Working Papers

FEDERAL RESERVE BANK OF BOSTON™

No. 23-17

The Impact of a Man-made Disaster on Consumer Credit Outcomes: Evidence from the 2018 Merrimack Valley Natural Gas Explosions

Bo Zhao

Abstract:
This paper is the first to empirically examine the impact of a man-made disaster on consumer credit outcomes. It uses the 2018 Merrimack Valley natural gas explosions as a quasi-random natural experiment and shows that the explosions had a temporary negative effect on debt balances, credit limits, and the number of delinquencies, and did not affect credit scores. The decreases in debt balances and credit limits were likely driven by a decline in credit demand when the affected individuals faced severe life disruption, great uncertainty, and negative financial shocks associated with the disaster. It took some time for the explosions to have an impact on delinquencies, suggesting that the affected individuals may have received short-term forbearance or used default as a last resort. The lack of large, long-lasting effects of the explosions likely reflects the critical role that external assistance to the affected communities played in mitigating the disaster's impact.

JEL Classifications: D14, G51, Q59
Keywords: Man-made disasters, household credit, consumer debt

---

Bo Zhao is a senior economist in the Federal Reserve Bank of Boston Research Department. His email address is Bo.Zhao@bos.frb.org.

The author thanks the Massachusetts Emergency Management Agency for data assistance and Eli Inkelas for excellent research assistance. He is grateful to the participants of the Federal Reserve Bank of Boston Research Department's internal seminar for their helpful comments.

This version: September 2023 **https://doi.org/10.29412/res.wp.2023.17**

I. Introduction

Disasters cause significant human and economic losses. Swiss Re Institute, a leading international insurance research organization, estimates that global economic losses caused by disasters increased from $200 billion in 2020 to close to $300 billion in 2022, and the number of people believed to have died or gone missing in disaster events increased from 7,993 in 2020 to 35,157 in 2022 (Bevere and Weigel 2021; Bevere and Remondi 2022; Banerjee et al. 2023).[1] While the lion's share of these estimated losses was attributable to natural catastrophes, man-made disasters, such as major fires and explosions and mining accidents, took an estimated toll of about $10 billion and claimed 2,000 to 3,600 victims each year in this three-year period.[2] These figures are certainly underestimates given how narrowly economic losses and man-made disasters are defined in these studies.

Greater access to credit bureau reporting data in recent years has led to more research on the impact of disasters on consumer credit outcomes. However, this emerging literature is still small, with only seven papers focused exclusively on natural disasters and mostly on major hurricanes. Gallagher and Hartley (2017) study Hurricane Katrina. Edmiston (2017) studies hurricanes that occurred during the 2000–2014 period. Billings, Gallagher, and Ricketts (2022) and Del Valle, Scharlemann, and Shore (forthcoming) study Hurricane Harvey. Gallagher,

[1] Swiss Re Institute defines *economic losses* as all the financial losses directly attributable to a major disaster (such as damage to buildings, infrastructure, and vehicles) and losses due to business interruption as a direct consequence of the property damage. This definition does not include indirect financial losses (such as loss of earnings by suppliers due to disabled businesses), estimated shortfalls in GDP, and non-economic losses (such as loss of reputation or impaired quality of life). The institute recommends considering its estimate of economic losses only as an indication of the general order of magnitude.
[2] Swiss Re Institute's definition of *man-made disasters* includes major fires and explosions, aviation and space disasters, shipping disasters, rail disasters, mining accidents, collapse of buildings and bridges, and miscellaneous major events, including terrorism. It excludes war, civil war, and war-like events.

Hartley, and Rohlin (2023) study 34 tornadoes. Tran and Sheldon (2017) and Ratcliffe et al. (2020) study Federal Emergency Management Agency (FEMA)–declared natural disasters, the most common types of which are hurricanes, tropical storms, and other severe storms. All these authors use credit bureau data, mainly from the Federal Reserve Bank of New York/Equifax Consumer Credit Panel (CCP), except for Del Valle, Scharlemann, and Shore (forthcoming), who use regulatory data for stress-testing (CCAR FR Y-14M). The studies differ in their geographic level of analysis, ranging from the census-block level to the county level.[3]

The existing research finds mixed results on the impact of natural disasters on credit. Some papers show an increase in the number of delinquencies and bankruptcies and a decrease in debt balances, credit supply, and credit scores after disasters. Other papers find no evidence of these negative effects. Taken as a whole, the literature suggests that natural disasters tend to have a temporary, modest impact on consumer credit outcomes (Del Valle, Scharlemann, and Shore, forthcoming). Studies attribute this finding to robust federal disaster aid (including Small Business Administration [SBA] loans and FEMA grants) and insurance payouts that mitigate the negative effects of natural disasters.

To the best of my knowledge, this is the first paper to study the impact of a man-made disaster on consumer credit outcomes. Man-made and natural disasters can differ in many dimensions, such as the causes; the scale, length, and type of impact; the source and size of external assistance that the affected areas receive; and the possible prevention mechanisms. Therefore, it is

[3] The geographic level in Tran and Sheldon (2017) is the county level. The geographic level in Ratcliffe et al. (2020) is the Zip code level. The geographic level in Del Valle, Scharlemann, and Shore (forthcoming) is the Zip+4 level. The geographic level in Edmiston (2017) is the US census–tract level. The geographic level in Gallagher and Hartley (2017), Billings, Gallagher, and Ricketts (2022), and Gallagher, Hartley, and Rohlin (2023) is the US census-block level.

not clear that the results and policy recommendations derived from studying natural disasters can necessarily apply to man-made disasters.

This paper uses the 2018 Merrimack Valley natural gas explosions as a quasi-random natural experiment. This event differs from previously studied natural disasters in several ways. It affected parts of three Massachusetts towns, an area that is much smaller than the impact area of hurricanes Katrina and Harvey. Despite the substantial damage it caused, this event did not receive a Presidential Disaster Declaration from FEMA. Therefore, the affected residents and towns were not eligible for federal disaster aid. In addition, property damage was less widespread and, on average, less severe than that caused by hurricanes Katrina and Harvey. Therefore, the negative financial shock associated with this event is likely much smaller than that associated with major natural disasters.

This paper studies five categories of credit outcomes: debt balances, credit limits, number of credit accounts, debt delinquencies, and Equifax Risk Scores. To the best of my knowledge, it is the first study to examine the impact of a disaster on consumer credit limits and the total number of accounts past due.

II. Brief Summary of the 2018 Merrimack Valley Natural Gas Explosions

The 2018 Merrimack Valley natural gas explosions were caused by an unexpected human error (National Safety Transportation Board 2019). During routine maintenance work in Lawrence, Massachusetts, on September 13, 2018, a Columbia Gas contractor made an error that resulted in high-pressure natural gas flowing into a low-pressure gas pipeline network serving parts of three towns in the Merrimack Valley region: Andover, Lawrence, and North Andover. The standard pressure in the low-pressure pipelines is roughly 0.5 pounds per square inch, but the pressure level of the mis-delivered gas reached about 75 pounds per square inch.

The immediate impact was significant, causing at least three explosions and 60 to 80 structure fires. One person was killed, and 25 peoples were injured. The explosions and the associated fires destroyed about 15 homes and left many more houses uninhabitable. It severely damaged more than 40 miles of underground gas lines and thousands of gas meters and gas-fueled home appliances. It was estimated to have affected about 7,300 residential units, 8,600 families, and 685 commercial units (Massachusetts Emergency Management Agency 2020).

Figure 1 shows the gas pipelines and the geographic area that were affected by the explosions. The area included the entirety of southern Lawrence (south of the Merrimack River) and northern parts of Andover and North Andover. Not shown in the figure are high-pressure gas pipelines located outside the area that were not connected to the affected low-pressure pipelines. Customers whose gas meters were connected to these high-pressure pipelines continued to receive gas service as usual after the explosions.

The effects on residents and businesses lasted much longer than the immediate impact. Appendix Figure 1 shows the overall event timeline. Many residents chose to remain in their homes until gas service could resume safely. Two large initiatives, Operation Space Heater and Operation Hot Plate, were intended to provide these people with interim means of heating their homes and cooking food. However, these initiatives had mixed results, leaving many families unsatisfied (Massachusetts Emergency Management Agency 2020). Families whose homes had no hot water had to use shower trailers and community shower facilities, which was a daily disruption and significant inconvenience.

Many people needed temporary housing because their homes became unsafe or uninhabitable. The options for temporary housing included apartments, trailers, hotel rooms, and, as a last resort, a shelter. The Massachusetts Emergency Management Agency (MEMA) reported that as many as 92 apartments and 376 trailers were used at peak occupancy and that as many as 187 people registered at the shelter at peak. The trailers were not sufficiently insulated for cold weather and were not ideal for families or those with disabilities (Massachusetts Emergency Management Agency 2020). Families and individuals with disabilities were sometimes placed farther away from their communities, where shelter could be coordinated. More than 65 families were permanently displaced because their homes were severely damaged. There were also reported cases of landlords refusing to accept tenants back into their original apartments.

Columbia Gas carried out two initiatives, Gas Ready and House Ready, to restart gas service in the affected areas. The company replaced more than 40 miles of damaged gas lines and repaired or replaced damaged house appliances in each affected unit, which was a long and complicated process. According to officials in the three towns, most families did not regain gas service until December. While the official end date of the event was marked as December 16, 2018—more than three months after the initial explosions—MEMA noted that the affected communities were still dealing with the consequences of the explosions more than a year later, in January 2020, as some debris and abandoned pipelines needed to be removed from streets (Massachusetts Emergency Management Agency 2020).

The affected communities received financial and nonfinancial assistance from the state, local governments, nonprofits, and community groups. Within days after the explosions, the state established a Recovery Resource Center, which combined many services, including housing assistance, food assistance, health and welfare services, and financial services. The program assisted more than 3,000 individuals in the first week after the explosions. In the following weeks and months, nonprofits such as the American Red Cross and the Salvation Army provided goods, services, gift cards, and store vouchers to many affected families. The Greater Lawrence Disaster Relief Fund (GLDRF) was established to serve the short- and medium-term shelter and sustenance needs of affected residents and businesses. By December 31, 2018, the GLDRF had delivered a total of more than $12 million in stipend checks and gift cards to over 10,000 households and businesses and had offered case-management support to more than 1,500 families. The Merrimack Valley Business Relief Fund was separately established to support the needs of the affected businesses. It planned to allocate $10 million over several years, including $2 million from an acute fund to address the most pressing needs of businesses.

However, the affected populations and towns did not receive federal disaster aid. Shortly after the explosions, the state made a request to FEMA for a Presidential Disaster Declaration. On December 10, 2018, FEMA informed Massachusetts officials that the agency was denying the request. As a result, the affected individuals and towns were not eligible to receive FEMA grants or loans from the US Small Business Administration (SBA).

III. Data

This paper uses data from multiple sources. Data on the gas explosions were provided by MEMA, which originally obtained the information from Columbia Gas of Massachusetts. The MEMA data include a map of the affected gas pipelines and a list of the names of the affected streets. Some of these streets are long and may therefore not be entirely affected by the gas explosions. MEMA also maintained a database of more than 200 properties that were inspected by the agency and the Massachusetts Office of Public Safety and Inspections (OPSI) in late September 2018 to assess the damage from the gas explosions. The inspectors recorded the property addresses and their damage levels.

Data on consumer credit outcomes are from the Federal Reserve Bank of New York/Equifax Consumer Credit Panel (CCP). This data set contains two anonymous samples. The "primary" sample is a nationally representative 5 percent anonymous random sample of all adults with a Social Security number (SSN) and a credit report that includes at least some credit history or public-record information. The "non-primary" sample includes anonymous consumers who live at the same address as "primary" members. Non-primary members are not required to have an SSN and may have inquiry-only files with limited credit information.

As its name suggests, the CCP comprises panel data in which each consumer is assigned a unique identifier that remains constant over time. Equifax provides the data on a quarterly frequency and typically pulls the credit bureau information on the last Tuesday of the final month of each quarter.[4] Thus, in the CCP data, the third quarter of 2018 is the first period after the gas explosions.

The CCP contains rich detail on consumer credit and debt. This paper focuses on five categories of credit outcomes: debt balance, credit limit, number of credit accounts, debt delinquency, and Equifax Risk Score. To avoid double counting associated with jointly held accounts, Lee and Van der Klaauw (2010) recommend that half of the debt balance, credit limit, and amount past due of a jointly held account be added to those of the individual accounts to

[4] Equifax temporarily changed the CCP data frequency from quarterly to monthly in 2020 and 2021. It typically pulled the credit report information on the last Tuesday of each month in those two years. However, it reverted to quarterly frequency in 2022. For this analysis, I use quarterly frequency in 2020 and 2021, which means that I use data from March, June, September, and December as the Q1, Q2, Q3, and Q4 data. Doing so maintains a consistent data frequency across the entire sample period. It also enables me to retain more individuals from the sample than I would otherwise because these people may have missing values in outcome variables in months other than March, June, September, and December in 2020 and 2021.

calculate the debt balance, credit limit, and amount past due for each individual who holds the joint account. I make this recommended data adjustment. The Equifax Risk Score is a composite indicator of overall credit health that predicts the probability of a consumer becoming severely delinquent (90-plus days past due) within the next 24 months. It differs from the FICO score because the two are created independently using different models; however, it is a close proxy for the FICO score. Equifax Risk Scores range from 280 to 850; FICO scores range from 300 to 850. In both systems, higher scores indicate lower credit risk.

To protect consumers' anonymity, their addresses are scrambled in the CCP and cannot be merged with data on the locations of the affected pipelines and streets. However, Equifax's geo-vendors match consumers' addresses, which are generally where creditors send account mail, to census blocks using the 2010 census-block boundaries. Blocks are the smallest geographic level in the census data. Because the three affected towns are relatively urban, many of the census blocks are single street blocks covering a compact area.

The CCP has several limitations. First, it does not include any socioeconomic or demographic information about individuals except their years of birth. There are no variables indicating the race, gender, income, or employment of each individual. However, credit bureaus state that they do not use the employment information in calculating credit scores or making credit decisions.[5] Instead, they use it as supplemental information to determine a consumer's identity. Second, the CCP does not report assets, savings, or consumption. Therefore, I cannot use the CCP to examine the impact of the gas explosions along these dimensions. Third, there are no consistent

[5] See Erica Sandberg, "Is Employment Listed in Your Credit Report?" Experion blog, February 5, 2020, https://www.experian.com/blogs/ask-experian/is-employment-listed-in-your-credit-report/.

unique household identifiers to enable tracking of the same households over time. Therefore, I cannot perform household-level panel regression analysis.

The 2010 census is the final data source used in this paper. I obtain block-level data for the three affected towns on racial and ethnic composition, status of household heads, and housing tenure. The 2010 census does not contain any information about income. While the American Community Survey (ACS) five-year estimates have income variables, they are at the block-group level. Block groups are too geographically large for my research design because each block group contains dozens of blocks and may encompass blocks in the treatment group as well as blocks in the control group.

To create the regression samples, I first pull all CCP individuals whose addresses in 2018:Q3 (that is, the end of September 2018) were matched by Equifax's geo-vendors to blocks in the three towns affected by the gas explosions. I include not only primary members but also non-primary members to pool a sufficiently large sample. Next, I track these individuals three years before and three years after the gas explosions. By doing so, I create a panel consisting of 25 quarters from 2015:Q3 through 2021:Q3. This sample period length is the same as that used by Gallagher and Hartley (2017), who argue that this length is "a compromise between longer length panels that would have limited the number of individuals in the sample [due to dropping more individuals with missing values over a longer period] and shorter length panels that would have only allowed us to estimate the relatively short-term effects" (p. 210).

Then, I merge this CCP sample with the geodata on the affected areas and the 2010 census data by the blocks of individuals' residences in 2018:Q3. I use the geographic information system (GIS) to measure the distance of each census block in the three affected towns from the affected pipelines and the affected streets. Finally, I apply some restrictions to create a balanced panel and to ensure the reliability of the data (see Appendix 1). The most binding restriction drops individuals with missing values in each dependent variable during the sample period. Other restrictions have only minor effects on the sample selection and amount to removing individuals with data-quality issues (for example, addresses flagged as non-residential or matched to blocks with zero reported population according to the 2010 census).

Because I do not observe the street address of each individual in the CPP, I define the treatment and control groups at the census-block level and assign that status to all individuals living in the same block. As a result, no variation within blocks is used for identification. My estimate is, in fact, the average treatment effect on the intended, not on the treated. Thus, it is an underestimate of the true effect of the gas explosions on the affected individuals whom I do not directly observe.

I define the treatment group to include individuals who in 2018:Q3 resided in census blocks that intersect both the affected pipelines and at least one street on the list of affected streets. Using this definition, I can reasonably assume that at least some people in these blocks were likely to be affected by the gas explosions. Depending on which outcome variable is studied, the number of blocks in the treatment group varies from 276 to 318, and the number of individuals in the treatment group varies from 1,000 to more than 1,600.

I define the control group to include individuals who in 2018:Q3 resided in census blocks that are within the three affected towns and whose boundaries are at least 30 meters from the nearest affected pipelines. I choose the control group from the same three affected towns as the treatment group because the two groups faced the same local economies and the same local government policies, and because their observable and unobservable characteristics are likely similar due to Tiebout sorting. I use 30 meters as the distance threshold because commercial utility gas service providers in Massachusetts offer customers free gas line installation if their gas meters are within 100 feet, or roughly 30 meters, of existing gas lines.[6] Given the extremely high cost of self-funding extra gas lines beyond 100 feet, households' gas meters were unlikely to connect to the affected pipelines if these households lived in census blocks that were more than 30 meters from any of those pipelines. They could instead have gas delivered through the unaffected high-pressure gas pipelines that were closer to their census blocks and outside the affected area. To test the sensitivity of the results to the 30-meter threshold, I later use 100 meters and 1,000 meters as alternative thresholds in choosing the control group. The regression results are almost the same as those using the 30-meter threshold. Depending on which outcome variable is studied, the number of blocks in the control group varies from 669 to 745, and the number of individuals in the control group varies from 2,924 to more than 4,400.

I assign an individual to the treatment or control group based on their census block in 2018:Q3. Then, I keep their treatment- or control-group status constant over time, allowing them to move in and out of the affected area before and after the gas explosions. Gallagher and Hartley (2017) and Billings, Gallagher, and Ricketts (2022) similarly allow individuals to migrate to and from the areas affected by hurricanes Katrina and Harvey before and after the disasters while holding each individual's treatment- or control-group status constant throughout their sample periods. One might prefer that the treatment-group individuals remain in the affected area and the

[6] For example, Eversource, which officially acquired Columbia Gas of Massachusetts in 2020 (see WBZ News, "Eversource Officially Acquires Columbia Gas as Result of 2018 Explosions," October 30, 2020), states on its website that it offers "(f)ree gas service line installation (a $5,000 value) when there is an existing gas main in the street in front of your home (up to 100 feet)"; see https://www.eversource.com/content/residential/services/connect-to-gas (accessed September 8, 2023). National Grid, which is another major piped-gas-service provider in the state, said in its marketing materials that "National Grid will also provide up to 100 feet of service line from the meter location to the street"; see https://www.nationalgridus.com/media/pdfs/resi-ways-to-save/gas-growth/pex_onlineguide_ma_12_1.pdf#:~:text=National%20Grid%20will%20provide%20up%20to%20100%20feet,quote%20to%20extend%20the%20main%20to%20your%20house (accessed September 8, 2023).

control-group individuals remain in the unaffected area so that a possibly cleaner estimate can be obtained and a potentially biasing of results toward zero can be avoided; however, similarly to Gallagher and Hartley (2017) and Billings, Gallagher, and Ricketts (2022), I consider that requirement unnecessary and costly. If I restrict the treatment group to include only individuals who never left the affected area during the 2017:Q3–2021:Q3 period, I would have to drop 30 to 40 percent of the original treatment-group individuals, depending on which outcome variable is studied, and therefore lose the power of statistical tests.

Another possible concern is that if more vulnerable populations were more likely to live in the areas serviced by the low-pressure gas pipelines before the explosions, this could result in an overestimate of the explosions' effects. However, there is no obvious reason to believe that individuals sorted themselves into different census blocks based on the high- versus low-pressure level of gas delivery. First, the difference between these two systems is technical and likely not known to customers in the area. Second, there was no difference in price and quality between the natural gas delivered through the low-pressure pipelines and the gas delivered through the high-pressure pipelines. Therefore, customers had no economic incentive to move so that they could have a different pressure level in the gas mains. In addition, the city of Lawrence offers a strong case against the hypothesis that more vulnerable people were likely to live in the affected area before the gas explosions. People living in the southern section of Lawrence, which was affected by the gas explosions, were, on average, wealthier and therefore could be considered less financially vulnerable than people living in the northern section of Lawrence, which was not affected by the explosions.

IV. Empirical Model

Following most of the literature, I use an event-study model to examine the dynamic effect of the gas explosions on consumer credit outcomes. I do not use the traditional static difference-in-differences model (DID) because it estimates the average treatment effect over the entire post-event period under an assumption that the treatment effect is relatively stable over time. However, both the descriptive trend figures and the regression results suggest that this assumption is not satisfied for many outcome variables. In the context of studying the impact of natural disasters, previous studies show that their data also fail to meet this assumption, as the effect proves to be short-lived (for example, Gallagher and Hartley 2017 and Edmiston 2017). Estimating the average effect over the entire post-event period would therefore mask the important dynamics of the effect over time.

The preferred model can be expressed as follows:

$$
\begin{aligned}
Y_{ibt} = {} & \beta_{-11} Time_{-11} Treat_{ib} + \beta_{-10} Time_{-10} Treat_{ib} + \cdots + \beta_{-1} Time_{-1} Treat_{ib} \\
& + \beta_1 Time_1 Treat_{ib} + \cdots + \beta_{13} Time_{13} Treat_{ib} + \gamma_1 age_{ibt} + \gamma_2 age_{ibt}^2 + \alpha_i + \alpha_t \\
& + \epsilon_{ibt} \ (1),
\end{aligned}
$$

where Y_{ibt} is a consumer credit outcome in year-quarter *t* for individual *i* living in census block *b* as of 2018:Q3, *Treat* is an indicator for the treatment-group status, $Time_x$ is a dummy variable for year-quarter that was x quarters from 2018:Q2 (the last year-quarter before the gas explosions), age is computed as the data year minus the year of birth, α_i is individual fixed effect, and α_t is year-quarter fixed effect. Age squared is added to account for nonlinear life-cycle patterns in consumer credit outcomes (Billings, Gallagher, and Ricketts 2022). Individual fixed effects control for unobserved time-invariant factors at the individual level that might influence consumer credit decisions, such as gender, race and ethnicity, and educational attainment. Because income tends to be slow moving, it can be at least partially accounted for by individual fixed effects. In addition, individual fixed effects account for characteristics of towns and blocks of individuals' residences in 2018:Q3. In an alternative specification, I replace individual fixed effects with fixed effects for block of residence in 2018:Q3; this specification is similar to an alternative model in Gallagher and Hartley (2017).[7] The results using this alternative specification are similar to the ones from my preferred model. Similarly to Gallagher and Hartley (2017), I prefer using individual fixed effects because they are at a lower level than census-block fixed effects and can control for more unobserved time-invariant factors than block fixed effects. Following Gallagher and Hartley (2017) and Billings, Gallagher, and Ricketts (2022), I cluster standard errors by block of residence in 2018:Q3 to allow for heteroskedasticity and correlation within blocks.

The coefficients of interest, β_x, measure the effect of the gas explosions on a consumer credit outcome variable in each time period. They are identified by comparing the change in the outcome variable from 2018:Q2 for treatment-group individuals with the same-period change for control-group individuals. The correct identification depends on the assumptions that the treatment and control groups are on parallel trends in the pre-explosions period and that no anticipation effect exists in that period. Because the explosions were caused by an accident during routine maintenance, the exact timing of the explosions could not be expected by residents of the three affected towns. Thus, there is no intuitive and logical basis for an anticipation effect in the pre-explosions period. Both assumptions will be empirically examined in the next section.

Based on data availability in the CPP, I study five categories of consumer credit outcomes. The first category is debt balances. The prediction for the impact of the gas explosions on consumer

[7] See Column 6 in Table 3 of Gallagher and Hartley (2017), in which census-block fixed effects are defined using the block of residence in 2005:Q3.

debt balance is ambiguous. On the one hand, consumers may borrow more to smooth consumption if they experienced a negative financial shock associated with the gas explosions. Some individuals may have lost a portion of their housing wealth and/or income (for example, rental income, labor income if their employers were temporarily shut down, or business income if their stores were closed). On the other hand, amid severe life disruption and uncertainty, consumers may have postponed, reduced, or eliminated major debt-financed purchases such as a house, cars or trucks, or major appliances. Previous research shows that people also use insurance payments to pay down a debt balance (Gallagher and Hartley 2017).

The second outcome-variable category is credit limits. Intuitively, it is unclear whether the gas explosions would have a positive or negative impact on consumers' credit limits. The affected individuals could have requested greater increases in their credit limits relative to the control group so they could borrow more to mitigate the negative financial shock associated with the explosions. However, they might not have needed to raise their credit limits as fast and as high as the control group if they postponed, reduced, or eliminated major debt-financed purchases amid the severe life disruption and uncertainty caused by the explosions. If they used aid and insurance payments to pay off existing debt, they might even have closed some loan accounts and had lower credit limits.

The third category is number of credit accounts. The prediction for the number of credit accounts is also ambiguous. Affected individuals could have opened more accounts to access more credit and borrow more. But if lenders considered lending in the affected area riskier after the explosions, they could have reduced the issuance of credit accounts to people living in that area.

The fourth category is debt delinquencies. Affected individuals might have had a reduced ability to make loan payments due to property damage and income loss caused by the disaster, which could have resulted in more delinquencies. Del Valle, Scharlemann, and Shore (forthcoming) also suggest that, after a negative financial shock, consumers might skip payments to effectively gain more liquidity.

The last outcome-variable category is Equifax Risk Scores. Previous research shows that natural disasters cause a decline in credit scores (Gallagher and Hartley 2017; Ratcliffe et al. 2020). Following a similar logic, the gas explosions might have worsened the overall credit health of the affected consumers and resulted in lower credit scores.

In addition, I explored other outcome variables in the CCP as shown in Appendix 2. Each either has a sample size that is too small due to many missing values or fails to satisfy the parallel-trends assumption. Therefore, I did not estimate Equation 1 on any of these variables.

V. Main Results

I present the results by the category of outcome variable.

1. Debt Balances

I examine total balance and two subcategories given the availability of usable data: bank card balance and student loan balance. I cannot analyze the balances of other types of credit accounts—including mortgage, home equity loan, home equity line of credit, auto bank loan, auto finance loan, consumer finance, retail trade, and other trades—because they have too many missing values to form a sample that is large enough to run a regression.[8]

[8] Gallagher and Hartley (2017) find that there was a spike in non-reporting of mortgage account information in New Orleans after Hurricane Katrina. There is no evidence of increasing non-reporting of mortgage accounts and other types of accounts after the gas explosions in this sample.

Figure 2 shows the simple average of the three debt-balance measures for the treatment and control groups separately in each period. The treatment and control groups appear to have parallel trends in the pre-explosions period regarding these debt-balance measures.

In Figure 3, I plot the coefficients β_x and their 95 percent confidence internals estimated from Equation 1. First, none of the coefficients in the pre-explosions period is statistically significant in all three panels. This provides further support for the parallel-trends assumption. In Panel (a), for total debt balance, the coefficients in the two periods immediately following the explosions are negative and statistically significant. Their point estimates are –$3,167 for 2018:Q3 and –$3,305 for 2018:Q4. They amount, respectively, to a 3.9 percent and 4.0 percent decrease from the four-quarter average debt balance of $82,129 immediately before the explosions (2017:Q3 through 2018:Q2) among treatment-group individuals. These effects are smaller than those of Hurricane Katrina. Gallagher and Hartley (2017) estimate that the average reduction in total debt balance in the three years after Hurricane Katrina ranged from –$4,489, or 9.5 percent of the pre-Katrina debt level, for individuals living in the least flooded blocks to –$11,092, or 23.6 percent of the pre-Katrina debt level, for individuals living in the most flooded blocks. This paper's estimates, in relative terms, are more aligned with those of Gallagher, Hartley, and Rohlin (2023), who find a 4 percent reduction in average quarterly credit card balances and auto debt balances and a 2 percent reduction in average quarterly mortgage balances for the three years after a tornado hit.

This short-term decline in total debt balance could be due to the treatment group's temporarily reducing debt-financed purchases relative to the control group. However, the coefficients after 2018:Q4 increase to nearly zero and are no longer significant. The lack of a symmetric increase in debt balance after the initial decline does not necessarily mean that temporal substitution did not occur. The affected individuals could still increase their debt balances in later periods. But because they could do so at any time instead of over only two quarters, as with the initial decline, the compensatory increase could be dispersed over a long stretch of time among treatment-group individuals. Therefore, it is difficult to detect a statistically significant increase in debt balance after the initial decline.

In Panels (b) and (c), no coefficients in the post-explosions period are significant. This suggests that the decline in total debt balance for the treatment group relative to the control group after the explosions was driven by types of accounts other than bank cards and student loans. I suspect that mortgages and auto loans were the most likely sources of the decline because they often hold the largest balance among credit accounts outside of bank cards and student loans. It is possible to imagine that the affected individuals reduced, postponed, or canceled borrowing for purchases of new homes or cars when they were facing a severe life disruption, great uncertainty, and possibly a negative financial shock. It warrants future research using another data set with a larger number of observations of mortgage and auto loan balances to test this hypothesis.

2. Credit Limits

I examine total credit limit and two subcategories: bank card credit limit and high credit for student loan.[9] I cannot analyze credit limits for mortgages, home equity loans, home equity lines of credit, auto bank loans, auto finance loans, consumer finance, retail trade, and other trades because they have too many missing values.

Based on the raw data, Figure 4 shows that trends in the measures of credit limits for the treatment and control groups were parallel in the pre-explosions period. Consistent with that

[9] Equifax defines *high credit* for a student loan as the balance at origination.

observation, Figure 5 shows that the regression coefficients in the pre-explosions period are never significant in all three panels.

Similar to the results for total debt balance, the coefficients in the first two periods after the explosions are negative and statistically significant in the regression for total credit limit (Panel a). Their estimated magnitudes are –\$4,624 for 2018:Q3 and –\$4,018 for 2018:Q4, which are \$700 to \$1,500 larger than the estimated average declines in total debt balance in these periods. They are equivalent, respectively, to a 3.8 percent and 3.3 percent decrease from the previous four-quarter average of \$121,560 in total credit limit among treatment-group individuals. For the periods after 2018:Q4, the estimated coefficients are not significant.

The coefficients in the post-explosions period in Panels (b) and (c) are never significant. I infer from these results that the decline in total credit limit for the treatment group relative to the control group was not driven by bank cards and student loans but possibly by mortgage and auto loan accounts given their prominent role in household debt portfolios. Testing this hypothesis with a more robust sample would be a valuable contribution to the literature.

3. Number of Credit Accounts

Was the decline in total credit limit and total debt balance caused by a decrease in credit demand or a decrease in credit supply after the gas explosions? To address this question, I create two proxy measures of credit supply following Gallagher and Hartley (2017) and Billings, Gallagher, and Ricketts (2022). One measure is the total number of credit accounts. The other is the number of newly opened credit accounts, which is defined as the change in the number of open accounts in the current quarter from the previous quarter if the change is positive. If there is no change or the change is negative, the number is zero. Figure 6 shows that the treatment and control groups were on parallel trends before the gas explosions for these two variables. Figure 7 shows that the coefficients are not significant in either the pre- or post-explosions period for both outcome measures. Therefore, there is no evidence suggesting an overall tightening of credit supply to the affected residents. This result is similar to that of Billings, Gallagher, and Ricketts (2022), who find no evidence of a retraction in credit supply after Hurricane Harvey. I have also not seen any anectodical evidence of a reduced credit supply in the affected area. In fact, it is well documented that financial institutions actively helped the affected residents after the gas explosions (Massachusetts Emergency Management Agency 2020).

While the CCP contains a variable for the number of credit inquiries initiated by the consumer within the previous three months, I cannot use that number as a direct measure of credit demand and run an event-study regression because the treatment and control groups were not on parallel trends in the pre-explosions period. The definition of this variable is also questionable because multiple inquiries in 30-day periods to the auto finance industry, for example, are counted as one inquiry in the CCP. Despite a lack of direct evidence of credit demand, I am inclined to infer that the decrease in overall credit limit and debt balance was more likely demand-driven than supply-driven.

4. Debt Delinquency

I construct three measures of debt delinquency: total number of accounts past due, an indicator for having at least one account 60-plus days delinquent, and an indicator for having a severe derogatory account. Equifax defines *severe derogatory* as having a charge-off to a bad loan, repossession, foreclosure, or a defaulted student loan. Balances in severe derogatory accounts may be in internal collections but not with third-party collection agencies. I also checked indicators for 90-plus days delinquent, 120-plus days delinquent, or in collections, and for bankruptcy. These delinquency measures for the treatment and control groups were not on parallel trends in the pre-explosions period. Therefore, they are not suitable for the event study. Ratcliffe et al. (2020) use an indicator for 60-plus days past due as the main measure for delinquent mortgages.

The delinquency measures after the gas explosions may be underreported in the CCP data. Mortgage lenders can offer a forbearance plan with an initial term of three months to homeowners affected by disasters.[10] The forbearance plan can be extended in three-month increments for as long as 12 months. During the forbearance period, regular payments are suspended while interest on debt balances continues to accrue. Consumers may make reduced payments, interest-only payments, or no payments. At the end of this period, borrowers are required to pay back any skipped payments plus the accrued interest over a short period of time unless they reach an agreement with lenders to recapitalize the missed payments and the accrued interests into the loan balance, which allows them to pay the debt over a longer period. During the forbearance period, lenders do not report the resulting delinquencies to credit bureaus (Billings, Gallagher, and Ricketts 2022). Depending on how many forbearance plans were in effect, my measures of debt delinquency could be an underestimate of the true delinquency rates after the gas explosions.

As shown in Figure 8, the delinquency measures for the treatment and control groups were largely on parallel trends before the explosions. The coefficients in the pre-explosions period in all three panels of Figure 9 are insignificant. For the total number of accounts past due, the results for 2019:Q1 and 2019:Q3 are significant, whereas the result for 2019Q2 is almost significant, with a p-value of 0.053. The estimated coefficients for 2019:Q1 and 2019:Q3 are 0.062 and 0.078, respectively. They are small in absolute terms, but they translate to increases of 16 percent and 20

[10] See Fannie Mae, *Servicing Guide: Fannie Mae Single Family*, July 12, 2023, https://singlefamily.fanniemae.com/media/36406/display.

percent, respectively, from the average number of accounts past due, 0.38, among treatment-group individuals in the four quarters before the explosions.

Unlike with the debt balances and credit limits, which change immediately after the explosions, the disaster appears to have taken some time to work through the delinquency channel. This suggests that after the gas explosions, the affected individuals might have received six-month forbearance from their lenders, then when the forbearance period ended, they did not have enough money to pay the amounts due and therefore defaulted. Gallagher and Hartley (2017) also find a delayed effect of Hurricane Katrina on debt delinquency. The increase in the share of the affected residents with at least one 90-day delinquent account does not become significant until four quarters after Katrina struck. The authors attribute the delay to a one-year forbearance plan for most mortgage holders in the flooded area. Another possible explanation is that defaulting on their debt may not have been the first choice for affected individuals; they may have resorted to default only after exhausting other options to cope with the negative shock (for example, reducing consumption and using savings).

While the results in Panel (a) are estimated using a linear regression, I employ a fixed effects Poisson model to re-estimate Equation 1, acknowledging that the number of accounts past due is a count measure. The results from the Poisson model are significant for 2019:Q1 through 2020:Q2 (Appendix Figure 2). Given that the linear model produces more conservative results with fewer significant periods and its results are easier to interpret compared with the incidence ratios from the Poisson model, the linear model remains preferred.

Panel (b) of Figure 9 shows that the regression for 60-plus days delinquency produces one significant coefficient in 2019:Q1. The estimate is 0.018, meaning that the treatment group had a 1.8 percentage point higher probability of having an account 60-plus days delinquent compared with the control group in that period.[11] That amounts to a 45 percent increase from the 4 percent average probability of having an account 60-plus days delinquent among treatment-group individuals in the four quarters before the explosions. This relative magnitude is similar to the ones estimated by Ratcliffe et al. (2020).

Panel (b) uses a linear probability model to estimate the likelihood of having an account 60-plus days delinquent. As a robustness check, I use a fixed effects logit model to re-estimate the equation. This alternative specification has at least three drawbacks. It drops all individuals that have no variation in the outcome variable over time, resulting in a substantially smaller sample. It includes only 1,263 individuals in the final sample compared with 5,435 individuals in the linear probability model sample and therefore has less testing power. The *xtlogit, fe* command in Stata does not allow for the clustering of standard errors. In addition, it does not converge after the default 100 iterations and shows no sign that it will ever converge. Among the reported results after the 100 iterations, the estimated odds ratio in 2019:Q1 is 1.57, but it is significant only at the 10 percent level. Given the drawbacks of the fixed effects logit model, the linear probability model remains the preferred specification for this outcome variable. Finally, all coefficients in Panel (c) are not significantly different from zero.

5. Equifax Risk Score

Figure 10 shows that the average Equifax Risk Scores for the treatment and control groups were on parallel trends before the gas explosions. While the estimated coefficients in the eight periods after the explosions are negative in Figure 11, they are not significant. No significant results are found, likely because credit bureaus considered individuals' long credit histories prior to the

[11] This is smaller than the 2.5 percentage point increase in the probability of having an account 90-plus days delinquent after Hurricane Katrina estimated by Gallagher and Hartley (2017).

explosions in computing credit scores, which could prevent the impact of the explosions from appearing statistically significant in the regression results. Similarly, Gallagher, Hartley, and Rohlin (2022) find no evidence for tornadoes having an impact on the Equifax Risk Scores of the affected communities.

6. Heterogeneity Analysis

I attempt to test whether more vulnerable populations and people who likely received more intense treatment were more affected by the gas explosions. I develop proxy measures of treatment intensity and individual vulnerability based on available data on individuals' ages, their pre-explosions credit records, and characteristics of their blocks of residence in 2018:Q3, including the percentage of streets that were affected, and the presence of properties with major damage from the explosions in each block. See Appendix 3 for the variables list.

I add interaction terms between one of these proxy variables and $Time_x Treat_{ib}$ and the interaction terms between the proxy variable and $Time_x$ to each new version of Equation 1. The interaction terms between the proxy variables and $Time_x Treat_{ib}$ are generally insignificant. When they occasionally become significant, they do so near the end of the sample period, which is counterintuitive because, before interaction terms are omitted from the model, the average effect is significant only in the earlier part of the post-explosions period. I am inclined to believe that these occasional significant interaction terms are a result of random noise in the data. Adding these interaction terms essentially divides the treatment group into smaller subgroups and then compares the impact of the explosions on one subgroup with the impact on the other. This approach requires that the treatment group is large so that the estimated interaction terms are reliable. Unfortunately, the treatment group in this analysis is small, consisting of 1,000 to 1,600 individuals, depending on which outcome variable is studied. In addition, the event-study framework exacerbates the small-sample problem. Even before adding the interaction terms, I compare only the difference between a single year-quarter and 2018:Q2 for the treatment group and the difference for the control group, which means that I already use small variations over time to identify the coefficients. After adding the interaction terms, I further compare the difference-in-differences between subgroups of the treatment group, which makes it entirely possible that some random noisy variation can drive significant results on the interaction terms. Therefore, I conclude that I do not have a large enough sample to conduct a reliable heterogeneity analysis. Future research using a larger data set to study the heterogeneous effect of man-made disasters is warranted.

VI. Robustness Checks

I conduct additional robustness checks surrounding mainly the definition of the control group. For simplicity, I present the results from these robustness checks for only the four outcome variables for which I find significant coefficients in the original analysis: total debt balance, total credit limit, number of accounts past due, and an indicator for an account 60-plus days delinquent. As with the original analysis, I do not find significant results for other outcome variables in these robustness checks. In addition, the parallel-trends assumption continues to be satisfied for the treatment and control groups defined in these robustness checks. The trend figures are available upon request.

First, I change the minimum distance from the affected gas pipelines from 30 meters to 100 meters when defining the control group. This robustness check serves two purposes: It tests how sensitive the results are to the distance threshold used, and it helps address a concern that the original control group might be contaminated because some people living slightly more than 30 meters from the affected pipelines may have paid to connect their gas meters to these low-pressure pipelines and were therefore directly affected by the gas explosions. However, this type of contamination, even if it occurred, would bias the results toward zero and result in an underestimate of the true effects of the gas explosions. Figure 12 shows that using the 100-meter threshold instead of the 30-meter threshold to define the control group does not change the results meaningfully.

While using the 100-meter threshold avoids including individuals who still might have been directly affected by the gas explosions in the control group, some people in the new control group may have been *indirectly* affected by the explosions, even if their gas meters were not connected to the affected pipelines. For example, they may have lost their jobs if their employers operated in the affected area and were negatively affected by the explosions. They may have lost their childcare if they sent their children to daycare providers in the affected area that had to shut down after the explosions. The number of such indirectly affected individuals is likely very small relative to the entire control group. Including them in the control group is unlikely to significantly affect the results and, if anything, is likely to bias the results toward zero.

To reduce the probability of including individuals indirectly affected by the gas explosions, I use a 1,000-meter threshold to define the control group. The underlying assumption is that most of the people who worked for employers or relied on daycare or other service providers in the affected area lived close to these businesses. Therefore, individuals living 1,000 meters or farther from the affected area were less likely to have been indirectly affected by the explosions compared with individuals living 30 to 1,000 meters from the affected area. Figure 13 shows that the results using this refined control group are similar to the results using the original control group.

Another possible concern is that the treatment and control groups may not be sufficiently similar in terms of factors that influence credit outcomes, and that including age, age squared, and individual fixed effects in the model may not be enough to control for the differences between the treatment and control groups, which could potentially result in biased results. However, the difference-in-differences approach does not require that the treatment and control groups are very similar. The approach is valid as long as the two groups were on parallel trends in the pre-event period and were expected to have continued on parallel trends in the post-event period if the shock had not occurred. In other words, the treatment and control groups can have significant differences in their characteristics, and those differences will not affect the results if their impact on the difference between the outcome variables for each group is constant over time. Indeed, the trends figures show that the treatment and control groups were clearly on parallel trends before the explosions.

I address this concern more directly by using the propensity score matching approach to select a control group that is more similar than the original control group to the treatment group. For each regression sample associated with one specific outcome variable, I first run a probit model with the treatment-group status regressed on age, Equifax Risk Score in 2018:Q2, and three characteristics of the census block of residence as of 2018:Q3 (percentage Black, percentage Hispanic, and percentage renters). Then, I use this probit model to calculate the propensity score for each individual, checking that the common support assumption is satisfied in each case. Next, I match members of the treatment group with “neighbors” from the original control group based on the similarity of their propensity scores. Appendix Table 1 shows that the differences in select individual and block characteristics between the new control group and the treatment group are smaller and less likely to be statistically significant compared with the differences between the old control group and the treatment group, except for the percentage Hispanic characteristic.[12] However, propensity score matching unavoidably contains arbitrary elements. Which observable characteristics to include in the matching regression, whether to use a probit or logit model to

[12] This table is based on the regression sample for total debt balance.

estimate the propensity score, how many closest "neighbors" to select for each treatment-group individual based on the propensity score, and whether to allow the replacement in the selection process are all subjective choices that affect the matched sample. In addition, selecting the control group based on propensity score matching can significantly shrink the sample size and reduce testing power. The number of individuals in the new control group is more than 38 percent smaller than the number of individuals in the old control group in Appendix Table 1.

I re-estimate Equation 1 using the matched sample. Figure 14 shows that the results are not fundamentally different from the original results. In the regression for total debt balance, the coefficient is still significant for 2018:Q4, whereas the coefficient is not significant for 2018:Q3. The results for total credit limit are similar to the original results. The regressions for the number of accounts past due and the indicator for an account 60-plus days delinquent yield more significant results than the regressions based on the original samples.

In Figure 15, I cluster standard errors by both year-quarter and census block of residence as of 2018:Q3. The two-way clustered standard errors turn out to be smaller than the one-way clustered standard errors. Not only do the original significant results remain significant, but more coefficients become significant, including some in the pre-explosions period and near the end of the sample period. This is consistent with the findings of Billings, Gallagher, and Ricketts (2022) in that standard errors clustered by both census block and time are less conservative than those clustered by census block only. A general rule of thumb is that 50 or more clusters are needed to reliably estimate clustered standard errors. With 25 year-quarters in my data, I may have too few time periods to cluster by time.

In addition, I conduct a placebo test in which I use individuals in the original control group to construct both a fake treatment group and a fake control group. The fake treatment group comprises individuals living in a census block in 2018:Q3 that is 30 to 999 meters from the nearest affected pipelines. The fake control group consists of individuals living in a census block in 2018:Q3 that is 1,000 meters or farther from the nearest affected pipeline. Figure 16 shows that the original significant results all disappear in this placebo test, indicating that the initial results are not spurious.

VII. Conclusion

This paper uses an event-study model and individual-level credit-reporting data to study the impact of the 2018 Merrimack Valley natural gas explosions. It helps fill a gap in the literature by providing the first empirical evidence of the effect of a man-made disaster on consumer credit outcomes. The analysis shows that the debt balance and credit limit of the treatment group, on average, decreased 3 to 4 percent more than those of the control group from their pre-disaster levels in the two quarters after the explosions. There is suggestive evidence that these decreases were more likely driven by a decline in credit demand than by a decline in credit supply. It is possible that the affected individuals might have reduced, delayed, or canceled new debt-financed major purchases, such as homes or cars, amid the severe life disruption, great uncertainty, and negative financial shock associated with the explosions.

This paper also finds that, relative to those of the control group, the treatment group's number of accounts past due and likelihood of being 60-plus days delinquent on loans, on average, increased temporarily, but not until two quarters after the explosions. This suggests that the affected individuals received short-term forbearance or used other options first rather than defaulting on their debt to cope with the negative shock. The gas explosions did not have a significant impact on credit scores.

Overall, the impact of the gas explosions, if any, was short-lived, lasting only one to two quarters. The lack of a large, long-lasting impact is likely due to the important role that external assistance played. Residents and businesses in the affected area received a significant amount of financial and nonfinancial assistance from the gas company, the state, local governments, nonprofits, and community organizations. Because I do not observe individual-level assistance and have no way to control for it, my estimates are essentially the impact of the gas explosions net of the effect of external assistance. They are underestimates of the potential impact of the disaster in the absence of assistance. These results highlight the importance of providing a broad range of assistance to communities affected by disasters as quickly and robustly as possible.

The results in this paper can be informative for areas that experience man-made disasters. The three towns affected by the 2018 Merrimack Valley natural gas explosions are a mix of different types of communities (Table 1). Andover and North Andover have a higher percentage of residents who are non-Hispanic white, higher income, and college-educated compared with Massachusetts and the United States as a whole, whereas Lawrence has a higher percentage of residents who are Hispanic, lower income, and not college-educated compared with the state and the country. Despite varying demographic compositions, the three towns combined have measures of income and educational attainment similar to those of the state and the United States. In this way, the area affected by this disaster is close to being representative of a typical American community while allowing for a wide range of socioeconomic and demographic characteristics within the overall area. In addition, the 2018 Merrimack Valley natural gas explosions are not a unique event. As the infrastructure systems including utility gas lines in the United States have further aged and deteriorated, gas explosions have become more common in recent decades.[13] These observations suggest that the findings of this paper could apply to many other communities and man-made disasters. A promising focus for future research would involve another man-made disaster in a different location with a different scale of damage, a different type of affected community, and different government responses. Comparing the results from such a study with those from this paper could enrich our understanding of the financial impact of man-made disasters.

[13] See US Department of Transportation, Pipeline and Hazardous Materials Safety Administration, "Pipeline Incident 20 Years Trend," November 15, 2022, https://www.phmsa.dot.gov/data-and-statistics/pipeline/pipeline-incident-20-year-trends (accessed September 26, 2023).

References

Banerjee, Chandan, Lucia Bevere, Thierry Corti, James Finucane, and Roman Lechner. 2023. "Natural Catastrophes and Inflation in 2022: A Perfect Storm." Swiss Re Institute.

Bevere, Lucia, and Federica Remondi. 2022. "Natural Catastrophes in 2021: The Floodgates Are Open." Swiss Re Institute.

Bevere, Lucia, and Andreas Weigel. 2021. "Natural Catastrophes in 2020: Secondary Perils in the Spotlight, but Don't Forget Primary-peril Risks." Swiss Re Institute.

Billings, Stephen B., Emily Gallagher, and Lowell Ricketts. 2022. "Let the Rich Be Flooded: The Distribution of Financial Aid and Distress after Hurricane Harvey." *Journal of Financial Economics* 146(2): 797–819.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin. 2020. "Fast Poisson Estimation with High-dimensional Fixed Effects." *The Stata Journal* 20(1): 95–115.

Del Valle Suarez, Alejandro, Therese Scharlemann, and Stephen H. Shore. forthcoming. "Household Financial Behavior after Hurricane Harvey." *Journal of Financial and Quantitative Analysis*.

Edmiston, Kelly D. 2017. "Financial Vulnerability and Personal Finance Outcomes of Natural Disasters." Federal Reserve Bank of Kansas City, Research Working Paper 17-09.

Gallagher, Justin, and Daniel Hartley. 2017. "Household Finance after a Natural Disaster: The Case of Hurricane Katrina." *American Economic Journal: Economic Policy* 9(3): 199–228.

Gallagher, Justin, Daniel Hartley, and Shawn Rohlin. 2023. "Weathering an Unexpected Financial Shock: The Role of Federal Disaster Assistance on Household Finance and Business Survival." *Journal of the Association of Environmental and Resource Economists* 10(2): 525–567.

Lee, Donghoon, and Wilbert Van der Klaauw. 2010. "An Introduction to the FRBNY Consumer Credit Panel." FRB of New York Staff Report (479).

Massachusetts Emergency Management Agency. 2020. *Merrimack Valley Natural Gas Explosions After Action Report, September 13–December 16, 2018*.

National Safety Transportation Board. 2019. *Overpressurization of Natural Gas Distribution System, Explosions, and Fires in Merrimack Valley,* *Massachusetts, September 13, 2018*. Pipeline Accident Report NTSB/PAR-19/02. Washington, D.C.

Ratcliffe, Caroline, William Congdon, Daniel Teles, Alexandra Stanczyk, and Carlos Martin. 2020. "From Bad to Worse: Natural Disasters and Financial Health." *Journal of Housing Research* 29(1): 25–53.

Tran, Brigitte Roth, and Tamara L. Sheldon. 2017. "Same Storm, Different Disasters: Consumer Credit Access, Income Inequality, and Natural Disaster Recovery." Unpublished manuscript.

Appendix 1. Sample Construction

To construct a balanced sample for each outcome variable, I keep only those individuals who did not have a missing value for this outcome variable. To ensure data accuracy, I also impose the following conditions, which were binding for only a small number of individuals.

- Year of birth is not missing
- Not older than 100 years of age during the sample period
- Not flagged for having died during the sample period
- Block of residence in 2018:Q3 not matched to a census block with zero population and household according to the 2010 census
- Address in 2018:Q3 not flagged for being non-residential (firm address, post office box, or general delivery)
- Individuals who must be in either the treatment group or the control group

Appendix 2. Additional Outcome Variables That Were Explored

Below is a list of variables that I explored but omitted from the regression analysis for one of two reasons: The use of each resulted in either a sample size that was too small due to missing values or a failure to satisfy a visual test of the parallel-trends assumption for the pre-explosions period.

- Mortgage balance
- Retail trade balance
- Auto bank balance
- Auto loan balance
- Other credit balance
- High credit for mortgage account
- High credit for retail trade account
- High credit for auto bank account
- High credit for auto loan account
- High credit for other credit account
- Number of credit inquiries initiated by the consumer within the past three months
- Total number of open credit accounts
- Total amount past due
- Bank card amount past due
- Amount past due for installment accounts
- Total amount past due as a share of total debt balance
- Number of accounts past due as a share of total number of trades
- Number of accounts 30-plus days delinquent as a share of total number of trades

- Number of accounts 60-plus days delinquent as a share of total number of trades
- Number of accounts 90-plus days delinquent as a share of total number of trades
- Number of accounts 120-plus days delinquent or in collections as a share of total number of trades
- Number of severe derogatory accounts as a share of total number of trades
- Indicator for having at least one account 90-plus days delinquent or in collections
- Indicator for having at least one account 120-plus days delinquent or in collections
- Indicator for bankruptcy
- Indicator for having moved to a different Zip code since previous quarter
- Indicator for having moved to a different census block since previous quarter
- Indicator for having moved to a different census tract since previous quarter

Appendix 3. Variables Used to Conduct Heterogeneity Analysis

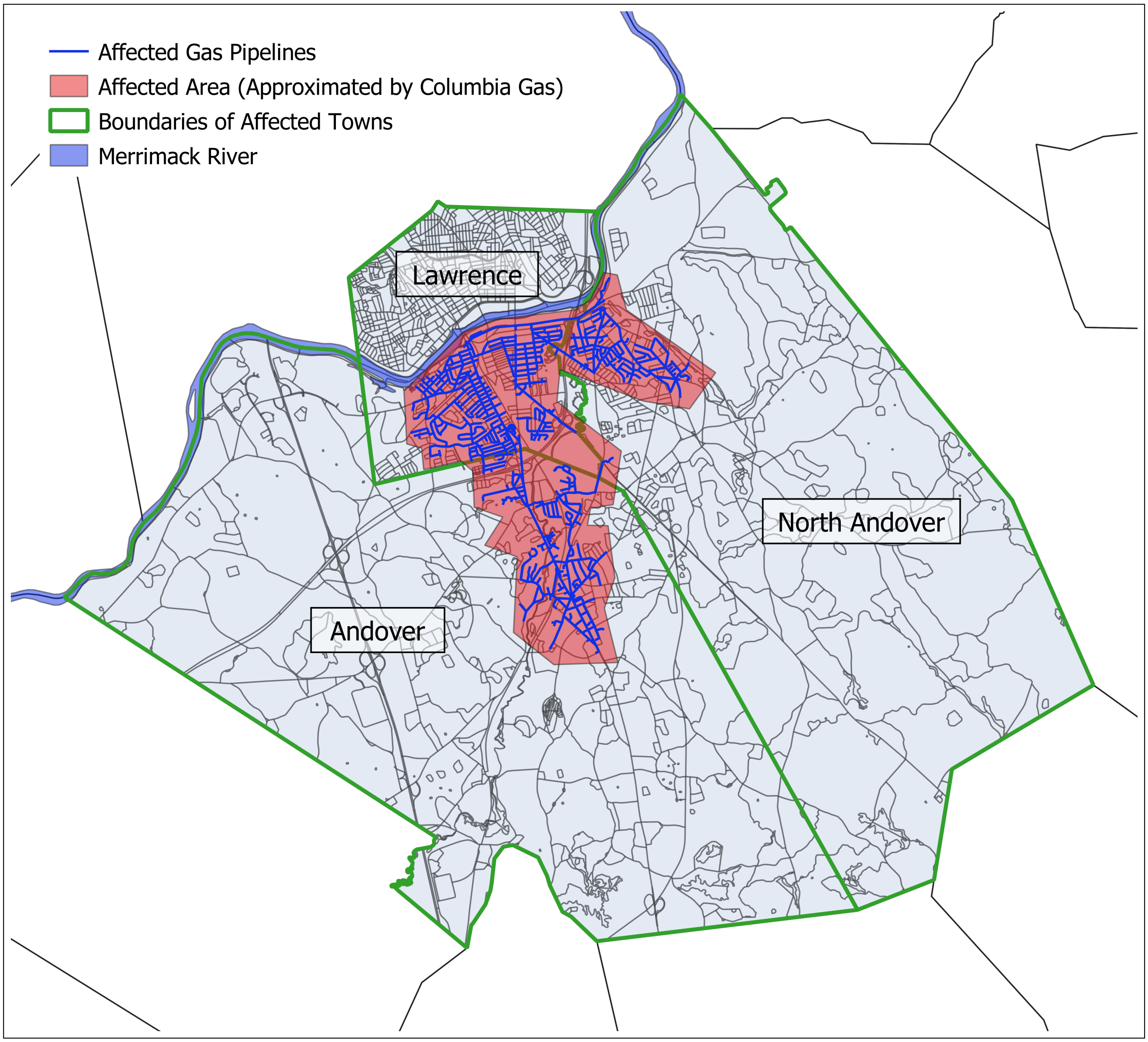
- Indicators for age groups: 16 to 24, 55 to 64, 65 and over (with 25 to 54 as the omitted category)
- Whether Black or Hispanic individuals made up more than half of the population in the block of residence in 2018:Q3
- Whether renters made up a majority of the population in the block of residence in 2018:Q3
- Whether single-mother-headed households made up more than 20 percent of the households in the block of residence in 2018:Q3
- Whether single-mother-headed households made up more than 25 percent of the households in the block of residence in 2018:Q3
- Dummy variable for living in Lawrence in 2018:Q3
- Whether individual's debt balance was 80 percent or more of their credit limit in 2018:Q2
- Whether individual's debt balance was 85 percent or more of their credit limit in 2018:Q2
- Whether individual's debt balance was 90 percent or more of their credit limit in 2018:Q2
- Whether the block of residence in 2018:Q3 contained any state-inspected properties that were deemed to have major damage or destroyed
- Whether the affected streets made up half or more of all streets in the block of residence in 2018:Q3

Table 1: Demographic and Socioeconomic Characteristics of Three Towns Affected by 2018 Merrimack Valley Natural Gas Explosions

	Andover	Lawrence	North Andover	Three Towns Combined	Massachusetts	New England	United States
Percentage Non-Hispanic White	72.9	14.1	81.1	42.7	70.8	74.3	60.1
Percentage Black	2.3	5.3	3.1	4.1	7.5	6.8	12.6
Percentage Hispanic	7.6	81.1	8.1	47.6	12.0	11.3	18.2
Percentage Asian or Pacific Islander	15.1	2.1	5.8	6.1	6.8	4.9	5.8
Percentage Owner-occupied	80.5	30.2	70.8	51.1	62.5	65.4	64.4
Median Household Income	153,315	46,871	113,916	87,163	84,385	78,565	64,994
Per Capita Income	70,222	21,655	56,552	40,953	45,555	43,138	35,384
Percentage High School Graduate or Less	11.3	63.0	17.0	39.8	32.5	34.4	38.1
Percentage Bachelor's Degree or More	74.3	12.5	61.2	38.9	44.5	40.8	32.9

Source: American Community Survey, 5-year Estimates (2016–2020)

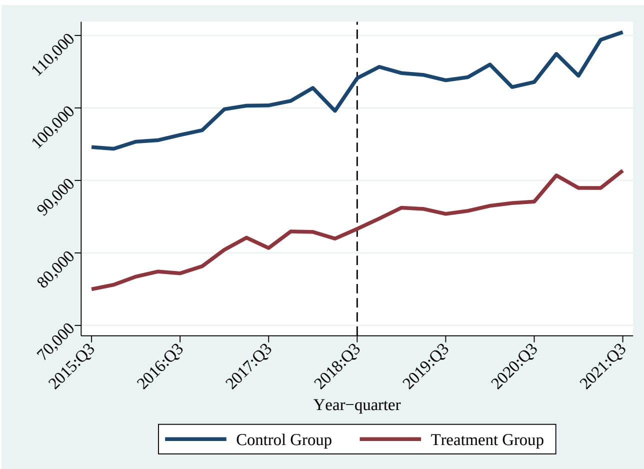
Figure 1: Gas Pipelines and Areas Affected by 2018 Merrimack Valley Natural Gas Explosions



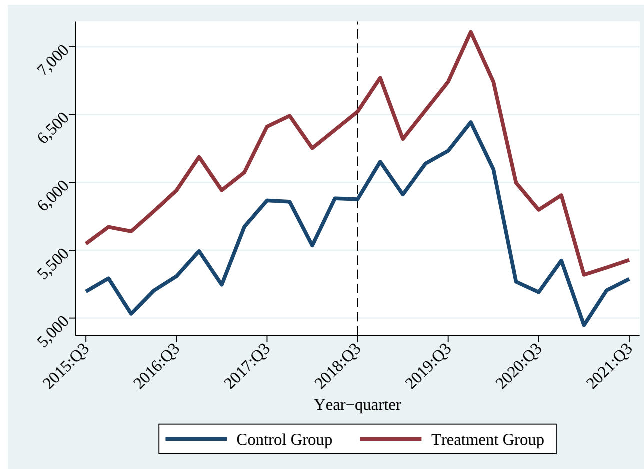
Source: Columbia Gas/Massachusetts Emergency Management Agency

Figure 2: Trends in Debt Balance

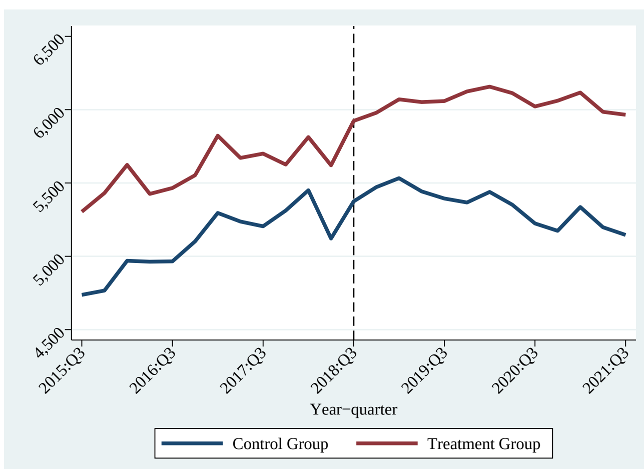
(a) Total Debt Balance ($)



(b) Bank Card Balance ($)



(c) Student Loan Balance ($)

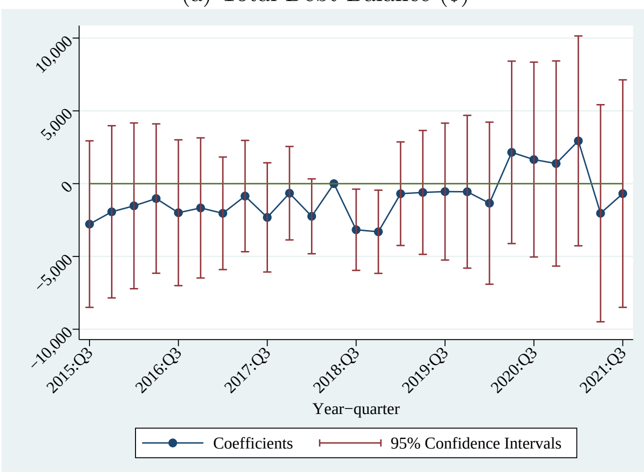


Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax

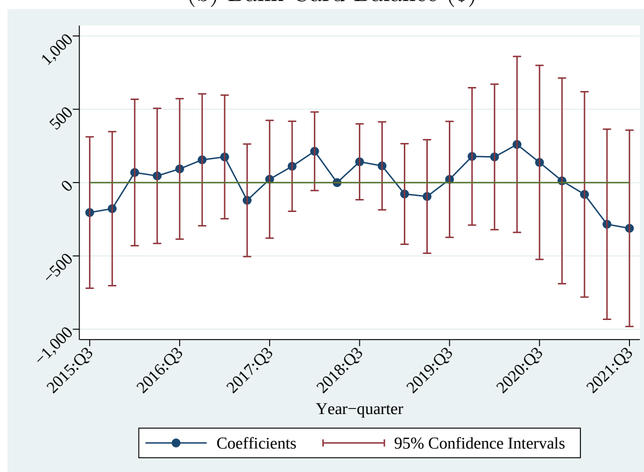
Notes: Figures report the simple average of the outcome variable within the treatment and control groups. The treatment group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover that intersected an affected street and intersected an affected gas pipeline. The control group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover located 30-plus meters from an affected gas pipeline.

Figure 3: Effect of Gas Explosions on Debt Balance

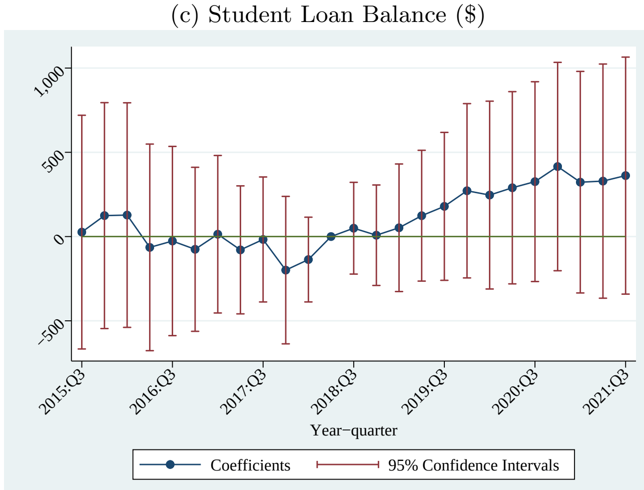
(a) Total Debt Balance ($)



(b) Bank Card Balance ($)



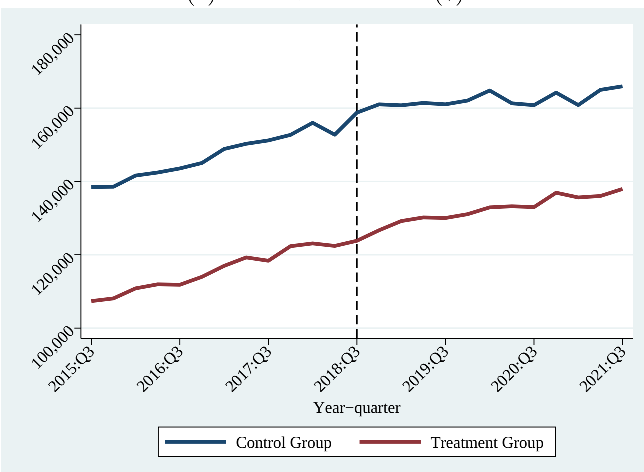
(c) Student Loan Balance ($)



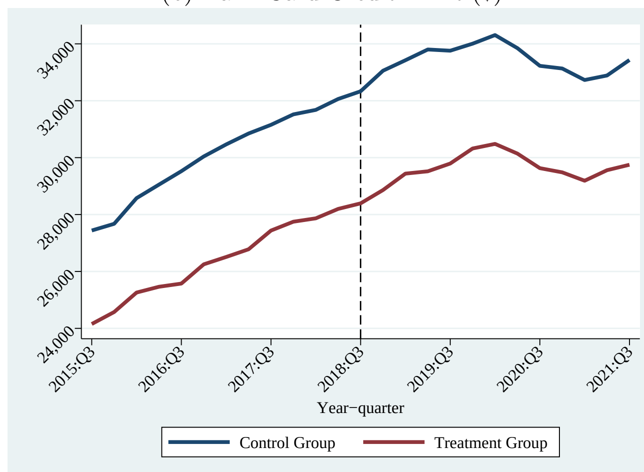
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. Coefficients and confidence intervals are estimated from Equation 1.

Figure 4: Trends in Credit Limit

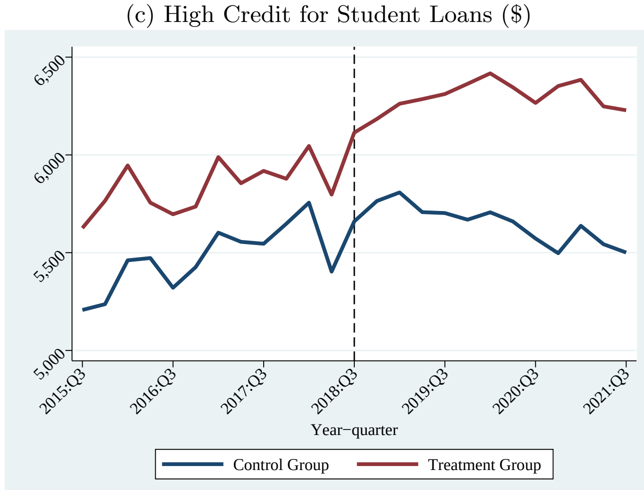
(a) Total Credit Limit ($)



(b) Bank Card Credit Limit ($)



(c) High Credit for Student Loans ($)

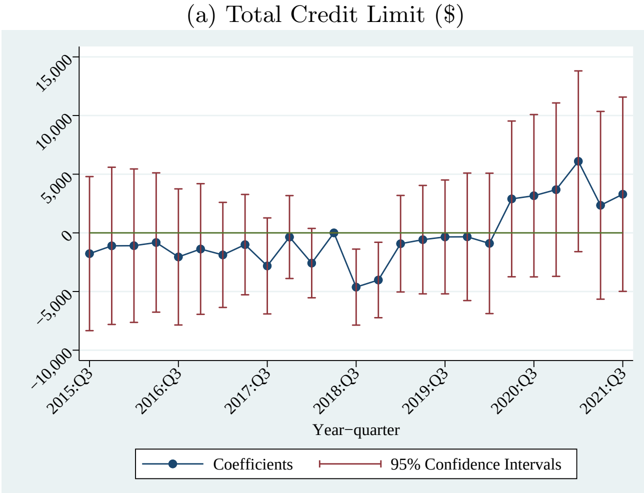


Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax

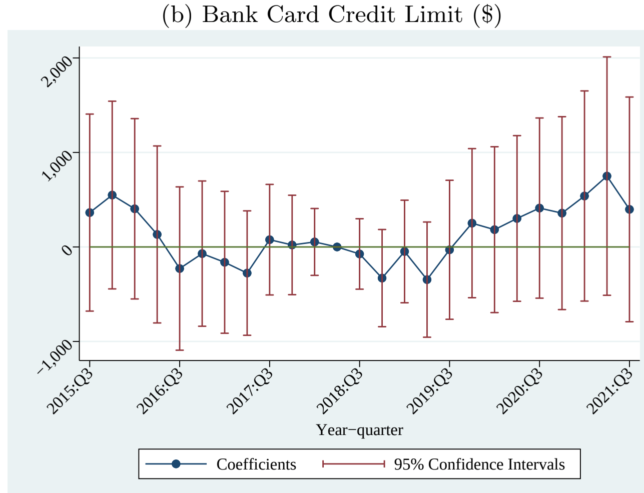
Notes: Figures report the simple average of the outcome variable within the treatment and control groups. The treatment group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover that intersected an affected street and intersected an affected gas pipeline. The control group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover located 30-plus meters from an affected gas pipeline.

Figure 5: Effect of Gas Explosions on Credit Limit

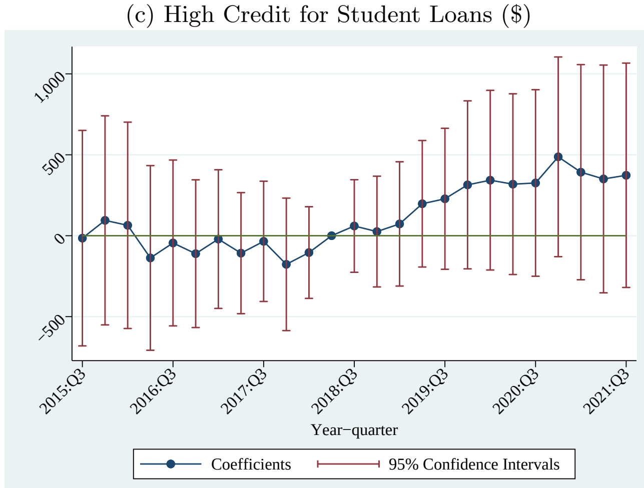
(a) Total Credit Limit ($)



(b) Bank Card Credit Limit ($)



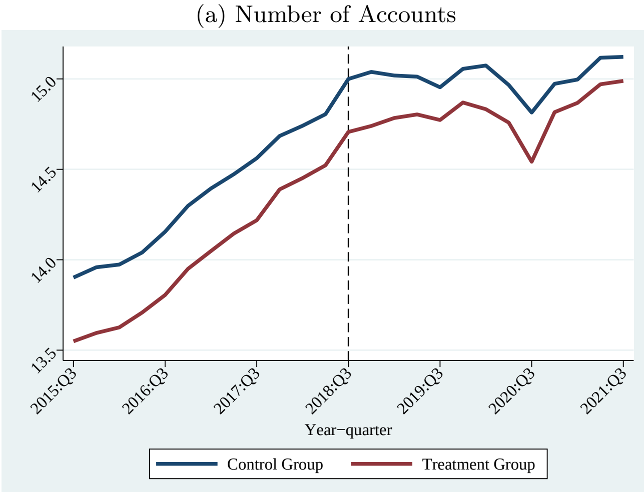
(c) High Credit for Student Loans ($)



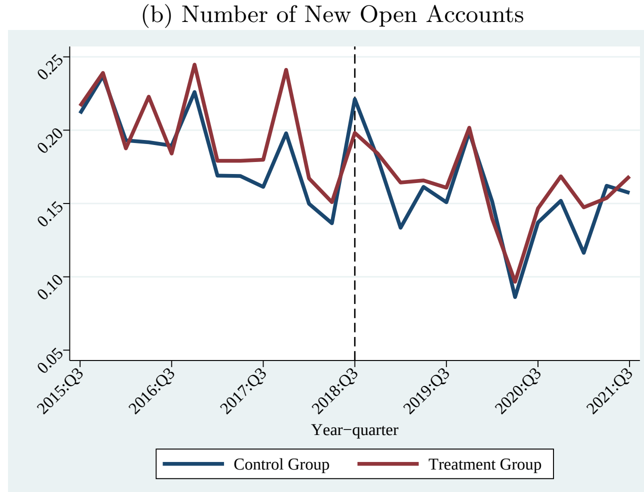
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. Coefficients and confidence intervals are estimated from Equation 1.

Figure 6: Trends in Number of Accounts

(a) Number of Accounts



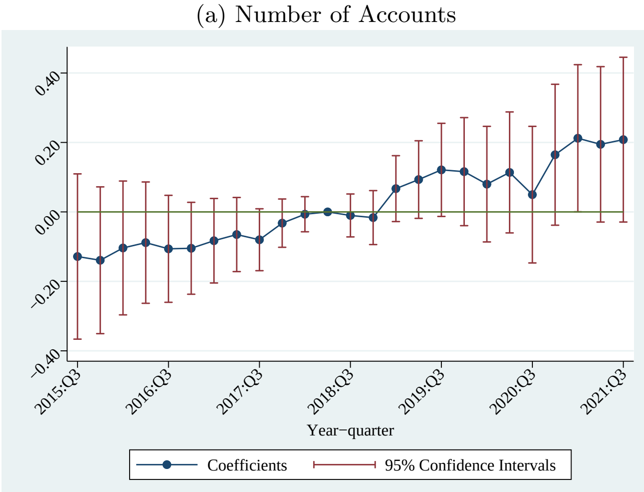
(b) Number of New Open Accounts



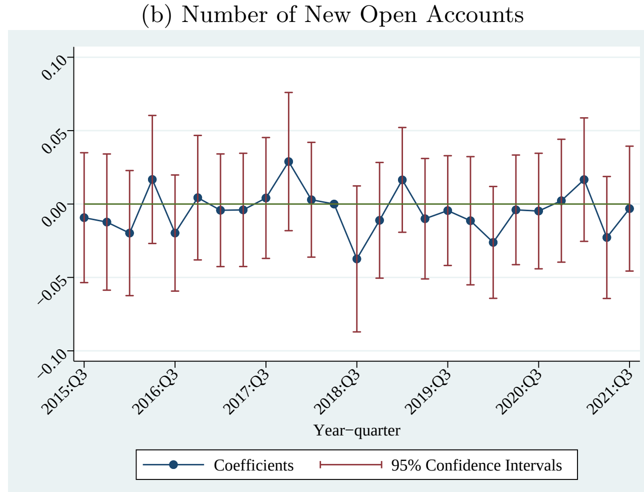
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax
Notes: Figures report the simple average of the outcome variable within the treatment and control groups. The treatment group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover that intersected an affected street and intersected an affected gas pipeline. The control group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover located 30-plus meters from an affected gas pipeline.

Figure 7: Effect of Gas Explosions on Number of Accounts

(a) Number of Accounts



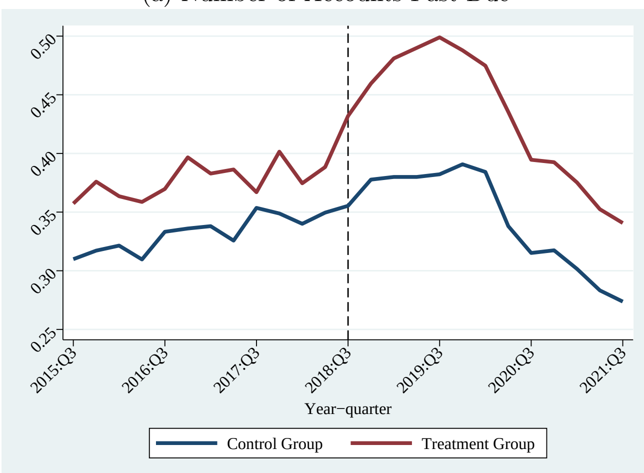
(b) Number of New Open Accounts



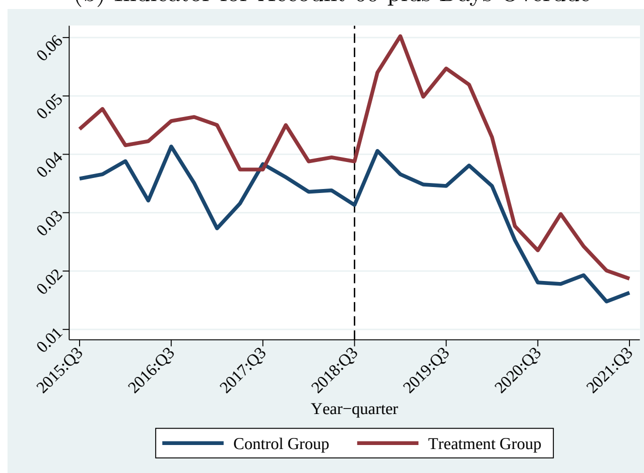
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. Coefficients and confidence intervals are estimated from Equation 1.

Figure 8: Trends in Debt Delinquency

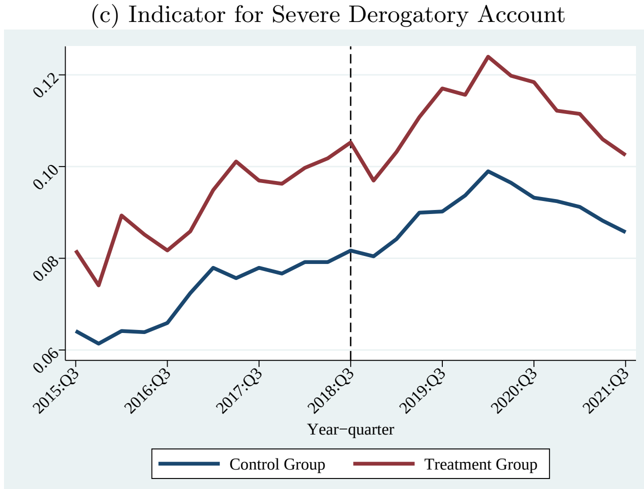
(a) Number of Accounts Past Due



(b) Indicator for Account 60-plus Days Overdue



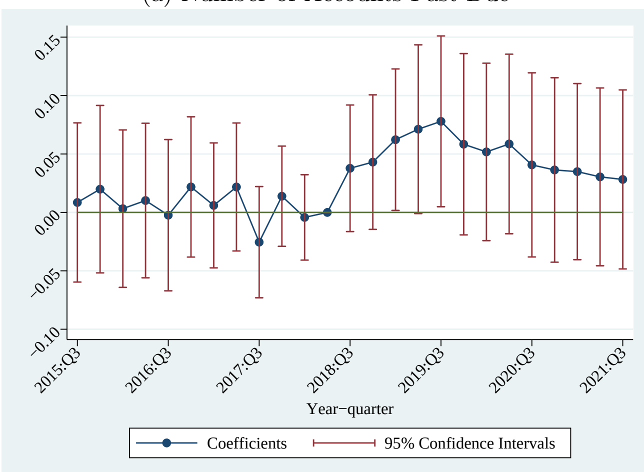
(c) Indicator for Severe Derogatory Account



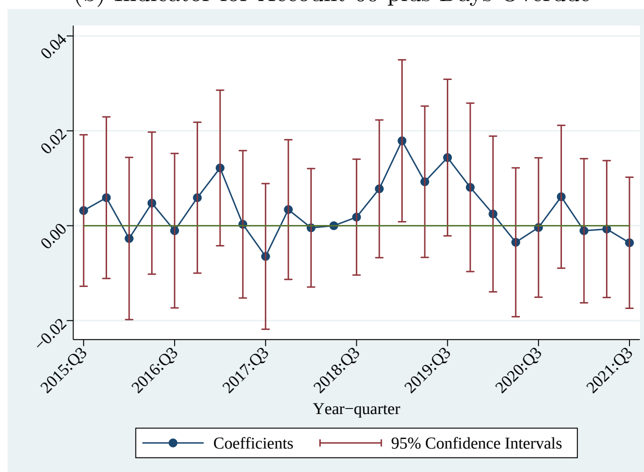
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax
Notes: Figures report the simple average of the outcome variable within the treatment and control groups. The treatment group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover that intersected an affected street and intersected an affected gas pipeline. The control group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover located 30-plus meters from an affected gas pipeline.

Figure 9: Effect of Gas Explosions on Debt Delinquency

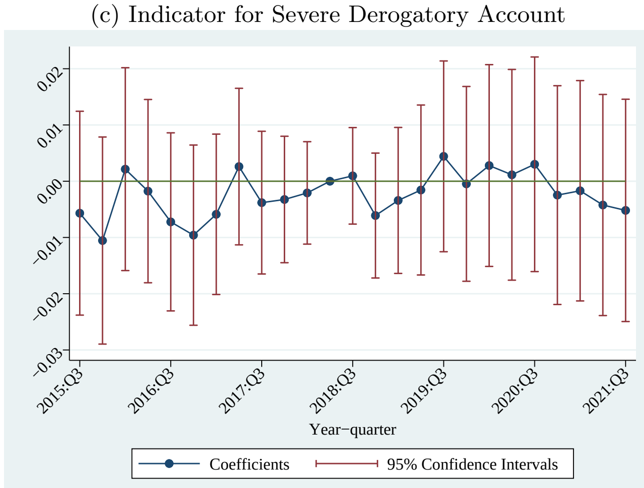
(a) Number of Accounts Past Due



(b) Indicator for Account 60-plus Days Overdue



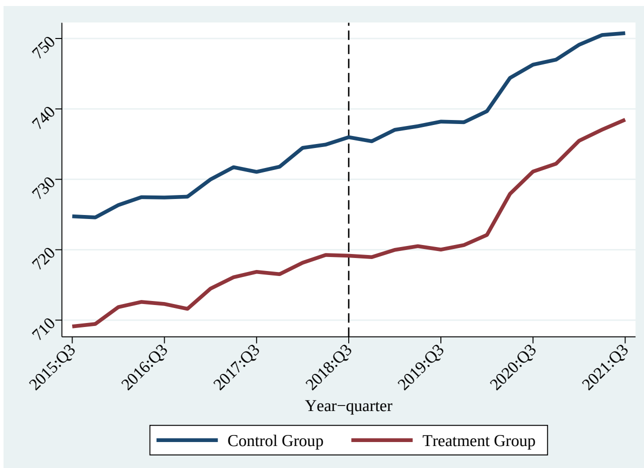
(c) Indicator for Severe Derogatory Account



Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations

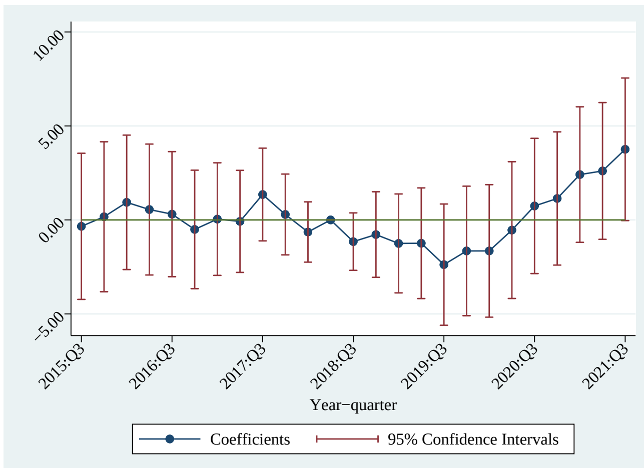
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. Coefficients and confidence intervals are estimated from Equation 1.

Figure 10: Trend in Equifax Risk Score



Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax
Notes: Figures report the simple average of the outcome variable within the treatment and control groups. The treatment group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover that intersected an affected street and intersected an affected gas pipeline. The control group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover located 30-plus meters from an affected gas pipeline.

Figure 11: Effect of Gas Explosions on Equifax Risk Score

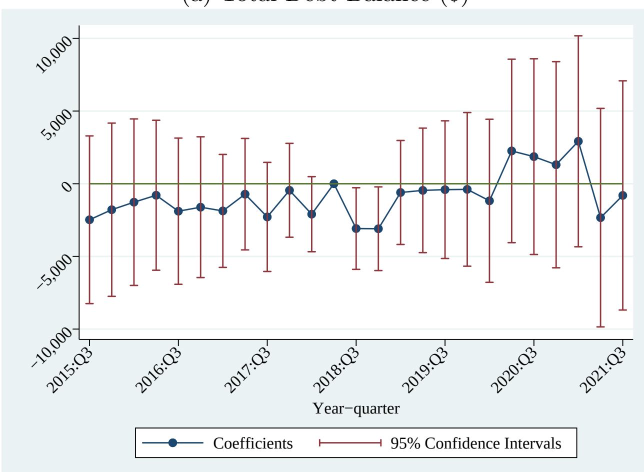


Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. Coefficients and confidence intervals are estimated from Equation 1.

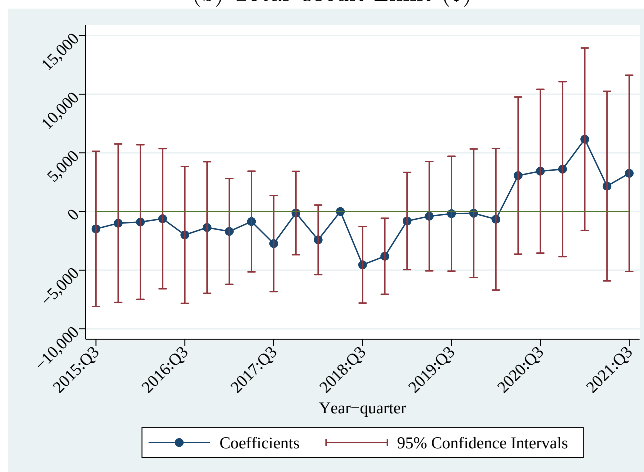
Figure 12: Robustness Checks for Effect of Gas Explosions on Consumer Credit Outcomes

100-meter Threshold for Control Group

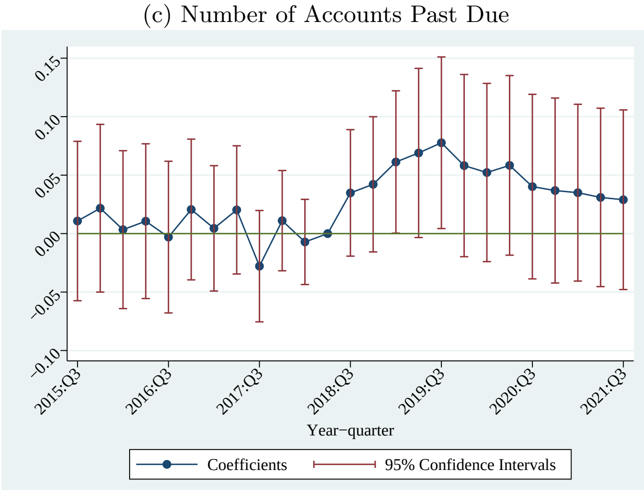
(a) Total Debt Balance ($)



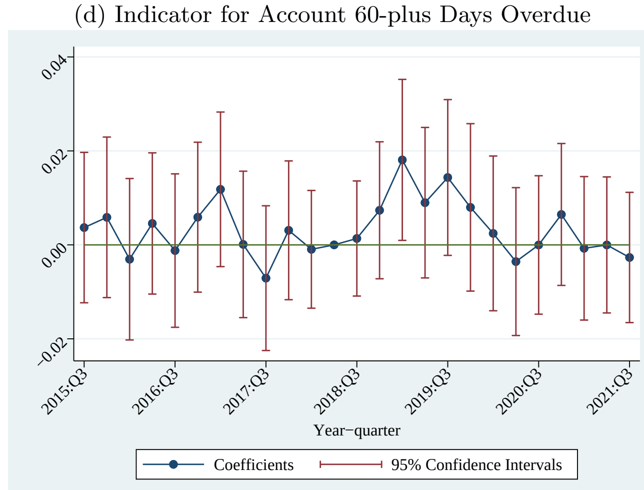
(b) Total Credit Limit ($)



(c) Number of Accounts Past Due



(d) Indicator for Account 60-plus Days Overdue

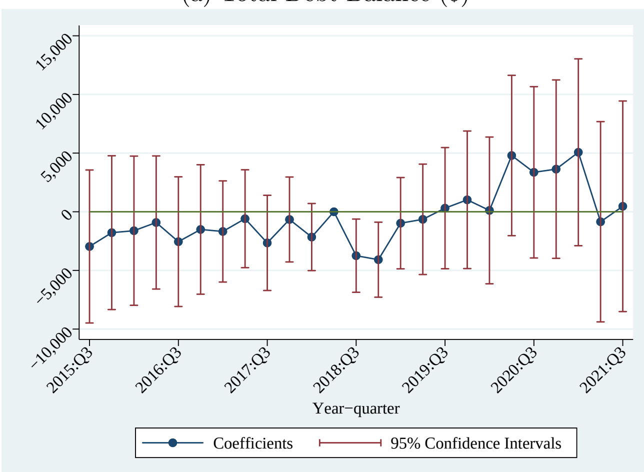


Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. The control group is defined as all individuals who, as of 2018:Q3, resided in census blocks within Andover, Lawrence, and North Andover that were located 100-plus meters from the affected gas pipelines. Coefficients and confidence intervals are estimated from Equation 1.

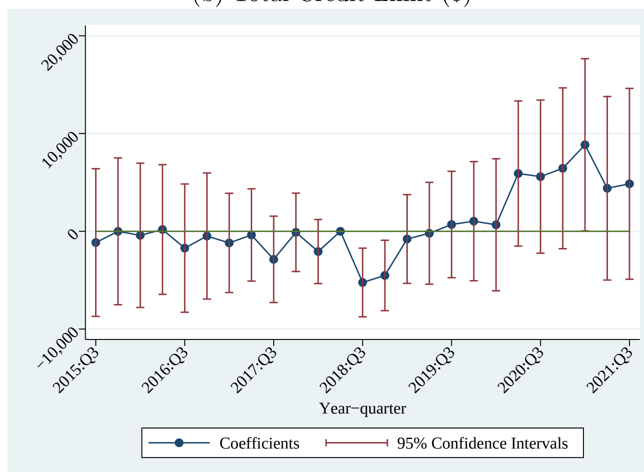
Figure 13: Robustness Checks for Effect of Gas Explosions on Consumer Credit Outcomes

1,000-meter Threshold for Control Group

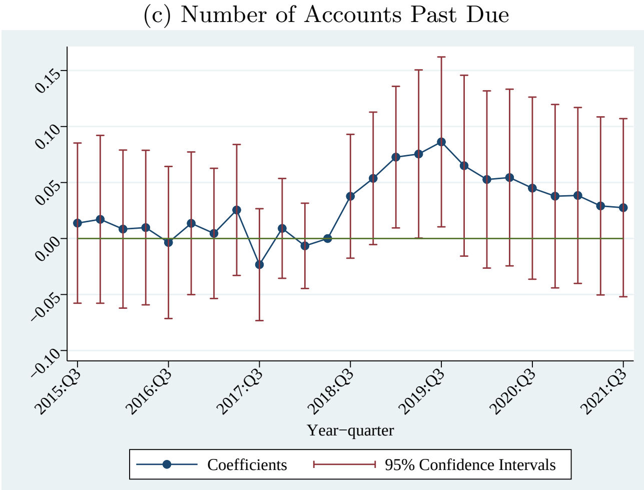
(a) Total Debt Balance ($)



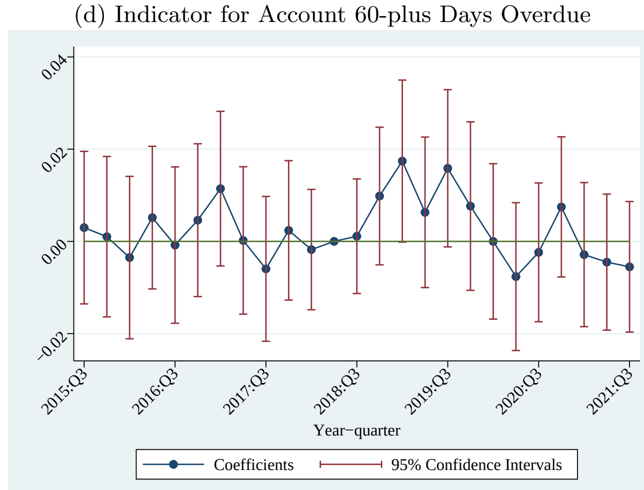
(b) Total Credit Limit ($)



(c) Number of Accounts Past Due



(d) Indicator for Account 60-plus Days Overdue



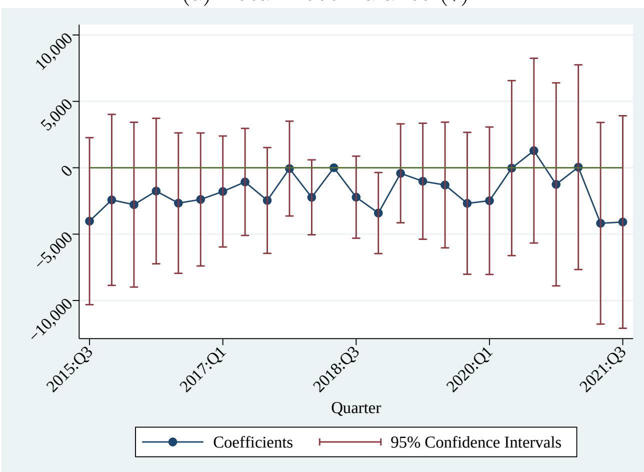
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations

Notes: Standard errors are clustered by census block of residence as of 2018:Q3. The control group is defined as all individuals who, as of 2018:Q3, resided in census blocks within Andover, Lawrence, and North Andover that were located 1,000-plus meters from the affected gas pipelines. Coefficients and confidence intervals are estimated from Equation 1.

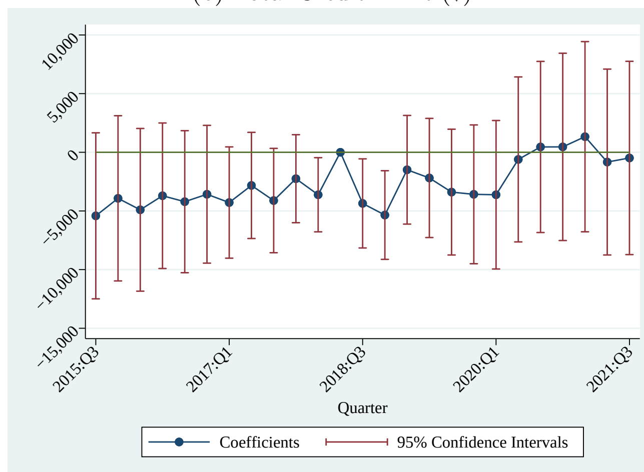
Figure 14: Robustness Checks for Effects of Gas Explosions on Consumer Credit Outcomes

Control Group Selected with Propensity Score Matching

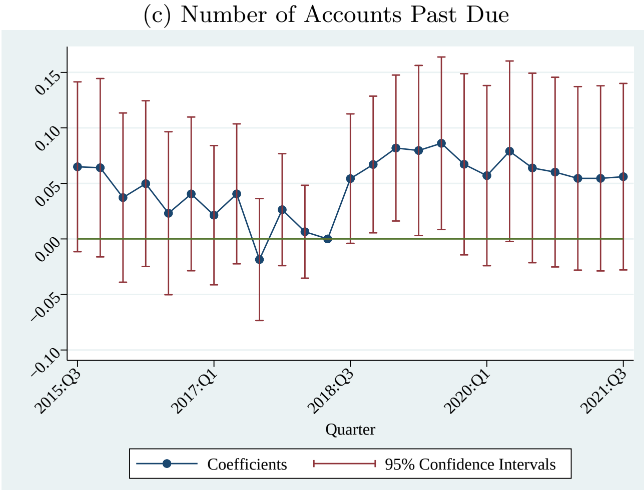
(a) Total Debt Balance ($)



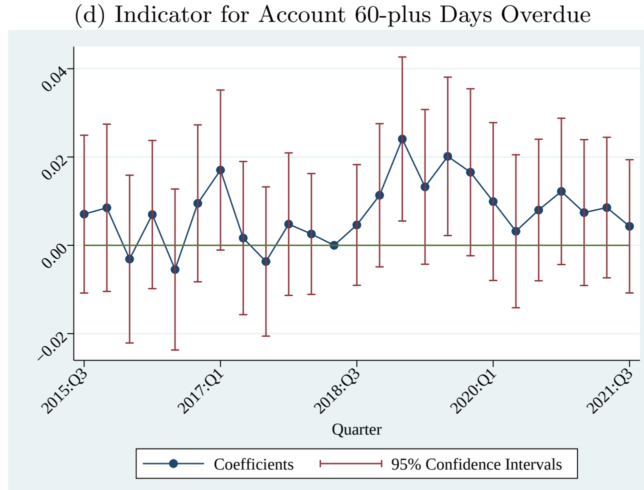
(b) Total Credit Limit ($)



(c) Number of Accounts Past Due



(d) Indicator for Account 60-plus Days Overdue

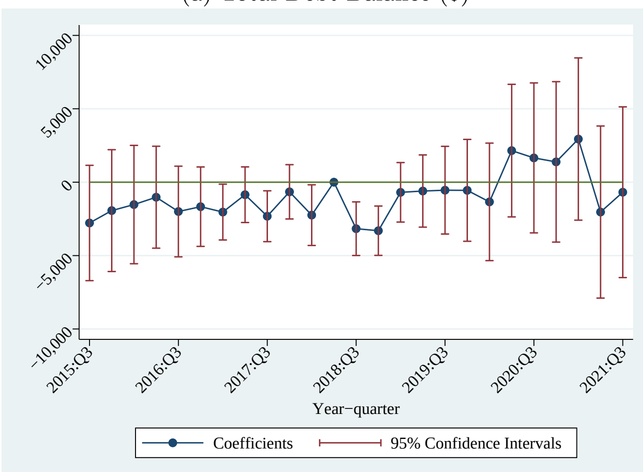


Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. The control group is selected using propensity score matching with the *psmatch2* command in Stata. Individuals' propensity scores are estimated from a probit model, using age and Equifax Risk Scores from 2018:Q2 and selected characteristics from their census blocks of residence as of 2018:Q3, including percentage Black, percentage Hispanic, and percentage renters. Based on the similarity of their propensity scores, individuals in the treatment group are then assigned four "neighbors," with replacement, selected from the original control group. This collection of neighbors is then used as the new control group to estimate Equation 1.

Figure 15: Robustness Checks for Effects of Gas Explosions on Consumer Credit Outcomes

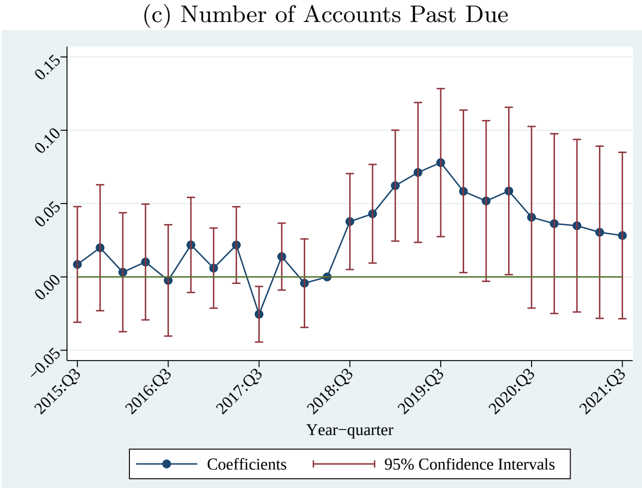
Two-way Clustered Standard Errors

(a) Total Debt Balance ($)

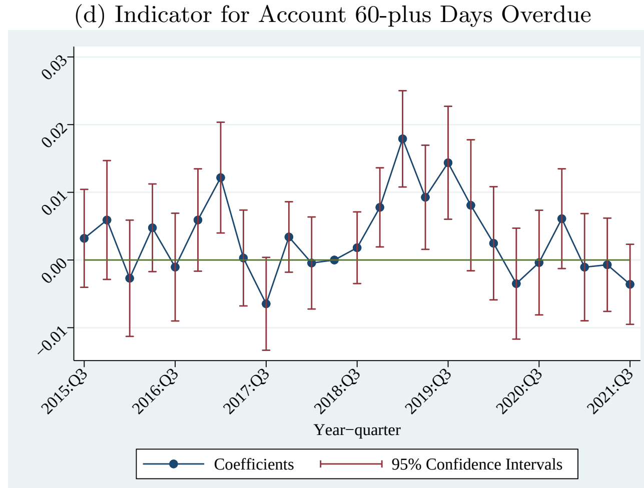


(b) Total Credit Limit ($)

(c) Number of Accounts Past Due



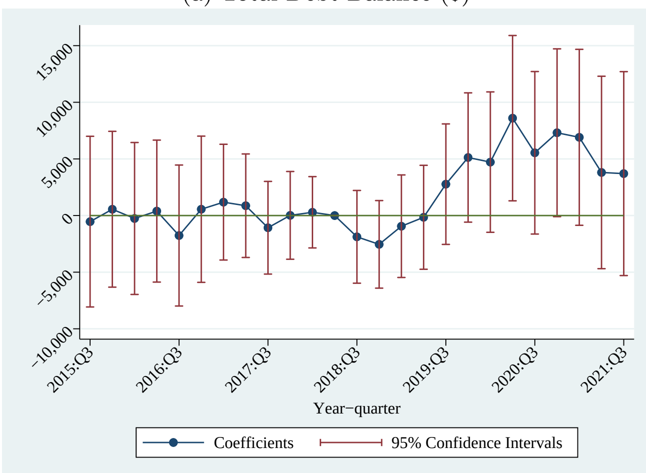
(d) Indicator for Account 60-plus Days Overdue



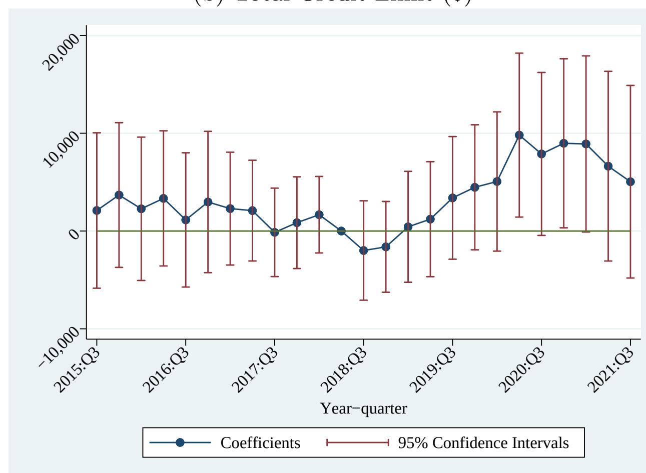
Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by both year-quarter and by census block of residence as of 2018:Q3. The treatment group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover that intersected an affected street and intersected an affected gas pipeline. The control group is defined as individuals who, as of 2018:Q3, resided in a census block in Andover, Lawrence, or North Andover located 30-plus meters from an affected gas pipeline.

Figure 16: Placebo Tests for Effect of Gas Explosions on Consumer Credit Outcomes

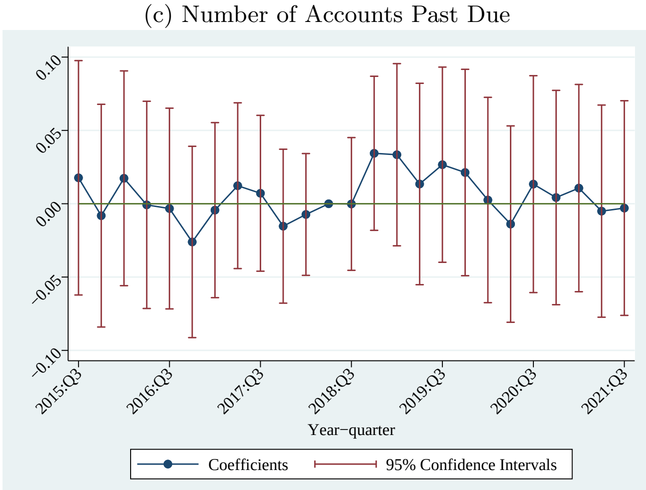
(a) Total Debt Balance ($)



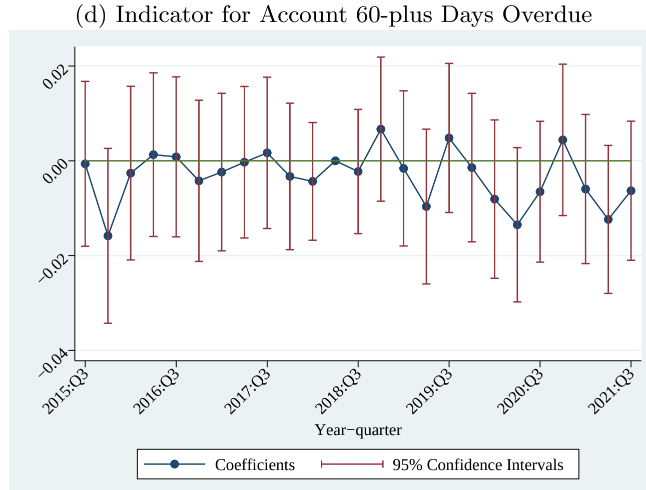
(b) Total Credit Limit ($)



(c) Number of Accounts Past Due

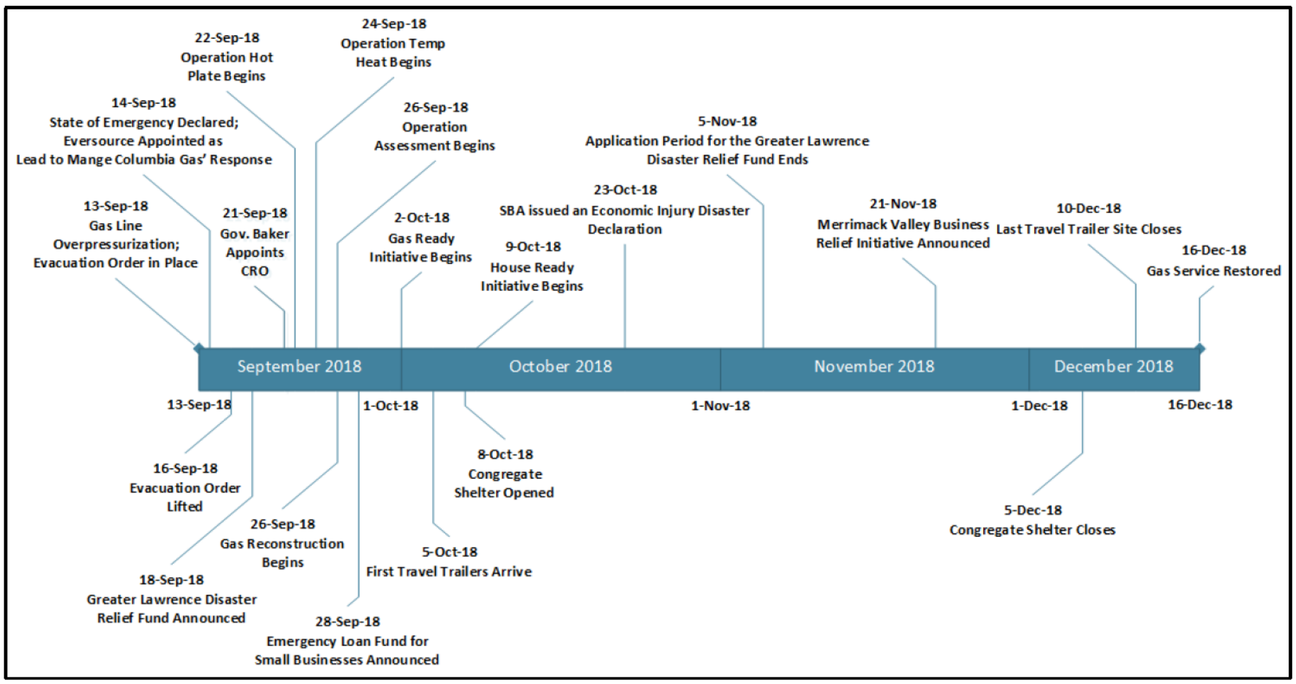


(d) Indicator for Account 60-plus Days Overdue



Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations
Notes: Standard errors are clustered by census block of residence as of 2018:Q3. The placebo treatment group is defined as all 2018:Q3 residents of census blocks within Andover, Lawrence, and North Andover that are 30 to 1,000 meters from the affected gas pipelines. The placebo control group is defined as all 2018:Q3 residents of census blocks within Andover, Lawrence, and North Andover located 1,000-plus meters from the affected gas pipelines. Coefficients and confidence intervals are estimated from Equation 1.

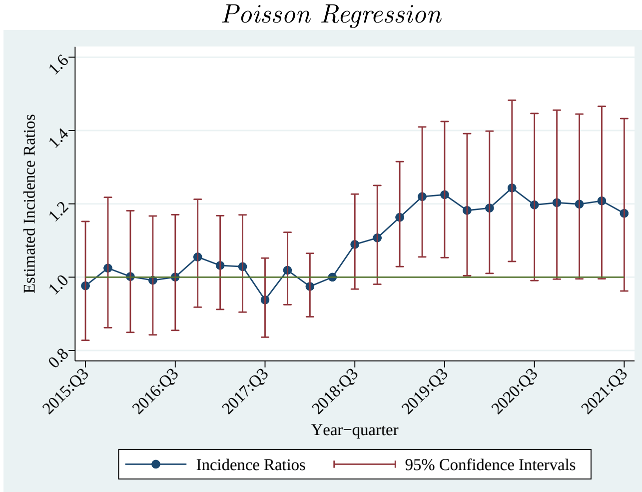
Appendix Figure 1: Overall Timeline of 2018 Merrimack Valley Natural Gas Explosions



Source: Massachusetts Emergency Management Agency (2020), Figure 8
Note: CRO stands for Chief Recovery Officer.

Appendix Figure 2: Robustness Check for Effect of Gas Explosions on Number of Accounts Past Due

Poisson Regression



Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, author's calculations

Notes: Standard errors are clustered by census block of residence as of 2018:Q3. Incidence ratios and confidence intervals are estimated from Equation 1 using the *ppmlhdfe* command in Stata. See Correia, Guimarães, and Zylkin (2019).

Appendix Table 1: Characteristics of Treatment and Control Groups

	Treatment Group	Control Group *30-meter Threshold*	Control Group *Propensity Score Matching*
Percentage Black	3.4	3.0*	3.3
Percentage Hispanic	24.5	25.3	28.2*
Percentage Renters	35.1	25.3*	30.3*
Age	51.9	50.9*	51.3
Equifax Risk Score	728.6	744.9*	733.8
N	1,267	3,517	2,164

Source: Federal Reserve Bank of New York Consumer Credit Panel/Equifax, 2010 census, author's calculations
Notes: This table is based on the regression sample for total debt balance. Values are simple averages within the groups from 2018:Q2. * indicates $p < 0.05$ in a two-group t-test compared with the treatment group, allowing for unequal variances between groups. Propensity score matching was conducted using the *psmatch2* command in Stata. Individuals' propensity scores are estimated from a probit model, using age and Equifax Risk Scores from 2018:Q2 and selected characteristics from their census blocks of residence as of 2018:Q3, including percentage Black, percentage Hispanic, and percentage renters. Based on the similarity of their propensity scores, individuals in the treatment group are then assigned four "neighbors," with replacement, selected from the original control group. This collection of neighbors is then used as the new control group to estimate Equation 1. The 2010 census is the data source for percentage Black, percentage Hispanic, and percentage renters.