Disaster Flags: Credit Reporting Relief from Natural Disasters

Benedict Guttman-Kenney*

30th July, 2025

Abstract

I study how informative delinquency signals are during natural disasters, and evaluate the consequences of masking them. Between 2015 and 2024, 68 million consumers had a "disaster flag" on their U.S. credit report. These disaster flags aim to provide temporary relief to consumers exposed to natural disasters but do not appear to increase consumers' credit access. I quantify the equity-efficiency trade-off of a counterfactual policy, which automatically masks all delinquencies in credit reports for all consumers exposed to natural disasters. This policy masks 2% to 33% of all delinquencies 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, Johannes Stroebel, 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, Shan Ge, Stephanie Johnson, Stephanie Moulton, Tess Scharlemann, Theresa Kuchler, Thomas Covert, Walter Zhang, Chicago Booth microeconomics and finance groups, industry participants, Consumer Financial Protection Bureau, Federal Reserve Bank of Philadelphia Consumer Finance Institute, FinRegLab, International Conference on Credit Risk Evaluation, NYU Stern Climate Finance Conference, American Finance Association Annual Meeting, and Association for Public Policy Analysis & Management 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). For consumer credit markets, it is ambiguous whether missing one or more payments ("delinquencies") on credit products during a natural disaster ("disaster delinquencies") is an informative signal of a consumer's credit risk, or contains signals that differ in their informativeness compared to non-disaster delinquencies.

In this paper, I assess the informativeness of delinquencies during natural disasters, and the potential to provide relief to consumers exposed to natural disasters through masking such delinquencies in credit reports. I do so in two ways. First, I document and evaluate the existing voluntary system of natural disaster credit reporting relief developed by the market. Second, I quantify the efficiency costs of an alternative system that could provide more equitable credit reporting relief from natural disasters.

I show how lenders currently respond to natural disasters by voluntarily 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. These flags temporarily mask negative information (i.e., delinquencies) about a consumer in the calculation of their VantageScore credit score but are ignored in the calculation of their FICO credit score. The only other study on this topic is a short report by the Consumer Financial Protection Bureau (Banko-Ferran and Ricks, 2018) on Hurricane Harvey that concludes that "more analysis is needed to better understand whether and how the furnishing [reporting/sharing] of information on natural disasters affects consumer credit". My paper addresses this research gap by providing a comprehensive study of this topic.

I use a 10% representative sample of monthly U.S. 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 un-

[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. See Botzen et al. (2019) for a broader review of the impacts of natural disasters.

til 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 an individual credit report tradeline (account) 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 delinquencies that also have disaster flags, "flagged delinquencies", compared to "unflagged delinquencies", delinquencies that do not also have disaster flags. I find that flagged delinquencies are slightly riskier signals than unflagged delinquencies. Although a model predicting a new credit delinquency that masks flagged delinquencies 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 temporarily mask delinquencies in VantageScore.

I evaluate the potential equity-efficiency trade-off of temporarily masking all delinquencies in credit reports for all consumers that reside in areas affected by natural disasters. Whether to mask the delinquencies that occur during natural disasters depends on how informative this information is in predicting future delinquencies. If disaster delinquencies are highly predictive, then masking this information is expected to be costly for lenders and reduce the market efficiency of lending. If disaster delinquencies offer limited predictive value, then a social planner may consider it worthwhile to equitably mask this information, as it would indicate that delinquencies arising during natural disasters are not informatively revealing a consumer's risk type. Such place-based policies may also be motivated by redistributive objectives (e.g., Gaubert et al., 2025), given the large and persistent geographic inequalities in credit access and financial distress across the U.S. (e.g., Keys et al., 2022; Bakker et al., 2025). 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 delinquencies 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 delinquencies of consumers exposed to natural disasters would mask 2% to 33% of all delinquencies of consumers with U.S. credit reports. I estimate the efficiency cost from masking all disaster delinquencies would be to reduce the ability to predict, as measured by the area under the receiver operating characteristic curve (AUC), a new credit delinquency by 0.1% to 0.5%, with more generous policies leading to larger predictive losses. This result is robust to other measures of predictive performance. These efficiency losses can be benchmarked relative to a model in which all delinquencies (i.e., irrespective of whether during a disaster or not) are masked, which would reduce the AUC by 1.0%. My results quantify the equity-efficiency trade-off that policymakers face. I find that offering more generous relief masking delinquencies for a longer period of time increasingly reduces predictive performance, and the costs depend on how delinquencies that initially occur during a disaster are recorded in credit reports after the disaster. The subset of consumers who have disaster delinquencies masked experience improved credit scores, while other consumers generally experience little-to-no changes in their credit scores. These results lead to redistribution of credit scores across geography.

My research informs the literature on the linkages between household finance, environmental, and public economics. A growing literature studies the effects of natural 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., 2025a; Del Valle 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 about an alternative policy that automatically provides credit reporting relief from natural disasters to inform recent public policy discussions (e.g., Banko-Ferran and Ricks, 2018; National Consumer Law Center, 2019; Urban Institute, 2019; FinRegLab, 2020). There is active policy interest in this topic. For example, following the damage caused by Hurricanes Helene and Milton, in October 2024 a coalition of 43 organizations wrote to regulators to urge them to encourage the lenders that they regulate to "refrain from supplying negative information" for consumers "within a presidentially declared disaster area" (National Consumer Law Center, 2024). This followed similar letters in 2023, 2022, and 2019 in response to earlier natural disasters. Separately, following the California wildfires in January 2025, lenders agreed to not report negative information on mortgages for 90 days (Office of Governor Gavin Newsom, 2025, also see California Bill AB 238 (Harabedian) proposed in June

2025). As discussed in FinRegLab (2020), bills have been proposed at the U.S. Federal level to prevent adverse information being reported during and after natural disasters.[2] Before my research paper, there was no evidence on how many consumers would be impacted by such a policy, the costs to lenders of temporarily masking this information, and how these would vary depending on the detail of how long delinquent information is masked for under such a policy.

My third contribution is to the literature on household finance that 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., 2025; Dinerstein et al., 2024; Katz, 2025; Kim et al., 2024). Much of this literature is focused on two crisis periods: the Great Recession and the COVID-19 pandemic, whereas my focus on natural disaster relief complements these studies by studying a broader time period. My contribution to this literature is to add evidence on a new complementary policy tool. Masking delinquencies 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 delinquencies appears in credit reports.

My fourth contribution is to advance the literature on the economics of credit information, reviewed in Gibbs et al. (2025). The 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., 2025), historical delinquencies (e.g., Bos and Nakamura, 2014; Bos et al., 2018; Liberman et al., 2020; Blattner et al., 2022), the implications of delayed credit market entry (e.g., Brown et al., 2019; Cookson et al., 2025b), and using alternative data sources (e.g., Berg et al., 2020; Chioda et al., 2025; Blattner and Nelson, 2024). 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; Guttman-Kenney et al., 2022), the credit risk value of information on medical debt in collections (Brevoort and Kambara, 2015; Duarte et al., 2025), and the implications of removing this information (Kluender et al., 2025). Information on medical debt is increasingly removed from credit reports and credit scores, partially driven by fairness concerns that this information, relative to other debts, arguably may arise more from bad luck rather than being an

[2] FinRegLab (2020): "Some policy-makers have pushed for a more categorical approach that would prohibit inclusion of negative information in consumer reports altogether. Such an approach could be structured in different ways, but the primary focus of debate at the federal level has been S. 3508 (Schatz & Brown) and H.R. 6321 (Waters & Sherman). Both bills were introduced prior to the CARES Act and would apply not only to financial hardships from the Covid-19 pandemic but to future major disasters". Urban Institute (2019) conclude a need for consideration of "Rules and guidance around how natural disasters and subsequent delinquencies are identified on consumers' credit reports and incorporated into credit scores."

informative negative signal of a consumer's type (Mahoney, 2025). A somewhat similar fairness concern may apply to the disaster delinquencies that I study.

I contribute to this literature by evaluating the predictive value of flagged and disaster delinquencies compared to non-disaster delinquencies. A consumer's credit report contains information on delinquencies from the last seven years, and therefore it is important to consider the implications of such information being shared as a delinquency can have persistent negative impacts on a consumer's access to credit. 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 U.S. credit reports, as described in Cherry et al. (2021), with similar policies implemented in other countries. The nature of the COVID-19 pandemic that widely disrupted all parts of society with many contemporaneous policies makes it challenging to evaluate individual policies, and the literature has not established the implications of preventing lenders from updating adverse information in credit reports. Understanding the implications of different information being shared is therefore increasingly important to study (e.g., Guttman-Kenney and Shahidinejad, 2025). Correlated shocks, such as delinquencies from natural disasters, can pose a particular challenge for lenders and the design of credit information markets. This is because a wave of delinquencies is a correlated shock to lenders' capital, and for consumers, it may persistently worsen the affected consumers' credit scores in a way that is not reflective of their future credit risk (or in the case of the COVID-19 pandemic lead to a wave of non-delinquencies that improve scores), and therefore impede efficient credit allocation.

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, suffering from poor health, or experiencing 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 fiscal 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).

A final broader implication of my research is to inform the literature on the determinants of defaults. Recent literature increasingly emphasizes the role of liquidity-driven rather than strategic defaults (e.g., Ganong and Noel, 2023; Low, 2023). My finding that much information on past delinquencies can be relatively uninformative, beyond other sources of information in credit reports, in predicting new delinquencies provides an independent source of evidence that appears consistent with this literature. It suggests that combining non-delinquency credit reporting data with other information sources that align more closely with consumers' time-varying liquidity needs (e.g., checking account data such as that studied in Babina et al., 2025; Guttman-Kenney et al., 2025) may be especially valuable for predicting defaults and help to allocate credit more efficiently, with more research into this being valuable to advance our understanding of the predictors of defaults.

The paper proceeds as follows. Section 2 provides a motivating framework and explains the data that I use. Section 3 provides institutional background on disaster flags, documents five new facts on their use, and considers the implications for consumers' credit access. Section 4 shows the characteristics of consumers with disaster flags and examines the information value of masking delinquencies. In Section 5, I examine the equity-efficiency trade-off of from automatically masking all delinquencies during natural disasters. Finally, Section 6 concludes.

2 Motivating Framework and Data

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

2.1 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 delinquency (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/delinquency, i.e., lower credit risk. Equation 1 shows a simple example, where a credit score predicts at time t, an outcome, $Y_{i,t+1}$, the likelihood that the consumer i will delinquency on a credit agreement one period in the future.

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

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 delinquency component, $d_{i,t}$, which takes a value of one if delinquencies and zero otherwise, and a vector of all other non-delinquency 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., 2025 for more general information on credit reporting data and credit scoring).

In such credit scoring models, past delinquencies are a strong predictor of future delinquencies with $\theta_1 > 0$. Consumers with past delinquencies 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 delinquency 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 delinquencies differ with natural disasters (e.g., wildfires, floods, hurricanes). Understanding the informativeness of delinquencies during natural disasters is important for lenders, as these events generate correlated delinquencies that can be a risk to their capital, whereas the risks of idiosyncratic delinquencies can be more easily diversified, also may be more informative of consumer types. Does the ability to predict future delinquencies vary depending on whether a delinquency occurs during a natural disaster or not? Equation 2 allows for this by including an interaction term between the binary delinquency term, $d_{i,t}$, and a binary variable, $N_{g(i,t),t}$, which takes a value of one if the consumer i resides in a geographical area g at time t where and when that area is exposed to a natural disaster, and if not, takes a value of zero.

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

The value of the π parameter in Equation 2 is informative of the marginal predictive value of delinquencies during natural disasters, "disaster delinquencies", compared to non-disaster delinquencies. It may be that $\pi < 0$, meaning that disaster delinquencies are lower risk than non-disaster delinquencies. 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 delinquency. In contrast, it may be that $\pi > 0$, meaning that disaster delinquencies may be higher risk than non-disaster delinquencies, 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 delinquencies from non-disaster delinquencies may not be informative to improve credit risk prediction. If disaster delinquencies are uninformative noise, then predictive performance may even be improved by masking such information. This theoretical ambiguity motivates my empirical analysis.

Given this framework, I can quantify the efficiency loss for the credit industry to mask disaster delinquencies in credit reporting data by adapting Equation 1 to Equation 3, where disaster delinquencies are masked to be recorded as not delinquent.

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

Comparison of the predictive performance of these two models can be informative. If the difference in predictive performance is small, the political economy may mean that the credit industry may even voluntarily agree to mask disaster delinquencies. 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 delinquency 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. Taking this framework to data enables me to quantify the potential equity-efficiency trade-off that policymakers face; masking delinquencies for more consumers is more equitable, however, may come at a cost of lower efficiency of evaluating credit risk. Although my study examines natural disaster delinquencies, this framework could be applied to evaluate other characteristics with richer data merged in, e.g., masking delinquencies linked to life events such as divorce, income shocks, or expenditure shocks.

2.2 Data

2.2.1 Consumer Credit Reporting Data

This research uses a large, anonymized, representative sample of U.S. 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., 2025 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. (2025).

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.

2.2.2 Natural Disasters Data

When a major disaster occurs, it is declared as such by the U.S. 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.[3] I restrict

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

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 Flags

I describe the institutional details of disaster flags in Section 3.1, document five new facts describing the use of disaster flags in the US in Section 3.2, and consider their implications for consumers in Section 3.3.

3.1 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.[4] 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 delinquencies in credit reports. Discussions with industry participants indicate that some lenders sometimes do not report new delinquencies during natural disasters, however, it is unclear how frequently such non-reported delinquencies are as they are, by definition, unobserved in credit reporting data.[5] Disaster flags may be ap-

[4]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

[5]Although not the focus of this study, such non-reporting of delinquencies may help to explain why the average effects of natural disasters on delinquencies, 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

plied instead of or in addition to changes in contract terms (e.g., deferring payments or offering forbearance) that may also be recorded in credit reports.[6]

Disaster flags mask negative information only on the flagged tradeline in the calculation of VantageScore credit scores.[7] 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".[8] Manual underwriters that review a consumer's credit report can observe disaster flags, along with other information including delinquencies and deferrals, 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 seven to ten years after bankruptcy, studied in many prior papers (e.g., Musto, 2004; Dobbie et al., 2020; Gross et al., 2020; Jansen et al., 2025). 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. One notable economic difference between the information hidden by these is that bankruptcy flag removal removes old information, seven or more years old, whereas disaster flag application masks recent information, which one might expect to be highly predictive and informative of repayment behaviors in the next few months or years.

3.2 Disaster Flag Flags

3.2.1 FACT 1. Growth in the use of disaster flags.

68.1 million consumers in the U.S. 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

the delinquency status in credit reports for their mortgages if they delinquency when affected by natural disasters, this was removed in 2020 following the CARES Act taking effect.

[6] 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

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

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

tradeline in their consumer credit report for at least one month.[9] 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 U.S. between 2015 and 2024.[10] 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.

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.

3.2.2 FACT 2. Broad geographic usage of flags.

The panels of Figure 2 display the fraction of consumers in U.S. 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 examples, 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, the use of disaster flags follows a path corresponding to the impacts of Hurricanes Milton and Helene.

[9] 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).

[10] 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.

3.2.3 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.

3.2.4 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 the 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, shown in Internet Appendix Figure A3.

3.2.5 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 there are frictions in the use of disaster flags, consistent with Kim et al. (2024) who show intermediation frictions in COVID-19 forbearance being applied to mortgages. Although I 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.

3.3 Implications for Consumers

Putting these institutional details, especially that disaster flags are ignored by FICO, together with facts 3, 4, and 5 suggest that disaster flags, in their current form, are unlikely to increase consumer's credit access. Internet Appendix B provides detailed evidence consistent with this explanation. I find no evidence of consumers' credit access, measured by the number of new account openings, increasing after disaster flags are applied. This finding holds across three methodologies and alternative ways to measure credit access. It holds among the subset of the most financially-distressed consumers, that experience temporary gains in VantageScore, who would have the largest potential to benefit from delinquencies being masked by disaster flags. Instead, there is some evidence of reduced credit access, that is potentially most consistent with flags sometimes being interpreted as a negative signal of a consumer's credit risk.

4 Information Costs of Disaster Flags Masking Delinquencies

Having documented how disaster flags are used, I now evaluate the information on consumers' 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 delinquencies, 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 delinquencies, 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, consumers flagged in 2017 typically have higher credit scores, whereas consumers 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 2.1 to my data to evaluate the information costs of disaster flags that mask delinquencies 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 U.S. 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 that my inputs and my outcome are not affected by the non-reporting of delinquencies that occurred during COVID-19. My outcome, Y_{t+24}, is a binary outcome for whether a consumer has any *new* delinquency in the 24 months after October 2017, measured as 90 or more days past due, regarded as a threshold for severe delinquency or default. A new delinquency is one where a tradeline was not in delinquency in October 2017 but is in delinquency 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 delinquency. I calculate my models in two steps. First, I construct a credit score ($S_{i,t}$) using only the vector of non-delinquency variables as inputs. Second, I use this credit score and evaluate how varying information on delinquencies as inputs to improves predictive performance.

In the first step, I want to capture predictive information in credit reports except for information on delinquencies. I therefore build a credit score without delinquency information using 171 non-delinquency variables as predictors. This uses 138 consumer-level attributes available in my dataset that do not use delinquencies information (and this does *not* include VantageScore credit score). I construct 33 additional consumer-level variables. 31 of these 33 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 an XGBoost (extreme gradient boosting) machine learning method to predict delinquency.[11] This is a methodology that has been previously shown to be highly predictive of delinquency in consumer credit applications (e.g., Blattner and Nelson, 2024; Fuster et al., 2022; Duarte et al., 2025), and is also used in industry (FinRegLab et al., 2022). This model without delinquency information predicts delinquencies

[11] 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.

well, with an area under the receiver operating characteristic curve (AUC) of 0.8902 (Table 4), where an AUC of 0.5 would mean that there is no information in the classification (e.g., a naive model that predicts all consumers would default) and an AUC of 1 would classify cases perfectly.

In the second step, I use either a logistic regression approach, enabling me to interpret coefficients on delinquency variables, or an XGBoost machine learning methodology to better assess the ability of delinquencies to improve credit risk prediction. I construct a vector of delinquency 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 delinquency over the last k months. These are calculated from tradeline-level data and aggregated to the consumer-level. $X_{i,t}$ denotes the non-delinquency 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 6, 12, 24, 36, and 84 months, as well my credit score variable created in the first step, and the 171 non-delinquency variables. I do not include information beyond 84 months because 84 months is the maximum duration for which delinquencies remain in credit reports under existing U.S. law (Gibbs et al., 2025).

My baseline predictive model is shown in Equation 4. This is a traditional credit score that includes historical delinquency information.

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

My second model is shown in Equation 5:

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

This second model adds a vector of "flagged delinquency" 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 delinquencies and disaster flags. For constructing $F'_{i,t}$, each consumer's tradeline delinquency status each month is assigned a value of one only if it is both in delinquency 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 delinquency and with disaster flags, e.g., $f^{12}_{i,t}$ denotes the number of accounts flagged delinquencies 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 delinquency variable.

My third model, "masked flagged delinquencies", shown in Equation 6, adjusts the input data to mask all delinquencies that also have disaster flags.

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

In Equation 6, 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 delinquencies" as not delinquent (see Appendix Table A3 for means of these variables). Comparing the predictive performance, measured by the AUC, of this third model with the baseline model shows the information costs of masking flagged delinquencies. I use AUC because it is a standard way to evaluate credit scoring models in the literature (e.g., Berg et al., 2020; Fuster et al., 2022; Blattner and Nelson, 2024; Chioda et al., 2025; Duarte et al., 2025), and also show robustness to a variety of other measures of predictive performance.

4.3 Predictive Results

Figure 6 shows the average marginal effects of an account in delinquency, the θ_2 in black, and of a flagged account in delinquency, 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 being in delinquency in the past increases the risk of a consumer being delinquent in the future. For flagged delinquencies in the last twelve months, the average marginal effects of $\pi > 0$, and are much greater than those of θ_2. There is some evidence that flagged delinquencies are economically and statistically significantly signals of higher risk. The average marginal effects of an additional account in delinquency in the last 84 months is 0.001266 (s.e. 0.000063) whereas the average marginal effects of an additional flagged account in delinquency 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 delinquency in the last 12, 24, and 36 months are all insignificant from zero, so the evidence is noisy. Masking flagged delinquencies reduces the number of delinquencies in the last 84 months by 0.03% (Table 3) with a 0.02% reduction in the number of consumers with any delinquency in the last 84 months (Internet Appendix Table A4).

Table 4 compares the out-of-sample predictive performance from XGBoost models measured by the AUC from the baseline credit risk model, 0.890236, to one masking flagged delinquencies, 0.890233. Masking flagged delinquencies therefore reduces predictive performance by a trivial amount 0.0003% (Table 4), whereas including flagged delinquencies 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 that are shown in Internet Appendix Tables A5 and 5: Accuracy, balanced accuracy, precision, true positive rate (also known as recall), true negative rate, F1 score ($2 \times \frac{Precision \times Recall}{Precision + Recall}$), and Brier score ($\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2$, and therefore the square root of the Brier score is the root mean squared error, RMSE).

Such an economically small information cost of flags masking delinquencies can help to understand why lenders voluntarily temporarily apply disaster flags. Given that this practice is not costly to lenders, and there is a lack of regulatory guidance on how to respond to disasters, it may arise for political economy reasons. It is one way that lenders can signal to government officials or regulators that lenders are supporting their customers, and is cheaper to the lender than loan modifications that would reduce their profits. Given a small cost, there could be small private benefits that would make such actions net beneficial to lenders. For example, voluntarily temporarily masking information may be one of many ways a lender aims to strengthen their relationship with their customers, making them more loyal and potentially increasing lifetime profits.

5 Equity-Efficiency Trade-Off From Masking All Disaster Delinquencies

The existing voluntary regime of disaster flags appears to have limited costs to lenders, shown in Section 4, without increasing consumer credit access, as summarized in Section 3.3 and shown in Internet Appendix B, for the selected subset of consumers with flags. In this section, I consider a feasible alternative form of credit reporting relief from natural disasters, with the methodology explained in Section 5.1 and the results presented in Section 5.2.

5.1 Methodology

I quantify the loss of information from a counterfactual government policy that automatically requires the masking of all delinquencies during disasters ("disaster delinquencies"). This counterfactual is designed to apply equitably to all consumers who are subject to a disaster: removing selection. Doing so pools consumers affected by disasters with those unaffected by disasters.[12] Applying such an automatic flagging approach removes

[12] Practically implementing this would mean disaster delinquencies are masked *before* they appear on credit reports so that disaster delinquencies (or disaster flags) cannot be observed by lenders in their credit decisions. This could be required of firms sharing information with credit reporting agencies or could

frictions for lenders and consumers. Such a counterfactual policy has been proposed to regulators by a coalition of 43 consumer organizations (National Consumer Law Center, 2024), however, the costs to lenders of removing such information from credit reports have not been estimated, nor how these costs vary depending on how generously such a policy implemented. I study a policy applied at the national-level, however, policies could also be applied at the state-level by their legislatures, as has occurred with medical debt reporting (Gibbs et al., 2025).

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.[13] 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. Blickle et al. (2022) shows that disasters increase loan demand, which offsets losses and actually increases profits at larger banks. A larger predictive loss from masking disaster delinquencies 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 delinquencies with disaster flags, I now vary the inclusion of information on all delinquencies that occur in areas and time periods exposed to natural disasters. I examine the marginal effects of a vector of "disaster delinquencies", $N_{i,t}^{\prime j} = [n_{i,t}^{12,j}, n_{i,t}^{24,j}, n_{i,t}^{36,j}, n_{i,t}^{84,j}]'$, in the regression model specified in Equation 7:

$$Pr(Y_{i,t+24} = 1) = f\left(X_{i,t}'\beta_4 + D_{i,t}'\theta_4 + N_{i,t}^{\prime j}\phi\right) \tag{7}$$

This regression specified in Equation 7 uses the same vectors of delinquencies, $D_{i,t}' = [d_{i,t}^{12}, d_{i,t}^{24}, d_{i,t}^{36}, d_{i,t}^{84}]'$, and non-delinquency variables, $X_{i,t}'$, as the earlier Equation 4. Disaster delinquencies are constructed from the interactions between delinquencies and natural disasters, over the same time horizons as $D_{i,t}$ as previously studied in Section 4.2. For constructing $N_{i,t}^{\prime j}$, I first construct a binary variable taking a value of one only if a con-

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 delinquencies, as enabled by flagging disaster delinquencies rather than non-reporting disaster delinquencies, 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.

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

sumer, i, has any delinquency 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 $N_{i,t}^{\prime j}$ for the number of disaster delinquency 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 delinquencies. This means that, for example, $n_{i,t}^{12,03}$ denotes the number of accounts that are both in delinquency and exposed to a natural disaster in the last twelve months, where disaster delinquencies are classified as such if they occur within three months of a FEMA event. The coefficients on $N_{i,t}^{\prime j}$ in a logistic model reveal whether disaster delinquencies are different from other delinquencies in being an informative signal of a consumer's future credit risk type.

Choosing a j threshold of six months would ensure that any delinquency during that six-month period only affects subsequent credit access if the delinquency is still present after six months. This limits the ability of temporary adverse shocks to propagate and have long-term impacts through credit reporting histories. 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 encouraging consumers to be less prepared for a disaster, or strategically defaulting during/after a disaster.[14] Important context for this concern is that the current literature shows that very few U.S. consumers appear to be adequately prepared for disasters in the existing information environment (e.g., Beatty et al., 2019; Dinerstein et al., 2025), and especially unprepared are low-income households because credit constraints limit their ability to prepare (e.g., van der Straten, 2025). The moral hazard concern would be expected to be more of a concern with a twelve-month threshold and less of a concern with a three-month threshold. This is because a potential undesirable outcome would be consumers in high-risk locations repeatedly exposed to disasters having their delinquencies perennially masked. One might even argue that a consumer temporarily becoming strategically delinquent on their debt during a natural disaster could increase welfare if it enables other consumption, given that natural disasters are a 'bad' economic state of the world, where the marginal utility of consumption is

[14]Recent empirical evidence shows the importance of life events rather than strategic defaults for driving delinquencies (e.g., Ganong and Noel, 2023; Low, 2023). 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). Even with delinquencies being temporarily masked, there remain other strong non-credit reporting financial incentives for consumers to repay debt on time that are expected to reduce moral hazard, such as the increased borrowing costs through late fees and additional interest, 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.

high. Consumers often move after a disaster (e.g., Gallagher and Hartley, 2017; Bleemer and van der Klaauw, 2019), and when they do, their credit history moves with them. This means that a potential benefit of an alternative policy masking disaster defaults could be to make consumers less constrained in their credit access, thus assisting them in moving away from a disaster area and in accessing credit in their new location, rather than their historic disaster defaults following them.

Analogously to my earlier masking of flagged delinquencies in Equation 6, masking disaster delinquencies takes the form shown in Equation 8, where the term $\tilde{D}'^{j}_{i,t}$ masks disaster delinquencies, based on the policy threshold j, which masks delinquencies based on whether exposed to a disaster in the one of three, six, or twelve months from a FEMA event.

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_5 + \tilde{D}'^{j}_{i,t}\theta_5\Big) \tag{8}$$

I examine how masking disaster delinquencies affects predictive performance using my XGBoost machine learning approach and varying the masking of disaster delinquencies. I benchmark the models' performance with masked delinquencies to the baseline model with all delinquencies, shown in Equation 4, and also to a model without delinquencies information, displayed in Equation 9.

$$Pr(Y_{i,t+24} = 1) = f\Big(X'_{i,t}\beta_6\Big) \tag{9}$$

5.2 Results

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

What happens when I mask disaster delinquencies? Rows three to five of Table 3 shows that using j thresholds of three to six to twelve months for masking disaster delinquencies reduces the mean number of accounts with delinquencies in the last 84 months by 2.38% to 5.16% to 14.03% respectively, and reduces the number of consumers with any delinquency in the last 84 months by 1.5% to 3.1% to 8.1% respectively, as shown in Internet Appendix Table A4. These are economically large removals of data from credit reports.

Figure 7 shows the ROC curves for masking disaster delinquencies ("Masked Disaster Delinquencies") under different duration thresholds. It is difficult to distinguish the different lines, though there is a clear gap to the no delinquencies model shown in orange, showing how some delinquencies information is clearly valuable for prediction. Table 4 shows how much predictive performance, measured by the 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 delinquencies. 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 delinquencies can be benchmarked against a counterfactual policy masking all delinquencies, i.e., disaster and non-disaster delinquencies, that significantly reduces predictive performance with the 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, precision, true positive rate (also known as recall), and true negative rate) as shown in Internet Appendix Table 5.

These reductions in predictive performance may appear small relative to the amount of information masked, and therefore policymakers may consider automatically masking disaster delinquencies 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 the 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 the AUC predicting non-mortgage delinquency from 0.835 to 0.879, an improvement of 5.0%.

An alternative way to implement a policy masking disaster delinquencies could be to adapt the approach used in the U.S. by the Coronavirus Aid, Relief, and Economic Security (CARES) Act of 2020, only masking delinquencies if the delinquency status worsens during the disaster. This "CARES Masked Disaster Delinquencies" approach means that accounts that were already in delinquency pre-disaster remain delinquent, whereas in the "Masked Disaster Delinquencies" approach they were temporarily masked. Rows six to eight of Table 3 (and Internet Appendix Table A4) show that this masks slightly less information, reducing mean delinquencies (relative to our all delinquencies baseline) by 2.33%, 4.88%, and 11.40% under three, six, and twelve month thresholds, than our earlier "Masked Disaster Delinquencies" approach. The changes in the 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, with ROC curves in Internet Appendix Figure A8.

However, the above approaches only temporarily mask delinquencies, masking delinquencies if they are present within j periods of a FEMA disaster. However, delinquencies 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 delinquencies that arise during natural disasters would affect prediction in my "PCARES Masked Disaster Delinquencies" approach. In this approach, I take each account's delinquency status in the month preceding a disaster, and mask the delinquencies on the accounts if their account status worsens during the disaster threshold, and, unlike the CARES approach, I keep masking the delinquency status on account months after the disaster.[15] I leave delinquency statuses of accounts that are opened after a disaster threshold unaffected, and the delinquency status of accounts whose status did not worsen during a disaster, and, of course, the delinquency 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 delinquency status may be earlier due to the consumer experiencing repeated disasters.

This "PCARES Masked Disaster Delinquencies" approach may be considered an upper bound on the amount of information that a policy masking disaster delinquencies may mask. Rows 9 to 11 of Table 3 show that this masks 10.3%, 18.7%, and 32.6% of delinquencies, depending on whether three, six, or twelve month thresholds are used, and 4%, 8%, and 17% fewer consumers have any delinquency in the last seven years (Internet Appendix Table A4). Rows 9 to 11 of Table 4 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 in Internet Appendix Figure A8). I quantify the trade-off that policymakers face, masking delinquencies for a longer period of time increasingly reduces predictive

[15] For example, an account not in delinquency at $t = -1$, and in delinquency 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 delinquencies at $t = 0$ but not $t + 1$, whereas the PCARES policy would mask delinquencies at both $t = 0$ and $t = +1$.

performance, and the costs depend upon how delinquencies that initially occur during disasters are recorded in credit reports after the disasters.

Predictive performance can be evaluated in a variety of ways. One may be interested in alternative measures of performance to AUC given that our outcome is a minority case with 16.9% becoming delinquent. Another reason for exploring alternative performance measures is that a lender may be more concerned with not classifying consumers that default as being high-risk than it may be about classifying consumers that do not default as being high-risk. I address such issues in Table 5 that shows that our results are robust to a variety of alternative ways to evaluate performance: accuracy (although this is a fairly uninformative measure given a low incidence of delinquency a naive model predicting delinquency for all consumers would have an "accuracy" of 0.8311), balanced accuracy (that addresses low incidence, so a naive model would yield 0.5), precision, true positive rate (also known as recall), true negative rate, F1 score, and Brier score.

Appendix Figure A9 presents the results by ventiles of credit scores that are calculated under the different models.[16] This ventile analysis shows that all of the credit scoring models are effective at differentiating consumers at risk of becoming delinquent, although the model without any delinquency information performs noticeably worse. Across all of the models, there are substantially higher delinquency rates for consumers predicted to be at higher risk. The three riskiest credit score ventiles contain over 50% of total delinquencies (i.e., delinquencies across all ventiles). As more delinquency information related to natural disasters is removed, these credit scores get worse at predicting delinquency, and the pattern of these results is consistent with our main results.

What are the distributional consequences of masking disaster delinquencies? Masking delinquencies may mean policymakers have to trade-off heterogeneous impacts across different types of consumers. It may improve credit access for consumers with masked delinquencies, who now appear less risky to lenders, but reduce credit access for those consumers without delinquencies, who instead now appear riskier having been pooled along consumers that have masked delinquencies.

I now study these distributional consequences of masking disaster delinquencies. Figure 8 shows the changes in probabilities of delinquencies, relative to the baseline model's predictions, for masking disaster defaults for six months, with Appendix Figure A11 showing this across different models. Figure 8 Panel A shows the changes in delinquencies by quintile of the baseline credit risk predictions. Masking disaster delinquencies leads to little change in predicted risk to low-risk consumers, being likely to have very

[16] Appendix Figure A10 shows that these models are well-calibrated with the predicted probability of delinquency closely aligning with the actual delinquency across these models.

small increases in their predicted credit risk. High-risk consumers are more likely to experience decreases in their predicted credit risk. The changes in delinquencies across credit score groups appear fairly small. Panel B shows that the changes in the predictions for the subgroup of consumers with any delinquency (measured without masking) are consistent with the credit score results.

Figure 8 Panel C shows remarkably little difference in the predictions for consumers irrespective of whether a consumer has experienced a natural disaster. Consistent with this, Panel D shows that fewer than 10% of consumers who have a disaster delinquency masked in the counterfactual, the subgroup expected to benefit the most from counterfactual policies, experience a 10 percentage point or larger decrease in the probability of delinquency. This is a group of consumers who benefit the most from decreases in their probability of delinquency, while other consumers generally experience no change, increases, or small decreases in their probabilities of delinquency.

Figure 9 shows the geographical implications of masking disaster delinquencies. Panel A shows the mean probabilities of delinquencies across counties in the baseline credit risk model, this is consistent with Bakker et al. (2025) using an industry credit score. Figure 9 Panels B to J show changes in these scores across different models, where darker green shows decreases in predicted credit risk (i.e., better credit scores) and black and darker green shows increase in risk (i.e., worse credit scores). Figure 9 shows that counties in the south-west of America, on average, experience reductions in predicted credit risk. The counties that experience increases in predicted credit risk are more scattered around the country, with a focus on the midwest and the north-east of America. Although, as discussed in the previous paragraph, the changes in such credit scores appear small relative to the amount of information masked.

A final way to evaluate these counterfactual models is to compare them with models that mask the same amount of delinquency information, but for a random selection of delinquent consumers. Appendix Figure A12 shows that random removal of more information on delinquencies fairly linearly reduces predictive performance. Counterfactual models that mask disaster delinquencies, shown in Table 4, produce slightly worse predictive performance than their respective model that randomly masks delinquency information, shown in Appendix Table A6. This finding is consistent with my other results that masking disaster defaults removes information that is predictive, but only marginally so.

6 Conclusions

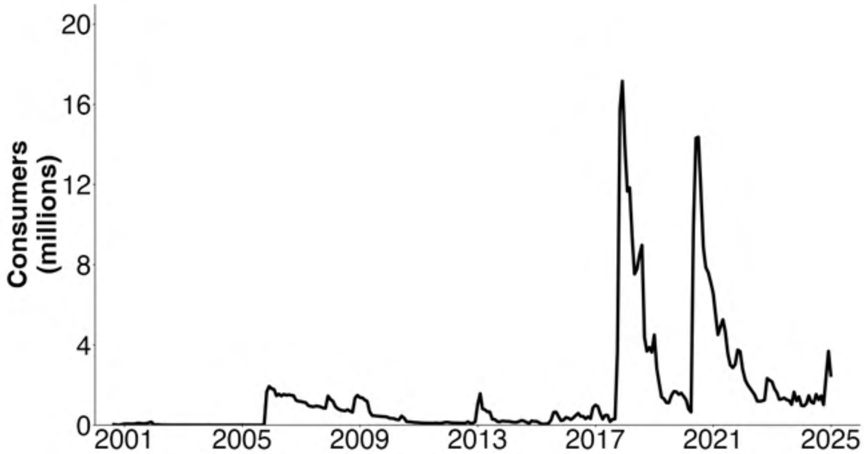
In this paper, I assess the informativeness of delinquencies during natural disasters, and the potential to provide relief to consumers exposed to natural disasters through masking such delinquencies in credit reports. I provide new facts documenting the widespread use of "disaster flags" on U.S. consumer credit reports. Disaster flags are intended to provide relief to consumers affected by natural disasters. I show that these flags are widely used, however, their institutional design makes them unlikely to increase credit access, and I find evidence consistent with this.

I quantify the trade-off of a counterfactual policy that automatically masks all delinquencies of consumers exposed to natural disasters. Doing so removes 2% to 33% of all delinquencies 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 equity-efficiency trade-off of masking disaster delinquencies can inform active policy discussions (e.g. National Consumer Law Center, 2024) on the design of credit information markets and how to alleviate financial distress from natural disasters.

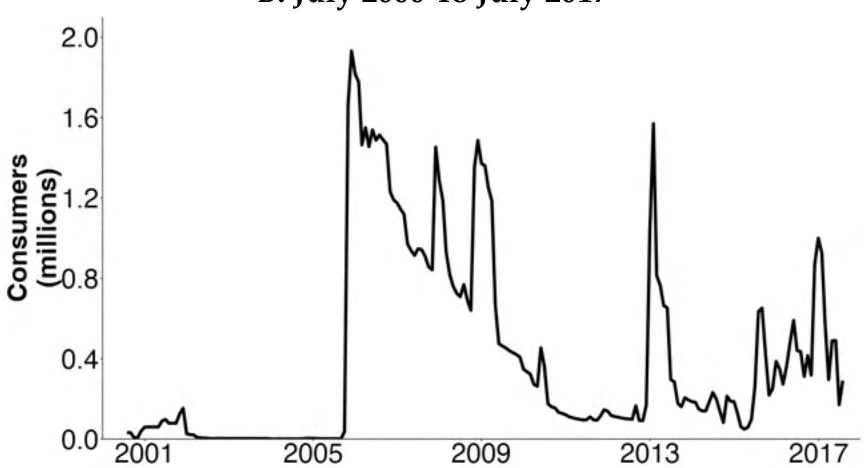
7 Figures and Tables

Figure 1: Consumers With Any Credit Report Disaster Flag

A. July 2000 To December 2024

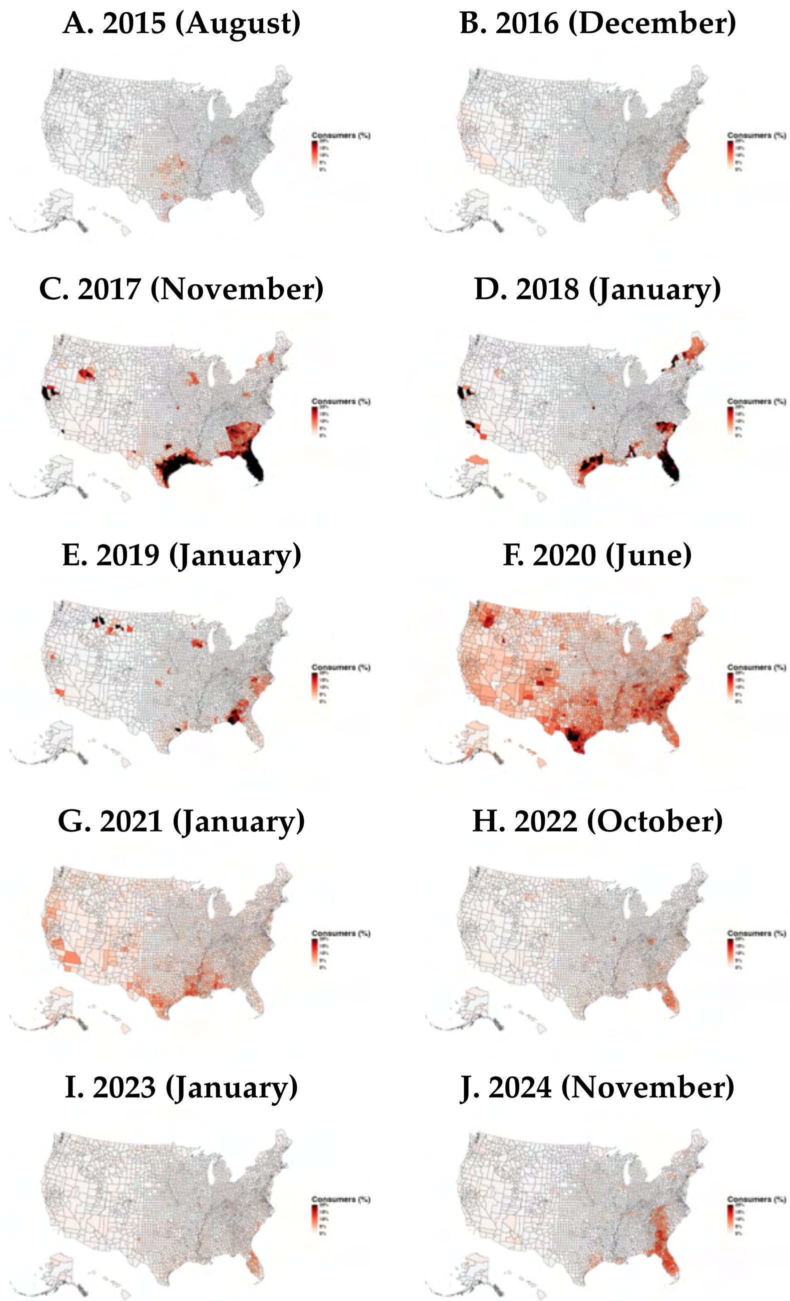


B. July 2000 To July 2017



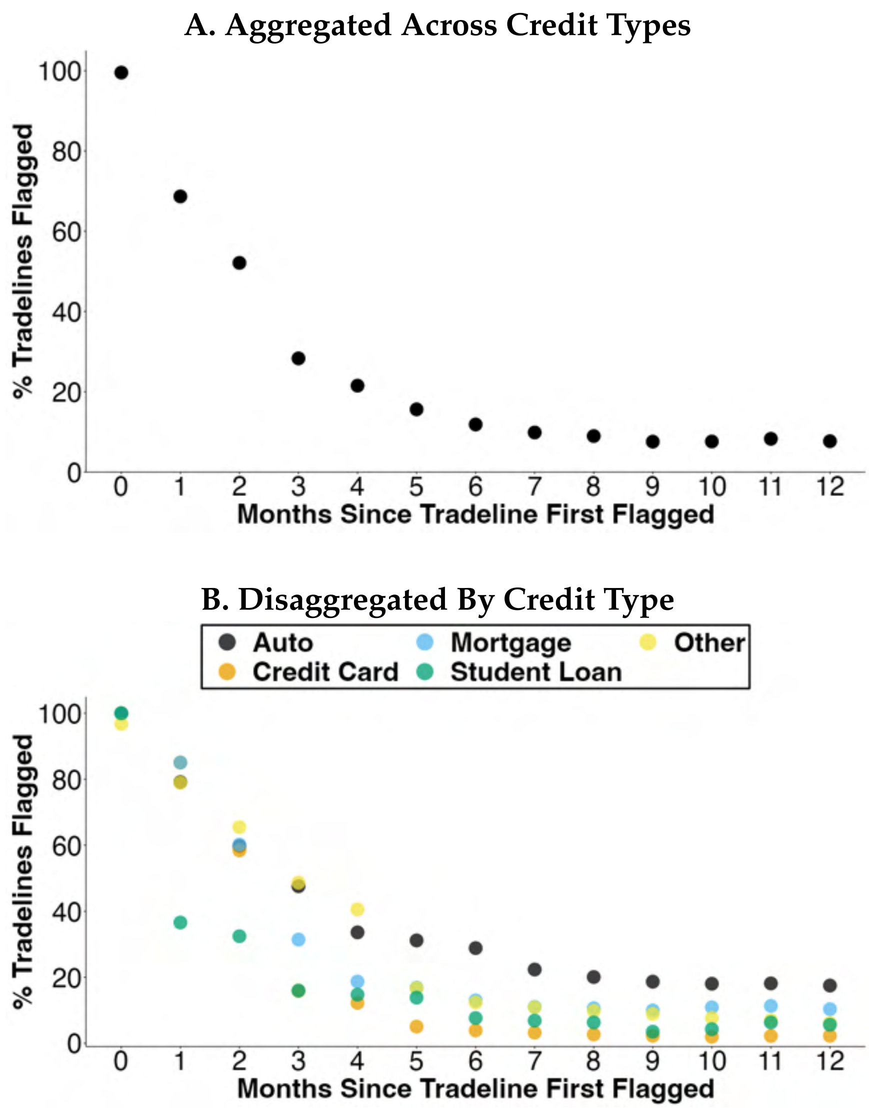
Notes: BTCCP data. Consumers with a credit report disaster flag on at least one open tradeline in their credit report. The number of consumers is extrapolated to population estimates from the BTCCP data's 10% sample of consumers.

Figure 2: Consumers (%) In A County With Any Disaster Flag



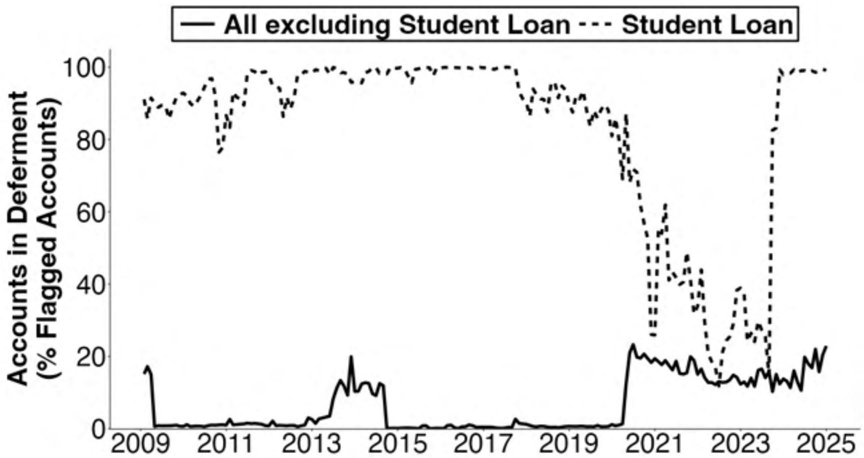
Notes: BTCCP data. Each panel shows the percentage of consumers in a county that have any disaster flag. The denominator in this calculation is the number of consumers with an open tradeline with a positive balance on their credit report in a county that month. The numerator in this calculation is the 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.

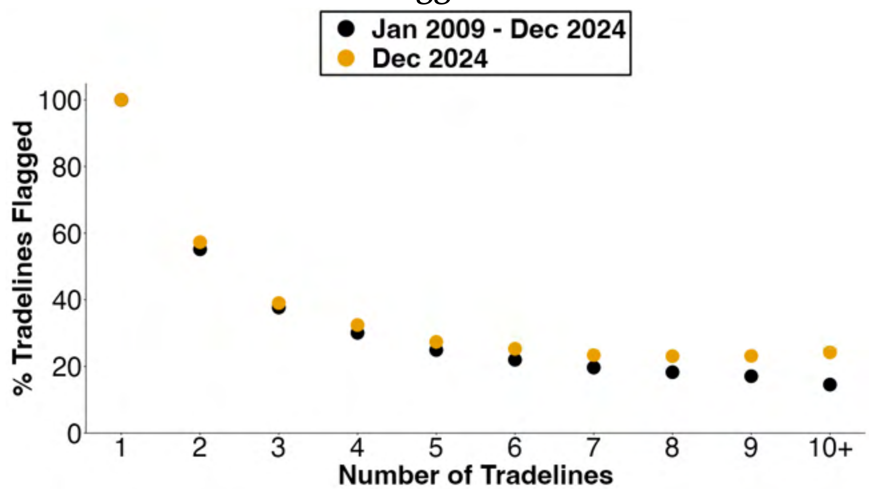
Figure 4: Tradelines With Credit Report Disaster Flags And Deferments Reported



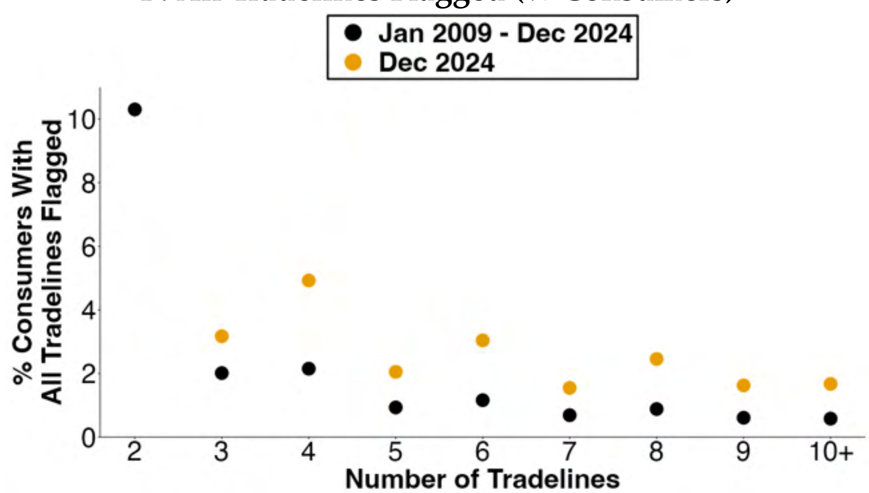
Notes: BTCCP data. This figure shows the fraction of accounts that have a disaster flag that also have a deferment reported. The denominator in this calculation is the number of open tradelines with a positive balance in their credit report with a credit report disaster flag. The numerator in this calculation 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

A. Tradelines Flagged (% Tradelines)



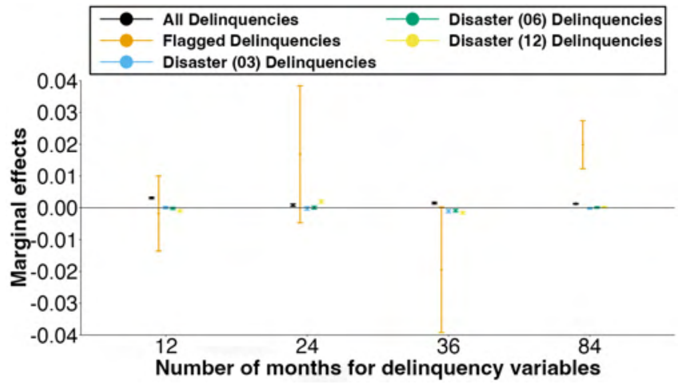
B. All Tradelines Flagged (% Consumers)



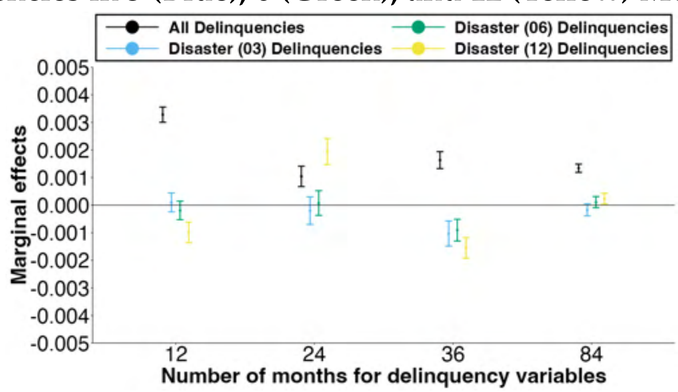
Notes: BTCCP data. Both Panels A and B restrict to consumer months where the consumer has a credit report disaster flag on at least one open tradeline with a positive balance in their credit report. Panel A shows the fraction of these consumers' tradelines that have a credit report disaster flag. Panel 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 Delinquencies Predicting Any New Delinquency

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

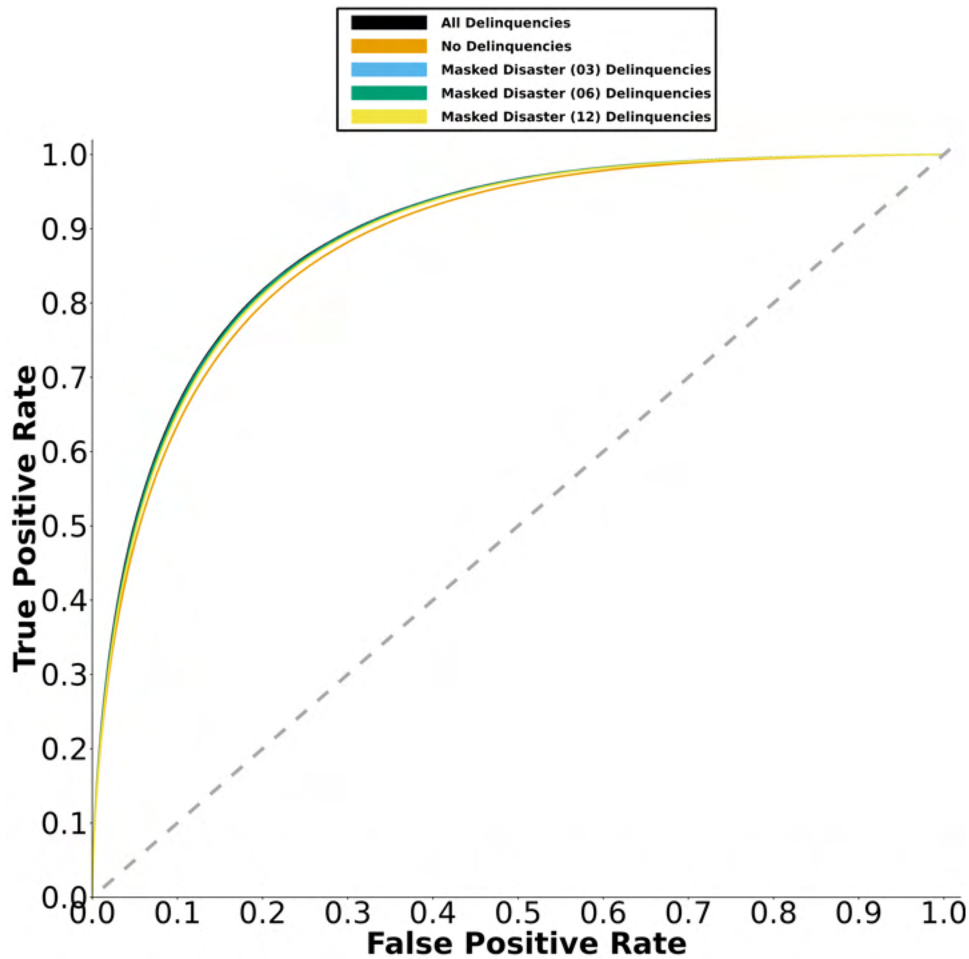


B. All Delinquencies (Black), Disaster Delinquencies in 3 (Blue), 6 (Green), and 12 (Yellow) Months of a Disaster



Notes: BTCCP data. The X axes show the predictive variables for delinquencies 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 delinquencies 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 delinquency terms ($D'_{i,t}$) from Equation 5 in Panel A (and the θ_4 coefficients on delinquency 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 delinquency 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 delinquency terms ($N'_{i,t}$) from Equation 7, using different definitions of disaster delinquencies: blue captures delinquencies 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 delinquency 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-delinquency variables as inputs.

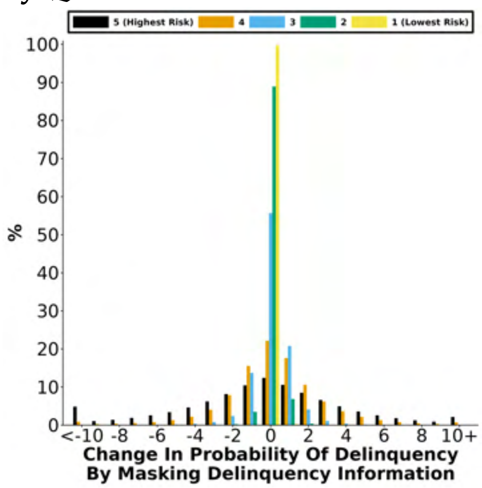
Figure 7: Receiver Operating Characteristic Curves (ROC) Showing The Predictive Performance For Models Masking Delinquency Information. Baseline Model Includes Information On Delinquency (Black), Compared To Models Masking Delinquency Over Three (Blue), Six (Green), and Twelve (Yellow) Months Of A Disaster, And A Model Including No Delinquency Information (Orange)



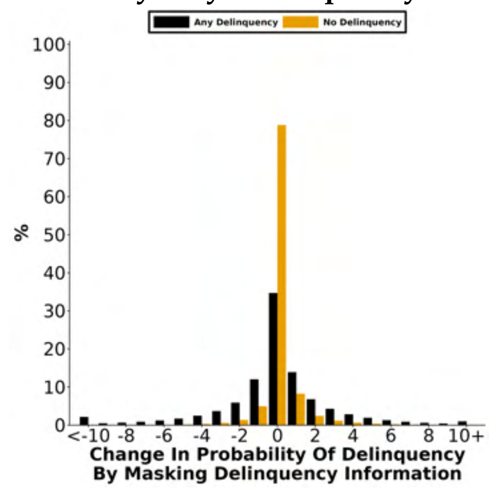
Notes: BTCCP data. Receiver Operating Characteristic Curves (ROC) from out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. The black line is the baseline model that includes delinquency information the prediction. The orange line respectively show performance from a model without delinquency information. The blue, green, and yellow lines show performance from models with delinquency information where delinquencies that occur within three (blue), six (green), and twelve (yellow) months of a disaster are masked. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables. True Positive Rate is defined as the fraction of consumers with predicted new delinquencies out of those with actual new delinquencies. False Positive Rate is defined as the fraction of consumers with predicted new delinquencies out of those with actual no new delinquencies. The dashed 45 degree line corresponds to an AUC of 0.5, as would be predicted by a naive model predicting default for all consumers or with random predictions.

Figure 8: Distribution Of Changes In Probability Of Delinquency, Calculated For CARES Model Masking Of Delinquency Information For Six Months, Relative To Baseline Model With Delinquency Information

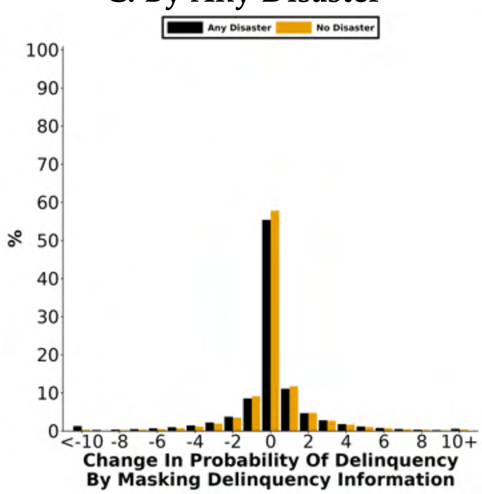
A. By Quintile of Baseline Credit Score



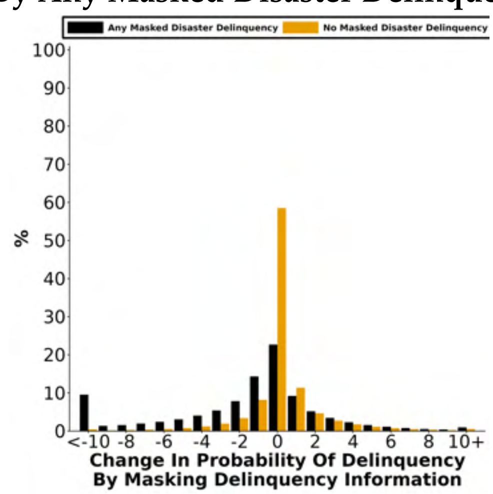
B. By Any Delinquency



C. By Any Disaster



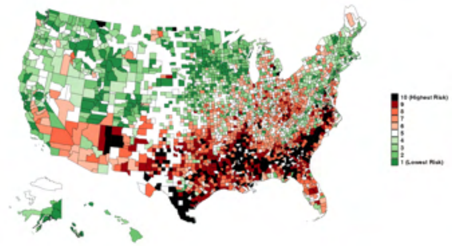
D. By Any Masked Disaster Delinquency



Notes: BTCCP data. Each panel shows the distributions of percentage point changes in the predicted probability of delinquency, under the CARES model masking disaster delinquencies for six months, compared to the baseline model with all delinquency information. Panel A shows results by quintile of baseline credit score, where group 1 is the lowest credit risk and 5 is the highest. Panel B shows results by whether a consumer has any delinquency. Panel C shows results by whether a consumer has any disaster. Panel D shows results by whether a consumer has any disaster delinquency that is masked in the counterfactual. Each bar shows 1 percentage point changes in probability, with the <10 bar capturing less than -9.5 percentage points changes, and the 10+ bar capturing greater than 9.5 percentage point changes. Probabilities of delinquency are calculated using out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Uses data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

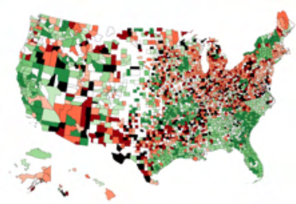
Figure 9: Geography of Probability Of Delinquency

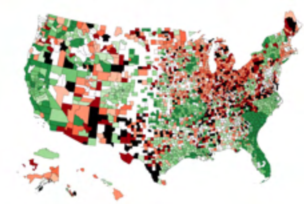
A. Credit Risk in Baseline Model



Changes in Credit Risk from Baseline Model to Models Masking Disaster Delinquencies:

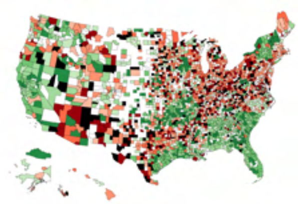
B. Masked Disaster (03) **C. Masked Disaster (06)** **D. Masked Disaster (12)**





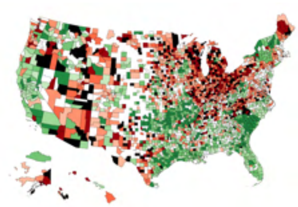
Changes in Credit Risk from Baseline Model to CARES Models:

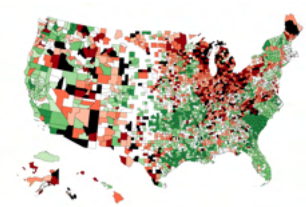
E. Masked Disaster (03) **F. Masked Disaster (06)** **G. Masked Disaster (12)**



Changes in Credit Risk from Baseline Model to PCARES Models:

H. Masked Disaster (03) **I. Masked Disaster (06)** **J. Masked Disaster (12)**





Notes: BTCCP data. Panel A shows the mean probability of delinquency in a county under the baseline model that contains all delinquency information. Each county is colored by decile of probability of delinquency, with a darker green showing the lowest credit risk counties and lower probability of delinquency (i.e., higher credit score) and a darker red showing the highest credit risk counties. Panels A to J show the changes in mean credit score in a county, relative to this baseline, under alternative models masking disaster delinquencies. In Panels A to J each county is colored by decile of changes in probability of delinquency, with a darker green showing the counties experiencing the largest decreases in mean credit risk counties (i.e., credit score increases) and a darker red showing the counties experiencing the largest increases in mean credit risk (i.e., credit score decreases). Probabilities of delinquency are calculated using out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Uses data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables. Counties with less than 100 observations in the testing sample are excluded.

Table 1: Summarizing Consumers With (1) Disaster Flags Compared To (2) Unflagged In Same Census Block Group × 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+ Delinquency (%)	9	5	6
30+ Delinquency (#)	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	124,168	60,068	59,624
Auto Balances	10645	5,333	4,610
Credit Card Balances	5,795	3,148	3,002
Mortgage Balances	95,433	46,286	47,477
Non Mortgage Balances	28,735	13,782	12147
Credit Card Limits	23,525	15,294	14,539

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 × 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 × zip code where other consumers did not have disaster flags at the same time.

Table 2: Predictive Performance Of Models Varying Inclusion Of Flagged Delinquencies And Disaster Delinquencies

Model	AUC	Change from Baseline
1. All Delinquencies	0.890236	
2. Flagged Delinquencies	0.890904	0.0751%
3. Disaster (03) Delinquencies	0.890927	0.0777%
4. Disaster (06) Delinquencies	0.890905	0.0751%
5. Disaster (12) Delinquencies	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 delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. Row 1 is the baseline model that includes delinquency information the prediction. Row 2 is performance of a model that includes inputs for delinquencies that have disaster flags. Rows 3, 4, and 5 show performance from models that include inputs for delinquencies within three, six, and twelve months of a disaster. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

Table 3: Quantity Of Delinquency Information Masked Across Models

Model	Mean Delinquencies	Change from Baseline
1. All Delinquencies	2.243	
2. Masked Flagged Delinquencies	2.242	-0.033%
3. Masked Disaster (03) Delinquencies	2.190	-2.375%
4. Masked Disaster (06) Delinquencies	2.127	-5.162%
5. Masked Disaster (12) Delinquencies	1.928	-14.026%
6. CARES Masked Disaster (03) Delinquencies	2.191	-2.327%
7. CARES Masked Disaster (06) Delinquencies	2.134	-4.875%
8. CARES Masked Disaster (12) Delinquencies	1.987	-11.396%
9. PCARES Masked Disaster (03) Delinquencies	2.011	-10.332%
10. PCARES Masked Disaster (06) Delinquencies	1.824	-18.663%
11. PCARES Masked Disaster (12) Delinquencies	1.511	-32.637%
12. No Delinquencies	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 delinquencies 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 delinquencies. Row 2 masks delinquencies that have disaster flags. Rows 3, 4, and 5 masks delinquencies 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 delinquencies.

Table 4: Predictive Performance Of Models Varying Masking Of Delinquencies

Model	AUC	Change from Baseline
1. All Delinquencies	0.890236	
2. Masked Flagged Delinquencies	0.890233	-0.0003%
3. Masked Disaster (03) Delinquencies	0.889206	-0.1157%
4. Masked Disaster (06) Delinquencies	0.888825	-0.1585%
5. Masked Disaster (12) Delinquencies	0.887044	-0.3586%
6. CARES Masked Disaster (03) Delinquencies	0.889176	-0.1191%
7. CARES Masked Disaster (06) Delinquencies	0.888824	-0.1586%
8. CARES Masked Disaster (12) Delinquencies	0.887381	-0.3207%
9. PCARES Masked Disaster (03) Delinquencies	0.888244	-0.2238%
10. PCARES Masked Disaster (06) Delinquencies	0.887343	-0.3249%
11. PCARES Masked Disaster (12) Delinquencies	0.885713	-0.5081%
12. No Delinquencies	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 delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. Each row shows performance under different models that vary the use of delinquencies data used as inputs. Row 1 is the baseline model that includes all delinquencies. Row 2 masks delinquencies that have disaster flags. Measures of delinquencies in rows 3, 4, and 5 masks delinquencies within three, six, and twelve months of a disaster are masked. 'CARES' measures of delinquencies 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 delinquencies 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 delinquency information. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

Table 5: Additional Measures Of Predictive Performance For Models Varying Masking Of Delinquency Information

Model	Accuracy	Balanced Accuracy	Precision	True Positive Rate	True Negative Rate	F1 Score	Brier Score
1. All Delinquencies	0.875008	0.718415	0.900711	0.954870	0.481959	0.927000	0.089813
2. Masked Flagged	0.874940	0.717854	0.900503	0.955054	0.480655	0.926977	0.089824
3. Masked Disaster (03)	0.874129	0.715893	0.899845	0.954830	0.476957	0.926522	0.090316
4. Masked Disaster (06)	0.873855	0.715866	0.899866	0.954429	0.477303	0.926345	0.090457
5. Masked Disaster (12)	0.872749	0.711754	0.898418	0.954857	0.468652	0.925778	0.091168
6. CARES Masked Disaster (03)	0.874032	0.715593	0.899741	0.954836	0.476350	0.926470	0.090324
7. CARES Masked Disaster (06)	0.873820	0.715408	0.899694	0.954610	0.476205	0.926339	0.090470
8. CARES Masked Disaster (12)	0.872908	0.713050	0.898896	0.954436	0.471664	0.925834	0.091052
9. PCARES Masked Disaster (03)	0.873521	0.714204	0.899267	0.954773	0.473634	0.926189	0.090714
10. PCARES Masked Disaster (06)	0.872810	0.712150	0.898563	0.954746	0.469553	0.925803	0.091082
11. PCARES Masked Disaster (12)	0.871998	0.709790	0.897754	0.954724	0.464856	0.925363	0.091683
12. No Delinquencies	0.870593	0.701915	0.894925	0.956618	0.447212	0.924744	0.092819

Notes: BTCCP data. Predictive performance measures from out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. Each row shows performance under different models that vary the use of delinquencies data used as inputs. Row 1 is the baseline model that includes all delinquencies. Row 2 masks delinquencies that have disaster flags. Measures of delinquencies in rows 3, 4, and 5 mask delinquencies within three, six, and twelve months of a disaster are masked. 'CARES' measures of delinquencies 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 delinquencies 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 delinquency information. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

Generative AI Disclosure: During the preparation of this work the author used ChatGPT, Gemini, and WriteFull in order to improve language and readability. After using these tools/services, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

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*.

Babina, T., Bahaj, S., Buchak, G., De Marco, F., Foulis, A., Gornall, W., Mazzola, F., and Yu, T. (2025). Customer data access and fintech entry: Early evidence from open banking. *Journal of Financial Economics*, 169:103950.

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.

Bakker, T. J., DeLuca, S., English, E. A., Fogel, J., Hendren, N., and Herbst, D. (2025). Credit access in the United States. *NBER Working Paper No. 34053*.

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.

Beatty, T. K., Shimshack, J. P., and Volpe, R. J. (2019). Disaster preparedness and disaster response: Evidence from sales of emergency supplies before and after hurricanes. *Journal of the Association of Environmental and Resource Economists*, 6(4):633–668.

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.

Berg, T., Burg, V., Gombović, A., and Puri, M. (2020). On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies*, 33(7):2845–2897.

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.

Bos, M. and Nakamura, L. I. (2014). Should defaults be forgotten? evidence from variation in removal of negative consumer credit information. *Federal Reserve Bank of Philadelphia Working Paper No. 14-21*.

Botzen, W. W., Deschenes, O., and Sanders, M. (2019). The economic impacts of natural disasters: A review of models and empirical studies. *Review of Environmental Economics and Policy*, 13(2):167–341.

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):73–97.

Brown, J. R., Cookson, J. A., and Heimer, R. Z. (2019). Growing up without finance. *Journal of Financial Economics*, 134(3):591–616.

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.

Chioda, L., Gertler, P., Higgins, S., and Medina, P. C. (2025). Fintech lending to borrowers with no credit history. *Working Paper*.

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. (2025a). Money to burn: Crowdfunding wildfire recovery. *Working Paper*.

Cookson, J. A., Guttman-Kenney, B., and Mullins, W. (2025b). Immigration and credit in America. *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.

Del Valle, A., Scharlemann, T., and Shore, S. (2024). Household financial decision-making after natural disasters: Evidence from hurricane harvey. *Journal of Financial and Quantitative Analysis*, 59(5):2459–2485.

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. (2025). Second chance: Life with less student debt. *Journal of Finance*, Forthcoming.

Dinerstein, M., Lucas, N., Nath, I., and Rayl, J. (2025). Private markets and public assistance for natural disaster supplies. *AEA Papers and Proceedings*, 115:385–390.

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.

Duarte, V., Fonseca, J., Kohli, D., and Reif, J. (2025). The effects of deleting medical debt from consumer credit reports. *NBER Working Paper No. 33644*.

Dube, A., Girardi, D., Jorda, O., and Taylor, A. M. (2025). A local projections approach to difference-in-differences event studies. *Journal of Applied Econometrics*, Forthcoming.

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. (2025). Place-based redistribution. *American Economic Review*, Forthcoming.

Gibbs, C., Guttman-Kenney, B., Lee, D., Nelson, S., van der Klauuw, W., and Wang, J. (2025). Consumer credit reporting data. *Journal of Economic Literature*, 63(2):598–636.

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., Adams, P., Hunt, S., Laibson, D., Stewart, N., and Leary, J. (2025). The semblance of success in nudging consumers to pay down credit card debt. *American Economic Journal: Economic Policy*, 17(4).

Guttman-Kenney, B., Kluender, R., Mahoney, N., Wong, F., Xia, X., and Yin, W. (2022). Trends in medical debt during the covid-19 pandemic. *JAMA Health Forum*, 3(5):e221031.

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. (2025). Data and welfare in credit markets. *Journal of Financial Economics*, Forthcoming.

Katz, J. (2025). 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):1–15.

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. (2025). The effects of medical debt relief: Evidence from two randomized experiments. *Quarterly Journal of Economics*, 140(2):1187–1241.

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.

Mahoney, N. (2025). Credit score creep. *Neale's Substack*.

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*.

National Consumer Law Center (2024). 2024 letter urging regulators to protect credit scores of survivors of hurricanes helene and milton. *Public Letter, October 31, 2024*.

Office of Governor Gavin Newsom (2025). Governor newsom announces commitments from state banks and credit unions to provide mortgage relief for firestorm survivors. *Press Release, January 23, 2025*.

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*.

van der Straten, Y. (2025). Flooded house or underwater mortgage? the macrofinancial implications of climate change and adaptation. *Working paper*.

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*.

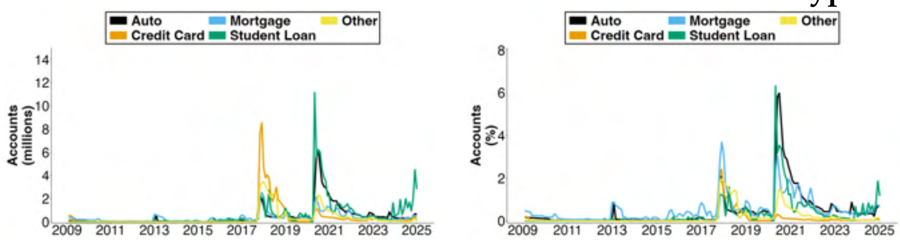
Wing, C., Freedman, S. M., and Hollingsworth, A. (2024). Stacked difference-in-differences. *NBER Working Paper No. 32054*.

Internet Appendix: *“Disaster Flags: Credit Reporting Relief from Natural Disasters”*

A. Supplementary Figures and Tables

Figure A1: Trades With Credit Report Disaster Flag By Credit Type, 2009 - 2024

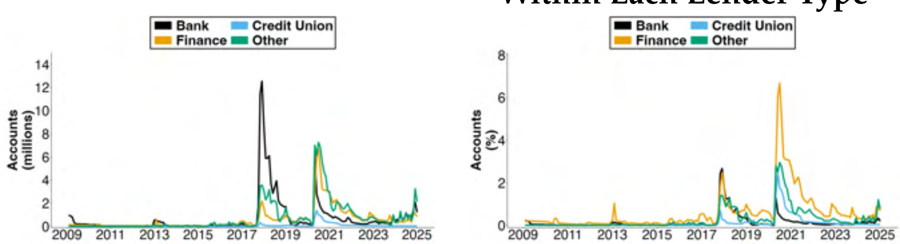
A. Number Of Accounts **B. Percent Of Accounts Within Each Credit Type**



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

Figure A2: Trades With Credit Report Disaster Flag By Lender Type, 2009 - 2024

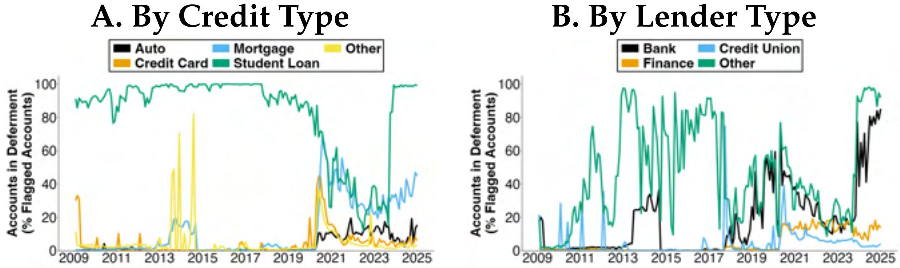
A. Number Of Accounts **B. Percent Of Accounts Within Each Lender Type**



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.

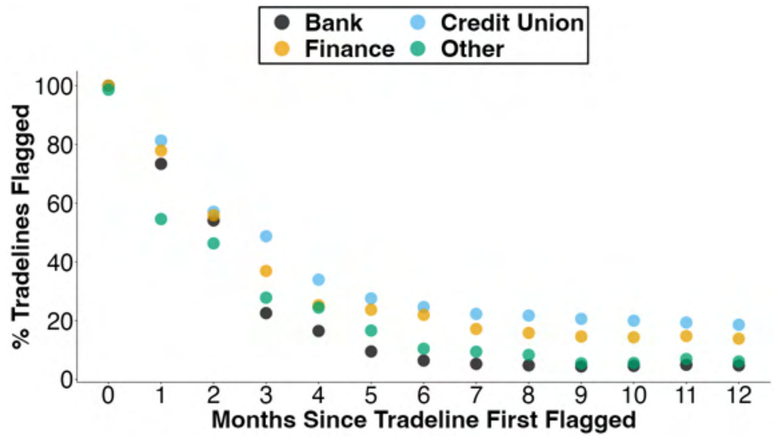
Figure A3: Trades With Both Credit Report Disaster Flag And Deferments, 2009 - 2024

A. By Credit Type **B. By Lender Type**



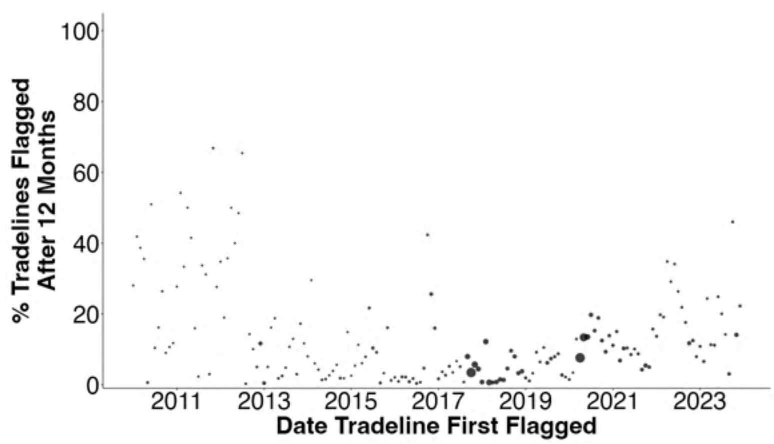
Notes: BTCCP data. Open tradelines with a positive balance and a credit report disaster flag. Lines show fractions of flagged tradelines that also have deferments.

Figure A4: Duration Of Disaster Flags Remaining On A Credit Report Tradeline, By Lender Type



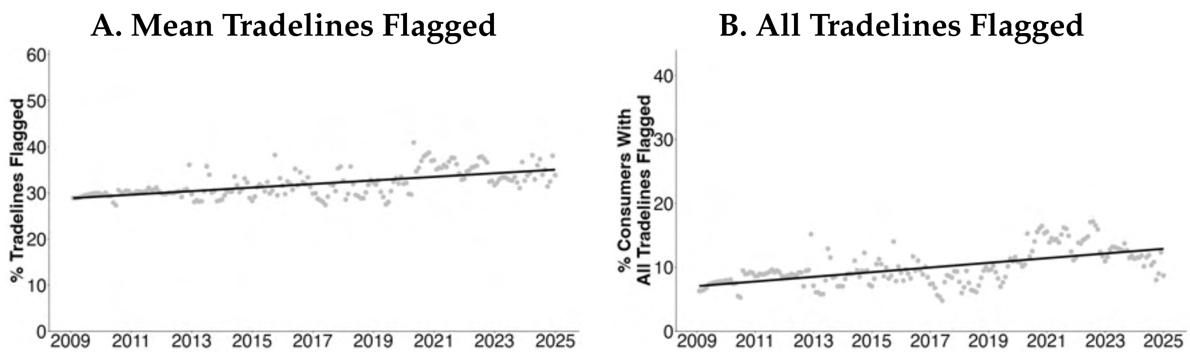
Notes: BTCCP data. This takes open credit report tradelines with positive balance that first have a disaster flag added between January 2010 and December 2023. Plots the fraction of these with disaster flags still present 1 to 12 months later. Colors are lender types.

Figure A5: Fraction Of Disaster Flags Remaining On A Credit Report Tradeline After Twelve Months, By Cohort



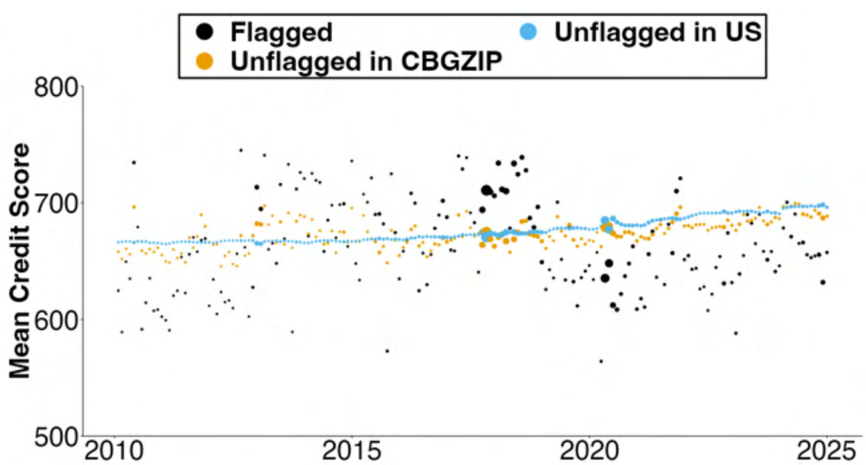
Notes: BTCCP data. This takes open credit report tradelines with positive balance that first have a disaster flag added between January 2010 and December 2023. Plots the fraction of these with disaster flags still present 12 months later for each cohort. X axis is cohort date when disaster flag first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



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.

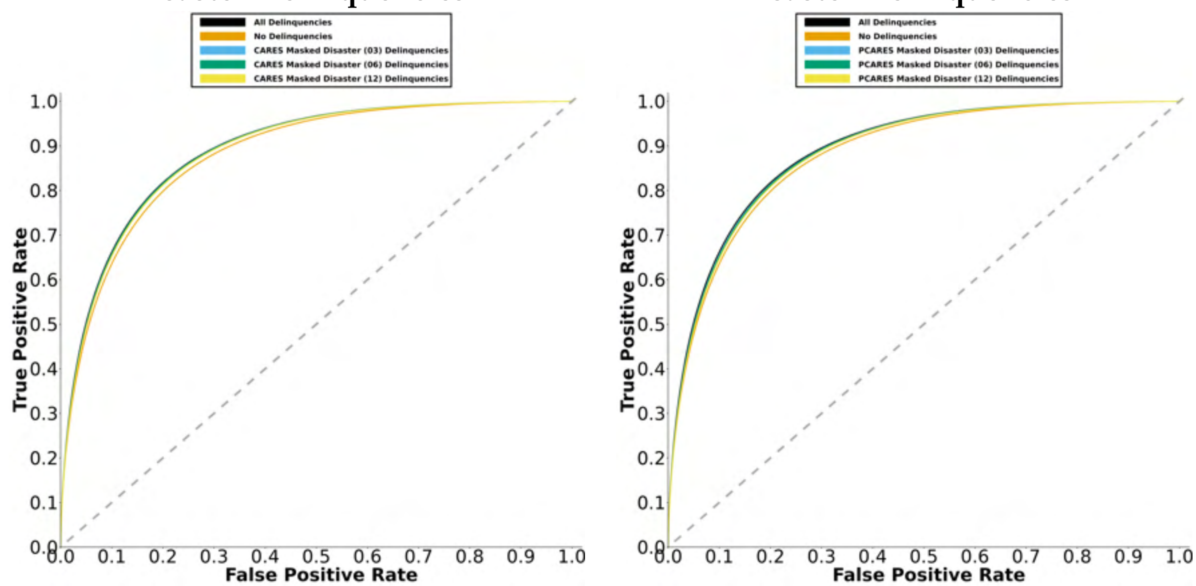
Figure A7: Mean Credit Score Of Cohorts Of Consumers First Flagged (Black) Compared To Unflagged Consumers In The Same Census Block Group × Zip Code (CBGZIP, Orange), And Unflagged In U.S. (Blue)



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: Receiver Operating Characteristic Curves (ROC) Showing The Predictive Performance Of Models Masking Delinquency Information. Baseline Model Includes Delinquencies (Black), For CARES (Panel A) and PCARES (Panel B) Masking Disaster Delinquencies Over Three (Blue), Six (Green), and Twelve (Yellow) Months, And No Delinquencies (Orange)

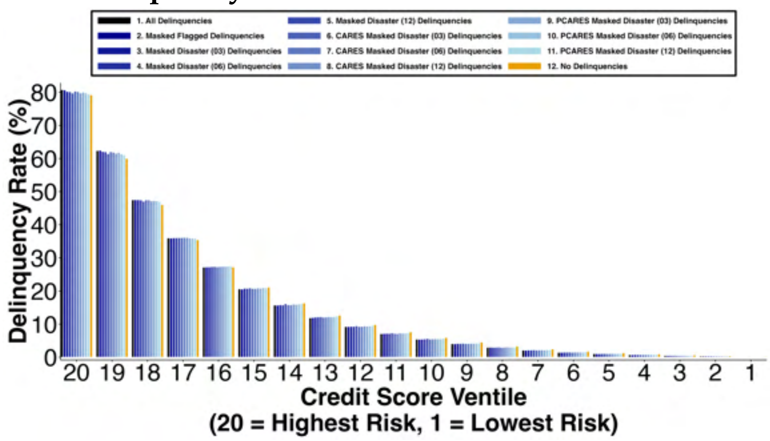
A. CARES Masking Disaster Delinquencies **B. PCARES Masking Disaster Delinquencies**



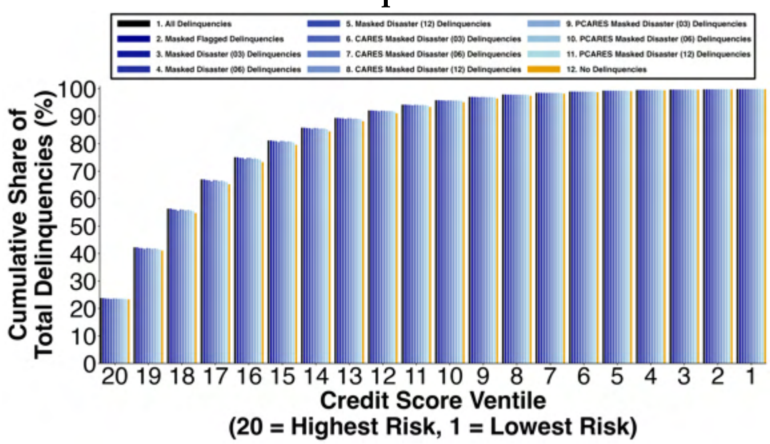
Notes: BTCCP data. Receiver Operating Characteristic curves (ROC) from out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. The black line is the baseline model that includes delinquency information the prediction. The orange line respectively show performance from a model without delinquency information. The blue, green, and yellow lines show performance from 'CARES' (Panel A) and 'PCARES' (Panel B) models with delinquency information that masks delinquencies 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, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables. True Positive Rate is defined as the fraction of consumers predicted new delinquency out of actual new delinquencies. False Positive Rate is defined as the fraction of predicted new delinquency out of actual no new delinquencies.

Figure A9: Performance Of Models Varying Masking Of Delinquency Information, By Credit Score Ventiles

A. Delinquency Rates Within Credit Score Ventiles



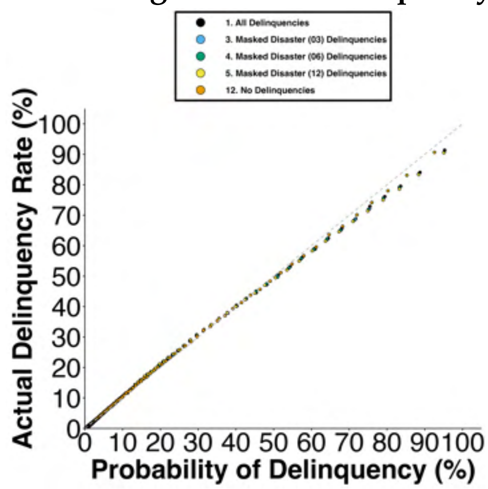
B. Cumulative Share Of Total Delinquencies Across Credit Score Ventiles



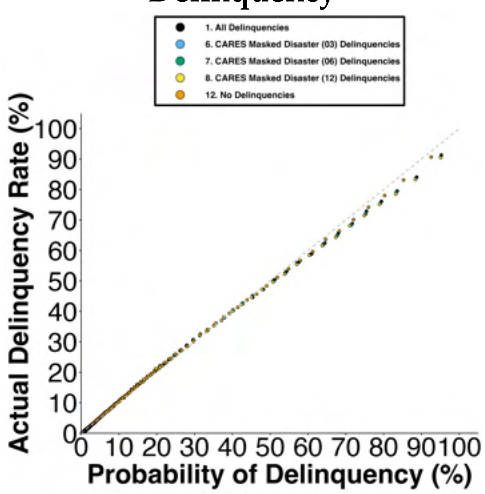
Notes: BTCCP data. Ventiles of credit scores are calculated using out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Ventile 20 is the 5% of consumers with highest probability of being delinquent, and ventile 1 is the 5% of consumers with the lowest probability of being delinquent. Uses data on 6.99 million consumers (the 30% testing sample) using data to October 2017. Panel A shows the fraction of consumers within each ventile that are delinquent. Panel B shows the cumulative fraction of total delinquencies (across all ventiles) in that ventile or higher risk ventiles. For example, the share in ventile 19 displays the sum of the number of delinquencies in ventiles 19 and 20 divided by the number of total delinquencies. Ventiles are shown for 12 different credit scoring models that vary the masking of information. The black bars are the baseline model ('1. All Delinquencies') that includes delinquency information the prediction. The orange bars are the model that includes no delinquency information ('12. No Delinquencies'). The blue bars display a variety of models that mask of delinquency information during natural disasters as explained in the main text and Table 5 notes. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

Figure A10: Calibration Plots Of Models Varying Masking Of Delinquency Information

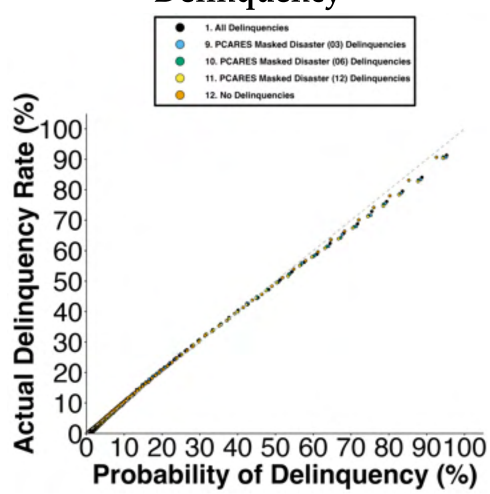
A. Masking Disaster Delinquency



B. CARES Masking Disaster Delinquency

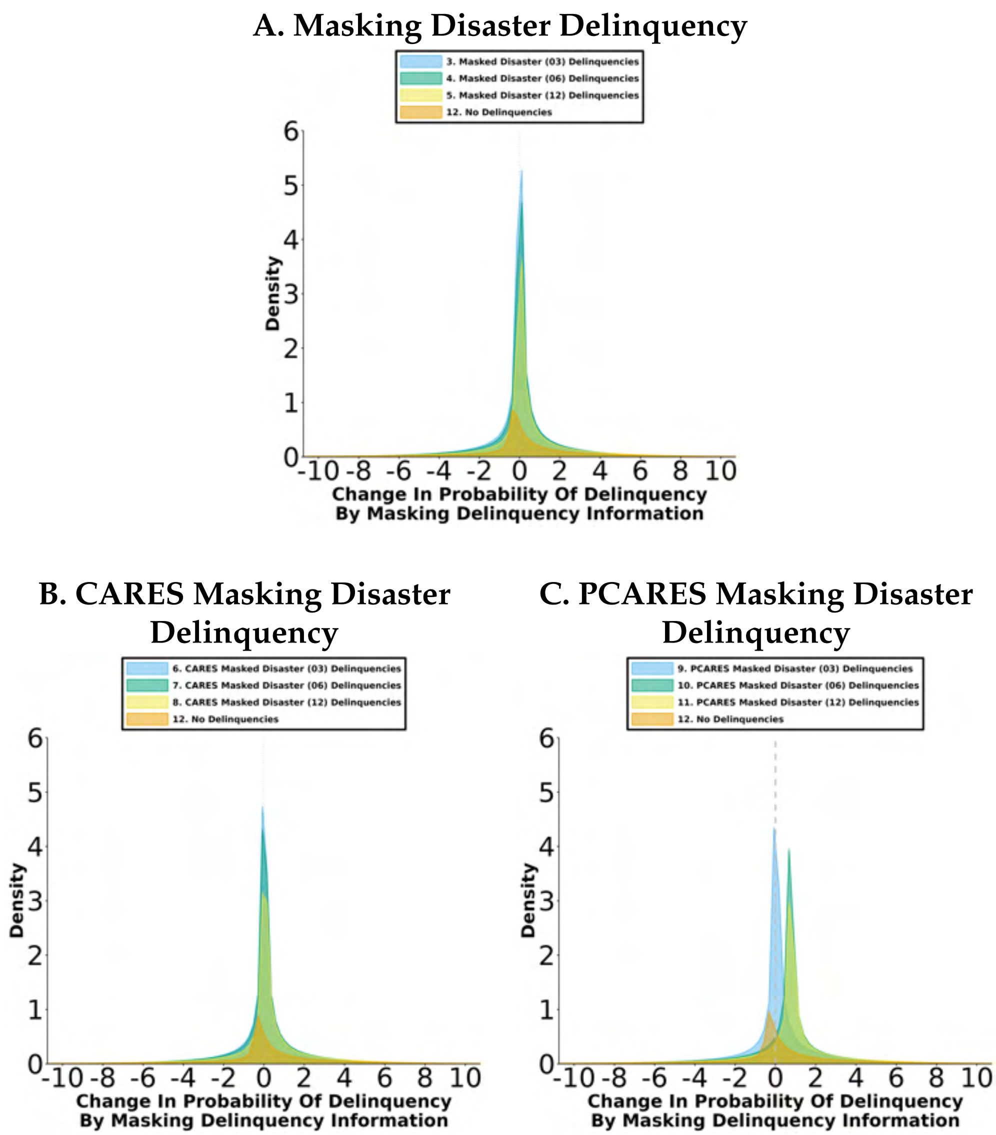


C. PCARES Masking Disaster Delinquency



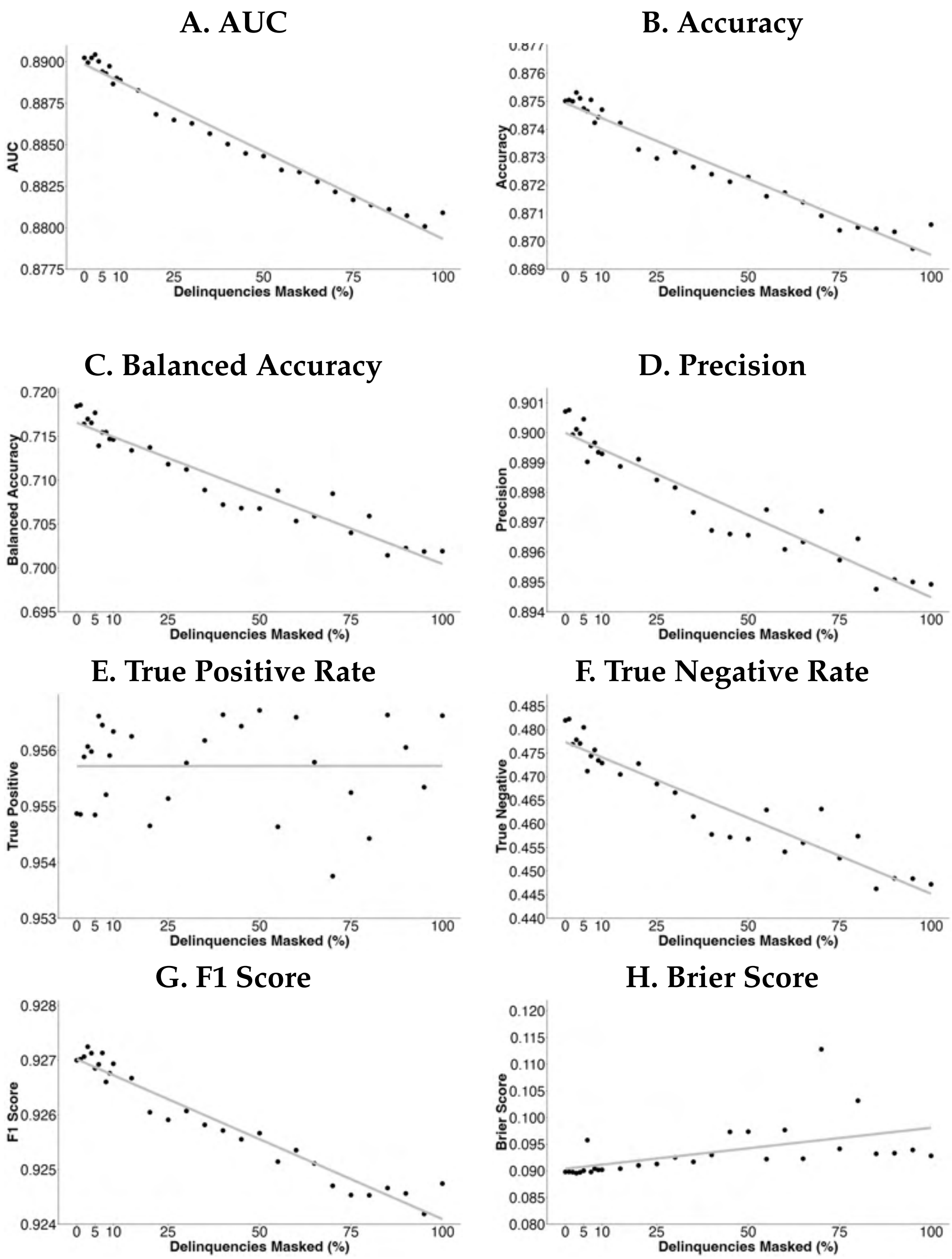
Notes: BTCCP data. Each panel shows the relationship between the predicted probability of delinquency on the x-axes compared to the actual delinquency rate on the y-axes. The gray dotted line shows the 45 degree line where predicted probabilities would perfectly match actual probabilities. Dots the mean values of these for 100 quantiles of credit scores that are calculated using out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Uses data on 6.99 million consumers (the 30% testing sample) using data to October 2017. The black dots are the baseline model ('1. All Delinquencies') that includes delinquency information the prediction. The orange dots are the model that includes no delinquency information ('12. No Delinquencies'). The blue, green, and yellow dots display models that mask of delinquency information during natural disasters for three, six, and twelve months respectively, with each panel showing a different methodology to calculate these, as explained in the main text and Table 5 notes. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

Figure A11: Distribution Of Changes In Probability Of Delinquency, Calculated For Models Varying Masking Of Delinquency Information, Relative To Baseline Model With Delinquency Information



Notes: BTCCP data. Each panel shows the distributions of changes in the predicted probability of delinquency compared to the baseline model with all delinquency information. Probabilities of delinquency are calculated using out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Uses data on 6.99 million consumers (the 30% testing sample) using data to October 2017. The orange bars shows how a model with no delinquency information ('12. No Delinquencies') changes predictions relative to the baseline models. The blue bars display a variety of models that mask of delinquency information during natural disasters as explained in the main text and Table 5 notes. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

Figure A12: Predictive Performance For Models Varying Randomly Masking Delinquency Information



Notes: BTCCP data. Each panel shows the relationship between a measure of predictive performance on the y-axes compared to the amount of delinquency information randomly masked on the x-axes. Predictive performance measures from out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. Each dot shows performance under a different model that varies the use of delinquencies data used as inputs. The gray line shows a linear line of best fit across all the models. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

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 Delinquency Variables As Predictors Of Any New Delinquency. Regressions Differentially Add Delinquency Information: (1) All Is Baseline, (2) Flagged Includes Terms For Disaster Flagged Delinquencies, (3),(4),(5) Disaster Include Terms For FEMA Natural Disaster Delinquencies Respectively Within 3, 6, and 12 Months

	(1)	(2)	(3)	(4)	(5)
All Delinquencies 12	0.003150	0.003096	0.003280	0.003305	0.003434
	(0.000135)	(0.000135)	(0.000141)	(0.000142)	(0.000150)
All Delinquencies 24	0.000851	0.000859	0.001042	0.000975	0.000424
	(0.000185)	(0.000185)	(0.000188)	(0.000189)	(0.000196)
All Delinquencies 36	0.001503	0.001496	0.001631	0.001664	0.001953
	(0.000154)	(0.000154)	(0.000156)	(0.000158)	(0.000164)
All Delinquencies 84	0.001270	0.001266	0.001338	0.001220	0.001141
	(0.000063)	(0.000063)	(0.000076)	(0.000079)	(0.000084)
Flagged Delinquencies 12		-0.001785			
		(0.006025)			
Flagged Delinquencies 24		0.016827			
		(0.010995)			
Flagged Delinquencies 36		-0.019482			
		(0.010060)			
Flagged Delinquencies 84		0.019841			
		(0.003856)			
Disaster Delinquencies 12			0.000099	-0.000193	-0.000989
			(0.000175)	(0.000171)	(0.000189)
Disaster Delinquencies 24			-0.000200	0.000075	0.001939
			(0.000255)	(0.000228)	(0.000240)
Disaster Delinquencies 36			-0.001029	-0.000908	-0.001553
			(0.000232)	(0.000202)	(0.000190)
Disaster Delinquencies 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 delinquency 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 delinquency terms ($D'_{i,t}$) from Equation 4. Column 2 shows the marginal effects from the θ_2 coefficients on the delinquency terms ($D'_{i,t}$) and the π coefficients on the flagged delinquency terms ($F'_{i,t}$) from Equation 5. Column 3 shows the marginal effects from the θ_3 coefficients on the delinquency terms ($D'_{i,t}$) and also the ϕ coefficients on disaster delinquency terms ($N'_{i,t}$) from Equation 7, when disaster delinquencies are measured as within 3 months of a disaster. Column 4 changes this definition of disaster delinquencies to be within 6 months of a disaster, and column 5 changes this definition of disaster delinquencies 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 delinquency 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-delinquency variables as inputs.

Table A3: Means For Predictive Outcome (Any New Delinquency In Next 24 Months) And Delinquency Variables

	Means
1. Outcome	0.168874
2. All Delinquencies 12	0.940522
3. All Delinquencies 24	1.222895
4. All Delinquencies 36	1.467395
5. All Delinquencies 84	2.242902
6. Flagged Delinquencies 12	0.004877
7. Flagged Delinquencies 24	0.005147
8. Flagged Delinquencies 36	0.005267
9. Flagged Delinquencies 84	0.005891
10. Disaster (03) Delinquencies 12	0.288569
11. Disaster (03) Delinquencies 24	0.410655
12. Disaster (03) Delinquencies 36	0.472400
13. Disaster (03) Delinquencies 84	0.831895
14. Disaster (06) Delinquencies 12	0.341319
15. Disaster (06) Delinquencies 24	0.467467
16. Disaster (06) Delinquencies 36	0.548176
17. Disaster (06) Delinquencies 84	0.978757
18. Disaster (12) Delinquencies 12	0.453650
19. Disaster (12) Delinquencies 24	0.579803
20. Disaster (12) Delinquencies 36	0.688245
21. Disaster (12) Delinquencies 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 delinquencies, 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 delinquencies. Rows 2 to 5 show the mean number of accounts in delinquency in the last 12, 24, 36, and 84 months respectively. Rows 6 to 9 show the mean number of accounts in delinquency, 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 delinquency, 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 A4: Consumers With Any Delinquency Across Models Varying Masking Delinquency Information

Model	Any Delinquency	Change from Baseline
1. All Delinquencies	52.209%	
2. Masked Flagged Delinquencies	52.199%	-0.020%
3. Masked Disaster (03) Delinquencies	51.412%	-1.528%
4. Masked Disaster (06) Delinquencies	50.583%	-3.115%
5. Masked Disaster (12) Delinquencies	47.975%	-8.110%
6. CARES Masked Disaster (03) Delinquencies	51.427%	-1.498%
7. CARES Masked Disaster (06) Delinquencies	50.665%	-2.958%
8. CARES Masked Disaster (12) Delinquencies	48.710%	-6.702%
9. PCARES Masked Disaster (03) Delinquencies	50.071%	-4.096%
10. PCARES Masked Disaster (06) Delinquencies	47.941%	-8.176%
11. PCARES Masked Disaster (12) Delinquencies	43.429%	-16.818%
12. No Delinquencies	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 delinquency 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 delinquencies. Row 2 masks delinquencies that have disaster flags. Rows 3, 4, and 5 mask delinquencies 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 delinquencies.

Table A5: Additional Measures Of Predictive Performance Of Models Varying Inclusion Of Flagged Delinquencies And Disaster Delinquencies

Model	Accuracy	Balanced Accuracy	Precision	True Positive Rate	True Negative Rate	F1 Score	Brier Score
1. All Delinquencies	0.875008	0.718415	0.900711	0.954870	0.481959	0.927000	0.089813
2. Flagged Delinquencies	0.875506	0.717824	0.900425	0.955923	0.479725	0.927344	0.089467
3. Disaster (03) Delinquencies	0.875499	0.718364	0.900633	0.955637	0.481091	0.927320	0.089444
4. Disaster (06) Delinquencies	0.875496	0.718135	0.900545	0.955750	0.480519	0.927327	0.089460
5. Disaster (12) Delinquencies	0.875489	0.718177	0.900562	0.955717	0.480637	0.927320	0.089462

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

Table A6: Measures Of Predictive Performance For Models Varying Randomly Masking Delinquency Information

Model	AUC	Accuracy	Balanced Accuracy	Precision	True Positive Rate	True Negative Rate	F1 Score	Brier Score
1. All Delinquencies	0.890236	0.875008	0.718415	0.900711	0.954870	0.481959	0.927000	0.089813
2R. Random Flagged	0.890443	0.875163	0.718498	0.900725	0.955062	0.481935	0.927098	0.089905
3R. Random (03)	0.890795	0.875582	0.716976	0.900090	0.956470	0.477482	0.927424	0.089489
4R. Random (06)	0.890215	0.875148	0.719070	0.900947	0.954748	0.483391	0.927068	0.089529
5R. Random (12)	0.890573	0.875502	0.716644	0.899971	0.956519	0.476769	0.927384	0.090806
6R. Random CARES (03)	0.890130	0.874933	0.718511	0.900757	0.954707	0.482315	0.926948	0.090001
7R. Random CARES (06)	0.890543	0.875120	0.717701	0.900423	0.955403	0.480000	0.927099	0.090314
8R. Random CARES (12)	0.890833	0.875706	0.716061	0.899724	0.957125	0.474997	0.927537	0.090115
9R. Random PCARES (03)	0.889991	0.874829	0.717540	0.900395	0.955047	0.480032	0.926916	0.089778
10R. Random PCARES (06)	0.890147	0.875044	0.718371	0.900690	0.954948	0.481794	0.927026	0.095882
11R. Random PCARES (12)	0.889535	0.874627	0.717456	0.900387	0.954784	0.480127	0.926788	0.090563
12. No Delinquencies	0.880902	0.870593	0.701915	0.894925	0.956618	0.447212	0.924744	0.092819

Notes: BTCCP data. Predictive performance measures from out-of-sample prediction from XGBoost models predicting any new delinquency in next 24 months. Performance uses data on 6.99 million consumers (the 30% testing sample) and information to October 2017. Each row shows performance under different models that vary the use of delinquencies data used as inputs. Row 1 is the baseline model that includes all delinquencies. Rows 2 to 11 are models that mask all delinquencies for a random subset of consumers that have delinquencies. The amount of delinquencies removed in the models shown in rows 2 to 11 respectively correspond to the same fraction of consumer delinquencies that are non-randomly masked in rows 2 to 11 of Table A4. Row 12 shows results for a model that does not use any delinquency information. Predictions use as inputs uses a credit score, that is built on the training sample using XGBoost machine learning algorithm with 171 non-delinquency variables as inputs, along with these 171 non-delinquency variables.

B. Relationship Between Disaster Flags And Credit Access

Does consumers' credit access increase after having disaster flags on their credit report? Across three methodologies, I find no evidence of disaster flags increasing consumers' credit access, measured by new account openings. Consumers with pre-disaster financial distress experience the largest, but temporary, VantageScore credit score increases from disaster flags, however, their credit access does not improve. My first methodology shows the unconditional means in outcomes for flagged consumers, relative to linear time trends as a benchmark that has been used in prior literature (Dobkin et al., 2018; Dobbie et al., 2020; Gross et al., 2020). This methodology is explained in Section B.1, and I then show my descriptive results in Section B.2. My second methodology uses a difference-in-differences approach to compare outcomes for flagged consumers to a control group of unflagged consumers in the same geographic area, also matched on pre-disaster characteristics. This methodology is explained in Section B.3 and I show my results in Section B.4. My third methodology uses a difference-in-differences approach exploiting exogenous variation in the timing of flags being added, using a control group of consumers in the same geographic area that are also flagged but their flags appear one-month later. This methodology is explained in Section B.5 and I show my results in Section B.6.

B.1 Descriptive Methodology

My first descriptive 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 visually displays changes in the unconditional means of a consumer's finances relative to a pre-disaster flag linear 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 groups 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 group without being affected by COVID-19 disruptions or contemporaneous deferments that more commonly appear in more recent groups. 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 event 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 unconditional means, 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 or natural disaster. The deviation from this linear trend may indicate the combined effect of a disaster and a disaster flag. This linear trend approach has been used to provide a counterfactual in the prior literature, for example Dobbie et al. (2020) and Gross et al. (2020) use it to study the effects of bankruptcy flag removals on household finances, and Dobkin et al. (2018) uses it to study the effects of hospital admissions on household finances.

I show descriptive results 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 delinquency (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 delinquency by this measure. These two measures are highly correlated; 79% of consumers with any delinquency have subprime credit scores, and 34% of consumers with subprime credit scores have any delinquency. 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; Urban Institute, 2019). 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 delinquencies that could be masked by flags.

B.2 Descriptive Results

Delinquencies. I examine the mechanism through which disaster flags can affect credit scores and credit access: masking delinquencies. Figure B1 shows the prevalence of delinquencies in credit reports of flagged consumers. The black line on Figure B1 Panel A shows the fraction of consumers with any delinquency, before masking by flags. This trend increases slightly over event time. The orange line masks delinquencies that occur on the tradeline months where the flags also appear. Flag masking immediately reduces the fraction of consumers with any delinquency appearing in their credit report by 1.5

percentage points, however, delinquencies do not go to zero but remain at 5.7 percent, which is a 21 percent decline in delinquencies. These consumers have delinquencies on their other tradelines without flags, and these remain unmasked. The two delinquency series quickly converge within twelve months, showing that any potential benefit of flags masking delinquencies is temporary.

These small average results are largely driven by consumers experiencing pre-disaster financial distress. Panels B and C of Figure B1 repeat this exercise for subsamples of the data by pre-disaster financial distress where $t-1$ values of delinquencies are normalized to zero. This shows that the temporary changes in delinquencies are concentrated among the minority of consumers who experience pre-disaster financial distress: subprime credit scores or those with any delinquency. Flags initially mask delinquencies for 6.4 percent of pre-disaster subprime consumers, a 24 percent decline in delinquencies, and 8.2 percent of consumers with any pre-disaster delinquency, a 23 percent decline in delinquencies. Masking of delinquencies is still only temporary for these groups. There is no discernible difference in delinquencies 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 B2 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., 2025). One might not expect the benefits of disaster flags to be as large as the benefits of removing bankruptcy flags for a couple of reasons. First, bankruptcy is one of the most negative signals about a consumer's credit risk, affecting highly financially distressed consumers who have high potential gains from these negative signals being removed to pool them with lower credit risk consumers. Second, an average increase in credit scores following disaster flags is expected to be a dilution of a larger increase given that only a small subset of consumers have delinquencies, whereas all bankrupt consumers have bankruptcy flags (before their removal).

Figure B2 Panels B and C show how the most financially distressed consumers—with subprime credit scores or any delinquency—receive the largest increases in their credit scores from disaster flags, with results for all credit score segments shown in Figure B5. 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 any delinquency, 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., 2025). 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 Delinquencies 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., 2025). 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 B3 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 B3 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 Figure B6, and the value of new credit card limits, shown in Figure B7.

B.3 Difference-in-Differences To Unflagged Consumers: Methodology

I find consistent results to my descriptive results when I use a difference-in-differences design to estimate the effects of disaster flags relative to a matched sample of unflagged consumers in the same geographic area, who are therefore also exposed to the same disaster. I estimate the 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., 2025; Cherry, 2024). Dube et al. (2025) 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 methodology exploits the *timing* of disaster flags being applied as a quasi-random function of natural disasters. I use a control group of consumers to difference out the contemporaneous effects of the natural disaster to leave only the effects of the disaster flag. The identifying assumption is that in the absence of the treated group having disaster flags added, that they would follow a similar trend to the control group. One aspect that helps my identification assumption is that the timing of disaster flags being added does not always perfectly align with the timing of the onset of a disaster, and therefore there is separate identifying variation in event time to the date of the disaster.

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 flagged group 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 delinquency, 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 descriptive dataset presented in the previous subsection.

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 10:

$$Y_{g,c,t} = \sum_{\tau \neq -1} \delta_\tau \Big(FLAG_c \times D_t^\tau \Big) + \gamma_{g,c} + \gamma_{g,t} + \varepsilon_{g,c,t} \tag{10}$$

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).

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. In my results, I will show a lack of pre-trends ($\delta_\tau \approx 0$ when $\tau < -1$), which provides some support that the consumers that are matched controls offer a reasonable counterfactual for the treated flagged consumers.

B.4 Difference-in-Differences To Unflagged Consumers: Results

Credit Scores. My difference-in-differences results for all consumers, and by subsamples of pre-disaster financial distress, are shown in Table B1 for the effects in the month of flagging and in Table B2 for the effects after twelve months (Tables B4 and B5 show additional results by all pre-disaster credit score groups). In line with my descriptive results, my difference-in-differences results show that disaster flags lead to 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 B1) that dissipate to being insignificant from zero within six months, as shown in Figure B11. 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 B2.

The average temporary increase in 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 delinquency at $t-12$. Figure B12 Panel A shows that at $t = 0$, consumers with any pre-disaster delinquency 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 delinquency 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 any pre-disaster delinquency 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 within twelve months, making them insignificant from zero. After twelve months, the estimated effects for those with any pre-disaster delinquency are -0.29, with a standard error of 0.83, and for those without any pre-disaster delinquency is -0.36, with a standard error of 0.33. Figure B14 Panel A shows the results for 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 evidence increasing credit access after disaster flags, even for the consumers who received the largest temporary boosts to their credit scores: those with any pre-disaster delinquency. There is no significant change on whether a consumer opened any new credit card account each month. Figure B12 Panel B shows credit access declines for both consumers with and without any pre-disaster delinquency. After twelve months the estimated effects for those with any pre-disaster delinquency are slightly negative -0.0036, with a standard error of 0.0022, shown in Table B2, which is a decrease of 10% relative to the baseline mean of 0.0346. As also shown in Table B2, the change for those without any pre-disaster delinquency 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 Figure B15.

There does not appear to be a change in the intensive margin. Figure B12 Panel B shows no increase in the value of new credit card limits. In this margin, consumers with any pre-disaster delinquency 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 B2, relative to their $t-1$ mean of \$64.6. Whereas consumers without any pre-disaster delinquency experience a significant decrease of 24%, with an estimate of $-\$108.1$ and a standard error of \$15.8, also shown in Table B2, relative to their $t-1$ mean of \$450.4. These results on credit access are consistent with segmenting by credit score (Figures B13 and B14) and when averaging across all consumers, as shown in Figure B11.

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. (2025). Applying a disaster flag has no positive effects on credit access, even for the most distressed group, and some significantly negative effects, despite temporarily boosting credit scores. This finding contrasts with the removal of bankruptcy flags that had positive effects on *both* credit scores and credit access (e.g., Dobbie et al., 2020; Gross et al., 2020; Jansen et al., 2025). 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, which 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., 2025), and therefore credit decisions taken using FICO scores would be unaffected. 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. The reduced credit access results could potentially be consistent with anecdotal reports of flags being viewed by some lenders that use manual underwriting or proprietary credit scores as negative signals of a consumers' credit risk (FinRegLab, 2020).[17]

One potential concern with interpreting these estimates as being causal is if these reflect a combination of both the effects of disaster flags and also the effects of natural disasters that co-move. This could occur if the consumers that experience a more severe impact of the natural disaster are selected into the disaster flag, relative to consumers even in the same narrow geographic area and matched on pre-disaster characteristics. If so, one would expect that flagged consumers experience decreases in both credit scores and credit access, however, that is inconsistent with my difference-in-differences and earlier descriptive results that show that credit scores *increase*. In the next section, I use a separate methodology to address this concern, and also find a consistent pattern of credit score increases and credit access decreases.

[17] FinRegLab (2020): "In connection with the pandemic, there are reports that some lenders are rejecting consumers for mortgage refinance applications because they have CP or AW codes on one or more accounts, even if they are in fact making full payments."

B.5 Exploiting Variation in Flag Timing: Methodology

I now abstract from selection into the use of disaster flags by using a separate difference-in-differences design that exploits variation in the *timing* of when disaster flags are first applied. In this design, both the treatment and control group are selected into disaster flags, potentially signaling that these consumers have been more directly impacted by a disaster than unflagged consumers. The key idea of this methodology is that a one-month difference in when a disaster flag is applied to a consumer's credit report is plausibly due to variations in the reporting policies of furnishers (who may be lenders or servicers, Gibbs et al., 2025) for how quickly this information is updated to the credit reporting agencies.[18] Given the lack of regulatory or industry guidance, it is highly likely that lenders vary in their reporting of disaster flags. To provide an example of this variation, I look at mortgages in Texas in September and October 2017 following Hurricane Harvey that made landfall on 24 August 2017. There are 189 mortgage furnishers that add flags to their mortgage accounts in either September or October 2017, of which 12 furnishers add all their flags in September and 77 furnishers add all their flags in October. Across all 189 furnishers, the mean and median percent of mortgages flagged in September or October that are flagged in September are 12.5% and 31.5% respectively, showing substantial variation. It is also possible that a one-month variation in disaster flags being applied is partially due to small differences in the timing of consumers contacting their lenders in response to the same disaster.

In this difference-in-differences design, the treatment group is the consumers who first have a disaster flag added to their credit report at time $t = 0$, and the control group is the consumers who first have a disaster flag added to their credit report one month later ($t = 1$). I take my event study dataset (from Section B.1), and restrict to cases where there are both treatment and control units in the same combination of geographic area—the same census block group $\times$ zip code—, credit score group, any delinquency, 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. By comparing flagged consumers in the same geographic area at the same time, I can control for the contemporaneous effects

[18]Furnishers vary in the time of month when they report new information with credit reporting agencies, and within-lender there is also variation on the time in a month when information is updated for different borrowers. For example, credit card lenders typically assign credit cards to different statement dates in a month for operational reasons, and report information on a credit card to the credit bureau shortly after a statement has been issued. Whereas mortgages may be reported shortly after the date the mortgage payment is due, which may be a function of the day in a month when a consumer bought a house or refinanced years previously. Consumers also often have multiple credit obligations, so the timing of whether the credit account that is flagged is scheduled to be updated early or late in a month is plausibly exogenous, leading to one-month variation.

of the natural disaster.

I calculate the first difference in outcomes at the consumer-level. This first difference accounts for time-invariant unobserved differences between consumers (e.g., financial sophistication, disaster preparedness). This first difference is calculated for the same calendar year-months for both the treated units and their respective control units. This means that for the treated units the difference is calculated at their event time (i.e., relative to first being flagged) $Y_{i,0} - Y_{i,-1}$, and for the control units (first flagged one month later than the treated units) at their event time $Y_{i,-1} - Y_{i,-2}$.

After taking first differences, I have a dataset of 1.7 million observations containing 1.4 million unique consumers. Each consumer can only appear as a treated unit for one observation in this data. A treated consumer sometimes appears in the dataset as a second observation when it is a control unit for another treated consumer flagged one month prior. With this dataset, I estimate Equation 11:

$$\Delta Y_{i,s,t} = \rho TREAT_{i,s,t} + \gamma_{s,t} + \Delta\varepsilon_{i,s,t} \tag{11}$$

In this regression, i denotes a consumer, t a calendar year-month, and s, t a ''set" of treated and control units that have the same combination of geographic area—the same census block group $\times$ zip code—, credit score group, any delinquency, and any mortgage debt, and where the control unit is first flagged one month later than the treated units. Set-by-time fixed effects, $\gamma_{s,t}$, are included that allow for heterogeneity across natural disaster locations, time periods, and consumer groups. This enables me to focus on within-set differences in outcomes that are attributable to flagging.

The coefficient ρ is my parameter of interest, showing the second difference in my difference-in-differences approach. This is the coefficient on an indicator, $TREAT_{i,s,t}$, for a consumer being in the treated group at time t. This shows the difference in outcomes due to a consumer being flagged one month earlier in the same geographic area at the same time. By comparing the first difference in outcomes between treated and control units and including the group-by-time fixed effects (in a similar way to the earlier stacked difference-in-difference approach to avoid known issues with staggered difference-in-differences), I remove time effects that are attributable to disasters or other observed or unobserved factors. For estimates to be given a causal interpretation, my identification only requires that a common trend assumption that, in the absence of a disaster flag being added consumers in the treated group would have trended similarly to the control group of consumers that are flagged one-month later, holds for one month. I cluster standard errors at the year-month level given correlated shocks across units, and my conclusions are unaffected by alternative clustering approaches.

B.6 Exploiting Variation in Flag Timing: Results

Table B6 shows my results exploiting a one-month difference in the timing of disaster flags being added, comparing flagged consumers in the same geographic area and other characteristics at the same time. My results are consistent with my previous two methodologies.

I find that after flagging VantageScore credit scores significantly increase across all heterogeneous groups of consumers. The average increase in credit score from a disaster flag being applied is 1.53 points, with a standard error of 0.31 points. There are significant increases in credit scores across all heterogeneous groups of consumers, with the largest increases concentrated among the consumers that experience pre-disaster financial distress. Credit scores for consumers with any pre-disaster delinquency experience a 10.07 points increase, with a standard error of 0.96 points.

However, as found in my previous results, such increases in credit scores do not lead to any significant increases in credit access. For example, there is a statistically significant 0.02 average decline in the number of new accounts opened. The lack of increases in credit access applies even for the consumers experiencing pre-disaster financial distress, where potential benefits are highest given that they experience the largest VantageScore increases and are the most credit constrained group. For example, the estimate for the group experiencing any pre-disaster delinquency shows a significant 0.03 decline in the number of new accounts opened.

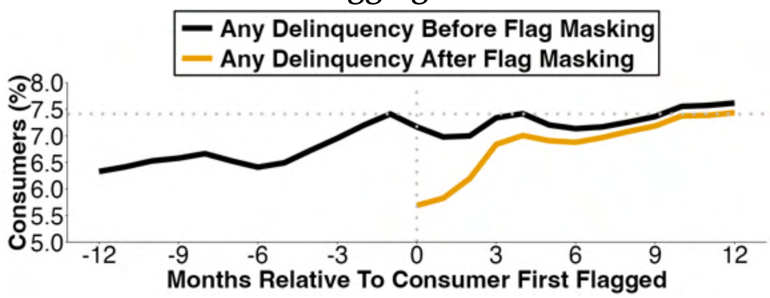
My conclusions hold for all three measures of credit access: the number of new account openings, the number of new credit cards, and the value of the new credit card limits opened. Across all these outcomes and all heterogeneous groups, a total of 36 estimates, I always find negative point estimates, suggesting a decrease in credit access. 35 of these 36 estimates show statistically significant decreases in credit access, with the estimate for the effect on the value of new credit card limits for the Near Prime subgroup being the only estimate that is insignificant from zero.

As a robustness exercise, I also estimate results using consumers in the same "set" (as defined in Section B.5) that are first flagged two- or three-months later rather than one-month. Over these longer-month differences, the variation in flag timing becomes less likely to be exogenous. The sample sizes decrease because there are fewer suitable controls, decreasing the sample from 1.7 million for the one-month variation to 0.43 million in the two-month variation and 0.09 million in the three-month variation. The benefits of these longer-than-one-month differences are that they use a separate group of consumers as control units to the one-month estimations so offer a useful robustness test, and they also allow for more time periods when the control units are untreated to trace out dynamic effects of being flagged.

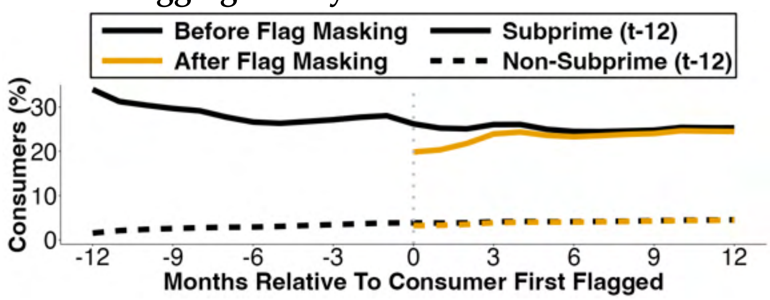
Figure B16 shows the results of this robustness exercise. In each of these panels, the black coefficients show the estimates using the one-month latter flagged group as a control group. These are similar to the orange estimates that use two-month later and the blue estimates that use the three-month later estimates as controls. The x-axes of each panels shows the dynamic effects produced by the two- and three-month latter estimation is also consistent with my earlier findings.

Figure B1: Unconditional Means Of Consumers (%) With Any Delinquency On Credit Reports Before (Black) And After (Orange) Masking By Disaster Flags

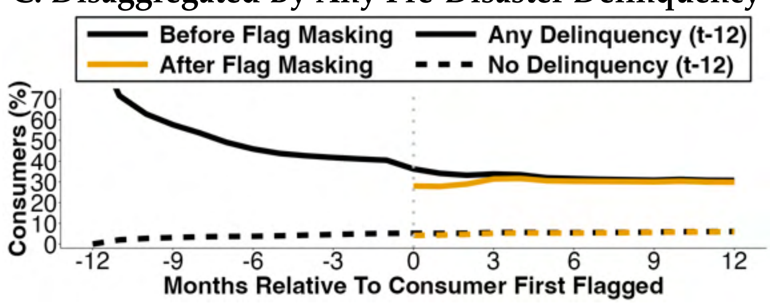
A. Aggregated



B. Disaggregated By Pre-Disaster Credit Score



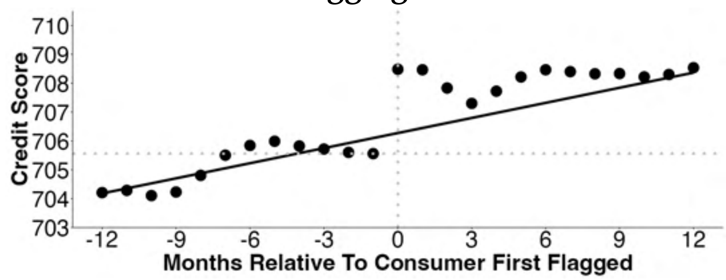
C. Disaggregated By Any Pre-Disaster Delinquency



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 delinquency, measured as 30 or more days past due, before (black) and after (orange) tradeline months where delinquencies 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 delinquency twelve months prior to first being flagged ($t - 12$). In Panel C, 6.3% of consumers have any delinquency twelve months prior to first being flagged.

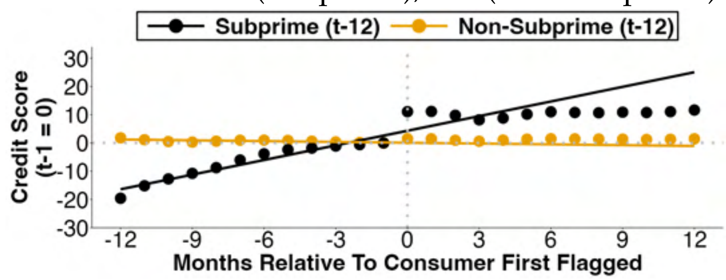
Figure B2: Unconditional Means Of VantageScore Credit Scores Relative To Linear Pre-Trends

A. Aggregated



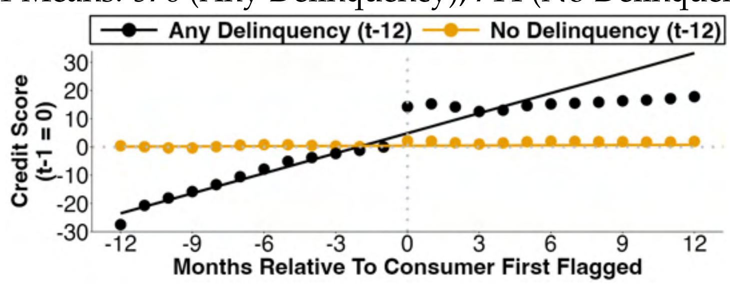
B. Disaggregated By Pre-Disaster Credit Score ($t-1=0$ for Each Subgroup)

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



C. Disaggregated By Any Pre-Disaster Delinquency ($t-1=0$ for Each Subgroup)

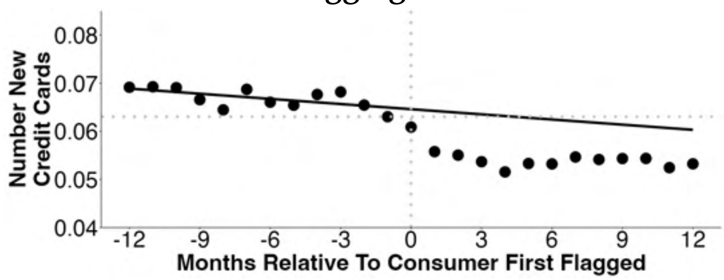
$t-1$ Means: 576 (Any Delinquency), 714 (No Delinquency)



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 delinquency twelve months prior to first being flagged ($t-12$). In Panel C, 6.3% of consumers have any delinquency twelve months prior to first being flagged. Panels B and C normalize credit scores for each subgroup to $t-1=0$.

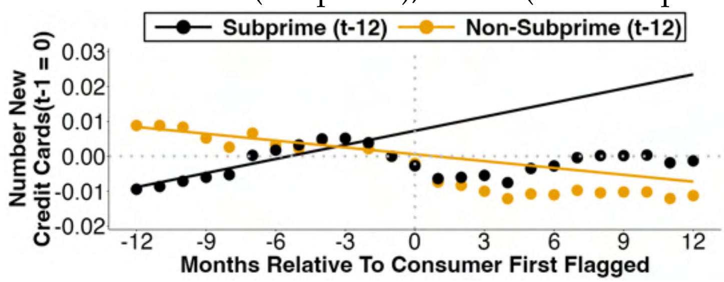
Figure B3: Unconditional Means Of New Credit Card Account Openings Relative To Linear Pre-Trends

A. Aggregated



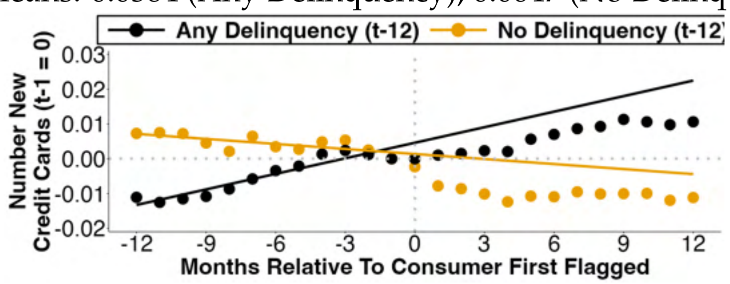
B. Disaggregated By Pre-Disaster Credit Score ($t-1=0$ for Each Subgroup)

$t-1$ Means: 0.0558 (Subprime), 0.0643 (Non-Subprime)



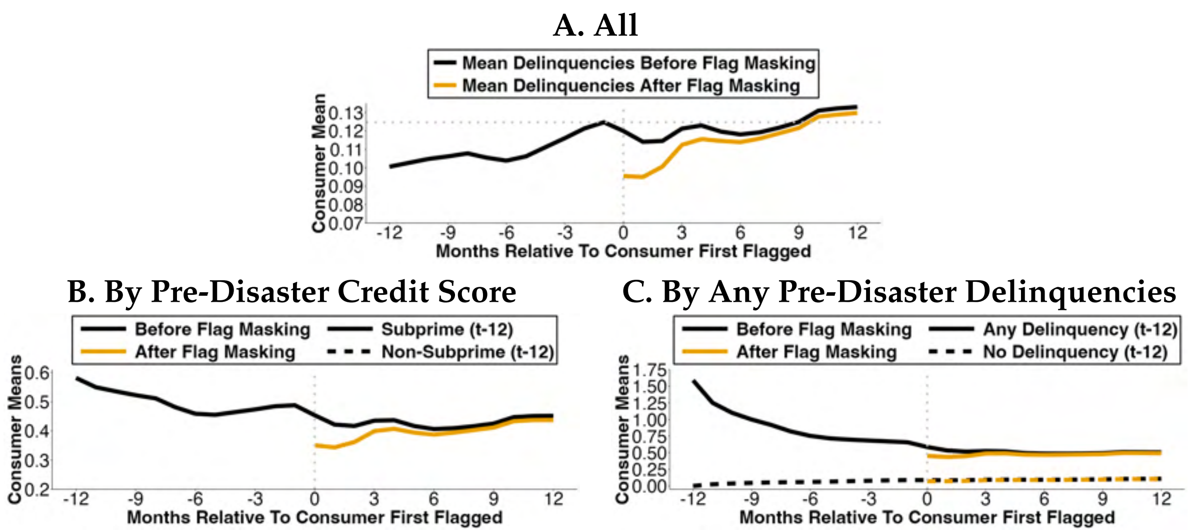
C. Disaggregated By Any Pre-Disaster Delinquency ($t-1=0$ for Each Subgroup)

$t-1$ Means: 0.0384 (Any Delinquency), 0.0647 (No Delinquency)



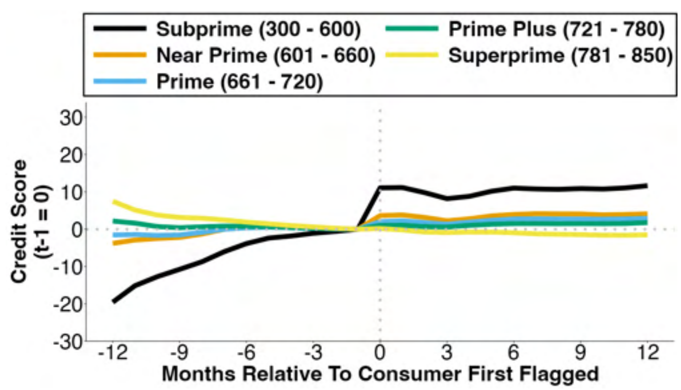
Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied between January 2010 and December 2018. The solid lines are linear pre-trends from OLS regressions on data $t-12$ to $t-1$. The X axes 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 delinquency twelve months prior to first being flagged ($t-12$). In Panel C, 6.3% of consumers have any delinquency twelve months prior to first being flagged. Panels B and C normalize new account openings for each subgroup to $t-1=0$.

Figure B4: Unconditional Means For The Number Of Delinquencies 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 Delinquencies



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 delinquencies on consumers' credit reports before (black) and after (orange) tradeline months where delinquencies 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 delinquency twelve months prior to first being flagged ($t-12$). Panels B and C normalize each series to $t-1=0$.

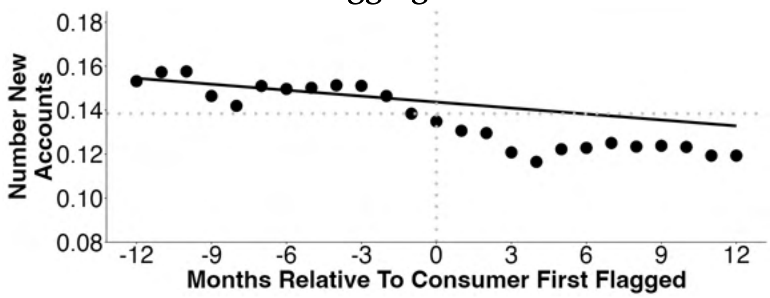
Figure B5: Unconditional Means of VantageScore Credit Scores by Pre-Disaster ($t-12$) Credit Score ($t-1=0$ For Each Subgroup)



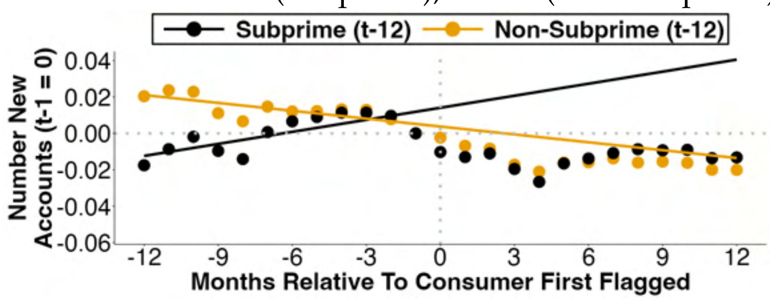
Notes: BTCCP data. Solid lines are unconditional means for 25 months of observations for a balanced panel constructed of 2.8 million consumers with disaster flags first applied January 2010 to December 2018. X axis shows months since consumer first flagged. Y axis is credit score. Splits by consumer's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

Figure B6: Unconditional Means Of Number Of New Accounts Relative To Linear Pre-Trend

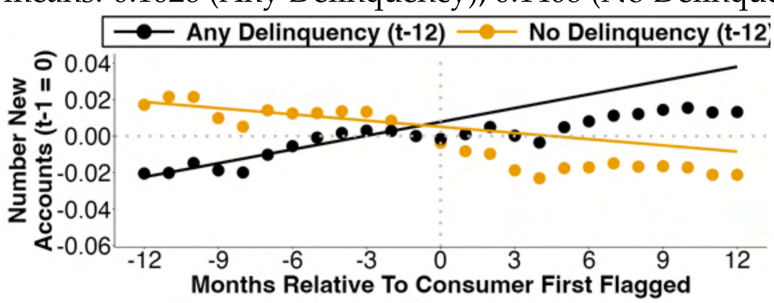
A. Aggregated



B. Disaggregated By Pre-Disaster Credit Score ($t-1=0$ for Each Subgroup)
$t-1$ means: 0.1399 (Subprime), 0.1382 (Non-Subprime)



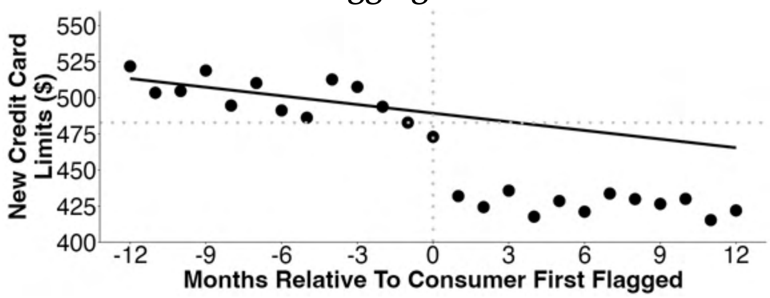
C. Disaggregated By Any Pre-Disaster Delinquency ($t-1=0$ for Each Subgroup)
$t-1$ means: 0.1026 (Any Delinquency), 0.1408 (No Delinquency)



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 delinquency twelve months prior to first being flagged ($t-12$). Panels B and C normalize new account openings for each subgroup to $t-1=0$.

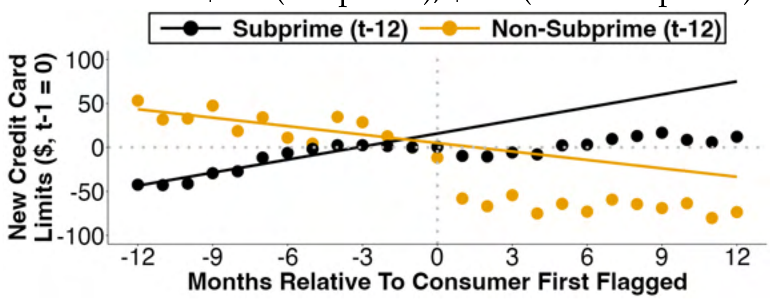
Figure B7: Unconditional Means of Value Of New Credit Card Limits Relative To Linear Pre-Trend

A. Aggregated



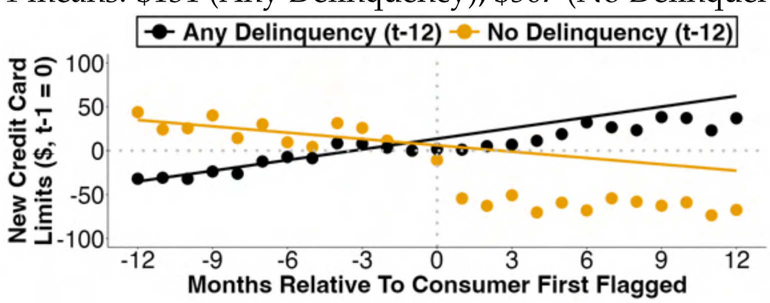
B. Disaggregated By Pre-Disaster Credit Score ($t-1=0$ for Each Subgroup)

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



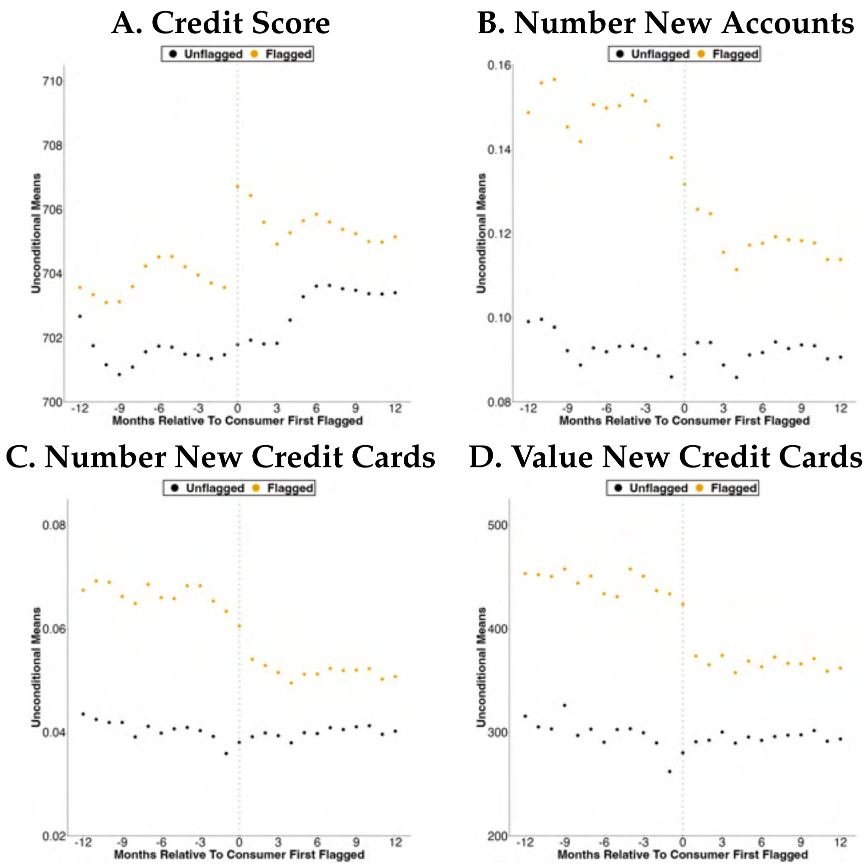
C. Disaggregated By Any Pre-Disaster Delinquency ($t-1=0$ for Each Subgroup)

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



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 delinquency 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 B8: Unconditional Means Of Outcomes For Difference-in-Differences Sample

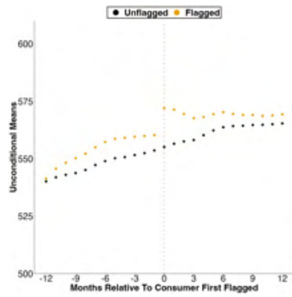


Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel constructed of 3.2 million consumers, consisting of a flagged group in yellow and a matched unflagged group in black. X axis shows months since consumer first flagged. Panels A to D show means for different outcomes: Panel A is credit score (VantageScore 3.0), Panel B shows the number of new accounts opened that month, Panel C shows the number of new credit card accounts opened that month, and Panel D shows the value of new credit card limits opened that month ($).

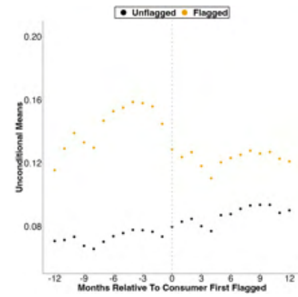
Figure B9: Unconditional Means Of Outcomes For Difference-In-Differences Samples With: (I.) With Subprime and (II.) Non-Subprime Pre-Disaster Credit Scores

I. Sample With Subprime Pre-Disaster Credit Score

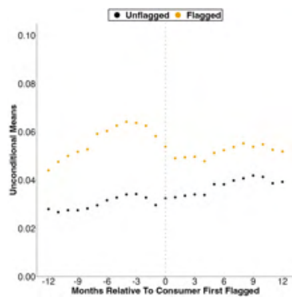
A. Credit Score



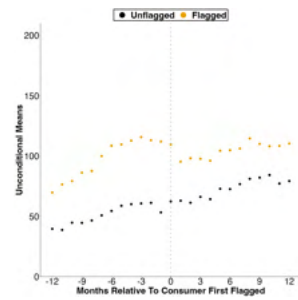
B. Number New Accounts



C. Number New Credit Cards

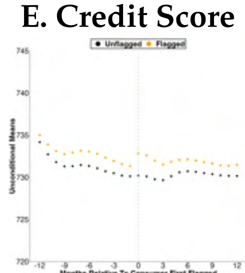


D. Value New Credit Cards

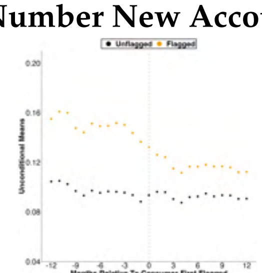


II. Sample With Non-Subprime Pre-Disaster Credit Score

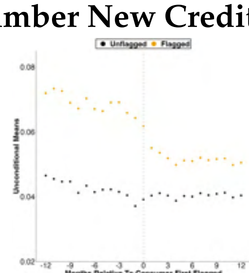
E. Credit Score



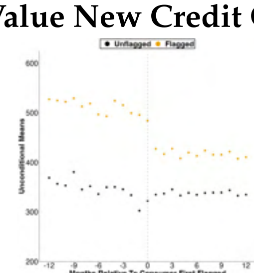
F. Number New Accounts



G. Number New Credit Cards



H. Value New Credit Cards

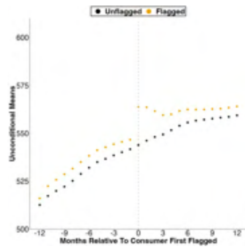


Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel, consisting of a flagged group in yellow and a matched unflagged group in black. Panels A to D show results for 0.52 million consumers with subprime (600 or below) credit scores twelve months pre-disaster. Panels E to H show results for 2.68 million consumers with non-subprime (601 or above) credit scores twelve months pre-disaster. X axis shows months since consumer first flagged. Panels show means for different outcomes: Panels A and E show credit score (VantageScore 3.0), Panels B and F show the number of new accounts opened that month, Panels C and G show the number of new credit card accounts opened that month, and Panels D and H show the value of new credit card limits opened that month ($).

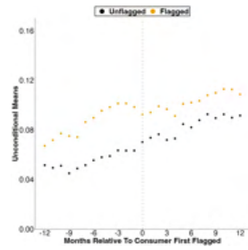
Figure B10: Unconditional Means Of Outcomes For Difference-In-Differences Samples With (I.) Any Delinquency and (II.) No Delinquency Pre-Disaster

I. Sample With Any Delinquency Pre-Disaster

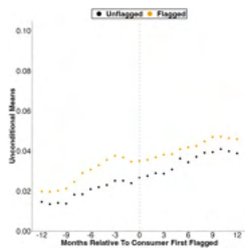
A. Credit Score



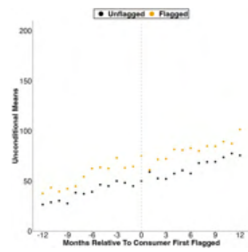
B. Number New Accounts



C. Number New Credit Cards

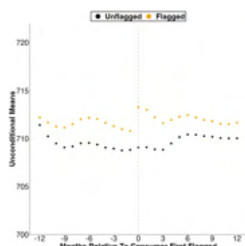


D. Value New Credit Cards

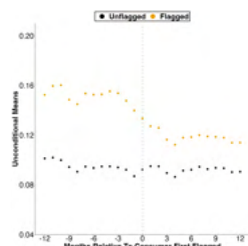


II. Sample With No Delinquency Pre-Disaster

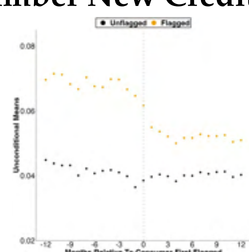
E. Credit Score



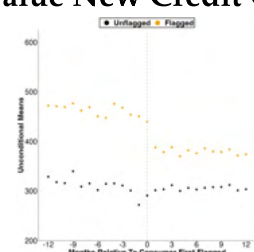
F. Number New Accounts



G. Number New Credit Cards



H. Value New Credit Cards

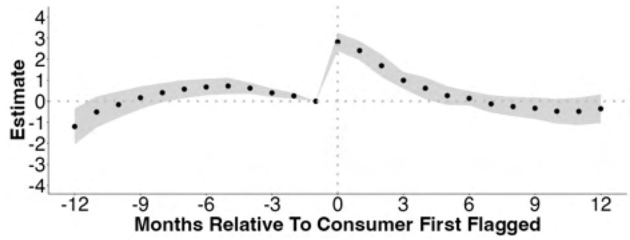


Notes: BTCCP data. Dots are unconditional means for 25 months of observations for a balanced panel, consisting of a flagged group in yellow and a matched unflagged group in black. Panels A to D show results for 0.14 million consumers with any delinquency twelve months pre-disaster. Panels E to H show results for 3.06 million consumers with no delinquency twelve months pre-disaster. X axis shows months since consumer first flagged. Panels show means for different outcomes: Panels A and E show credit score (VantageScore 3.0), Panels B and F show the number of new accounts opened that month, Panels C and G show the number of new credit card accounts opened that month, and Panels D and H show the value of new credit card limits opened that month ($).

Figure B11: Difference-In-Differences Estimates Of Average Effects on (A) Credit Score, (B) Any New Credit Card Openings, And (C) Value Of New Credit Card Limits

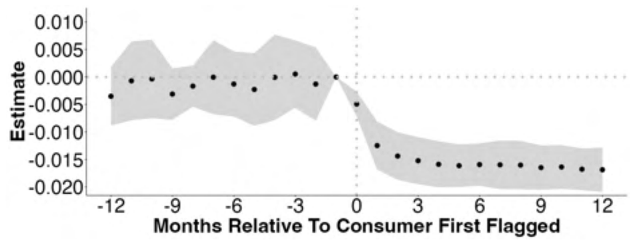
A. Credit Score

$t-1$ Mean: 704



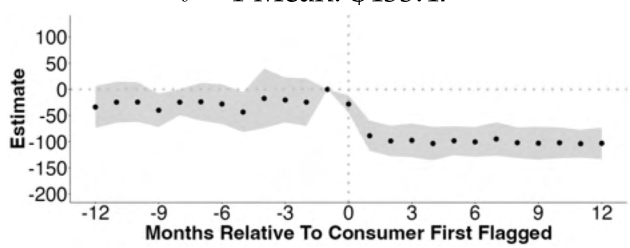
B. Number New Credit Card Openings

$t-1$ Mean: 0.0633



C. Value New Credit Card Limits

$t-1$ Mean: \$433.47

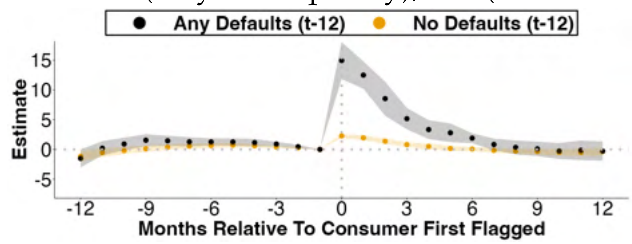


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 10's δ_τ. δ_τ are the coefficients on the interaction 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 B12: Difference-In-Differences Estimates Of Heterogeneous Treatment Effects Of Disaster Flags On Credit Scores And Credit Openings, By Any Pre-Disaster Delinquency

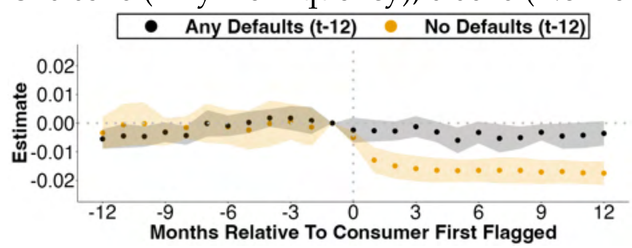
A. Credit Score

$t-1$ Means: 547 (Any Delinquency), 711 (No Delinquency)



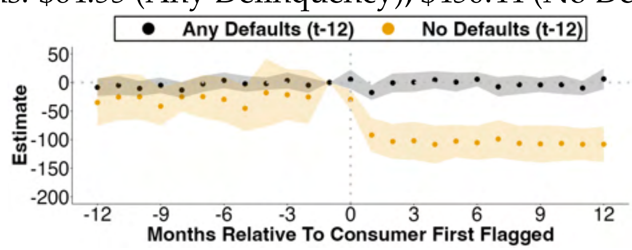
B. Number New Credit Cards

$t-1$ Means: 0.0346 (Any Delinquency), 0.0646 (No Delinquency)



C. Value New Credit Card Limits

$t-1$ Means: \$64.55 (Any Delinquency), \$450.44 (No Delinquency)

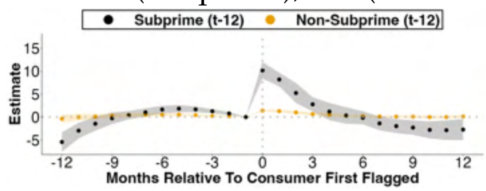


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 10's δ_τ. δ_τ are the coefficients on the interaction 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 delinquency 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 delinquency ($t-12$), and the orange estimates use data for the 95.6% (3.06 million) subsample without any pre-disaster delinquency.

Figure B13: 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

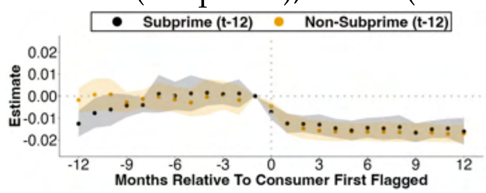
A. Credit Score

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



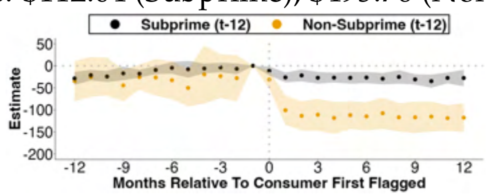
B. Number New Credit Card Openings

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



C. Value New Credit Card Limits

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

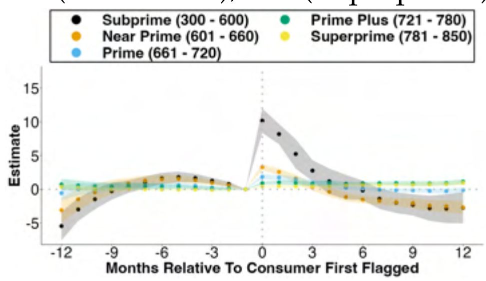


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 10's δ_τ. δ_τ are the coefficients on the interaction 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 B14: 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

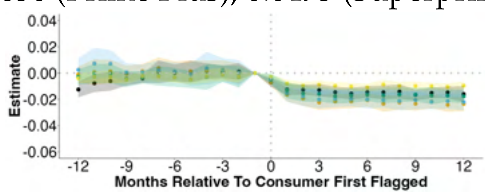
A. Credit Score

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



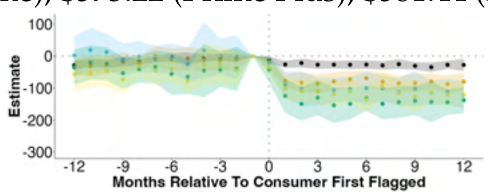
B. Number New Credit Card Openings

$t-1$ 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)

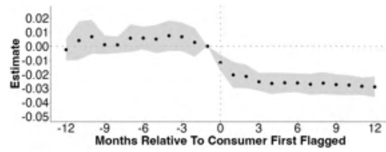


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 10's δ_τ. δ_τ are the coefficients on the interaction 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 B15: 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 Delinquency, And (C) Pre-Disaster Credit Score

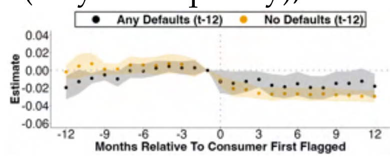
A. All

$t-1$ Mean: 0.1380



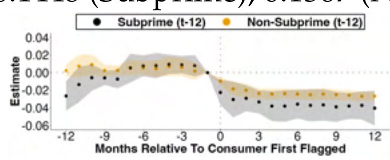
B. By Any Pre-Disaster Delinquency

$t-1$ Means: 0.0984 (Any Delinquency), 0.1398 (No Delinquency)



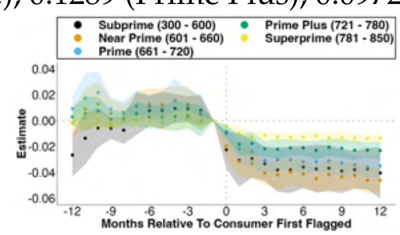
C. By Pre-Disaster Credit Score

$t-1$ 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)

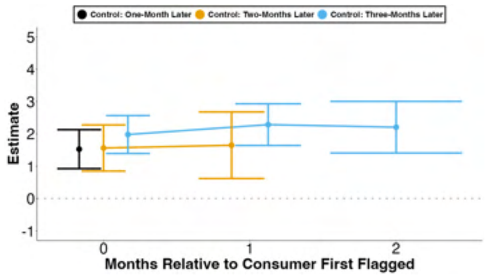


Notes: BTCCP data. Stacked difference-in-differences regressions from a balanced panel of 3.2 million consumers, where each flagged consumer is matched to an unflagged consumer, containing 25 monthly observations. Plots show estimates of Equation 10's δ_τ. δ_τ are the coefficients on the interaction 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 delinquency, and the orange estimates for the 95.6% (3.06 million) consumers without any pre-disaster delinquency. 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.

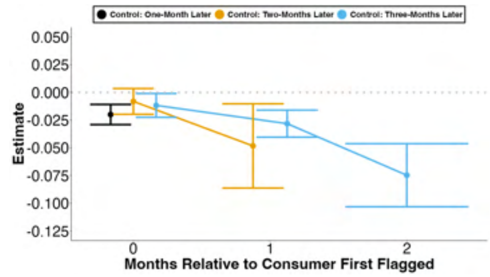
Figure B16: Estimates Of The Effects Of Disaster Flags On Consumers From Difference-In-Differences Exploiting One, Two, and Three Month Variation in Flag Timing

I. All

A. Credit Score

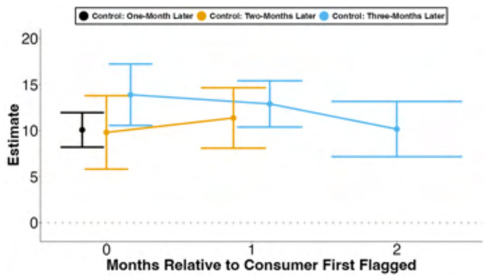


B. Number New Accounts

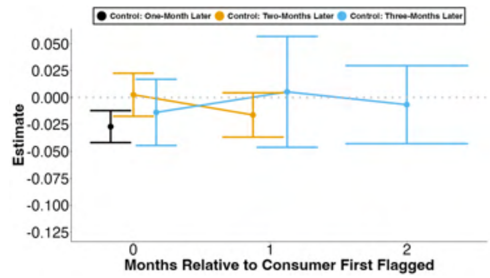


II. Sample With Any Delinquency Pre-Disaster

C. Credit Score

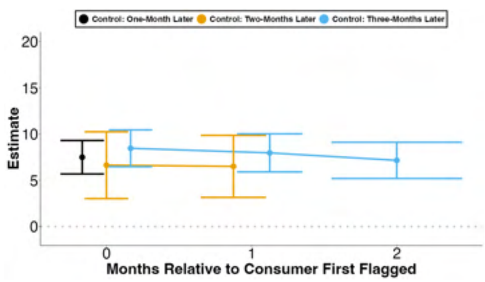


D. Number New Accounts

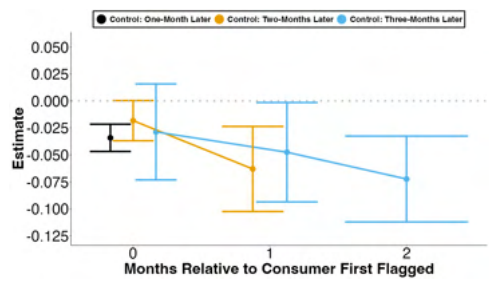


III. Sample With Subprime Credit Score Pre-Disaster

E. Credit Score



F. Number New Accounts



Notes: BTCCP data. Table show estimates of Equation 11's ρ with standard errors in parentheses. ρ is the coefficient on the treatment indicator, in the month a consumer first has a disaster flag. The treated group are consumers first flagged and the control group are first flagged one-month (black coefficients), two-months (orange coefficients), or three-months (blue coefficients) later, in the same set of geographic area and other characteristics. These one-, two-, and there-month outcomes are estimated in separate regressions. Regressions estimated also contain set-by-time fixed effects as explained in detail in Section B.5. 95% confidence intervals shown where standard errors are clustered at the calendar year-month level. Each column shows results from separate regressions with different outcomes. Each outcome is a first difference relative to the month before the treated group is first flagged. The outcome in Panels A, C, and E is the credit score (VantageScore 3.0). The outcome in Panels B, D, and F is the number of accounts opened that month. Panels A and B use data for 1,669,391 observations for one-month, 429,945 observations for two-month, and 90,503 observations for three-month variation in flag timing. Panels C to F are for subsamples of consumers by their pre-disaster characteristics. Panels C and D use data for the observations with any pre-disaster delinquency: This is 39,559 observations for one-month, 9,080 observations for two-month, and 2,573 observations for three-month variation in flag timing. Panels E and F use data for the observations with subprime (300 to 600) pre-disaster credit scores: This is 172,055 observations for one-month, 40,384 observations for two-month, and 9,897 observations for three-month variation in flag timing.

Table B1: 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 Delinquency	14.93	-0.0024	5.5
	(1.59)	(0.0022)	(8.6)
C. No Delinquency	2.27	-0.0050	-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	1.40	-0.0045	-31.4
	(0.09)	(0.0012)	(10.0)

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 10'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 Delinquency' uses data for the 4.4% (0.14 million) consumers with any pre-disaster delinquency. Row 'C. No Delinquency' uses data for the 95.6% (3.06 million) consumers without any pre-disaster delinquency. 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 B2: 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 Delinquency	-0.29	-0.0036	6.3
	(0.83)	(0.0022)	(9.3)
C. No Delinquency	-0.36	-0.0175	-108.1
	(0.33)	(0.0021)	(15.8)
D. Subprime	-2.74	-0.0160	-27.7
	(1.18)	(0.0031)	(9.6)
E. Non-Subprime	0.11	-0.0170	-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 10'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 Delinquency' uses data for the 4.4% (0.14 million) consumers with any pre-disaster delinquency. Row 'C. No Delinquency' uses data for the 95.6% (3.06 million) consumers without any pre-disaster delinquency. 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 B3: Difference-In-Differences Estimates (S.E.) Of Effects of Disaster Flags On New Account Openings In (1) Month First Applied ($t = 0$) And (2) After Twelve Months ($t = 12$)

Sample	(1) t=0	(2) t=12
A. All	-0.0117	-0.0289
	(0.0029)	(0.0037)
B. No Delinquency	-0.0116	-0.0294
	(0.0028)	(0.0037)
C. Any Delinquency	-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 10'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 Delinquency' uses data for the 95.6% (3.06 million) consumers without any pre-disaster delinquency. Row 'C. Any Delinquency' uses data for the 4.4% (0.14 million) consumers with any pre-disaster delinquency. 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 B4: Difference-In-Differences Estimates (S.E.) Of Effects Of Disaster Flags On Credit Access Outcomes 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)	(13.0)
D. Prime Plus	0.91	-0.0045	-41.7
	(0.11)	(0.0020)	(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 10'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 B5: Difference-In-Differences Estimates (S.E.) Of Effects Of Disaster Flags On Credit Access Outcomes 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	-0.0160	-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)	(21.0)
E. Superprime	0.97	-0.0096	-120.0
	(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 10'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 B6: Estimates (S.E.) Of The Effects of Disaster Flags In The Month First Applied On Consumers From Difference-In-Differences Exploiting One-Month Variation In Flag Timing

Sample	(1) Credit Score	(2) Number New Accounts	(3) Number New Credit Cards	(4) Value New Credit Card Limits
A. All	1.53	-0.0200	-0.0176	-153.2
	(0.31)	(0.0046)	(0.0051)	(42.7)
B. No Delinquency	1.29	-0.0198	-0.0179	-156.8
	(0.27)	(0.0048)	(0.0052)	(43.7)
C. Any Delinquency	10.07	-0.0270	-0.0095	-22.7
	(0.96)	(0.0076)	(0.0026)	(10.6)
D. Non-Subprime	0.81	-0.0183	-0.0173	-164.4
	(0.10)	(0.0053)	(0.0054)	(44.9)
E. Subprime	7.50	-0.0342	-0.0205	-59.5
	(0.92)	(0.0065)	(0.0035)	(14.4)
F. Near Prime	2.10	-0.0211	-0.0164	-71.8
	(0.35)	(0.0082)	(0.0087)	(57.1)
G. Prime	1.04	-0.0207	-0.0191	-141.9
	(0.12)	(0.0067)	(0.0070)	(53.1)
H. Prime Plus	0.57	-0.0206	-0.0194	-203.7
	(0.08)	(0.0059)	(0.0058)	(50.5)
I. Superprime	0.25	-0.0137	-0.0150	-191.0
	(0.08)	(0.0033)	(0.0031)	(39.0)

Notes: BTCCP data. Difference-in-differences regressions on a dataset of 1.7 million observations. Table show estimates of Equation 11's ρ with standard errors in parentheses. ρ is the coefficient on the treatment indicator, in the month a consumer first has a disaster flag. The treated group are consumers first flagged and the control group are first flagged one month later, in the same set of geographic area and other characteristics. Regressions estimated also contain set-by-time fixed effects as explained in detail in Section B.5. 95% confidence intervals shown where standard errors are clustered at the calendar year-month level. Each column shows results from separate regressions with different outcomes. Each outcome is a first difference relative to the month before the treated group is first flagged. The outcome in column (1) is credit score (VantageScore 3.0). The outcome in column (2) is the number of accounts opened that month. The outcome in column (3) is the number of new credit card opened that month. The outcome in column (4) is the value of new credit card limits opened that month ($). Row 'A. All' uses data for all 1.7 million observations. Rows B to E are for subsamples of consumers by their pre-disaster characteristics. Row 'B. No Delinquency' uses data for the 97.6% (1.63 million) observations without any pre-disaster delinquency. Row 'C. Any Delinquency' uses data for the 2.4% (0.04 million) observations with any pre-disaster delinquency. Row 'D. Non-Subprime' uses data for the 89.7% (1.50 million) observations with non-subprime (601 to 850) pre-disaster credit scores. Row 'E. Subprime' uses data for the 10.3% (0.17 million) observations with subprime (300 to 600) pre-disaster credit scores. 'F. Near Prime' uses data for the 15.2% (0.25 million) observations with near prime (601 to 660) credit scores. Row 'G. Prime' uses data for the 18.3% (0.31 million) observations prime (661 to 720) credit scores. Row 'H. Prime Plus' uses data for the 24.7% (0.41 million) observations with prime plus (721 to 780) credit scores. Row 'I. Superprime' uses data for the 31.5% (0.53 million) observations with superprime (891 to 850) credit scores.