# Unemployment Insurance Fraud in the Debit Card Market

Umang Khetan* Jetson Leder-Luis† Jialan Wang‡ Yunrong Zhou§

October 2023

PRELIMINARY - PLEASE DO NOT CITE OR CIRCULATE

**Abstract**

We study fraud in the unemployment insurance (UI) system using a dataset of 35 million debit card transactions. We apply machine learning techniques to group cards into clusters corresponding to varying levels of suspicious or potentially fraudulent activity. We then conduct an analysis based on the staggered adoption of state-level identity verification systems between 2020 and 2021 to assess the effectiveness of technological solutions for reducing fraud. Our findings suggest that identity verification reduced payouts to suspicious cards by 40% relative to non-suspicious cards, which were largely unaffected by these technologies. Our results indicate that identity screening of new and continuing UI applicants may be an effective mechanism for mitigating fraud in the UI system.

*Keywords*: Unemployment insurance, fraud, identity theft, COVID-19, screening, debit cards, machine learning

*JEL classification*: J65, G51, K42, H53

*University of Iowa

†Boston University & NBER. Corresponding Author. Email: jetson@bu.edu

‡University of Illinois at Urbana-Champaign & NBER

§Purdue University

We thank Cameron Ellis, Felipe Netto, and seminar participants at the University of Iowa for helpful comments. We are extremely grateful to Jonathan Chin and Facteus for providing data access and technical support, to Filipe Correia for making this research possible with labor-intensive work on data intake and analytics, and to the AWS Greg Gulick Honorary Award for generous support.

## 1. Introduction

Unemployment insurance (UI) is one of the largest financial support programs in the United States, serving millions of individuals who lose their jobs each year and boosting the aggregate economy during economic downturns. The onset of the COVID-19 pandemic and sharp economic contraction led to a large expansion in unemployment insurance, with state and federal governments disbursing over $800 billion between 2020-2021. However, the rapid UI expansion prioritizing speed over security was also susceptible to fraud, with frequent news reports claiming diversion of funds away from the intended recipients.[1]

Like other benefit programs, unemployment insurance faces a trade-off between hassle costs and more precise targeting toward intended recipients (Nichols and Zeckhauser, 1982). In the case of UI fraud, identity verification represents a potential hassle cost, as all applicants and continuing recipients need to undergo screening with some time and potentially monetary cost borne by both the applicant and UI system. Strict screening creates an operational burden on the system and could potentially delay payments or even prevent eligible applicants from receiving benefits. However, a more permissive system could potentially lead to waste and fraud, allowing non-eligible or nefarious applicants to receive payments they are not entitled to. Poor targeting of UI benefits due to fraud could have significant economic consequences, as fraudsters are unlikely to have the same labor market and spending responses as eligible recipients. Therefore, effective anti-fraud policy in the UI system is an important area for both public administration and for understanding the broader economic effects of UI.

In this paper, we examine the effectiveness of anti-fraud screening measures in reducing the misallocation of unemployment benefits. In response to growing concerns about fraud, 29 states adopted identity verification technologies using third-party vendors between March 2020 and September 2021. Although the specific protocols varied by state, the vendors typically verified each applicant's identity through photo or video-based authentication. The staggered adoption of these policies provides quasi-experimental variation that we use to identify their effects on disbursement of UI to potentially fraudulent recipients. We find that identity verification led to a 40% reduction in UI disbursed to suspicious cards, with minimal negative impact on most other UI recipients. These results suggest that the identity verification technologies implemented during the pandemic were likely to be relatively efficient and cost-effective measures to reduce fraud.

Measurement is a natural challenge when studying fraud, as those who commit fraud try to conceal it. We identify indicators of potentially fraudulent activity by analyzing the income and spending patterns in a data set of tens of millions of debit card transactions, where we can observe both UI deposits and consumer spending behavior. We deploy an unsupervised machine learning

[1]The US Government Accountability Office (GAO) released an assessment report in 2023 highlighting a high risk of fraud (United States Government Accountability Office, 2023). It can be accessed here. Other media and policy reports highlighting fraud instances include Axios, ProPublica, CNBC, the Attorney General of Pennsylvania, and the US Department of Labor.

algorithm on the transactions in our database and group these cards into clusters representing varying levels of suspicious behavior. The key features we use in the clustering algorithm include the amount of UI received by each card, cash withdrawals as a proportion of spending, the speed with which money is withdrawn as cash from ATMs, the amount and proportion of funds wired outside of the United States, and concurrent income that might indicate ineligibility for UI. Some of these features, such as the quantity and speed of cash withdrawal, have been previously documented in the fraud-detection literature as strong signals of suspicious activity (Wu, Xu, and Li, 2019). Our clustering algorithm produces five groups of cards, representing behavior ranging from least to most consistent with fraudulent activity.

We estimate the effectiveness of anti-fraud measures by comparing the change in UI disbursements to suspicious cards before and after these measures were introduced, with the control group defined as the cluster least consistent with fraudulent activity. We create a panel of state-level identity verification measures using data from Freedom of Information Act requests we conducted on all 50 state (and the District of Columbia) unemployment agencies. We then estimate the effects of identity verification policies using a staggered triple difference-in-differences design. Our identification technique captures common time trends across states, state-specific changes in unemployment insurance payments that affect all clusters, and cluster fixed effects. Our results are robust to critiques of modern two-way-fixed-effects difference-in-difference designs.

Our findings suggest that identity verification measures are effective in reducing payouts to cards that display characteristics most consistent with fraud. Specifically, we find that the policy intervention reduced payout to cards in the most suspicious clusters by 40% relative to the control cluster. Moreover, we do not find a significant reduction in the UI paid to non-suspicious cards after the policy was introduced, indicating that the frictions introduced by the ID verification system had a limited negative impact on non-fraudulent applicants. Traditional difference-in-difference estimates looking at the staggered policy variation on each cluster individually further confirm the pattern that UI disbursed to suspicious cards drops, while the control group is unaffected.

In assessing the economic magnitude of our findings, we extrapolate from our sample to estimate that $28 billion was disbursed to suspicious recipients between March 2020 and September 2021 nationwide. Our treatment effects imply that, if all the states had identity verification measures in place before the start of the pandemic, then this policy could have resulted in a nationwide savings of $11.5 billion over the course of our sample period. This assessment is based on an estimated $283 million per month reduction in the UI disbursed by treated states after the policy was implemented, subject to assumptions including generalizability of our treatment effects across states and the representativeness our data.

Our work shows the value of identity verification in eliminating fraud in public benefits. Screening and regulation eliminate the need to claw-back ill-gotten gains, which can be more effective in circumstances where fraud is heavily diffuse and therefore hard to prosecute (Mookherjee and Png, 1992, Polinsky and Shavell, 2000). Moreover, when those committing fraud have limited liability,

preventing the disbursement of funds is necessary because ex-post prosecution would be unlikely to recover funds (Eliason et al., 2021).

Our analysis also provides important insights into future anti-fraud efforts. Identity verification is a low-cost, high benefit program, yet it took high levels of fraud to build the political momentum to implement it. This speaks both to the limited incentives of bureaucrats to reduce fraud in public programs, as well as the value of screening mechanisms in eliminating wasteful or fraudulent government spending.

Our work also reflects on the recent literature on the effects of unemployment insurance during the pandemic and in general, discussed below. While extensive research has shown the effects of UI on household consumption, these effects must be balanced with the caveat that a substantial share of UI spending may be lost to fraud. Moreover, future policy reforms to the UI system should consider implementing effective verification metrics, as we show they have little negative impact on legitimate users but are effective at limiting fraud.

**Related Literature**: Our work connects the literatures on unemployment insurance during the COVID-19 pandemic, fraud in public programs, and consumer financial markets.

A large literature has examined unemployment insurance, including recent papers examining UI during the COVID-19 pandemic. In particular, Ganong, Greig, Noel, Sullivan, and Vavra (2022) show that pandemic-era expansions of UI benefits had large impacts on spending but small impacts on job search, and Dube (2021) similarly shows that the expiration of the Federal Pandemic Unemployment Compensation (FPUC) had little impact on job finding.

Beyond the COVID-19 pandemic, household finance scholars have examined the relationship between UI and consumer finances more broadly. Ganong and Noel (2019) find that spending drops sharply when UI benefits predictably expire, consistent with behavioral models. Hsu, Matsa, and Melzer (2018) explore the interaction between UI and mortgage markets, and find that UI expansions during the Great Recession prevented more than one million foreclosures. While these papers examine the effects of UI funds on households assumed to have been qualified and seeking employment, our paper attempts to quantify the extent to which funds went to individuals and organizations on a fraudulent basis and whose spending, credit, and labor market responses may be very different from that of qualified recipients.

Fraud against the government has generated substantial interest in the literature, but little work has examined unemployment insurance fraud per se. Most closely related to our work, Griffin, Kruger, and Mahajan (2023) estimate the size of fraud in the pandemic paycheck protection (PPP) program, which provided grants to small businesses during the pandemic. Aman-Rana, Gingerich, and Sukhtankar (2022) similarly find that additional paperwork requirements for second-round PPP loans reduced blatant fraud. Fuller, Ravikumar, and Zhang (2015) discuss theoretically optimal monitoring of unemployment insurance fraud; our paper complements that work by examining real-world policies.

Other complementary research has examined waste and fraud in other US benefit programs, including federal health insurance (Leder-Luis, 2020, Howard and McCarthy, 2021) and public procurement (Liebman and Mahoney, 2017). Eliason, League, Leder-Luis, McDevitt, and Roberts (2021) show that up-front paperwork requirements are effective at eliminating fraud in unnecessary federally-funded ambulance rides, with a mechanism analogous to the screening we study in this paper. In seminal work, Nichols and Zeckhauser (1982) show that ordeals can be useful for targeting appropriate beneficiaries, and a long literature has discussed how various forms of administrative ordeals not dissimilar to identity verification have played a role in appropriately targeting beneficiaries.

A growing literature on fraud in consumer financial markets highlights its impact across domains, including investment decision-making and corporate governance. Dimmock and Gerken (2012) demonstrate the predictability of investment fraud, revealing that investors could avoid more than 40% of total dollar losses by avoiding the riskiest 5% of firms. Extending this line of research, Dimmock, Gerken, and Graham (2018) examine the social dynamics of fraudulent behavior, showing that misconduct can be contagious as coworkers significantly influence an advisor's propensity to engage in misconduct. On the regulatory front, Gao, Pacelli, Schneemeier, and Wu (2020) delve into Suspicious Activity Reports (SARs) filed by banks, offering critical insights into institutional reporting mechanisms. More recently, Bian, Pagel, and Tang (2023) show the effect of data protections on financial fraud against consumers. Griffin and Kruger (2023) encourage more investigation in forensic finance, which attempts to detect and understand the economic consequences of these behaviors.

Extensive work in machine learning (ML) has attempted to detect different types of fraud, such as public insurance fraud, e-commerce fraud, and credit card fraud. Shekhar, Leder-Luis, and Akoglu (2023) develop novel machine-learning tools to identify fraud against federal health insurers. Their unsupervised approach is able to provide an 8-fold lift over random targeting. Nanduri, Jia, Oka, Beaver, and Liu (2020) show that Microsoft, one of the largest e-commerce merchants in the world, uses customized long-term and short-term sequential ML models to detect both historical and emerging fraud patterns. Recent work on ML in credit card fraud mostly focuses on supervised learning (Sadineni, 2020, Melo-Acosta, Duitama-Munoz, and Arias-Londoño, 2017). Khatri, Arora, and Agrawal (2020) and Jain, Agrawal, and Kumar (2020) present comparisons of established supervised learning algorithms to differentiate between genuine and fraudulent transactions.

This paper proceeds as follows. Section 2 discusses the background and institutional details. Section 3 discusses the data, and Section 4 discusses the methodology of measuring suspicious cards. Section 5 presents the methodology and effects of identity verification, and Section 6 presents a broader discussion of these findings in the context of anti-fraud policy. Section 7 concludes.

## 2. Background and Institutional Details

Unemployment insurance (UI) is a social safety net program that provides temporary financial assistance to eligible individuals who lose their jobs through no fault of their own. The benefits are meant to provide income support for workers who are laid off or furloughed while they search for new employment opportunities. These programs are generally administered by state governments, and applicants file for claims using their state's predominantly online application portal. States verify eligibility using the demographic and economic data supplied by the applicants, which includes their identity, social security number, date of birth, and address. This system is susceptible to fraud by various means, including identity theft, wherein criminals use attributes such as the social security number of a potentially eligible recipient to apply for and divert UI funds.

Unemployment insurance took two forms during the COVID-19 pandemic: regular state UI programs, and new federal programs such as Pandemic Unemployment Assistance (PUA) that provided UI to gig-economy and other traditionally ineligible workers. While the disbursement of both the programs was administered by states, in general, PUA funds were subject to federal eligibility requirements. Other programs such as the Federal Pandemic Unemployment Compensation (FPUC) provided a top-up to existing state UI and were funded by the federal government.

Improper payments in financial support programs have been a problem historically. In 2019 for example, the Bureau of Labor Statistics reported a 9% average overpayment rate of UI, with some states as high as high as 32%.[2] UI fraud became an even more pressing concern in the wake of the COVID-19 pandemic, when unemployment insurance expanded significantly. According to the Bureau of Labor Statistics (BLS), the unemployment rate in the U.S. rose from 3.5% in February 2020 to 14.8% in April 2020, the highest rate since the Great Depression. Figure 1 plots the sharp increase in UI benefits paid starting in 2020, which is mirrored in debit card data we use. The unprecedented surge in unemployment claims, relaxed eligibility criteria, and the implementation of new relief programs acted together to make it easier for fraudsters to exploit the system and obtain benefits illicitly.[3]

In order to detect and deter fraudulent activity, many states implemented anti-fraud measures such as identity verification. Generally contracted through third party vendors such as ID.me, identity verification measures sought to eliminate fraudulent spending by ensuring that the applicant and recipient of the funds were not using a stolen identity to claim these benefits. Several states began the implementation of ID.me as early as September 2020, including Indiana and Georgia. Conversely, other states commenced this process at a later date. For instance, Massachusetts did not start using ID.me until March 2021. Several states chose to implement identity verification only for their regular (non-PUA) UI programs. Further, some states implemented this screening in a staggered manner (such as initially for new claimants and later to include continuing claimants).

[2]The Department of Labor (DOL) UI payment accuracy datasets are available here.

[3]See pages 5-7 of the US GAO assessment report here.

To determine the timing of state identity verification programs, we conducted Freedom of Information Act (FOIA) requests on all 50 state unemployment agencies to determine the nature and timing of their unemployment insurance identity verification policies and any external contracts to provide such services. We corroborated these dates with a Congressional report about UI fraud policies (Committee on Oversight and Accountability, 117th Congress, 2022) as well as news reports.[4]. Table A1 provides a timeline of ID verification measures, and Figure A1 presents the time variation in state-level adoption of ID verification measures as a map.

A typical identity verification process required a UI applicant to submit photos of themselves and their identity documents such as a driver's license or passport via a smartphone. When states implemented identity verification technology, it became a requirement before funds are disbursed. This made it more challenging for identity thieves to use others' stolen identity to apply for unemployment benefits. Figure A2 pictorially represents the steps involved in ID.me's verification process. A majority of claimants were able to complete this process within 10 minutes.[5]

## 3. Debit card transaction data

Our data come from Facteus, a fintech firm that partners with banks, card issuers, and payment processors to aggregate, standardize, and anonymize transaction-level information from debit cards.[6] The data on each transaction include the amount, transaction date, and description. Further details such as merchant category code (MCC) and card-holder ZIP code help us confirm that the data represent a wide variety of consumer transactions and a geographically representative sample.[7]

Facteus perturbs the raw data it obtains from its data partners to generate synthetic data in order to protect consumer privacy. For example, the transaction amounts we observe are perturbed by adding small mean zero noise to raw transaction amounts. Similarly, transaction time is perturbed by up to several hours around the original time. This mean zero noise is small and we ignore it in our analysis. The transaction description, MCC code, and cardholder ZIP that we rely on for key results are not perturbed by Facteus.

Debit card holders are a particularly interesting and poorly-understood population to study in the context of public benefits. Most states allow claimants to receive benefits on state-issued prepaid debit card or pre-existing prepaid card via direct deposit.[8] Recipients who receive UI on

[4]The Congressional report is part of a set of documents published with the press release in November 2022 and accessed by us in January 2023. A Reuters news article on the states using ID.me technology for UI verification can be found here

[5]Estimates from the US Treasury department are available here.

[6]For other papers that use data from Facteus, see Brave, Fogarty, Aaronson, Karger, and Krane (2021), Karger and Rajan (2020) and Zhou and Correia (2022).

[7]Using card-holder ZIP codes, Zhou and Correia (2022) show that the geographic distribution of cards in Facteus sample closely resembles the population density across the United States.

[8]The Consumer Financial Protection Bureau (CFPB) explains the methods by which UI funds can be received here. A specific example of a state-issued prepaid care from the state of Michigan can be found here.

debit cards can also engage in spending and transfers through those cards, as well as make deposits and receive other forms of income. We observe all of these types of transactions in our data. Consumers who receive government benefits via prepaid cards tend to be lower income and more likely to lack traditional bank accounts.

Hence debit cards are a particularly relevant context for policies aimed at the economically vulnerable, and also less likely to be observed through traditional data sources originating from banks and surveys. Furthermore, prepaid cards may be more susceptible to fraud than bank accounts, as criminals and fraudulent recipients can quickly open one or more card accounts for the express purposes of obtaining and siphoning funds, and the banking system has much more stringent measures for fraud and account opening compared with debit cards. One limitation of our data source is that we can only connect transactions at the card level, not the individual level. Therefore, if a single individual or entity obtains UI benefits on multiple different cards, we have no way of linking them and analyzing cross-card behavior.

We use a string matching technique on transaction descriptions to identify UI inflows from state agencies. Each state agency uses a consistent transaction description for UI payments, we source these descriptions from the Washington Bankers' Association.[9] We are able to identify UI inflows for 41 of the 50 states and the District of Columbia, and Table A1 provides details on our data coverage. In four states (Arkansas, Massachusetts, Ohio, and West Virginia), we are further able to distinguish inflows from regular UI versus the Pandemic Unemployment Assistance (PUA). For these four states, we aggregate state-level UI flows into PUA and non-PUA separately.

Based on the technique above, we observe state-level UI transactions on about 127,000 debit cards beginning in January 2019 and ending in September 2021. Out of these, we focus on 84,865 cards that received at least $1,000 of UI during the sample period. For these cards, our full sample contains 2 million UI inflow transactions, 3 million non-UI inflow transactions, and 30 million outflow transactions across various spending categories. Each transaction is tagged with the merchant's MCC code, which we use to categorize into types such as income from UI or other sources, spending via cash withdrawal from ATMs or bank branches, spending on groceries or other purchases, and wire transfers.[10] Additionally, we observe transaction timestamps which allow us to measure the time gap between receipt of UI and spending through cash withdrawals.

As a validation check, we compare UI flows from Facteus to administrative data from each state's Department of Labor. Figure 1 plots the time series of UI disbursements observed in our analysis sample with the corresponding aggregates from state administrative data, and shows that the two series follow a very consistent pattern over the sample period and have a correlation of 0.95. Table A2 shows the relationship between our data and the BLS data by state.

Table 1 shows card-level and card-month level summary statistics for our sample of cards that received any UI. On average, cards received a total of $10,000 of UI benefits across 19 disbursements,

[9]The list of transaction descriptions for each state can be found here.
[10]The list of MCC codes is available here.

at a rate of about 4 per month. Spending is broken into categories including cash withdrawals, wire transactions, and groceries. The average monthly income per card is $2,790, of which $2,126 comes from UI. Total spending has a mean of $1,807 per month, with cash spending at $449 (25%) and groceries at $341 (19%) of the total. We group spending on other categories such as liquor shops, gambling, tobacco, restaurants, and purchase of vehicles as discretionary spend and note that it represents 16% of the total.

## 4. Measuring Fraud with Machine Learning: K-Means Clustering

Machine learning techniques are finding increasing use in analyzing high-dimensional data for the purposes of prediction, classification, and clustering in finance and economics (Mullainathan and Spiess, 2017). Athey and Imbens (2019) highlight that machine learning methods are designed to maximize out-of-sample predictive accuracy and deliver better performance over traditional econometric approaches that emphasize parameter estimation. Fuster et al. (2022) show that machine learning delivers higher out-of-sample predictive accuracy for default rates but has implications on racial disparity in the distribution of these gains.

A growing literature also argues that consumer payments data contains valuable information on financial behavior and can be used to identify patterns for predicting future outcomes. Puri et al. (2017) provide evidence that payment data are predictive of loan default rates. Berg et al. (2020) show that signals from individuals' digital footprints can be as informative as credit bureau scores, and Berg et al. (2022) review the growth of FinTech lending and note the increasing use of machine learning to improve the screening of customers.

We deploy a popular unsupervised machine learning algorithm called k-means clustering on the income and spending patterns of cards to classify them into mutually exclusive and collective exhaustive groups. Clustering algorithms are machine learning techniques that group observations by their related values. Clustering is an unsupervised form of machine learning, meaning that it does not rely on labeled training data. Rather than extrapolating from known fraudulent patterns, unsupervised learning instead groups observations by similar patterns among variables chosen by the researcher. Unsupervised techniques have the benefit that they do not rely on successfully-identified fraud, which may be non-representative (Shekhar et al., 2023).

Clustering algorithms require the analyst to first choose which features, i.e. variables, are used to create groups. In related literature, Wu et al. (2019) discuss the use of feature construction techniques to improve the performance of machine learning models for fraudulent credit card cash-out detection. They note that patterns in cash spending and speed of cash withdrawal can meaningfully improve the accuracy of fraud detection. In order to extract signals of potential fraud from our dataset of 35 million debit card transactions , we create several features that are likely consistent with fraudulent behavior:

- Average unemployment insurance received per month
- Longest unbroken UI spell (months)
- Cash withdrawals as a fraction of monthly outflows
- Speed of cash withdrawals in hours after a UI inflow
- International wire transfers as a fraction of of monthly outflows
- Discretionary spending as fraction of monthly outflows
- Concurrent non-UI income (number of consecutive months)

Discretionary spending includes spending in liquor stores, gambling/casinos, tobacco stores, purchase of vehicles, and spending in restaurants. We measure concurrent non-UI income as the number of consecutive months a card receives at least as much income from non-UI sources as UI sources.

We use K-means clustering to partition the cards into several clusters that display distinct characteristics with respect to the features listed above. A K-means algorithm chooses centroids for each dimension and creates clusters based on the proximity of each card to the centroid along that dimension (Hartigan and Wong, 1979). In short, this algorithm allows us to group cards together based on similarities along the dimensions we identify as salient.

This method confers several advantages including: (i) it flags anomalous patterns in transactions without relying on a priori labeled training data, (ii) there is no manual specification of the number of cards that must belong to each cluster, and (iii) we do not set arbitrary cut-off points for any of the dimensions that separate one card from another.

We make a few considerations when designing the clustering algorithm. First, several features collectively indicate potential fraud, but no one feature can be used standalone as a basis for grouping cards. For example, the UI received on a card, even in excess of statutory limits, does not by itself convey sufficient information because it is possible that multiple people use the same card. For the same reason, overlap of UI with other income by itself does not indicate fraud because one person could still be employed while the other is a legitimate beneficiary of UI rolls. However, when unusually large UI inflows (for which a single person is unlikely to be eligible) are combined with concurrent income from other sources and a long unbroken spell, then these features collectively could indicate ineligibility for UI.

Second, some dimensions are informative only conditionally. For example, the speed of withdrawal is applicable to only those cards that withdraw cash after the receipt of UI. Therefore, we construct a multi-stage clustering algorithm that starts with broad clusters that separate cards along the provided dimensions and then creates more granular categories based on additional relevant dimensions. A hierarchical or iterative K-means clustering procedure can also help to locate within-cluster outliers better than a single step procedure. For example, Rao, Garai, Clarke, and Dey (2018) apply iterative K-means in two steps: first, over the entire dataset, and second, over cluster subsets to further elaborate any dimensions that contain outliers. This process allows the

uncovering of outliers within what might appear to be broadly homogeneous clusters in the initial step. The algorithm updates the centroids at each iteration to more accurately identify homogeneous clusters. As a result, the within cluster heterogeneity is small and the cross-cluster difference on informative features is large.

Figure A3 plots a flow chart of the sequence in which clusters are created. Our main clustering algorithm focuses on around 85,000 debit cards which receive at least $1,000 in UI. We start with clustering cards into those that receive unusually large amount of UI and those that do not. The relevant dimensions used by the algorithm for the first set of clusters include the average UI received by a card each month, and the longest unbroken spell of being on UI rolls (in months). The first step creates two clusters: High UI with about 24% of cards and Low UI with the rest. High UI cards are considered for further sub-clustering, with the rest included in the control category.

On the cards with large UI inflows, we apply the clustering algorithm using spending patterns simultaneously encompassing cash withdrawals, wire out of money outside the US, and spending on non-essentials. This process generates one cluster each for cards that withdraw an abnormally large proportion of funds in cash, those that spend a large proportion of wealth on categories such as alcohol and gambling, and all other cards. Within the cluster of cards with abnormally high cash withdrawals, the algorithm further splits them into those with fast or slow withdrawals after the receipt of UI. We calculate the speed of withdrawal from ATMs using the timestamps in our data and then apply this to the card cluster that withdraws a lot of cash to split it into two.

Finally, for the third cluster of remaining high UI cards, we apply the clustering algorithm to detect potentially ineligible recipients using overlap of their UI income with concurrent other income on the same card. Note that the dimension of international wires, while potentially informative, does not produce a unique cluster because very few cards in our sample both receive high UI and transmit an abnormally large proportion of their income abroad.

The machine learning algorithm produces five clusters that group cards along economically similar behavior patterns. Figure A4 shows bi-variate plots with color groups representing clusters along two dimensions at a time. Table 2 shows the means and standard deviations of the feature variables associated with each cluster. Based on their characteristics, we name each of the clusters, which we use to discuss our results going forward.

1. "Suspicious (Fast Cash)": abnormally large UI, and extremely large and quick cash withdrawals. This may reflect organized criminal behavior.
2. "Suspicious (Slow Cash)": abnormally large UI and cash withdrawals, but not at the same speed as Suspicious (Fast Cash).
3. "Concurrent Income": concurrent income from other sources and high levels of UI.
4. "Discretionary Spending": high amounts of UI received, and large spending on non-necessities, such as liquor stores.
5. "Control": cards that do not fall into one of the other clusters, generically reflecting genuine,

non-fraudulent behaviors.

The labels we give to each cluster reflect our interpretation of the behavior of these cards. The "Suspicious (Fast Cash)" and "Suspicious (Slow Cash)" clusters receive much higher UI than other cards, and withdraw 66% and 61% (respectively) of their money in cash. Suspicious (Fast Cash) is distinguished by having a mean time to cash withdrawal of only 14 hours, which may be consistent with organized criminal activity. Withdrawing cash removes it from the banking system and makes future transactions untraceable, as well as lowering the probability of any future clawback measures. The remaining two clusters represent either spending on non-essentials or concurrent income from other sources, indicating some likelihood of ineligibility to receive UI. All other cards with the least fraudulent behavior form part of the control cluster.

Cards belonging to Suspicious (Fast Cash) and Suspicious (Slow Cash) clusters display behaviors most consistent with those likely to be exhibited by claimants who would conduct identity theft. For example, one in these clusters receives UI in excess of $20,000 in a single day, and withdraws over 80% by cash. Another card in these clusters receives UI in excess of $80,000 in under 6 months, and withdraws $60,000 in cash from various ATMs. Suspicious (Fast Cash) is also characterized by same-day withdrawals on average, consistent with the idea that converting illegitimate gains into cash makes it much harder for authorities to claw back the disbursed funds. While we cannot claim that every card in these clusters is necessarily fraudulent, the behaviors displayed by cards in this cluster are most consistent with fraudulent claims. Furthermore, using the card-level zip codes in our data, we map each card to a metropolitan statistical area (MSA) and calculate the proportion of potentially fraudulent cards in each MSA. The MSAs with some of the highest concentration of fraudulent cards in our data include Reno-Sparks (NV), Las Vegas-Paradise (NV), Boston-Cambridge-Quincy (MA-NH), Springfield (MA), Seattle-Tacoma-Bellevue (WA) and Miami-Fort Lauderdale-Pompano Beach (FL). Many of these show up among the MSAs with above-average rates of identity theft in publicly available administrative data.[11]

Table A3 shows the amount and share of UI that went to each cluster before and after the onset of the COVID-19 pandemic. In line with media reports, the share of UI that went to the clusters most consistent with fraud increased during the pandemic. Likewise, for the four states where we can identify PUA and non-PUA inflows separately, the share of UI that went to suspicious clusters was higher under the PUA program, as shown in Table A4. This is consistent with reports that the PUA program was more prone to fraud (United States Government Accountability Office, 2023)

## 5. The Effects of Identity Verification on UI Fraud

We construct a difference-in-differences model to estimate the effect of identity verification on state-level UI disbursement. Our main outcome variable is the UI inflows to cards based on their cluster,

[11]The data from Federal Trade Commission is available here.

as measure of fraudulent or suspicious flows. We also consider the number of cards to whom UI was disbursed within a cluster as an alternate measure.

Our sample consists of the 41 states in our data where we can identify UI payments. Between January 2019 and September 2021, 22 states in our sample contracted with third-party agencies such as ID.me and LexisNexis to screen the identity of UI applicants; these states are considered treated.[12] Within these 22 treated states, the timing of adoption differs, starting from September 2020 for Georgia and Indiana to June 2021 for Delaware.[13]

We implement two specifications, each examining 6 months before and after the implementation of identity verification. First, we consider a triple-difference design, comparing the Suspicious cluster to the control (non-suspicious) cluster, before versus after the implementing of identity verification, with treated and control states. Next, we consider the classic difference-in-difference design, looking at each cluster separately. Both of these designs use the two-way-fixed-effects framework, saturating the model with state and time fixed effects. These designs allow us to examine if the more likely fraudulent cards were disproportionately affected by identity verification measures.

Finally, we consider critiques of modern two-way fixed-effects estimators, and implement a number of robustness checks. Our estimates are robust to different specifications and are not contaminated with issues such as negative weighting.

### *5.1. Triple-difference specification*

We estimate the model

$$y_{sct} = \sum_{cl \neq \text{“}Control\text{”}} \beta_{cl} \times \mathbb{I}_{[c=cl]} \times Post + \delta Post + \alpha_s + \alpha_c + \alpha_t + \alpha_{cs} + \alpha_{ct} + \varepsilon_{cst}, \tag{1}$$

where the dependent variable $y_{sct}$ is the log unemployment insurance dollars paid by state $s$ to cluster $c$ in month $t$. $Post$ is an indicator variable covering six months after a state $s$ adopts identify verification technology, with six months before the implementation date being the pre-period for comparison. The treatment effect of identity verification is identified by the three $\beta_{cl}$ parameters where the base cluster is “Control”.

This comparison is relative to clusters with most likely genuine cards and therefore can be thought of as a triple-differences specification. The specification accounts for cluster, state, month,

[12] Washington state implemented only a pilot version of identity verification on a small sample of recipients but did not extend it to the general population. We consider Washington as untreated.

[13] Some states implemented these measures for one program only (i.e. PUA or non-PUA, not both). We are able to distinguish between PUA and non-PUA inflows of UI for four states - Arkansas, Massachusetts, Ohio, and West Virginia. For those states, we construct the panel identifier as state-program rather than state.

post, cluster-state, and cluster-month fixed effects. Standard errors are clustered by state and observations are weighted by state population. We estimate this model for Suspicious (combination of Suspicious (Fast Cash) and Suspicious (Slow Cash) clusters), Concurrent Income, and Discretionary Spending clusters. Table 3 reports the estimation result. Column (1) of Table 3 uses the log of UI dollars as the baseline dependent variable, and column (2) repeats the estimation on the log of number of cards that received UI within the cluster.

Identity verification measures significantly impacted cards that displayed spending behavior most consistent with fraud. Column (1) of Table 3 shows that the treatment effect of this policy on Suspicious cluster was -0.52 log points, or a 40.5% decline, relative to the control cluster over a period of six months. This effect is observed after controlling for geographical and time-series variations in UI paid to card clusters. On the other hand, we do not observe a statistically significant impact on the other two (Concurrent Income and Discretionary Spending) clusters, indicating that these measures were targeted towards reducing specific kinds of fraud. Column (2) of Table 3 also shows a comparable decline in the number of unique cards receiving UI in the Suspicious cluster.

We next estimate a dynamic specification to validate that our results are not driven by pre-trends. The key identifying assumption is the standard parallel trends assumption – that the only differential change between suspicious and non-suspicious clusters at the time of transition that would impact UI payouts is the transition itself. We estimate the model

(2)
$$y_{sct} = \sum_{\substack{\tau \in -6,6, \\ \tau \neq -1}} \sum_{cl \neq \text{“}Control\text{”}} \beta_{cl,\tau} \times \mathbb{I}_{[c=cl]} \times Reltime_{\tau} + \delta Reltime_{\tau} + \alpha_s + \alpha_c + \alpha_t + \alpha_{cs} + \alpha_{ct} + \varepsilon_{cst},$$

where the dependent variable $y_{sct}$ is the log unemployment insurance dollars paid by state $s$ to cluster $c$ in month $t$. $Reltime_{\tau}$ refers to the number of months until treatment takes place, starting and ending six months about the month when identity verification was implemented. The treatment effect each month for each cluster is estimated using month “-1” and the “Control” cluster as base levels. Fixed effects for cluster, state, month and relative-month are included. Standard errors are clustered by state and observations are population-weighted.

Figure 2 shows the effect of identity verification on the UI received by the Suspicious cluster, relative to the control cluster. Panel (a) of Figure 2 uses log UI as the dependent variable and panel (b) uses the log of number of cards. The differences in UI payouts manifest themselves immediately upon the introduction of identity verification in (relative) month zero. ID verification reduced the UI received by the most suspicious cards by over 20% relative to the level of the control (non-suspicious) cluster immediately after its introduction.[14] This effect grows over subsequent months

[14] Figure A5 shows the estimation results when the Suspicious (Fast Cash) and Suspicious (Slow Cash) clusters are kept separate, and shows similar trends.

as more claimants are brought under the scope and payouts reduce to suspicious cards by as much as 70% in later months compared to the month before these measures were introduced. We also note the validity of parallel trends assumption before the treatment month and rule out the observed impact to any pre-existing trends.

We do not observe a similarly sharp change in the UI paid out to other clusters in Figure A6. Furthermore, we cannot rule out a downward-sloping pre-existing trend in the UI paid to the Concurrent Income cluster. These result suggest that the policy was successful in targeting claimants that sought to remove funds from the system soon after their receipt rather than individuals who may have claimed monies in excess of their eligibility or spent them on non-essential needs, but may not have conducted identity theft in the process of claiming them.

### *5.2. Within-cluster difference-in-differences design*

Our triple difference specification compares the outcomes of the Suspicious cluster with respect to the control cluster; however, it does not tell us whether and how much the control cluster itself was impacted by this policy. We consider the possibility that both the Suspicious and control clusters may have experienced a decline in UI due to the treatment. For example, if ID verification introduced costly system-wide frictions, that would also affect genuine applicants. In contrast, if genuine applicants are unaffected, this would indicate UI identity verification was an effective form of targeting.

We conduct a classic difference-in-differences analysis to assess cluster-specific treatment effects. We estimate the following model on each of the four clusters separately:

$$y_{st} = \sum_{\substack{\tau \in -6,6, \\ \tau \neq -1}} \beta_\tau \times Reltime_\tau + \alpha_s + \alpha_t + \varepsilon_{st}. \tag{3}$$

This model is a cluster-specific version of Equation 2 and controls for state and time fixed effects. Parameter $\beta_\tau$ identifies the impact of identity verification on the UI paid by treated states to that cluster for each of the six months before and after implementation, based on the UI paid in month -1. Standard errors are clustered by state and observations are population-weighted. We estimate this model in turn for Suspicious, Concurrent Income, Discretionary Spending, and control clusters and show the resulting event plots in Figure 3. (Figure 4 reports the estimation using log of number of cards in each cluster that received UI instead of the log of UI dollars.)

Figure 3 show an immediate and persistent decline in UI paid to the Suspicious cluster of about 0.5 log points. This magnitude is comparable to the triple-difference specification, which points to a real decline in the UI disbursed to potentially fraudulent recipients.

In contrast, Panels (b) and (c) show that ID verification does not seem to have impacted cards

that received UI despite potential concurrent income or cards that spend disproportionately more on discretionary categories. This suggests that even if such cards were not necessarily eligible for UI, their primary means of claiming UI may not have been through identity theft.

Importantly, for the control cluster we do not observe a statistically significant decline in UI in any of the post-treatment months. Panel (d) of Figure 3 shows that the control cluster continued to receive UI comparable to month -1, with an upward trajectory in later months. This points to the idea that identity verification did not lower UI receipiency among non-fraudulent cardholders.

### *5.3. Robustness*

We conduct three robustness checks to address critiques of modern two-way fixed effects designs in settings with staggered adoption of policy and potentially heterogeneous treatment effects.

#### *5.3.1. Negative weights.*

A TWFE design can give negative weights to already treated states if they are included as controls for states treated later during the sample period. This is particularly a concern when there is heterogeneity in treatment effects over time (Borusyak, Jaravel, and Spiess, 2021). In our setting, treatment effects may be heterogeneous across both time and jurisdiction. For instance, some states introduced ID verification initially for a subset of claimants (such as new applicants only) and extended it to all claimants over the following months. Similarly, the initial level of fraud in some states may have been higher than others, which could also affect the post-treatment outcomes. Roth, Sant'Anna, Bilinski, and Poe (2023) suggest that if treatment effects vary across both time and units, some observations may get negative weight in the TWFE estimand.

Following Roth et al. (2023), we check and rule out the presence of negative weights in our estimation. Under the common trends assumption, all 123 (41 states $\times$ 3 clusters) average treatment effects on the treated carry a positive weight. Consequently, the sum of the positive weights is equal to 1 and that of negative weights is equal to 0.

#### *5.3.2. Alternative estimation of triple-differences design.*

The potential inclusion of "already treated" states as control states for later treatments can lead to the violation of common trends assumptions. This is because the treated states may no longer follow the same trends as the never treated ones, due to the heterogeneity in treatment effects and a "phase-in" of treatment over time. In addition to checking for negative weights, we re-estimate both Equation 1 and Equation 2 using a "stacked" difference-in-differences approach to mitigate this concern.

To do this, we follow Cengiz, Dube, Lindner, and Zipperer (2019) in creating a stacked data set that constructs appropriate control groups for each treatment date. We follow four steps: (i) we define a window of 6 months before and after the treatment month to be included in each sub-experiment, (ii) we enumerate a total of 10 unique treatment months (or experiments) when

all 22 states in our sample implemented ID verification, (iii) we define the inclusion criteria for control states in each sub-experiment to comprise of never treated and not treated during the post six month time window of that experiment, and (iv) we append the individual experiment data and re-estimate both Equation 1 and Equation 2 on this stacked data set.

The advantage of this approach is that the already treated states or the soon-to-be-treated states are not included in the evaluation of any given experiment, by design. This ensures that the estimated treatment effect is free from potential biases arising from violation of common trends. Each of these data sets is appended such that each transitioning jurisdiction appears once while each jurisdiction may appear as a control multiple times (although with different time values). In order to account for duplication of control states in the stacked dataset, we cluster standard errors by state. The resulting specification averages all of the time-varying effects into a single effect.

Table A5 shows the estimation of Equation 1 for the Suspicious, Concurrent Income, and Discretionary Spending clusters based on month -1 and the control cluster. We note that both the point estimates and the statistical significance are comparable to the baseline results reported in Table 3.

Figure A7 shows the dynamic event-study plots for the three clusters as in Equation 2. Again, we obtain trends consistent with the baseline TWFE version and interpret the results as causal evidence of identity verification on lower UI payout to the most likely fraudulent cards.

*5.3.3. Alternative estimation of within-cluster difference-in-differences design.*

We re-estimate Equation 3 using the semi-parametric group-time average treatment effect estimator proposed in Callaway and Sant'Anna (2021). This method estimates the effect of treatment separately for each group of states treated at the same time, using only those states that are never treated or not-yet treated as the control group. Using only valid comparisons for the estimation of average treatment effects, this method avoids weighting problems associated with TWFE specifications when there are multiple time-periods and variation in treatment timing.

Figure A8 shows the event study plots for all four clusters using the Callaway and Sant'Anna (2021) estimator. We continue to observe a decline in UI paid to the Suspicious cluster after the introduction of ID verification. The pre-trends for the Suspicious cluster confirm that there was no decline observed before the implementation of policy. Further, the control cluster cards did not see a decline in the post period, consistent with the baseline results in Figure 3.

## 6. Estimation of Economic Magnitude

Our results suggest that identity verification reduced payouts to cards that displayed characteristics most consistent with fraud. In this section, we provide an estimate of the economic magnitude of fraud and the dollar benefit of ID verification by extrapolating our results to the universe of UI payments in the United States. We also discuss the assumptions that underlie these estimates.

We start by noting that the share of total UI received by the Suspicious clusters between March 2020 and September 2021 was 10.17%, as per Table A3. Based on administrative data sourced from the US Department of Labor, the total UI paid under the PUA and non-PUA programs by these 41 states was $192.8 billion during the same period. Under the assumption that the share of fraud observed in our data is representative, the amount of UI allocated to potentially fraudulent cards amounts to $19.6 billion among these 41 states. Furthermore, we can extrapolate this number to include the remaining 10 states that do not constitute our sample. With the same assumption of representativeness in mind, we scale up the estimate using their population. This brings us to an estimated $28 billion paid to cards over the 19 months since March 2020.

Next, we estimate the dollar savings from ID verification programs to treated states and then extrapolate it to all states. Identity verification declined the monthly UI paid out to the Suspicious cluster by 40.5%. Based on the share of Suspicious cluster in total UI of 10.17%, this represents an overall decline of 4.1% in the monthly sums paid out by states. (The percentage of UI to Suspicious cluster is nearly identical in treated and non-treated states). The actual UI paid by all treated states as per administrative data in the month prior to treatment was $6.9 billion. Thus, ID verification saved these states about $283 million per month after implementation of the policy. Extrapolating this estimate to 19 months in our sample, we arrive at a counterfactual estimated savings of $5.4 billion for treated states if they had ID verification measures in place before the start of the pandemic. Further, applying the same counterfactual exercise to include all states (both, control states and those that do not form part of our sample), we arrive at an estimated total savings of $11.5 billion assuming that the treatment effect uniformly applies to the remaining states.

Notably, this estimate is based on the level of spending in the month before the policy was implemented, as is standard in difference-in-difference estimates. This assumes that the pre-treatment share of fraud clusters would have remained at the same level had ID verification not been introduced. Figure 3 suggests that fraud was on a rising trend; hence it is more likely that the proportion of fraud would have increased over time in the absence of this policy intervention, and the cost savings of the policy were therefore even greater.

We stress that these estimates rely on assumptions about our data coverage and the degree to which treatment effects observed in the states in our sample can be equally applied to all states. These are strong assumptions, particularly if there is any selection into the debit card market which we cannot observe. For example, the share of potential fraud observed in the debit card data may not equal the share of fraud through, for example, direct bank deposits and paper checks. However, in our favor, many states used debit cards to issue UI payments, regardless of the consumer's access to direct deposit, and therefore we can think of the debit card market as more representative of the population than it would be in other contexts.

Furthermore, we base our estimates on the two main UI programs active during the pandemic – regular state UI and the Pandemic Unemployment Assistance (PUA) for informal sector workers.

The pandemic also led to an array of additional programs and enhancements to existing programs. For example, the Federal Pandemic Unemployment Compensation (FPUC) provided a federal top-up to existing benefits and was active from March through July 2020, and then again from January 2021. Since we do not observe the exact time-series of this and other programs, we do not include them in our estimation. Our estimate is therefore likely a lower bound of both potential fraud and the savings from the identity verification programs.

Our estimates are presented as a point of comparison. Our conservative estimate of $11.5 billion in savings hints at the magnitude of economic gains that may accrue from adoption of modern technologies to screen UI applications. The direct cost of these technologies, on the other hand, is a fraction of these savings: the Congressional documents we cite indicate that the sum total of contracts signed between ID.me and state agencies was under $50 million (Committee on Oversight and Accountability, 117th Congress, 2022). Overall, our analysis indicates that ID verification was a highly effective screening tool with significant positive economic benefits and low costs to the government .

## 7. Conclusion

Fraud is a major problem in public programs, and it is both challenging to measure as well as to eliminate. Through the COVID-19 pandemic, unemployment insurance saw a historic expansion that was plagued by fraud in the form of identity theft. This fraud threatened to divert funds from necessary recipients to potential fraudsters and to waste valuable resources in mitigating the economic effects of the recession.

We provide a first assessment of the effectiveness of identity verification in reducing fraud in UI. First, we leverage machine learning tools to identify fraud, clustering cards together on salient features such as how quickly cash was withdrawn, or whether the recipient had other income. Then, using data from a set of new FOIA requests, we show that the rollout of identity verification procedures was effective at reducing fraud, while sparing cards that did not engage in suspicious behavior. In all, these results suggest that identity verification measures were effective at targeting UI to needy recipients and mitigating fraud.

This study further provides new insights on how to measure fraud in benefits programs. Our clustering algorithm allows us to identify potentially fraudulent spending, which is valuable for understanding cost-benefit trade-offs in benefit expansion or the implementation of new regulations. A primary challenge in the study of fraud is that it is concealed; machine learning tools may prove broadly useful in future work to overcome these challenges.

Our work also opens a number of questions about the political economy of benefits programs, which may prove fruitful for future work. The technology required for UI verification was available before the pandemic, but it took potentially billions of dollars of fraud to occur before it was implemented. This highlights agency issues in the administration of UI fraud and other benefits

programs, where technological adoption has largely been inefficiently slow (Pahlka, 2023). Given that identity verification was so valuable, our study highlights the fact that it could have been implemented before the expansion of UI, eliminating even more fraud.

## References


Aman-Rana, S., D. Gingerich, and S. Sukhtankar (2022). Screen now, save later? the trade-off between administrative ordeals and fraud. *The Trade-Off between Administrative Ordeals and Fraud (August 18, 2022)*. [4]

Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economics 11*, 685–725. [9]

Berg, T., V. Burg, A. Gombović, and M. Puri (2020). On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies 33*(7), 2845–2897. [9]

Berg, T., A. Fuster, and M. Puri (2022). Fintech lending. *Annual Review of Financial Economics 14*, 187–207. [9]

Bian, B., M. Pagel, and H. Tang (2023). Consumer surveillance and financial fraud. Technical report, mimeo. [5]

Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*. [16]

Brave, S. A., M. Fogarty, D. Aaronson, E. Karger, and S. D. Krane (2021). Tracking us consumers in real time with a new weekly index of retail trade. [7]

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230. [17, 44]

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134*(3), 1405–1454. [16, 43, 45]

Committee on Oversight and Accountability, 117th Congress (2022, November 17). Congressional Press Release. [7, 19]

Dimmock, S. G. and W. C. Gerken (2012). Predicting fraud by investment managers. *Journal of Financial Economics 105*(1), 153–173. [5]

Dimmock, S. G., W. C. Gerken, and N. P. Graham (2018). Is fraud contagious? coworker influence on misconduct by financial advisors. *The Journal of Finance 73*(3), 1417–1450. [5]

Dube, A. (2021). Aggregate employment effects of unemployment benefits during deep downturns: Evidence from the expiration of the federal pandemic unemployment compensation. Technical report, National Bureau of Economic Research. [4]

Eliason, P. J., R. J. League, J. Leder-Luis, R. C. McDevitt, and J. W. Roberts (2021). Ambulance taxis: the impact of regulation and litigation on health care fraud. Technical report, National Bureau of Economic Research. [4, 5]

Fuller, D. L., B. Ravikumar, and Y. Zhang (2015). Unemployment insurance fraud and optimal monitoring. *American Economic Journal: Macroeconomics 7*(2), 249–290. [4]

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2022). Predictably unequal? the effects of machine learning on credit markets. *The Journal of Finance 77*(1), 5–47. [9]

Ganong, P., F. E. Greig, P. J. Noel, D. M. Sullivan, and J. S. Vavra (2022). Spending and job-finding impacts of expanded unemployment benefits: Evidence from administrative micro data. Technical report, National Bureau of Economic Research. [4]

Ganong, P. and P. Noel (2019). Consumer spending during unemployment: Positive and normative implications. *American economic review 109*(7), 2383–2424. [4]

Gao, J., J. Pacelli, J. Schneemeier, and Y. Wu (2020). Dirty money: How banks influence financial crime. *Available at SSRN 3722342*. [5]

Griffin, J. M. and S. Kruger (2023). What is forensic finance? *Available at SSRN 4490028*. [5]

Griffin, J. M., S. Kruger, and P. Mahajan (2023). Did fintech lenders facilitate ppp fraud? *The Journal of Finance*. [4]

Hartigan, J. A. and M. A. Wong (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics) 28*(1), 100–108. [10]

Howard, D. H. and I. McCarthy (2021). Deterrence effects of antifraud and abuse enforcement in health care. *Journal of Health Economics 75*, 102405. [5]

Hsu, J. W., D. A. Matsa, and B. T. Melzer (2018). Unemployment insurance as a housing market stabilizer. *American Economic Review 108*(1), 49–81. [4]

Jain, V., M. Agrawal, and A. Kumar (2020). Performance analysis of machine learning algorithms in credit card fraud detection. In *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, pp. 86–88. IEEE. [5]

Karger, E. and A. Rajan (2020). Heterogeneity in the marginal propensity to consume: evidence from covid-19 stimulus payments. [7]

Khatri, S., A. Arora, and A. P. Agrawal (2020). Supervised machine learning algorithms for credit card fraud detection: a comparison. In *2020 10th international conference on cloud computing, data science & engineering (confluence)*, pp. 680–683. IEEE. [5]

Leder-Luis, J. (2020). Can whistleblowers root out public expenditure fraud? evidence from medicare. [5]

Liebman, J. B. and N. Mahoney (2017). Do expiring budgets lead to wasteful year-end spending? evidence from federal procurement. *American Economic Review 107*(11), 3510–3549. [5]

Melo-Acosta, G. E., F. Duitama-Munoz, and J. D. Arias-Londoño (2017). Fraud detection in big data using supervised and semi-supervised learning techniques. In *2017 IEEE Colombian conference on communications and computing (COLCOM)*, pp. 1–6. IEEE. [5]

Mookherjee, D. and I. P. Png (1992). Monitoring vis-a-vis investigation in enforcement of law. *The American Economic Review*, 556–565. [3]

Mullainathan, S. and J. Spiess (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106. [9]

Nanduri, J., Y. Jia, A. Oka, J. Beaver, and Y.-W. Liu (2020). Microsoft uses machine learning and optimization to reduce e-commerce fraud. *INFORMS Journal on Applied Analytics 50*(1), 64–79. [5]

Nichols, A. L. and R. J. Zeckhauser (1982). Targeting transfers through restrictions on recipients. *The American Economic Review 72*(2), 372–377. [2, 5]

Pahlka, J. (2023). *Recoding America*. Metropolitan Books. [20]

Polinsky, A. M. and S. Shavell (2000). The economic theory of public enforcement of law. *Journal of economic literature 38*(1), 45–76. [3]

Puri, M., J. Rocholl, and S. Steffen (2017). What do a million observations have to say about loan defaults? opening the black box of relationships. *Journal of Financial Intermediation 31*, 1–15. [9]

Rao, A. R., S. Garai, D. Clarke, and S. Dey (2018). A system for exploring big data: an iterative k-means searchlight for outlier detection on open health data. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. [10]

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*. [16]

Sadineni, P. K. (2020). Detection of fraudulent transactions in credit card using machine learning algorithms. In *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, pp. 659–660. IEEE. [5]

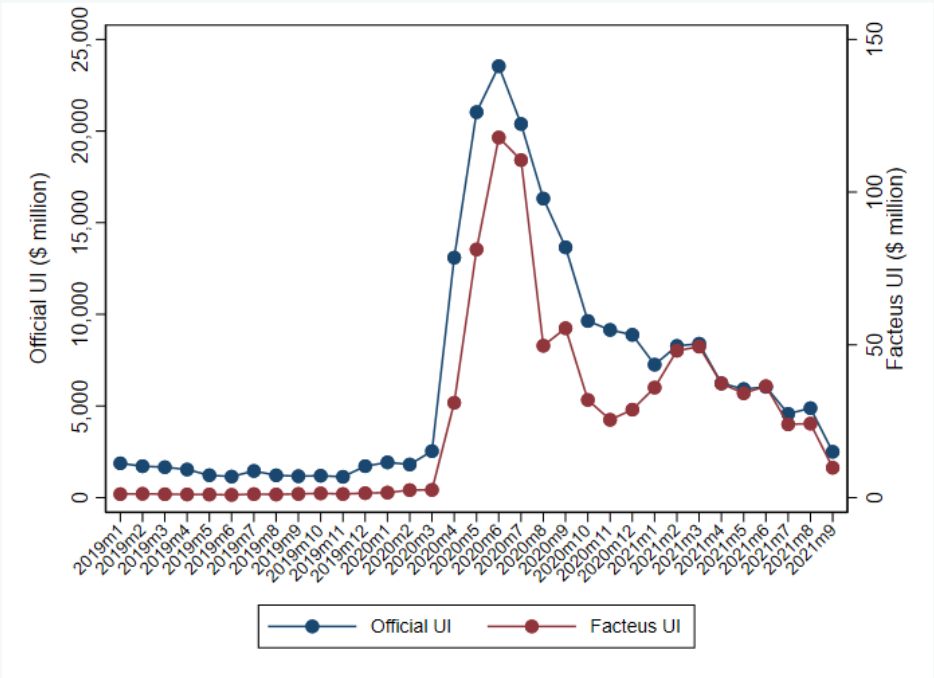
Shekhar, S., J. Leder-Luis, and L. Akoglu (2023). Unsupervised machine learning for explainable health care fraud detection. [5, 9]

United States Government Accountability Office (2023). Unemployment insurance: DOL needs to address substantial pandemic UI fraud and reduce persistent risks. Technical Report GAO-23-106586. [2, 12]

Wu, Y., Y. Xu, and J. Li (2019). Feature construction for fraudulent credit card cash-out detection. *Decision Support Systems 127*, 113155. [3, 9, 46]

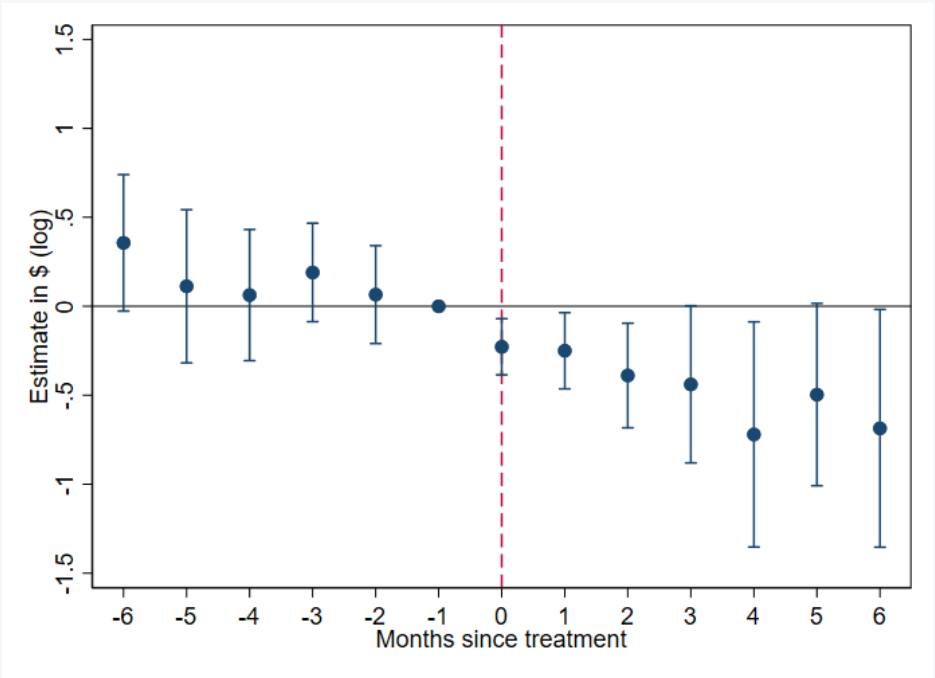
Zhou, Y. and F. Correia (2022). Are we friends? cross-predictability of stock returns. [7]

**Figure 1: Monthly UI Disbursements in Administrative and Facteus Data**

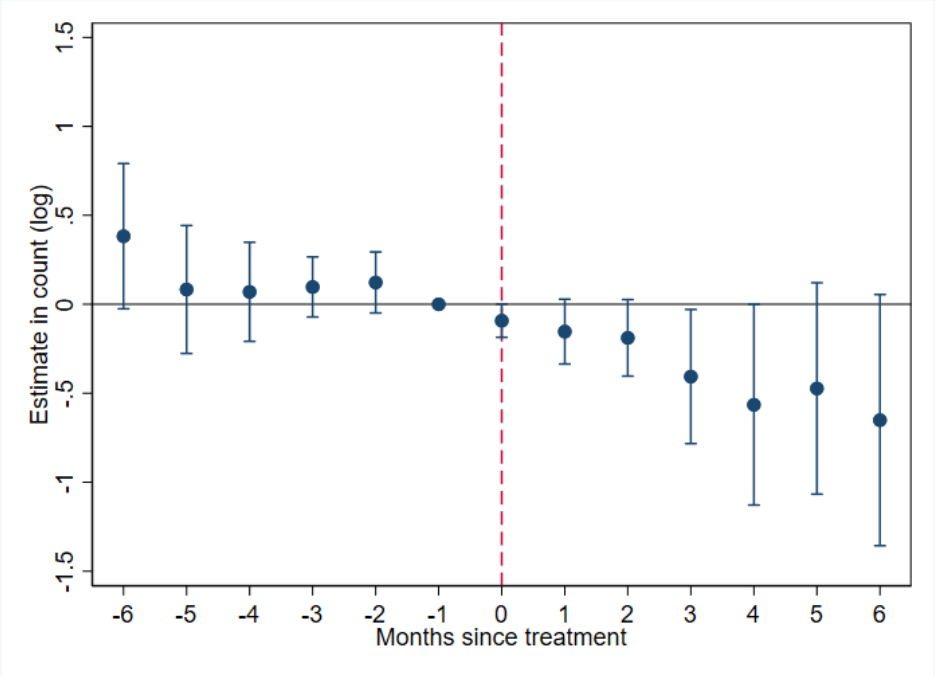



*Notes*: This figure plots total monthly UI disbursements from administrative data (left axis) and from Facteus debit card data (right axis). The sample period runs from January 2019 through September 2021, and comprises 40 states and the District of Columbia. Administrative data are sourced from the US Department of Labor and can be accessed here.

**Figure 2: Impact of Identity Verification on Suspicious Cluster**



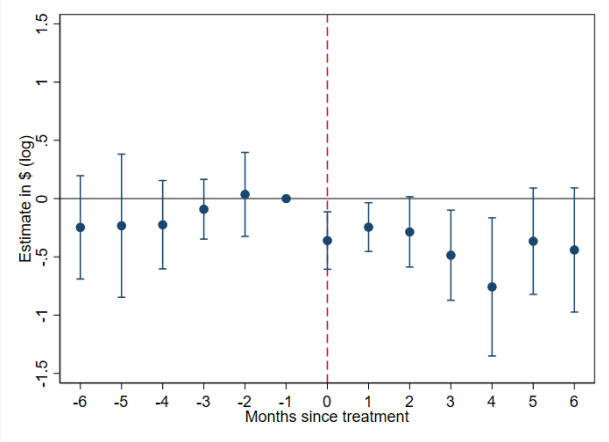

(a) UI dollars



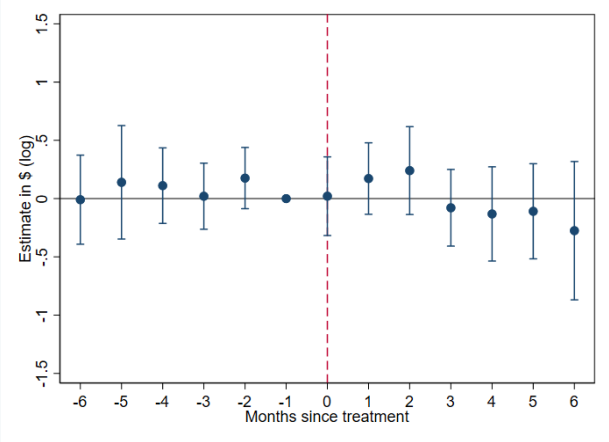

(b) Number of cards

*Notes*: This figure plots an event-study version of the triple difference Equation 2 for the suspicious cluster. Treatment refers to the implementation of identity verification measure at a state level in month 0. The treatment effect is compared to month -1 and to the control cluster. The dependent variable in plot (a) is log(UI) and in plot (b) is the log of number of cards in each cluster.

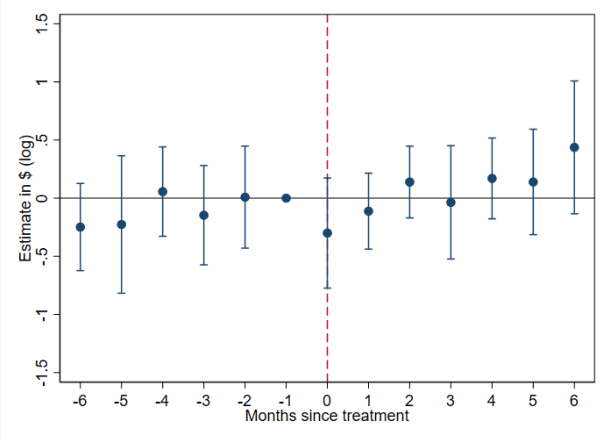
**Figure 3: Within-cluster Impact of Identity Verification on UI dollars**



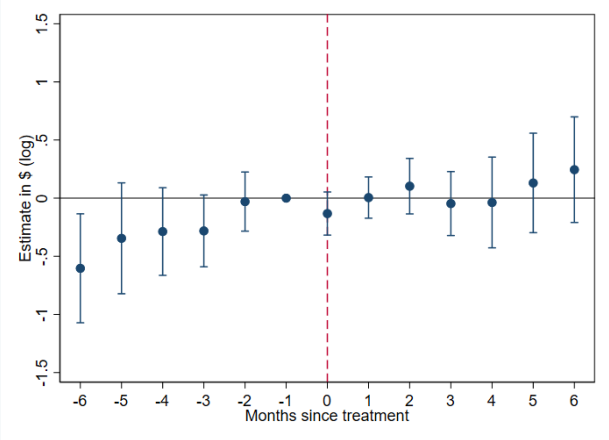

(a) Suspicious




(b) Concurrent Income



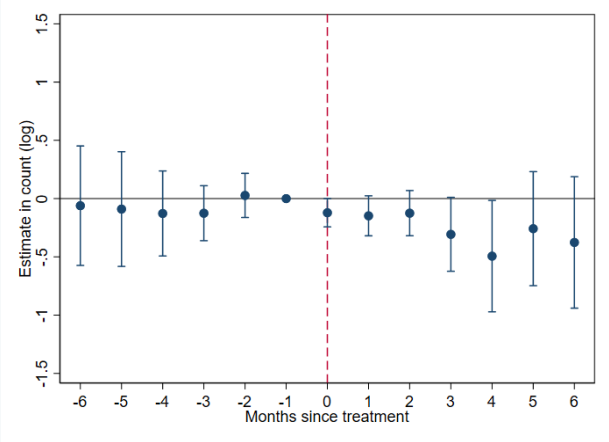

(c) Discretionary Spending



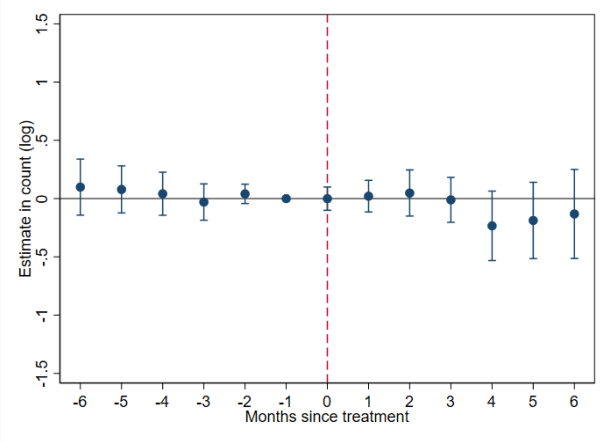

(d) Control

*Notes*: This figure plots the event study version of Equation 3, the difference-in-difference specification, for all four clusters individually. While our clusters reflecting concurrent income, discretionary spending, or control (non-suspicious) spending are unaffected, the suspicious cluster shows an immediate and persistent drop in UI disbursement following identity verification.

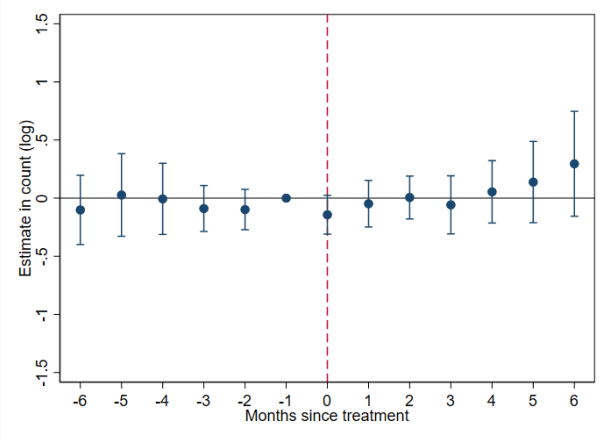
**Figure 4: Within-cluster Impact of Identity Verification on Number of Cards**



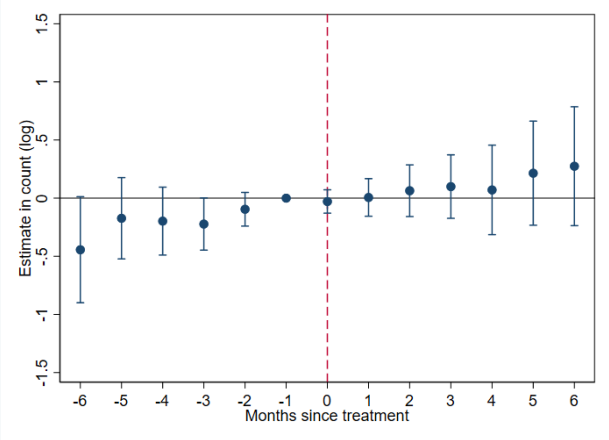

(a) Suspicious




(b) Concurrent Income




(c) Discretionary Spending




(d) Control

*Notes*: This figure plots the event-study version of Equation 3, the difference-in-difference specification, for all four clusters individually. The dependent variable in all plots is the log of number of cards in each cluster to whom these benefits were paid by each state in a month. While our clusters reflecting concurrent income, discretionary spending, or control (non-suspicious) spending are unaffected, the suspicious cluster shows an immediate and persistent drop in the number of cards to whom UI benefits were disbursed following identity verification.

**Table 1: Descriptive Statistics of Debit Card Sample**

| **Panel A: Card-level** | Mean | SD | p25 | p50 | p75 | N |
|---|---|---|---|---|---|---|
| Total UI ($) | 10,104 | 8,957 | 3,254 | 7,512 | 14,093 | 84,865 |
| # UI (count) | 19 | 21 | 4 | 12 | 26 | 84,865 |
| UI per trans ($) | 1,008 | 1,519 | 382 | 575 | 949 | 84,865 |
| UI as % of income | 0.63 | 0.32 | 0.34 | 0.68 | 0.97 | 84,865 |
| Total income ($) | 22,942 | 23,810 | 6,675 | 15,406 | 31,161 | 84,865 |
| Total spend ($) | 14,956 | 15,915 | 3,602 | 10,090 | 21,032 | 84,865 |
| Cash spend ($) | 2,875 | 5,347 | 0 | 343 | 3,594 | 84,865 |
| Wire spend ($) | 371 | 1,794 | 0 | 0 | 0 | 84,865 |
| International wire ($) | 40 | 731 | 0 | 0 | 0 | 84,865 |
| Groceries ($) | 3,176 | 5,464 | 252 | 1,271 | 3,738 | 84,865 |
| Discretionary spend ($) | 2,570 | 3,589 | 287 | 1,275 | 3,469 | 84,865 |
| **Panel B: Card-month level** | Mean | SD | p25 | p50 | p75 | N |
| Total UI ($) | 2,126 | 2,104 | 825 | 1,703 | 2,700 | 403,349 |
| # UI (count) | 4 | 3 | 2 | 4 | 5 | 403,349 |
| UI per trans ($) | 697 | 1,005 | 261 | 451 | 769 | 403,349 |
| UI as % of income | 0.82 | 0.26 | 0.68 | 0.98 | 1.00 | 403,349 |
| Total income ($) | 2,790 | 2,589 | 1,205 | 2,196 | 3,476 | 403,349 |
| Total spend ($) | 1,807 | 1,849 | 563 | 1,374 | 2,468 | 403,349 |
| Cash spend ($) | 449 | 945 | 0 | 0 | 524 | 403,349 |
| Wire spend ($) | 52 | 310 | 0 | 0 | 0 | 403,349 |
| International wire ($) | 4 | 92 | 0 | 0 | 0 | 403,349 |
| Groceries ($) | 341 | 641 | 11 | 133 | 416 | 403,349 |
| Discretionary spend ($) | 288 | 413 | 40 | 166 | 391 | 403,349 |

*Notes*: This table shows card-level (panel A) and card-month level (panel B) income and spending statistics from our sample of transaction-level debit card data conditional on receiving at least $1,000 of total unemployment insurance (UI) during the sample period from January 2019 to September 2021.

**Table 2: Summary Statistics by Cluster**

| | Suspicious (Fast Cash) | Suspicious (Slow Cash) | Concurrent Income | Discretionary Spending | Control |
|---|---|---|---|---|---|
| Unemployment Insurance ($/month) | 3,122 | 3,015 | 2,132 | 2,618 | 1,200 |
| | (2,412) | (1,823) | (458) | (1,473) | (1,527) |
| Longest UI spell (months) | 4.7 | 4.5 | 10.9 | 5.0 | 4.2 |
| | (3.7) | (3.7) | (4.1) | (4.1) | (3.5) |
| Cash withdrawal (share of spend) | 0.66 | 0.61 | 0.09 | 0.02 | 0.13 |
| | (0.18) | (0.19) | (0.11) | (0.05) | (0.21) |
| International wire (share of spend) | 0.002 | 0.001 | 0.003 | 0.0004 | 0.002 |
| | (0.02) | (0.02) | (0.03) | (0.01) | (0.02) |
| Non-essentials (share of spend) | 0.07 | 0.07 | 0.19 | 0.52 | 0.22 |
| | (0.06) | (0.06) | (0.09) | (0.13) | (0.15) |
| Time to cash (hours) | 14 | 78 | 412 | 1,371 | 322 |
| | (6) | (173) | (754) | (921) | (738) |
| Concurrent non-UI income (months) | 0.2 | 0.3 | 4.7 | 0.5 | 0.6 |
| | (0.8) | (1) | (2.1) | (1.2) | (1.2) |
| Stimulus checks ($, thousands) | 0.2 | 0.2 | 2.1 | 0.2 | 0.6 |
| | (0.9) | (0.9) | (2.6) | (0.8) | (1.5) |
| Tax refunds ($, thousands) | 0.2 | 0.2 | 1.4 | 0.2 | 0.8 |
| | (0.7) | (1) | (2.7) | (0.8) | (2.3) |

*Notes*: This table shows the mean and standard deviations of card features at a cluster level. Clusters were created using an unsupervised K-means algorithm. Standard deviations are reported in parentheses below the means.

**Table 3: Treatment Effect of Identity Verification**

| | log(UI) | log(number of cards) |
|---|---|---|
| | (1) | (2) |
| Suspicious × Post | -0.520** | -0.441** |
| | (0.240) | (0.217) |
| Concurrent Income × Post | -0.309 | -0.331 |
| | (0.205) | (0.211) |
| Discretionary Spending × Post | -0.045 | -0.137 |
| | (0.137) | (0.176) |
| Observations | 2,271 | 2,268 |
| Adj. $R^2$ | 0.91 | 0.95 |
| Cluster, State, Month, Post, Cluster-State, Cluster-Month FE | Y | Y |

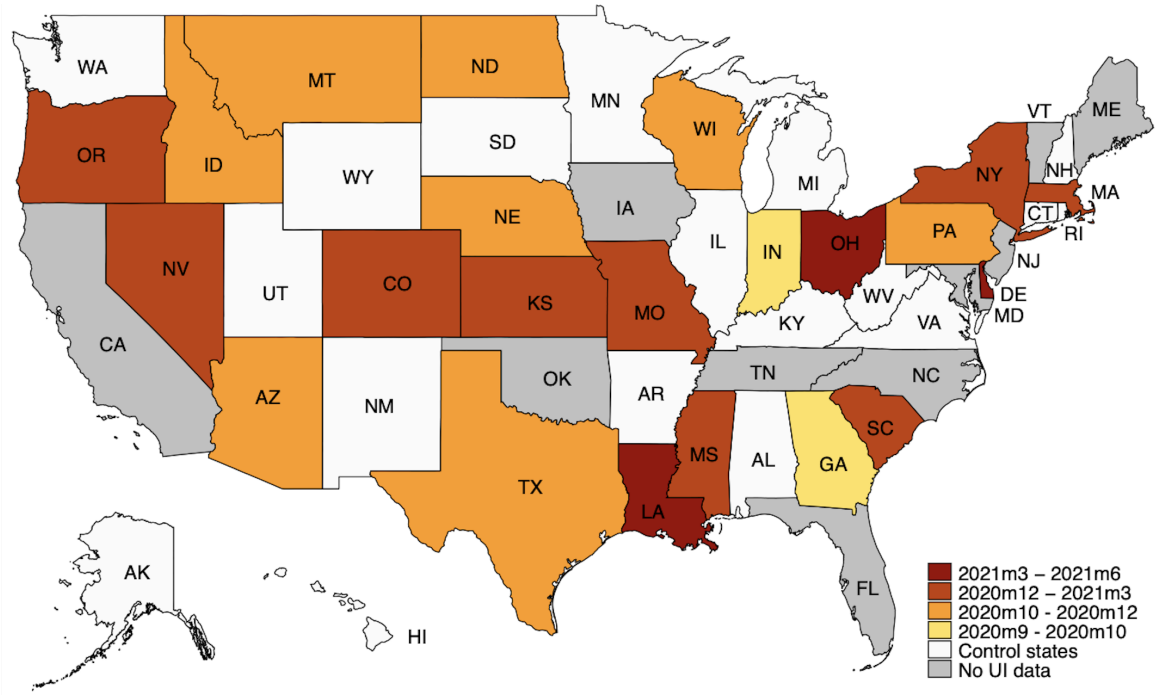
*Notes*: This table reports the results from a triple difference estimation from Equation 1 of the impact of identity verification measures on the UI received by various card clusters. In column (1), the dependent variable is log(UI), the log of unemployment benefits paid by each state to a cluster in a month, and in column (2) it is the log of number of cards in each cluster to whom these benefits were paid. The control group is the cluster with non-suspicious cards. The post period includes the six months after this policy was introduced for the respective state. Standard errors are clustered by state, and reported in parentheses. Identity verification reduces UI disbursed to the Suspicious cluster. Figure 2 shows this table's event-study version for the Suspicious cluster and Figure A6 in Appendix C for other clusters. $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

# Appendix for "Unemployment Insurance Fraud in the Debit Card Market"

Umang Khetan Jetson Leder-Luis Jialan Wang Yunrong Zhou

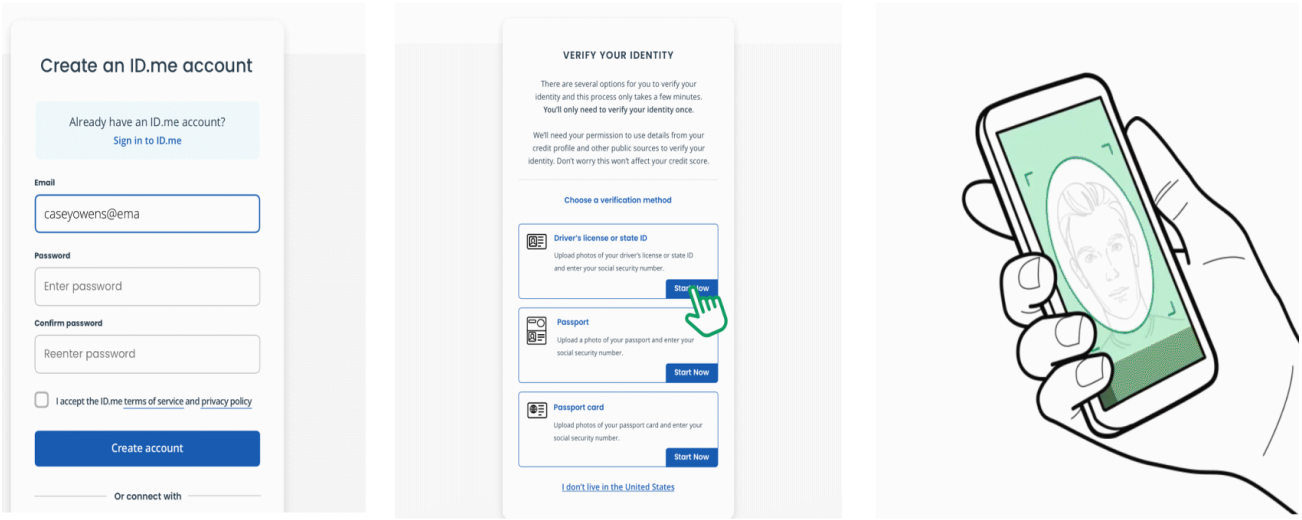
October 2023

**Figure A1: Treatment Month for Each State**




*Notes*: This figure shows treated states, control states, and the states for which we cannot identify UI or the timing of ID verification. This figure also shows the timing of the adoption of identity verification by each treated state.

## Figure A2: Identity Verification Process




*Notes*: This figure shows three general steps involved in verifying an applicant's identity when claiming UI benefits. The figure has been adapted from ID.me's website.

**Table A1: Timing of Identity Verification Adoption by State**

| Treated state | Treatment | Agency/vendor | Data source |
|---|---|---|---|
| Arizona (AZ) | Oct-20 | ID.me | FOIA |
| Colorado (CO) | Jan-21 | ID.me | FOIA |
| Delaware (DE) | Jun-21 | ID.me | FOIA |
| Georgia (GA) | Sep-20 | ID.me | Congress/Public |
| Idaho (ID) | Dec-20 | ID.me | FOIA |
| Indiana (IN) | Sep-20 | ID.me | FOIA/Congress |
| Kansas (KS) | Feb-21 | LexisNexis | FOIA |
| Louisiana (LA) | May-21 | ID.me | Congress/Public |
| Massachusetts (MA) | Mar-21 | ID.me | FOIA |
| Mississippi (MS) | Mar-21 | ID.me | Congress/Public |
| Missouri (MO) | Mar-21 | ID.me | FOIA |
| Montana (MT) | Nov-20 | ID.me | Congress |
| Nebraska (NE) | Dec-20 | GIACT | FOIA |
| Nevada (NV) | Mar-21 | ID.me | FOIA |
| New York (NY) | Feb-21 | ID.me | Congress |
| North Dakota (ND) | Dec-20 | ID.me | Congress |
| Ohio (OH) | Apr-21 | LexisNexis | FOIA/Public |
| Oregon (OR) | Mar-21 | ID.me | Congress/Public |
| Pennsylvania (PA) | Oct-20 | ID.me | FOIA |
| South Carolina (SC) | Mar-21 | ID.me | Congress/Public |
| Texas (TX) | Nov-20 | ID.me | FOIA |
| Wisconsin (WI) | Nov-20 | Google Analytics | FOIA/Public |
| Untreated states | | | |
| Alabama (AL) | Hawaii (HI) | New Hampshire (NH) | Virginia (VA) |
| Alaska (AK) | Illinois (IL) | New Mexico (NM) | Washington (WA) |
| Arkansas (AR) | Kentucky (KY) | Rhode Island (RI) | West Virginia (WV) |
| Connecticut (CT) | Michigan (MI) | South Dakota (SD) | Wyoming (WY) |
| District of Columbia (DC) | Minnesota (MN) | Utah (UT) | |

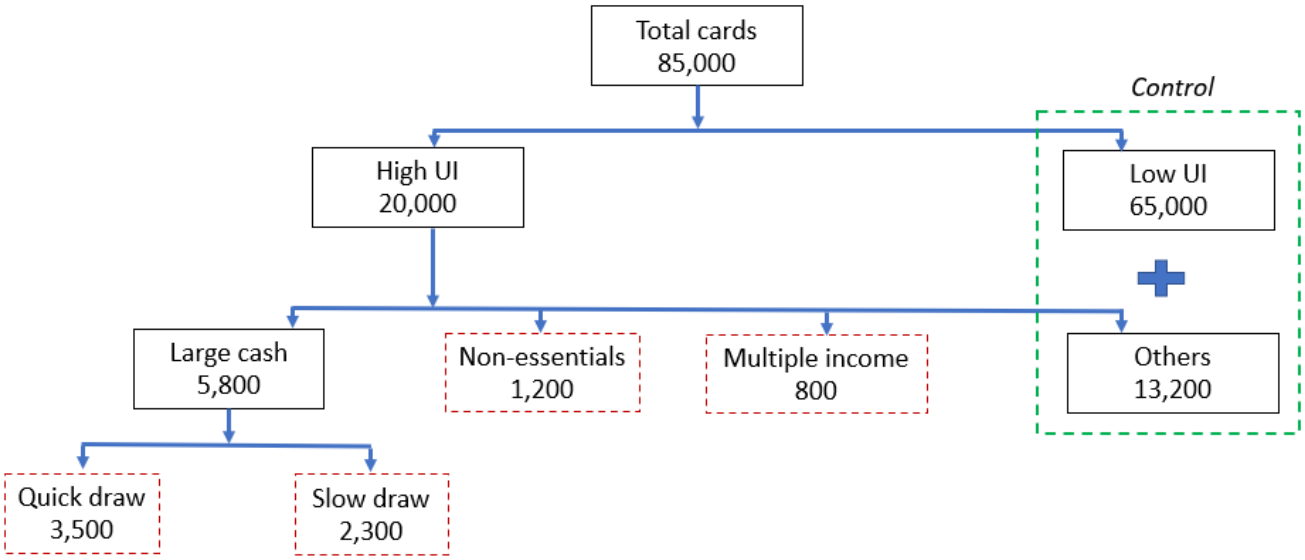
*Notes*: This table lists states for which we can identify both UI disbursements in the Facteus data and the timing of ID verification adoption based on FOIA, publicly-available information, and congressional records. For states that implemented multiple programs, we use the earliest date of implementation.

**Table A2: Disaggregation of UI Data by State**

| | Correlation | Share in % | | Correlation | Share in % |
|---|---|---|---|---|---|
| Federal (41 states) | 0.95 | 0.40 | Mississippi (MS ) | 0.75 | 0.02 |
| Alaska (AK) | 0.83 | 0.37 | Montana (MT) | 0.89 | 1.17 |
| Alabama (AL) | 0.78 | 1.15 | North Dakota (ND) | 0.86 | 0.45 |
| Arkansas (AR) | 0.83 | 1.04 | Nebraska (NE) | 0.82 | 0.60 |
| Arizona (AZ) | 0.74 | 0.08 | New Hampshire (NH) | 0.93 | 0.26 |
| Colorado (CO) | 0.91 | 0.30 | New Mexico (NM) | 0.83 | 0.14 |
| Connecticut (CT) | 0.89 | 0.45 | Nevada (NV) | 0.72 | 0.20 |
| District of Columbia (DC) | 0.55 | 0.18 | New York (NY) | 0.87 | 0.22 |
| Delaware (DE) | 0.91 | 0.33 | Ohio (OH) | 0.83 | 1.10 |
| Georgia (GA) | 0.73 | 0.05 | Oregon (OR) | 0.56 | 0.14 |
| Hawaii (HI) | 0.80 | 0.22 | Pennsylvania (PA) | 0.91 | 0.16 |
| Idaho (ID) | 0.81 | 0.57 | Rhode Island (RI) | 0.86 | 0.59 |
| Illinois (IL) | 0.84 | 0.44 | South Carolina (SC) | 0.88 | 1.21 |
| Indiana (IN) | 0.88 | 1.34 | South Dakota (SD) | 0.92 | 0.47 |
| Kansas (KS) | 0.87 | 0.85 | Texas (TX) | 0.92 | 0.31 |
| Kentucky (KY) | 0.87 | 0.92 | Utah (UT) | 0.89 | 0.36 |
| Louisiana (LA) | 0.89 | 0.84 | Virginia (VA) | 0.95 | 0.40 |
| Massachusetts (MA) | 0.93 | 0.23 | Washington (WA) | 0.89 | 0.27 |
| Michigan (MI) | 0.88 | 0.71 | Wisconsin (WI) | 0.80 | 0.84 |
| Minnesota (MN) | 0.86 | 0.66 | West Virginia (WV) | 0.93 | 0.51 |
| Missouri (MO) | 0.91 | 0.60 | Wyoming (WY) | 0.93 | 0.37 |

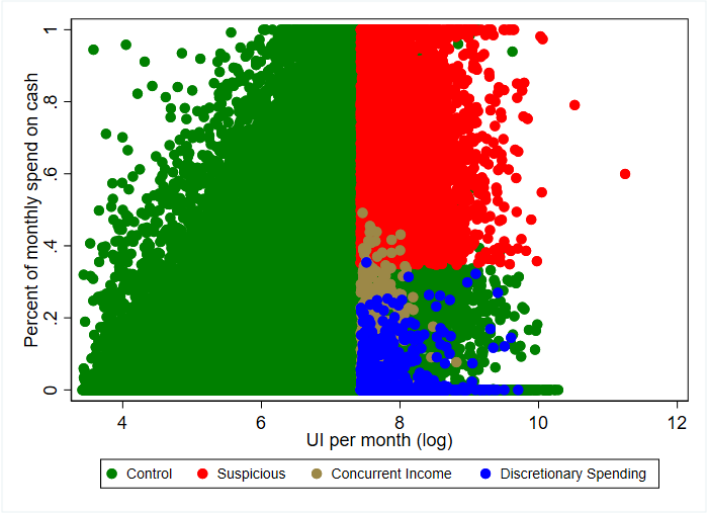
*Notes*: This table characterizes our UI data by state, showing the correlation of our time series data with that state's official UI data, as well as the estimated share of UI data we can observe from that state.

**Figure A3: Summary of K-Means Clustering**

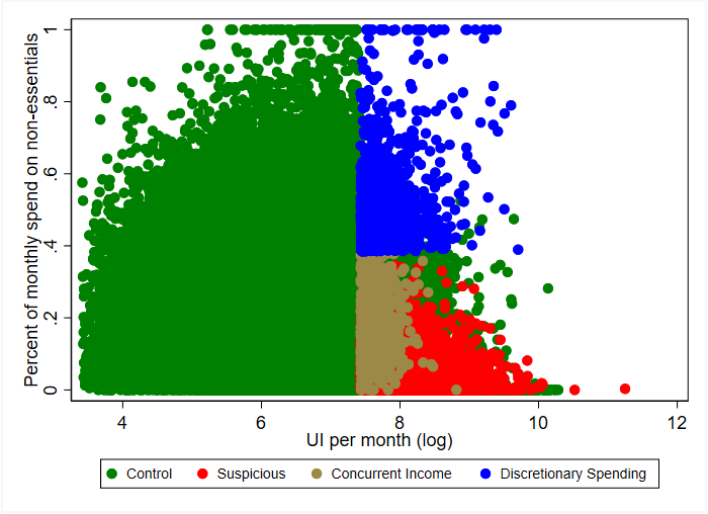



*Notes*: This figure depicts the order in which cards are grouped into clusters representing varying levels of fraudulent behavior. The top branch splits about 85,000 cards into those with high and low UI using monthly UI received by each card and the longest spell of being on UI rolls. Within cards receiving high levels of UI, the algorithm uses spending on cash withdrawals, non-essential categories such as alcohol and gambling, and concurrent non-UI income to create three clusters. Cards with large cash spending are further split based on the speed of withdrawal after the receipt of UI. Cards with low UI or high UI but without outsized spending in the above-mentioned categories are collected under the control cluster. The labels show the approximate number of cards in each split.

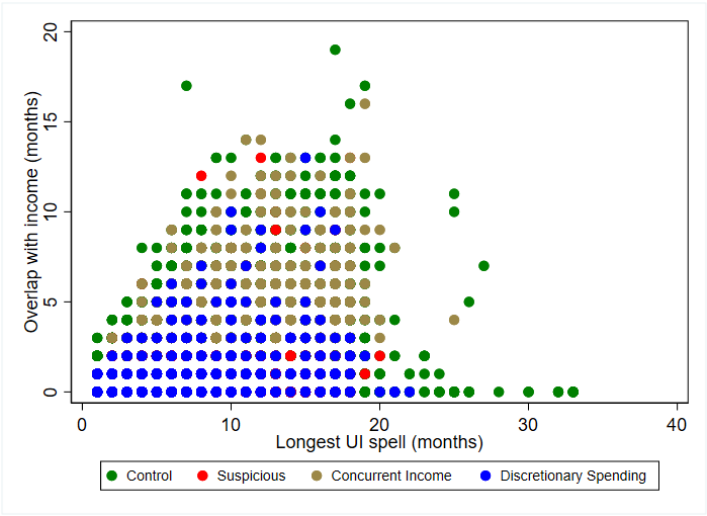
## Figure A4: Bivariate Plots of K-Means Clusters



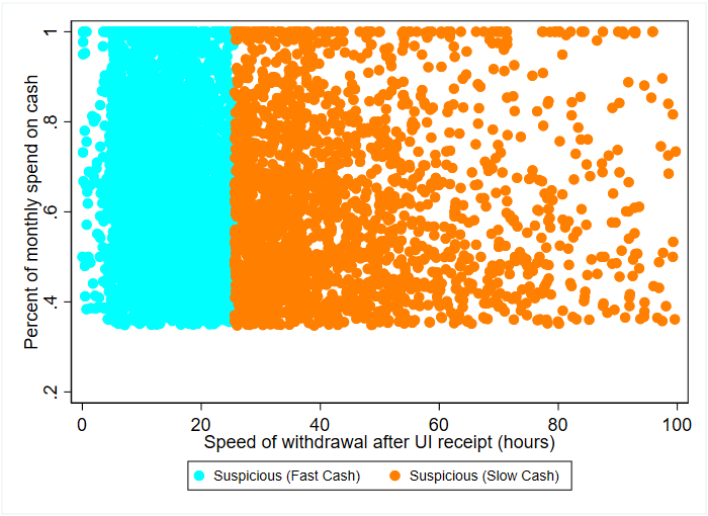

(a) Cash withdrawals (%) v/s UI per month (log)




(b) Non-essential spend (%) v/s UI per month (log)




(c) Other income overlap v/s longest UI spell




(d) Cash withdrawals (%) v/s Time to cash (hours)

*Notes*: This figure plots each card and the cluster it belongs to using two dimensions at a time.

**Table A3: UI $ and % Share to Clusters Before and After the Onset of COVID-19 Pandemic**

| | UI received ($, million) | | Share of UI received (%) | |
|---|---|---|---|---|
| Category | Jan 2019-Feb 2020 | Mar 2020-Sep 2021 | Jan 2019-Feb 2020 | Mar 2020-Sep 2021 |
| Suspicious (Fast Cash) | 0.01 | 50.62 | 0.08 | 6.13 |
| Suspicious (Slow Cash) | 0.01 | 33.39 | 0.08 | 4.04 |
| Concurrent Income | 0.20 | 22.26 | 1.19 | 2.69 |
| Discretionary Spending | 0.11 | 19.38 | 0.64 | 2.35 |
| Control | 16.2 | 700.55 | 98.00 | 84.79 |

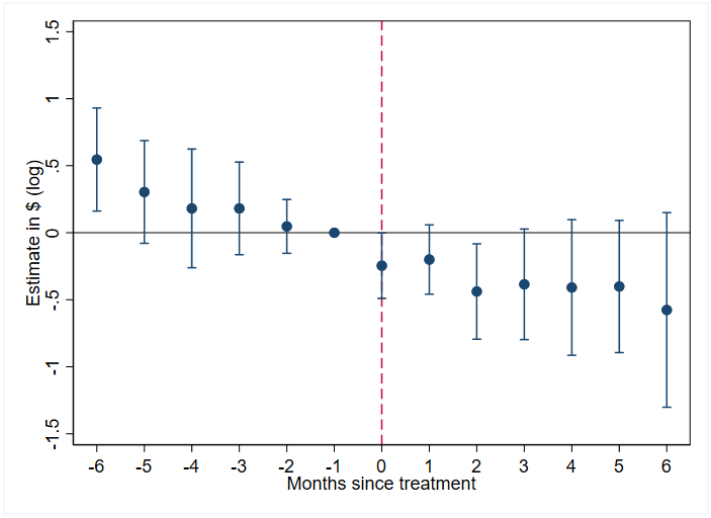
*Notes*: This table describes UI inflows by cluster, including the total dollars and share of total dollars paid to each cluster before and after the onset of COVID-19 pandemic.

**Table A4: UI $ and % Share to Clusters Under PUA and Non-PUA Programs**

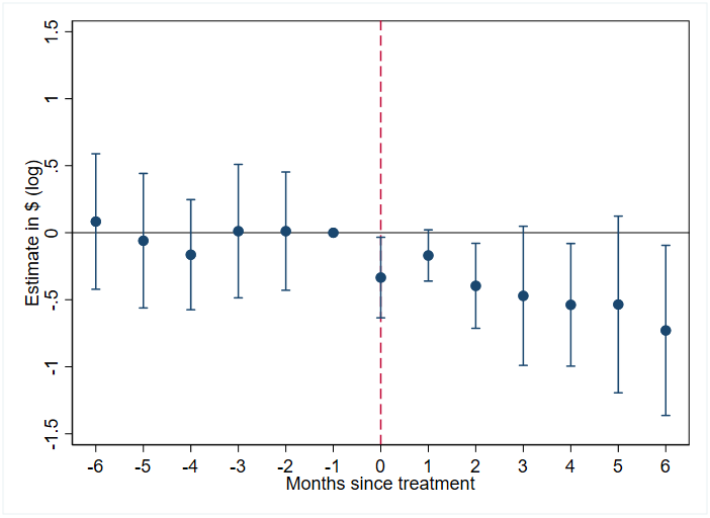
| | PUA | | Non-PUA | |
|---|---|---|---|---|
| Category | $ million | % of total | $ million | % of total |
| Suspicious (Fast Cash) | 11.78 | 13.4% | 1.68 | 4.1% |
| Suspicious (Slow Cash) | 6.27 | 7.2% | 0.58 | 1.4% |
| Concurrent Income | 1.02 | 1.2% | 1.02 | 2.5% |
| Discretionary Spending | 1.77 | 2.0% | 0.57 | 1.4% |
| Control | 66.79 | 76.2% | 37.54 | 90.7% |

*Notes*: This table reports the cluster-level dollars and share of total dollars paid under the PUA and non-PUA programs in our data for four states where we can separate the two flows: Arkansas, Massachusetts, Ohio, and West Virginia. The sample period for this comparison runs from April 2020 through May 2021, when both programs were fully active.

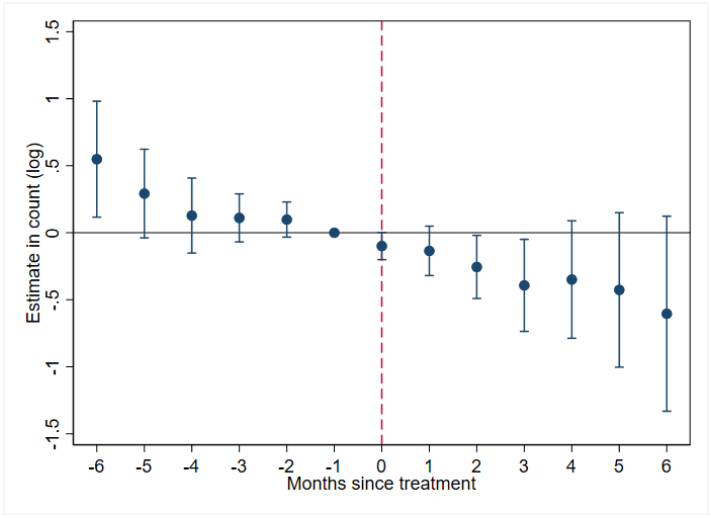
**Figure A5: Event-study for Suspicious (Fast Cash) and Suspicious (Slow Cash) Clusters**



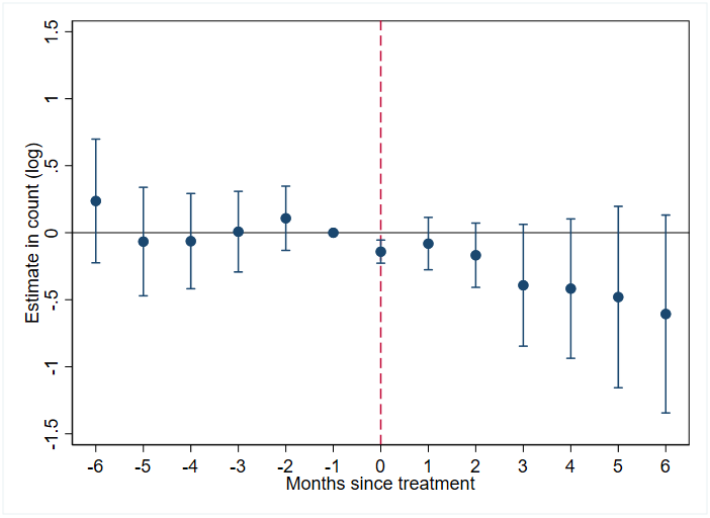

(a) Suspicious (Fast Cash): UI dollars




(b) Suspicious (Slow Cash): UI dollars



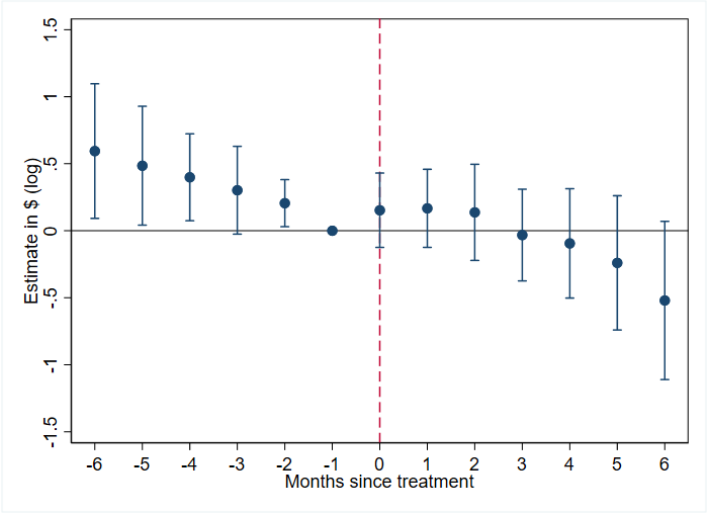

(c) Suspicious (Fast Cash): number of cards



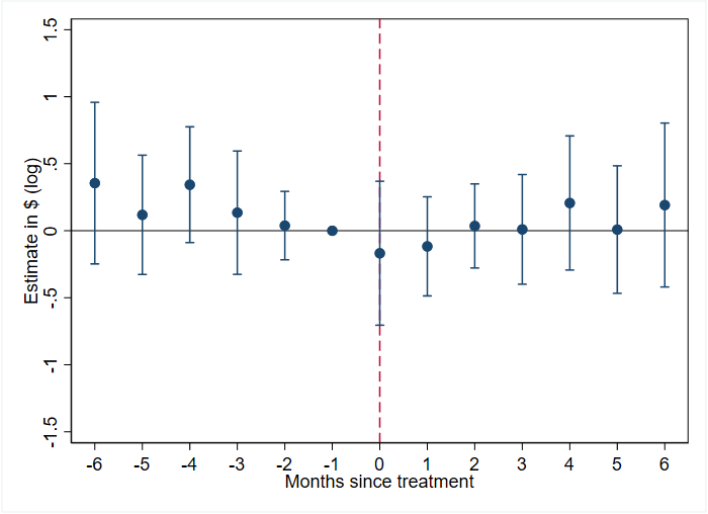

(d) Suspicious (Slow Cash): number of cards

*Notes*: This figure reports the results of Equation 2 for Suspicious (Fast Cash) and Suspicious (Slow Cash) clusters separately. The dependent variable in plots (a) and (b) is log(UI), the log of unemployment benefits paid by each state to a cluster in a month. The dependent variable in plots (c) and (d) is the log of number of cards in each cluster to whom these benefits were paid by each state in a month.

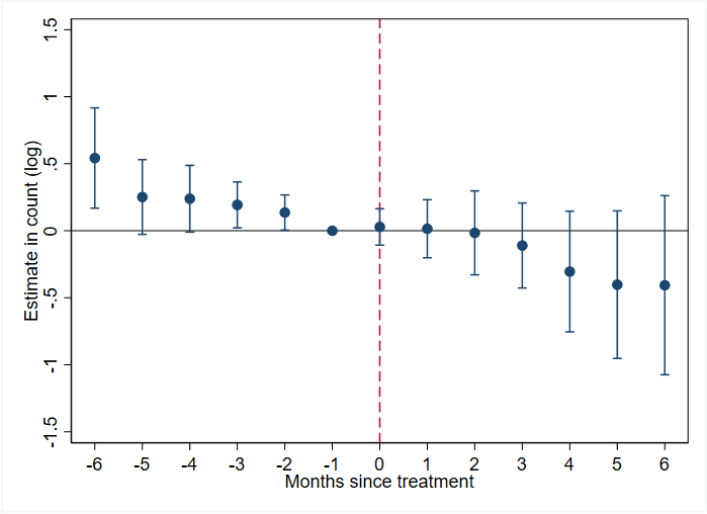
**Figure A6: Event-study for Concurrent Income and Discretionary Spending Clusters**



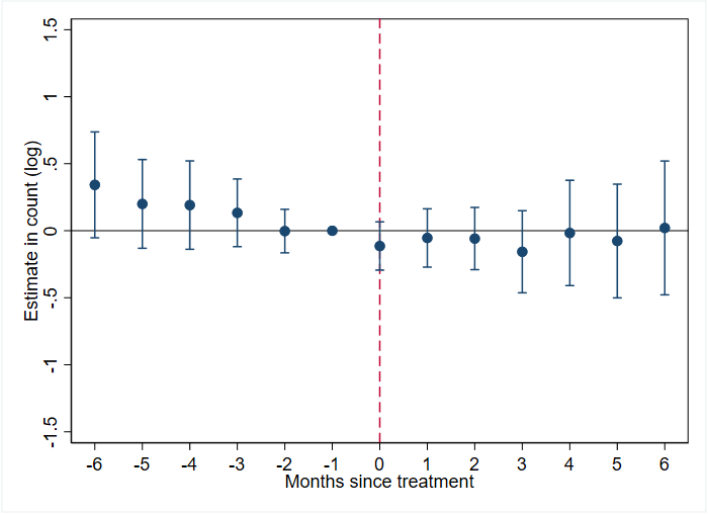

(a) Concurrent Income: UI dollars




(b) Discretionary Spending: UI dollars



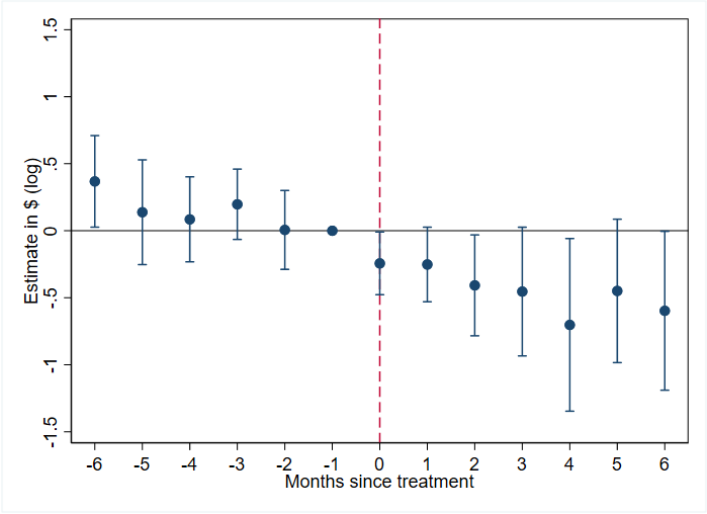

(c) Concurrent Income: number of cards



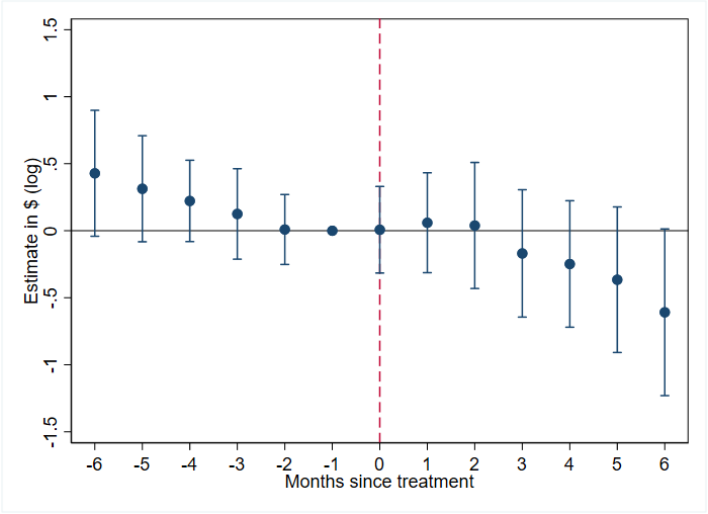

(d) Discretionary Spending: number of cards

*Notes*: This figure reports the results of Equation 2 for Concurrent Income and Discretionary Spending clusters. The dependent variable in plots (a) and (b) is log(UI), the log of unemployment benefits paid by each state to a cluster in a month. The dependent variable in plots (c) and (d) is the log of number of cards in each cluster to whom these benefits were paid by each state in a month.

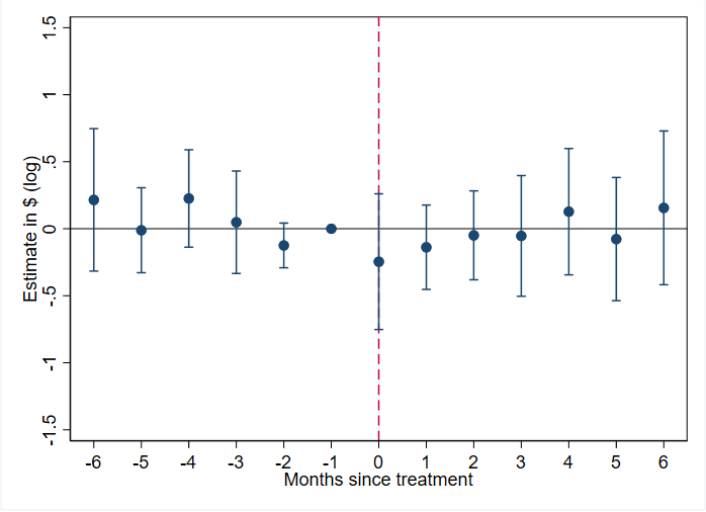
**Figure A7: Event-study using Stacked Difference-in-differences**




(a) Suspicious



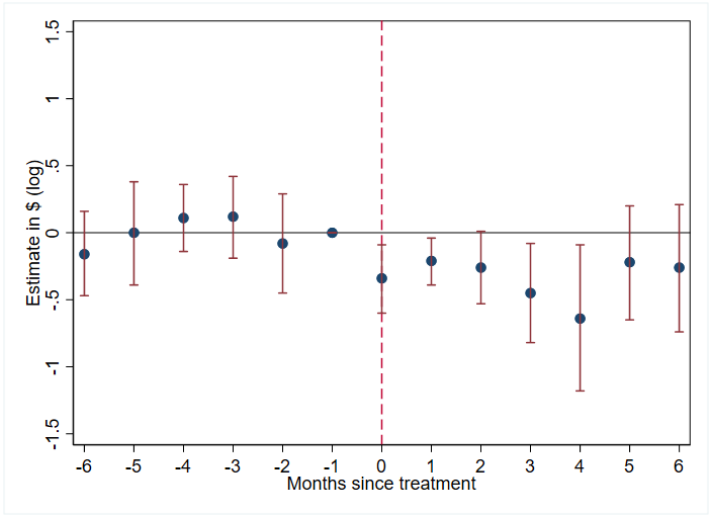

(b) Concurrent Income



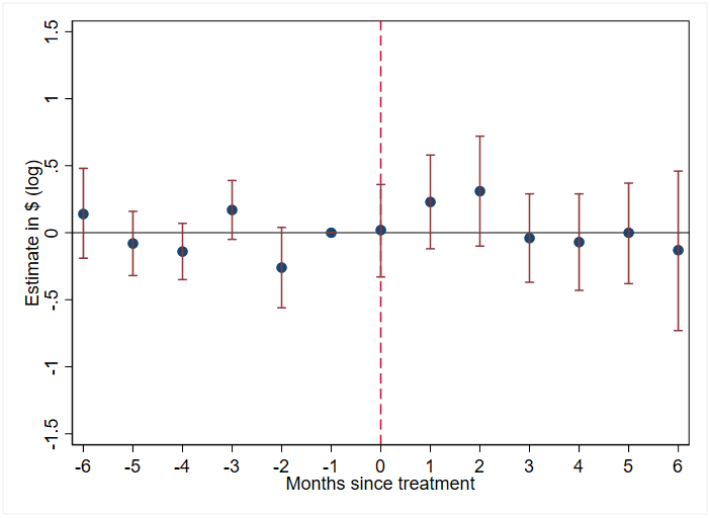

(c) Discretionary Spending

*Notes*: This figure plots the event-study of the form in Equation 2 estimated using the stacked difference-in-differences method as in Cengiz et al. (2019). The dependent variable in each plot is log(UI), the log of unemployment benefits paid by each state to a cluster in a month.

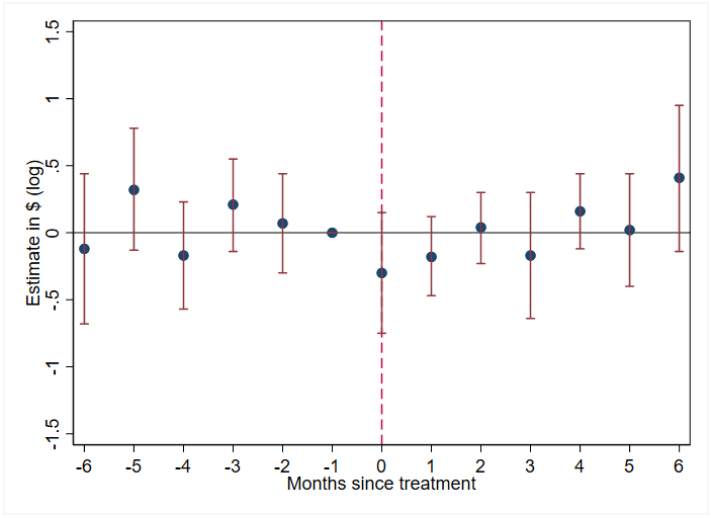
**Figure A8: Within-cluster Impact of Identity Verification - Alternative Estimation**



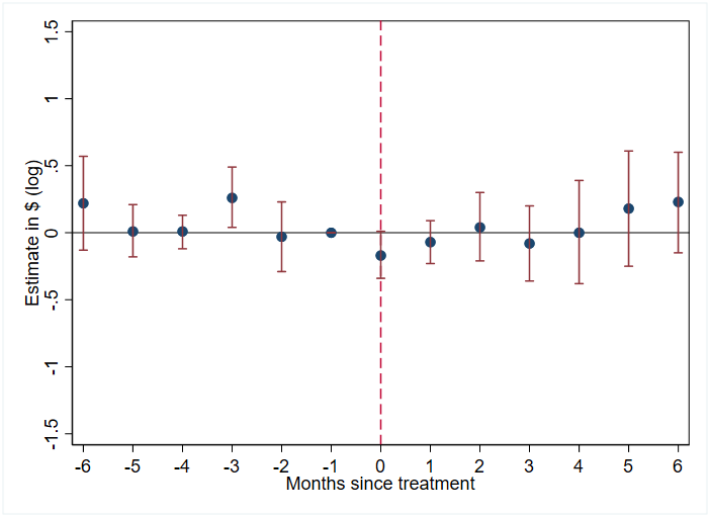

(a) Suspicious




(b) Concurrent Income




(c) Discretionary Spending




(d) Control

*Notes*: This figure plots the estimation of event studies of the form in Equation 3 based on the estimation technique in Callaway and Sant'Anna (2021). The dependent variable in each plot is log(UI), the log of unemployment benefits paid by each state to a cluster in a month.

**Table A5: Treatment Effects of Identity Verification - Stacked Difference-in-differences**

| | log(UI) | log(number of cards) |
|---|---|---|
| | (1) | (2) |
| Suspicious × Post | -0.527** | -0.460* |
| | (0.244) | (0.238) |
| Concurrent Income × Post | -0.323 | -0.350 |
| | (0.226) | (0.236) |
| Discretionary Spending × Post | -0.020 | -0.119 |
| | (0.146) | (0.187) |
| Observations | 9,739 | 9,727 |
| Adj. $R^2$ | 0.92 | 0.94 |
| Cluster, State, Month, Post, Cluster-State, Cluster-Month FE | Y | Y |

*Notes*: This table reports the results from a triple difference-in-differences estimation of the form in Equation 1 using the estimation method in Cengiz et al. (2019), where stacked datasets are created for each treatment date. In column (1), the dependent variable is log(UI), the log of unemployment benefits paid by each state to a cluster in a month, and in column (2) it is the log of number of cards in each cluster to whom these benefits were paid. The control group is the cluster with non-suspicious cards. Standard errors clustered by state are reported in parentheses. Figure A7 shows this table's event-study version. $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

## D. Innovative Household Finance Data used in this paper

This paper makes use of transaction-level data covering inflows and spending of about 85,000 debit cards across the United States. Our data come from Facteus, a fintech firm that partners with banks, card issuers, and payment processors to aggregate, standardize, and anonymize transaction-level information from debit cards. The data on each transaction include the amount, transaction date, and description. Further details such as merchant category code (MCC) and card-holder ZIP code help us confirm that the data represent a wide variety of consumer transactions and a geographically representative sample. We observe 35 million transactions for cards that each received at least $1,000 of unemployment insurance cumulatively between January 2019 and September 2021.

These data are particularly suited to study fraud in unemployment insurance for at least three reasons. First, we observe both inflows into and spending out of these cards. We identify inflows representing receipt of insurance benefits using a string matching technique that locates standard ACH descriptions that states use when remitting benefits. This provides us with the dependent variable for testing the impact of anti-fraud policy on benefits disbursed by states. Further, we leverage the spending data to group cards into clusters representing varying likelihood of committing fraud. Our data set includes spending dimensions that are informative about the propensity to commit fraud, such as the quantity and speed of cash withdrawal from ATMs (Wu et al., 2019).

Second, debit card holders are a particularly interesting and poorly-understood population to study in the context of public benefits. Most states allow claimants to receive benefits on state-issued prepaid debit card or pre-existing prepaid card via direct deposit. Consumers who receive government benefits via prepaid cards tend to be lower income and more likely to lack traditional bank accounts. Furthermore, prepaid cards may be more susceptible to fraud than bank accounts, as criminals and fraudulent recipients can quickly open one or more card accounts for the express purposes of obtaining and siphoning funds, while the banking system has much more stringent anti-fraud measures.

Finally, we estimate that about a third of unemployment insurance is paid into debit cards. Hence, debit cards are a particularly relevant context for policies aimed at the economically vulnerable, and also less likely to be observed through traditional data sources originating from banks and surveys. Our paper therefore contributes to an understanding of the behaviors exhibited by a large and under-studied population.

Unlike corporate or institutional fraud, mis-allocation of funds in consumer markets is particularly hard to evaluate due to limited data on consumer behavior. Our granular and multi-dimensional data overcome this hurdle by enabling us to apply machine learning techniques to extract signals of potential fraud, with minimal researcher subjectivity or arbitrarily determined cut-offs. With these innovative data, we make progress on answering important questions surrounding the mis-allocation of costly public funds and the effectiveness of anti-fraud measures.