	0.954830	0.476957
4. Masked Disaster (06) Defaults	0.873855	0.715866	0.899866	0.954429	0.954429	0.477303
5. Masked Disaster (12) Defaults	0.872749	0.711754	0.898418	0.954857	0.954857	0.468652
6. CARES Masked Disaster (03) Defaults	0.874032	0.715593	0.899741	0.954836	0.954836	0.476350
7. CARES Masked Disaster (06) Defaults	0.873820	0.715408	0.899694	0.954610	0.954610	0.476205
8. CARES Masked Disaster (12) Defaults	0.872908	0.713050	0.898896	0.954436	0.954436	0.471664
9. PCARES Masked Disaster (03) Defaults	0.873521	0.714204	0.899267	0.954773	0.954773	0.473634
10. PCARES Masked Disaster (06) Defaults	0.872810	0.712150	0.898563	0.954746	0.954746	0.469553
11. PCARES Masked Disaster (12) Defaults	0.871998	0.709790	0.897754	0.954724	0.954724	0.464856
12. No Defaults	0.870593	0.701915	0.894925	0.956618	0.956618	0.447212

Notes: BTCCP data. Predictive performance measures from out-of-sample prediction from XGBoost models predicting any new default in next 24 months. Performance using data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Each row shows performance under different models that vary the use of defaults data used as inputs. Row 1 is the baseline model that includes all defaults. Row 2 masks defaults that have disaster flags. Measures of defaults in rows 3, 4, and 5 masks defaults within three, six, and twelve months of a disaster are masked. 'CARES' measures of defaults in rows 6, 7, and 8 temporarily keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. 'PCARES' measures of defaults in rows 9, 10, and 11 permanently keep accounts at their pre-disaster delinquency status for three, six, and twelve months from a disaster starting. Row 12 does not use any defaults information. Predictions use as inputs uses a credit score built on the training sample using XGBoost machine learning algorithm with 171 non-default variables as inputs, along with these 171 non-default variables.