

Did FinTech Lenders Facilitate PPP Fraud?^{*}

John M. Griffin[†]

Samuel Kruger[‡]

Prateek Mahajan[§]

December 6, 2021

Abstract

In the distribution of the Paycheck Protection Program’s (PPP) \$803 billion in funds, FinTech lenders began minimally but ramped up their market share to over 80% of originated loans by the end of the program. We examine metrics related to potential misreporting including non-registered businesses, multiple businesses at residential addresses, abnormally high implied compensation per employee, and large inconsistencies in jobs reported with another government program. We assess these four metrics with five supplemental measures and extensive supporting analysis. Suspicious loans exhibit sharp and discontinuous increases in misreporting at maximum loan thresholds and round loan amounts with discontinuities more pronounced among FinTechs. FinTech loans are more than 3.5 times as likely to be initiated by someone with a felony record, strongly cluster in industry-county pairs to a degree that is infeasible based on U.S. Census data, and frequently exhibit similar loan features within lender-county pairs. Differences across lenders are persistent with certain FinTech lenders seemingly specializing in questionable loans. Few of these loans have been prosecuted by authorities or repaid. FinTech lenders in round three rapidly increased both their market share and the fraction of their loans with potential misreporting, particularly in zip codes with the highest levels of questionable lending in earlier rounds. From the first round of the program in April 2020 to the last month of the program in May 2021, the amount of potential misreporting increased more than four-fold. While FinTech lenders likely expand PPP access, this may come at the cost of facilitating fraudulent credit.

JEL classification: G21, G23, G28, H12

keywords: FinTech, Paycheck Protection Program (PPP), Misreporting, Fraud.

^{*}We are thankful for comments from seminar and conference participants at the Federal Reserve Bank of Atlanta, Hong Kong Baptist University, the Lone Star Finance Conference, the Office of the Comptroller of the Currency, the Texas Finance Festival, the U.S. Government Accountability Office, The University of Texas at Austin, and the UT Dallas Fall Finance Conference. We also thank Andres Almazan, Bruce Carlin, Jonathan Cohn, William Fuchs, Benjamin Herbert, Zack Liu, Alex Priest, Alberto Rossi, Laura Starks, Brian Wolfe, Yao Zeng, Harold Zhang, Capital Plus, Blue Acorn, Itria Ventures, and many others for helpful comments. We thank Integra Research Group for research support and the University of Texas McCombs Research Excellence Grant. Griffin is an owner of Integra Research Group, Integra FEC, and Integra REC, which engage in research, financial consulting, and recovery on a variety of issues related to investigation of financial fraud.

[†]McCombs School of Business, University of Texas at Austin. Email: John.Griffin@utexas.edu.

[‡]McCombs School of Business, University of Texas at Austin. Email: Sam.Kruger@mcombs.utexas.edu.

[§]McCombs School of Business, University of Texas at Austin. Email: Prateek.Mahajan@mcombs.utexas.edu.

Did FinTech Lenders Facilitate PPP Fraud?

December 6, 2021

Abstract

In the distribution of the Paycheck Protection Program’s (PPP) \$803 billion in funds, FinTech lenders began minimally but ramped up their market share to over 80% of originated loans by the end of the program. We examine metrics related to potential misreporting including non-registered businesses, multiple businesses at residential addresses, abnormally high implied compensation per employee, and large inconsistencies in jobs reported with another government program. We assess these four metrics with five supplemental measures and extensive supporting analysis. Suspicious loans exhibit sharp and discontinuous increases in misreporting at maximum loan thresholds and round loan amounts with discontinuities more pronounced among FinTechs. FinTech loans are more than 3.5 times as likely to be initiated by someone with a felony record, strongly cluster in industry-county pairs to a degree that is infeasible based on U.S. Census data, and frequently exhibit similar loan features within lender-county pairs. Differences across lenders are persistent with certain FinTech lenders seemingly specializing in questionable loans. Few of these loans have been prosecuted by authorities or repaid. FinTech lenders in round three rapidly increased both their market share and the fraction of their loans with potential misreporting, particularly in zip codes with the highest levels of questionable lending in earlier rounds. From the first round of the program in April 2020 to the last month of the program in May 2021, the amount of potential misreporting increased more than four-fold. While FinTech lenders likely expand PPP access, this may come at the cost of facilitating fraudulent credit.

JEL classification: G21, G23, G28, H12

keywords: FinTech, Paycheck Protection Program (PPP), Misreporting, Fraud.

The melding of financial technology and banking, also known as FinTech lending, has emerged at a rapid pace in the aftermath of the financial crisis. [Buchak et al. \(2018\)](#) find that an increase in regulatory burdens for traditional banks is the predominant driver in the rise of FinTech lending. A large aspect of the scrutiny and regulation of traditional banking was their perceived role in the financial crisis, which included facilitating wide-scale mortgage fraud as partially evidenced by over \$137 billion in government fines and settlements. FinTech lenders offer a new banking model that replaces traditional lending relationships with online advertisements, app interfaces, and loan screening algorithms. Are FinTech lenders able to harness the power of technology to reduce loan maleficence?

The Paycheck Protection Program (PPP), a historic COVID-19 relief program for businesses, rapidly distributed over \$803 billion in funds through 11.7 million loans in three short rounds spread between April 2020 and May 2021. Although FinTech lenders began with a slow start with less than 10% of loans in round 1, they ramped up their market share to over 80% of loans by May 2021, highlighting their growing importance. FinTech lending was recognized for broadening access to PPP loans, particularly to smaller firms without pre-existing lending relationships with traditional banks, and for facilitating quick and efficient lending at a time when many small businesses were in dire need due to the COVID-19 pandemic. However, the rapid expansion of FinTech lending may have come at the expense of underwriting standards. Whereas traditional banks have established borrower relationships and extensive Bank Secrecy Act (BSA) compliance programs, many FinTech lenders had few established relationships and may have been less diligent when establishing formal procedures with little reputation to protect.

Alternatively, FinTech lenders have been shown to use financial data with increased speed and accuracy. [Fuster et al. \(2019\)](#) find that FinTech mortgage lenders not only process government agency loans faster than traditional banks but have fewer defaults, indicating potentially superior loan screening. Peer-to-peer FinTech platforms utilize a rich set of alternative data and machine learning to optimize credit decisions ([Jagtiani and Lemieux 2019](#)). If used effectively, this enhanced technology and increased data access may be able to detect and prevent PPP applications from fictitious businesses and individuals. Do traditional or FinTech loans exhibit more features consistent with potential PPP misreporting? And how does this potential misreporting vary across individual traditional and FinTech lenders?

To investigate these questions, we perform a big data analysis of loan features on over eleven million PPP loans with six disparate datasets. We introduce four primary and five secondary indicators of whether a loan is potentially misstated. Each indicator creates an inference that a loan is suspicious but is not proof of misreporting on its own. The four primary

measures are non-registered businesses, multiple loans at a residential address, abnormally high implied compensation relative to industry and CBSA averages, and large inconsistencies with differences as large as tenfold between the jobs reported by a borrower on its PPP application and jobs reported to another contemporaneous government program application with a different incentive structure. The five secondary measures are discontinuities around the maximum PPP compensation level of \$100,000, rounded loan amounts, overrepresentation of PPP loans relative to U.S. Census data on the number of business establishments in a particular industry and county, clustering of loans with similar features within lender-county pairs, and criminal records for PPP borrowers.

We assess each of the four primary indicators with multiple discontinuity and comparative analyses. First, there is substantial cross-validation across the four main indicators. For example, loans that report abnormally high compensation relative to the U.S. Census’s average compensation in the loan’s industry and CBSA also have higher incidences of non-registered businesses and multiple loans at the same address. These patterns are substantially elevated in FinTech lenders but are also present for traditional lenders, indicating that misreporting is not just confined to FinTech. Second, someone receiving a fictitious loan might wish to maximize or come close to maximizing their proceeds. We find monotonically increasing levels of indicators when approaching the perceived maximum compensation threshold, and a sharp discontinuity at the threshold with lower levels of suspicious lending just above the threshold. These patterns are present for all four main indicators; discontinuities are present both for traditional and FinTech lenders, but are much more pronounced for FinTechs with an increase in potential misreporting of seven times when approaching the threshold from below. Third, even though PPP loan amounts were supposed to be based on historical past compensation with requirements for detailed supporting documentation, loans cluster at rounded monthly compensation values, and these spikes coincide with higher levels of each of the four primary misreporting indicators. These patterns are most pronounced for FinTech lenders.

Fourth, PPP lending at the industry-county level frequently exceeds the number of establishments listed for that industry and county in U.S. Census data. For FinTech lenders, 40.4% of loans exceed industry-county establishment counts, and 33.7% of loans exceed industry-county establishment counts by a factor of more than two.¹ Measures of misreporting monotonically increase as the ratio of PPP loans to businesses documented by the U.S. Census increases, particularly for FinTech lenders. Fifth, based on the idea that networks in a region may use recurring loan features, we construct a concentration ratio to measure

¹The corresponding figures for traditional lenders are 14.9% and 8.8%, respectively.

clustering in loan amounts, number of jobs, and industries within each lender-county pair. Like the other secondary measures, FinTech lenders have higher levels of clustering along loan features, and clustering is monotonically associated with higher levels of potential misreporting. Finally, we collect criminal background data for a sample of 150,000 individuals. FinTech borrowers are more than 3.5 times as likely to have a felony record, and borrowers flagged for potential misreporting based on the primary and other secondary measures are also more likely to have felony records.

Overall, we find more than 1.51 million questionable loans representing over \$68.9 billion in capital. FinTech loans are more than 3.17 times as likely to have at least one primary indicator of misreporting and 6.16 times as likely to have a primary indicator that is confirmed by an additional primary or secondary indicator. There is also substantial geographic heterogeneity in the rates of suspicious lending across counties and across zip codes within the same county. Suspicious lending rates also vary substantially across lenders, with potential misreporting rates in excess of 32% for two large FinTech lenders. Moreover, potential misreporting increases over time with particularly high rates in the last month of round 3 (25.8%), even after the Office of the Inspector General for the Small Business Administration (SBA) flagged PPP fraud as a concern. Network graphs of the lending space indicate that many FinTech portals switch lenders and utilize multiple lenders even for the same individual. Several of the FinTech lenders with the highest suspicious loan rates are new lenders that did not start making PPP loans until round 3, and there is no evidence that lenders attempted to decrease misreporting over time. Instead, second-draw loans to borrowers with suspicious first-draw loans by the same lender are common, and lenders with high rates of misreporting in rounds 1 and 2 increased both their misreporting rates and their loan volume in round 3. For example, the largest four FinTech lenders, Prestamos, Cross River, Capital Plus, and Harvest exhibited high and increasing rates of misreporting and lending volume while receiving over a billion dollars in processing fees each. Finally, FinTech lenders often doubled, tripled, or even quadrupled their lending in zip codes with high levels of potential misreporting in rounds 1 and 2 while also substantially increasing their misreporting percentages.

Our work is related to four main literatures. First, there is a rapidly emerging literature on FinTech lending that highlights its growing importance and positive economic effects through filling gaps left by traditional banks in both residential ([Buchak et al. 2018](#)) and business lending ([Gopal and Schnabl 2020](#)). [Fuster et al. \(2019\)](#) find that FinTech mortgage lenders process loans faster and increase the odds of borrowers refinancing their loans at lower rates, all with fewer defaults, indicating that FinTechs are not simply engaged in lax

screening as was the case for securitized lending in the run-up to the financial crisis (Keys et al. 2010; Purnanandam 2011). Erel and Liebersohn (2021) examine FinTech lending in the PPP and finds that FinTech lenders increased access to the PPP by lending more in zip codes with fewer traditional banks, lower incomes, and higher minority percentages. Chernenko and Scharfstein (2021) find that black- and Hispanic-owned firms were less likely to receive PPP loans. Howell et al. (2021) find that FinTechs were more likely to provide PPP loans to black-owned businesses, and Atkins et al. (2021) find that FinTechs helped close a gap in loan size between black- and white-owned businesses.² With respect to FinTech lending before the PPP, Gopal and Schnabl (2020) find that FinTech lenders have positive economic effects by filling in gaps in lending to small businesses left by traditional banks following the financial crisis. While most of the FinTech literature finds benefits to FinTech lending such as increased competition, broader financial access, less discrimination, faster lending speed, and lower defaults, our paper analyzes a potential cost of FinTech expansion and differential practices across FinTechs. We are not anti- or pro-FinTech and leave overall welfare analysis to future research.

Second, regarding the efficacy of the PPP, Chetty et al. (2020) find that the PPP increased employment at participating firms by only 2% at a cost of \$377,000 per job saved, and Autor et al. (2020) find only slightly higher employment benefits of 2% to 4.5%. Granja et al. (2020) find small employment effects due to the PPP and a low correlation between regional COVID variation and PPP funding allocation.³ In contrast, Faulkender et al. (2021) finds that the program was much more effective with an estimated 18.6 million jobs saved at an average cost of \$28,000. Additionally, there is evidence of differential access to the PPP based on knowledge of the program, distance to the closest bank branch, banking relationships, and personal banking connections (Amiram and Rabetti 2020; Bartik et al. 2020; Neilson et al. 2020; Duchin et al. 2021; Glancy 2021; Li and Strahan 2021). Our evidence adds an additional concern regarding the PPP’s efficacy and fairness. We are the first academic paper to examine wide-scale potential PPP loan misreporting, but there have been interesting press and investigative reports regarding suspicious PPP loans (Miami Herald 2020; The Wall Street Journal 2020; Bloomberg Businessweek 2020; Project on Government Oversight 2020; ProPublica 2021), some of which feature FinTech lenders.⁴

Third, assessment of the PPP program also relates to a broader literature on fraud,

²In contrast, Bartlett et al. (2021) find that FinTech algorithms charge higher interest rates to minorities in residential mortgage lending but price discriminate less than traditional banks. Begly et al. (2021) find that the SBA disaster-relief home loan program denies more loans to minorities and subprime borrowers due to the program’s risk-insensitive pricing.

³Relatedly, Meirer and Smith (2021) find that the PPP was a windfall for some firms.

⁴Concerns about PPP fraud have been flagged by the Office of the Inspector General for the Small Business Administration (SBA) (see report [here](#)). Beggs and Harvison (2021) find that among the 2,999 registered investment advisors who took PPP loans, those with a history of financial misconduct received unusually large PPP loan allocations.

waste, and abuse in government programs. While most of this literature focuses on active corruption within the government (e.g., [Shleifer and Vishny 1993](#); [Svensson 2003](#); [Glaeser and Goldin 2006](#); [Avis et al. 2018](#)), waste due to poor program design or administration can also be costly ([Bandiera et al. 2009](#)). [Hart et al. \(1997\)](#) model tradeoffs between internal and external provision of government services, and [Deflo \(2017\)](#) emphasizes the importance of program details. In the PPP program, lack of direct tools for validating eligibility and limited incentives for high-quality underwriting by lenders with no skin in the game may have significantly increased fraud and abuse. [Hanson et al. \(2020\)](#) argue that direct relief for small businesses from the Internal Revenue Service could have been more targeted and efficient than the PPP’s external lender model.

Finally, our work relates to forensic economics and loan misreporting. [Zitzewitz \(2012\)](#) surveys the literature on forensic economics, noting that a common thread in this literature is quantifying activity about which there was previously only anecdotal evidence, in large part because agents have an incentive to keep it hidden. Widescale mortgage fraud and misreporting in securitized mortgages prior to the financial crisis included second-lien and owner-occupancy status misreporting ([Piskorski et al. 2015](#); [Griffin and Maturana 2016](#)), misreported income ([Jiang et al. 2014](#); [Mian and Sufi 2017](#)), misreported assets ([Garmaise 2015](#)), and inflated appraisals ([Ben-David 2011](#); [Kruger and Maturana 2020](#)). This fraud involved both smaller, less-known mortgage originators and large bank underwriters who knowingly passed along these misrepresentations in mortgage-backed securities. FinTech lending emerged and grew against this backdrop as related regulation increased for traditional banks ([Buchak et al. 2018](#)). Our findings indicate that replacing traditional lending with FinTech lending may amplify misreporting problems, at least with respect to the PPP.

Our findings also have important practical implications regarding the extent and nature of PPP misreporting, the expanding role of FinTech lending, waste in the PPP, the proliferation of fictitious lending, and the insufficient deterrence of current policies and enforcement. The potential policy and practical implications of these findings are further discussed in the conclusion.

1 Data and Summary Statistics

1.1 Data Sources

The basis for our sample is loan-level PPP data released on June 30, 2021 by the Small Business Administration (SBA). This dataset covers all PPP loans issued from the start of the program on April 3, 2020 through the end of the program on June 30, 2021 that had

not been repaid as of June 30, 2021. At the loan-level, the data include business name, address, business type (e.g., corporation, LLC, self-employed, etc.), NAICS code (industry), loan amount, number of employees, date approved, loan draw (i.e., initial, first-draw loan or repeat, second draw loan), and lender for 11,768,689 loans originated by 4,890 different lenders and with a total value of \$803 billion. We follow [Erel and Liebersohn \(2021\)](#) and classify lenders as either traditional or FinTech (consisting of online banks and non-bank lenders) based on automated name matching with bank identifiers from the Federal Financial Institutions Examination Council (FFIEC) with hand matching for remaining lenders.⁵

Concurrently with the PPP, the SBA provided businesses and individuals with the ability to receive an Economic Injury Disaster Loan (EIDL), with forgivable advances of up to \$10k. EIDL Advance loan-level data was released on December 1, 2020 and covers all EIDL Advance issued in 2020.⁶ To check for inconsistencies between the information borrowers provided on their PPP and the EIDL applications, we match borrowers in the PPP and EIDL loan-level datasets based on business name and zip code.

We also match PPP borrowers with business registry data from OpenCorporates, a non-profit that maintains a database of companies around the world. OpenCorporates collects its data directly from state governments and covers 76 million businesses across all US states except Illinois. The data include incorporation dates, dissolution dates (if applicable), and, implicitly, whether the business has ever been registered. We match OpenCorporates data to the PPP loan-level data based on business name and state.

To examine previous criminal and financial activity, we collect criminal background data from LexisNexis based on the borrower’s name and address for a random sample of 150,000 round 1 and 2 loans made to individuals (12.9% of rounds 1 and 2 PPP loans made to individuals).⁷

Finally, we use several U.S. governmental data sources for address and demographic information. We standardize addresses and distinguish residential and commercial addresses from one another based on the Address Validation Application Programming Interface from the United States Postal Service. For data on the number of establishments and average

⁵We use [Erel and Liebersohn’s \(2021\)](#) classifications for lenders that were active in rounds 1 and 2 (the sample period for [Erel and Liebersohn \(2021\)](#)), and we use the same methodology for classifying round 3 lenders that were not active enough to be classified in the earlier rounds. We also note that the classification of FinTech lenders can be difficult because traditional banks with multiple branch locations may also originate loans from other lenders or online portals. See the Internet Appendix for additional details.

⁶The SBA has not released updated EIDL Advance data for 2021.

⁷Because the LexisNexis searches require an individual’s name, only loans with an individual name listed as the borrower (rather than a business name) and where the business type is a self-employed individual, an independent contractor, or a sole-proprietor are included in this criminal search. The criminal records data is collected only from rounds 1 and 2 loans because round 3 data was released after the criminal records data was collected.

compensation, we use the 2019 County Business Patterns (CBP) data from the US Census Bureau, aggregated by region (either core-based statistical area (CBSA) or county) and North American Industry Classification (NAICS) code. The CBP data include the number of establishments, number of employees, and total wages for a given industry in a county or CBSA. Similarly, for data on total receipts by non-employer businesses, we use data from the Nonemployer Statistics (NES) data from the US Census. Matching between the loan-level data and the CBP and NES datasets is based on the business’s zip code and the first four digits of its NAICS code.

1.2 Summary Statistics

Panel A of Figure 1 shows the number of loans originated on the left axis and the total amount lent on the right axis by each of the top 75 PPP lenders. FinTech lenders are highlighted in red (non-bank FinTech lenders) and cream color (online banks). Six of the ten top lenders by number of loans are FinTech, with Prestamos, Cross River, and Capital Plus in the top five alongside Bank of America and JP Morgan Chase.⁸ Due to their larger average loan size (Erel and Liebersohn 2021), dollar lending volume tends to be higher for traditional banks.

Panel B of Figure 1 shows the total FinTech market share during each week throughout the three rounds of PPP lending. Total FinTech market share grew from only 1.4% of loans in the first week of round 1 to 7.6% in the last week of round 1. Round 2 continued the PPP after a short break of ten days in May 2020 with new funding for borrowers who did not receive a loan in round 1. By the end of round 2 in August 2020, FinTech market share grew to 49.0% of loans, over the last two weeks, for an overall market share of 4.8% in round 1 and 20.4% in round 2. Round 3 of the PPP, which includes both first-draw loans for new borrowers and second-draw loans for borrowers that already obtained loans in round 1 or 2, started in January 2021 with a low FinTech market share of less than 20% for the first three weeks as traditional lenders were once again the fastest to originate PPP loans. However, FinTech share grew rapidly during round 3, reaching over 80% of loans by the end of May 2021 for an overall round 3 market share of 47.9%.⁹

Table I reports summary statistics for the 4.0 million FinTech and 7.8 million traditional bank loans in our sample. FinTech loans have an average loan amount of \$25 thousand compared to \$90 thousand for traditional bank loans. Despite these large differences in means, the median loan sizes of \$19 and \$21 thousand are similar. The average FinTech

⁸Comparing this figure to Panel A of Figure IA.1 shows how the top lenders differ between the entire sample and solely rounds 1 and 2. In particular, the growth of Capital Plus and Harvest in round 3 is apparent.

⁹Panel B of Figure IA.1 shows the number of loans originated each week of the PPP by type of lender.

loan reports supporting 2.4 jobs compared to 10.3 for traditional banks. After normalizing loan size relative to reported jobs, FinTech loans have higher average (\$64 thousand) and median implied compensation (\$71 thousand) than traditional bank loans (\$47 thousand average and \$39 thousand median). For FinTech loans, 18.5% of borrowers are organized as corporations, S-corporations, or limited liability companies (LLC) compared to 64.5% for traditional banks lenders. FinTech loans were also less likely to be repeat loans, with 26.8% of round 3 FinTech loans going to borrowers with previous PPP loans, compared to 59.4% for round 3 traditional bank loans.

2 Suspicious Loan Measures

We introduce four primary indicators that a loan is potentially misstated. In this section, we define and introduce the indicators. Each indicator creates an inference that a loan is suspicious but is not definitive proof of misreporting on its own. In subsequent sections, we validate the measures and explore how they relate to one another and other misreporting indicators.

2.1 Business Registry Flag

Businesses organized as corporations, S-corporations, and LLCs are required to file an article of incorporation or LLC filing with a state, either as a domestic company in their home state or as a foreign company in another state. Further, the SBA required businesses to be “in operation on February 15, 2020... [and] not permanently closed.”¹⁰ Based on these requirements, we check the following conditions for all corporation, S-corporation, and LLC borrowers:

1. Is the business found in the business registry for its home state or in another state while listing an address in its home state? (“Missing Business”)
2. Was the business dissolved and inactive before being approved for a PPP loan? (“Dissolved Business”)¹¹
3. Is the earliest incorporation or initial filing date for the business after February 15, 2020? (“Late Incorporation/Filing”)

These three subflags are combined to form an overall business registry flag.¹²

¹⁰See loan application [here](#).

¹¹To be flagged, the dissolution date of the business must be before the PPP loan approval date and, to screen out businesses that may be administratively dissolved (e.g., for not filling some paperwork), the business status must be listed as inactive.

¹²As external validation of this flag for a smaller sample, we also compare PPP borrower names to data from the Florida Department of Business and Professional Regulation following [Chernenko and Scharfstein \(2021\)](#) (see Figure [IA.2](#)). Loans flagged as a missing business based on the overall business registry are over 6.7 times less likely to have a potential match in

Panel A of Figure 2 plots the proportion of corporate and LLC borrowers with missing, dissolved, or late business registrations. The flag is plotted as a percent of corporation, S-corporation, and LLC loans because other PPP business entities, such as sole-proprietorships, partnerships, and independent contractors, do not require business registrations.¹³ Missing registrations are the most common type of business registry flag, representing 4.2% of corporate and LLC loans. Another 0.7% of corporate and LLC loans are to dissolved entities, and 0.2% have late registrations for a total business registry flag percentage of 5.1%. Nine of the ten lenders with the highest rates of business registry flags are FinTech lenders. These lenders have 7.6% to 26.7% of their corporate and LLC loans flagged for one of the three business registry issues, with the vast majority of the flagged loans simply not appearing in the business registry data. It is possible that there are errors in the data or that some businesses have names that are difficult to match; which may explain why all of the lenders have at least some missing registrations, with business registry flag rates of one to five percent common across many lenders. However, there is not an obvious explanation for why certain lenders, who also have elevated levels of other indicators, would have disproportionately high matching issues.

2.2 Multiple Loan Flag

While it is possible that a business owner may have multiple businesses registered to the same address, the presence of multiple loans at a residential address during the same draw is also a potential sign of fictitious operations. Using the business address disclosed in the PPP loan-level data, we identify individual residential addresses associated with three or more loans during the same draw. To do so, we first standardize addresses and identify addresses that are known business or central addresses (e.g., office and apartment buildings) using the Address Validation Application Programming Interface from the United States Postal Service. Then, we find residential (i.e., non-business, non-central) standardized addresses with three or more loans within the same draw.

As an example, Panel A of Exhibit 1 shows 14 loans given to a single address, all with colorful business names, almost all in the same industry, most with the same loan amount, and all backing ten jobs. The address associated with all 14 loans is a modest single-family home in suburban Chicago (estimated to have a value of \$170k per Zillow). The borrower associated with the first loan is an LLC that was registered in 2018, but the 13 subsequent loans during July and August of 2020 are to LLCs that were registered only shortly before the

the restaurant data as compared to loans that are not flagged. Loans flagged as having missing business registrations that also have another flag (primary or secondary) are over 12.5 times less likely to have a potential match in the restaurant data.

¹³Businesses in Illinois are excluded because Illinois is missing from the business registry data due to restrictive terms and conditions (see regulation [here](#)).

loans were approved, well after the February 15 eligibility cutoff. Detailed internet searches did not produce information for any of the other 13 business names or any indication of employees other than the owner. Panel B of Exhibit 1 shows another multiple-loan example, this one involving loans to four people in the same household, again in a modest suburban Chicago home, all of whom received loans for the same amount, \$20,833, which corresponds to the PPP’s maximum annual compensation of \$100,000.¹⁴ This income is at the top of the spectrum for the indicated industries, which have average compensation \$25-46k in the Chicago CBSA according to the US Census CBP. Random loan-level inspections of the data reveal numerous other examples of multiple suspicious loans flowing to addresses that do not seem to be the locations of identifiable businesses. The multiple loan flag functions as a way to systematically analyze these loans.

Panel B of Figure 2 shows the percentage of PPP loans that involve at least three loans to the same residential address in the same draw by lender. Nine of the ten lenders with the highest multiple loan flag rates are FinTech lenders. For these lenders, 3.7% to 5.1% of their loans involve multiple loans to the same residential address, and most of their flagged loans are to individual borrowers identified as independent contractors, self-employed, or sole-proprietors. This contrasts with traditional banks, which have fewer flagged loans (1.0% on average).¹⁵ Interestingly, three FinTech lenders, Capital One, Square, and Intuit, have lower than median levels of multiple loans at the same address.

2.3 High Implied Compensation Flag

PPP loan size is limited to 2.5 times a business’s average monthly payroll expenses, including up to \$100,000 in annual compensation per employee. PPP loan applications report how many employees the business has based on the same time period used to calculate average payroll expenses (2019 in most cases). Using loan size and number of reported employees, we are able to impute implied average annual compensation. Implied compensation at the borrower level is strongly related to average compensation in the borrower’s industry (NAICS 4-digit) and CBSA (e.g., see Figure IA.4).¹⁶

¹⁴All four of these individuals also received second draw loans for the same amount. SBA guidelines asks the borrower for their business address. The industries themselves are also somewhat suspicious in that two are equipment manufacturing and one is auto repair despite no evidence of these businesses in photos of the property. Further, the borrower in the nail salon industry does not appear to have an Illinois nail technician license. One of the equipment manufacturing borrowers also switched to the nail salon industry during round 3 despite also not having a nail technician license.

¹⁵Large differences between traditional and FinTech lenders hold even if we limit the sample to only loans with residential addresses and are even larger if we flag loans if there are at least two loans at the same residential address with the same draw (see Figure IA.3). Flagged loans for traditional banks are mainly to formally registered corporations, S-corporations, and LLCs, consistent with the incentives for owners of multiple legitimate businesses to formally register their businesses for tax and limited liability purposes. We control for differences in loan composition across lenders in subsequent regression analysis.

¹⁶Schedule C filers also had the option to use gross income instead of net income for owner compensation after March 3, 2021. To conservatively account for the option, we compare implied compensation for sole proprietor, independent contractor,

To capture abnormally high compensation, we examine the kernel density of implied average compensation for the borrower normalized by mean compensation across all firms in the borrower’s industry and CBSA based on US Census Bureau CBP data separately for FinTech and traditional loans in Panel A of Figure 3. FinTech borrowers have a much fatter right tail of the abnormal compensation distribution. Specifically, in the lower plot, which includes all loans, 17.8% of FinTech borrowers have normalized compensation above 3, compared to 3.5% of traditional borrowers.¹⁷

It is instructive to examine how normalized compensation relates to our first two suspicious loan flags. On the right axis of Panel A of Figure 3, we plot the percentage of loans with the business registry and multiple loan flags across the distribution of normalized compensation separately for FinTech and traditional lenders.¹⁸ Both flags increase significantly as normalized compensation increases for loans made by FinTech lenders. Whereas 7.8% of corporate and LLC FinTech loans with normalized compensation below one have the business registry flag, 26.7% of loans with normalized compensation above three have the flag. Similarly, the multiple loan flag increases from 2.7% for FinTech loans with normalized compensation below one to 5.1% when normalized compensation is above three. Importantly, while FinTech loans exhibit a stronger relation between normalized compensation and the other loan flags, this pattern is not limited to FinTechs. Traditional bank loans also have more business registry and multiple loan flags when normalized compensation is higher suggesting that they also have loan misreporting issues, though at a much lower scale. Overall, the results show that while some variation in normalized compensation across firms is to be expected, high implied compensation is strongly related to other suspicious loan characteristics, particularly for FinTech loans.

For our main measure of high implied compensation, we conservatively only flag loans where the implied compensation per job reported is more than three times the industry-CBSA average compensation/receipts (“high implied compensation”). Because compensation is censored at \$100,000 for most borrowers, this flag is only possible in industry-CBSA pairs with average annual compensation/receipts below \$33,333.33.¹⁹ Within this set of

self-employed, and single member LLC loans after March 3, 2021 to the greater of industry/CBSA average compensation and industry/CBSA average receipts for single-employee firms. This adjustment is also used in calculating the high implied compensation measure below.

¹⁷Most of this is due to round 3 FinTech loans, as evident in Figure IA.4, Panel A. FinTech implied compensation is much higher round 3 and appears to be almost completely disconnected from average industry-CBSA compensation and receipts. Panel B of Figure IA.4 shows that this differential pattern for fintech and traditional lenders in round 3 is also evident even when the sample is restricted to Schedule C borrowers, both before and after Schedule C borrowers were permitted to use gross income, starting on March 3, 2021.

¹⁸The business registry flag plot includes only corporation, S-corporation, and LLC loans because the business registry flag can only be determined for these business types.

¹⁹Some loans are also outside of a CBSA or in an industry-CBSA pair that is too small to be included in the US Census CBP data. In total, 3,416,620 loans are in industry-CBSA pairs with average annual compensation/receipts below \$33,333.33.

industry-CBSA pairs, 48.9% of FinTech loans and 9.4% of traditional loans have normalized compensation above three. Panel B of Figure 3 plots how the percentage of loans with the high compensation flag varies across lenders. Eight of the ten lenders with the highest abnormal compensation percentages are FinTech. For all of these lenders, more than 50% of applicable loans have the high implied compensation flag. By contrast, 46 of the 71 largest lenders have less than 10% of their loans flagged, and 43 of these lenders are traditional banks. Although most of the FinTech lenders cluster with high rates of abnormally high compensation, Capital One and Square are near the very low end of flagged loans, perhaps indicating these FinTechs were relying on existing relationships with the borrowers.

2.4 EIDL Advance Jobs > PPP Jobs Flag

Concurrently with the PPP, the SBA provided businesses and individuals with the ability to receive a forgivable Economic Injury Disaster Loan (EIDL) Advance of up to \$10,000.²⁰ For all EIDL Advances issued in 2020, the advance amount was calculated as \$1,000 per employee (up to the \$10,000 maximum).²¹ Thus, there was an incentive for borrowers to inflate the number of jobs reported on their EIDL applications.²² We focus on cases where EIDL jobs exceed PPP jobs because the job inflation incentive is provided by the EIDL Advance program (PPP loans are based on total payroll as opposed to number of jobs).

Panel A of Figure 4 plots the distribution of differences between EIDL and PPP jobs. Three patterns stand out. First, consistent with the incentive to inflate EIDL jobs as opposed to PPP jobs, EIDL exceeds PPP by three or more jobs 9.1% of the time, whereas PPP exceeds EIDL by three or more jobs only 3.6% of the time. Second, the most common discrepancy between the programs is a difference of nine jobs, which implies that the borrower claimed 10 or more jobs and took out the maximum EIDL Advance of \$10,000 despite only reporting one PPP job. Third, EIDL job inflation is much more pronounced in FinTech loans than in traditional loans. In particular, EIDL data exceeds PPP data by nine jobs 14.5% of the time for FinTech loans compared to 0.6% for traditional loans.

Panel B of Figure 4 shows the prevalence of EIDL job inflation by lender as a percentage of PPP loans with matching EIDL Advances. To be conservative, we only include cases

²⁰While the EIDL Advance program was billed as a forgivable advance with the potential for a larger non-forgivable loan, 65.8% of EIDL Advances involved no additional EIDL loan. EIDL Advances were immediately forgiven by the SBA.

²¹The EIDL Advance rules changed for 2021 to: A) provide the entire \$10,000 regardless of employee count, and B) to target the advances to low-income communities and those with a demonstrated decrease in revenue. The SBA has not yet reported data on 2021 EIDL Advances.

²²For borrowers who take out the maximum EIDL Advance of \$10,000, we can infer that the borrower claimed at least 10 employees on their EIDL Advance application. Subsequent to the original public version of this paper, an October 7, 2021 report by the SBA OIG found that over 700,000 EIDL recipients applied for and received advances for multiple employees even though they only had a single employee, resulting in \$4.5 billion of improper EIDL Advance payments (see report [here](#)).

where the EIDL implied number of jobs is at least three more than the PPP reported number of jobs. All ten lenders with the most frequent discrepancies between EIDL and PPP jobs are FinTech. In particular, Capital Plus, Prestamos, Lendistry, Harvest, Benworth, Fountainhead, Itria, MBE, Cross River, and Leader Bank (all FinTechs) have job reporting inconsistencies ranging from 12.8% to 54.6%. For all of these lenders except MBE and Leader Bank, most of the inconsistencies are a full nine jobs. Five FinTech lenders (Intuit, WebBank, Square, Capital One, and Live Oak) have levels of inconsistencies that are similar to traditional banks, with nine-job differences only in rare cases.

2.5 Are the Suspicious Loan Flags Related to One Another?

If the above indicators of potential misreporting are due to random data errors or honest mistakes, one might expect different types of indicators to occur randomly across loans and lenders. Therefore, multiple flags for the same loan create a heightened misreporting inference, and high lender flag rates across multiple indicators may be due to policies and practices that facilitate more misreporting.

In Table II, we examine how the four flags relate to one another by reporting odds ratios between each pair of flags. The odds ratios are calculated based on loans for which data to calculate both flags are available (e.g., corporate and LLC loans for the business registry flag and loans with matched EIDL Advances for the EIDL > PPP Jobs flag), with z -statistics calculated based on standard errors double-clustered by zip code and lender in parentheses. Panel A reports odds ratios for the full sample, all of which are above 1.50 and highly significant. In particular, the odds ratio between the high implied compensation and EIDL > PPP jobs flags is 14.75 and has a z -statistic of 19.90. Panel B reports odds ratios separately for FinTech and traditional loans with FinTech loans in the lower triangle and traditional loans in the upper triangle. The odds ratios are all positive and highly significant with consistently higher ratios for FinTech loans. To check that these relations are independent of one another and not explained by loan characteristics, Table IA.I regresses each of the flags jointly on the other flags, controlling for loan size and number of jobs with zip code, business type, and industry \times CBSA fixed effects, with and without lender fixed effects. Except for the relation between the business registry flag and EIDL > PPP Jobs flags, the coefficients between the flags are all positive, economically large relative to the mean flag rates, and highly statistically significant.

We also find that flag rates are significantly correlated with one another across lenders, and the same lenders frequently have high flag rates across all four indicators (as shown in Figure IA.5). In particular, FinTech lenders Capital Plus, Prestamos, MBE, and Harvest

have flagged rates in the top 10 for all four flags, and Itria and Benworth are in the top ten for three of the four flags. In contrast, no traditional lender is consistently in the top 10 for more than two flags. This pattern is exactly what we would expect if some lenders have looser underwriting standards and is difficult to explain with random mistakes or errors in the data.

As an additional validation of the primary flags, we also compare them to direct evidence of loan size inflation for nonprofits based on comparing loan sizes to non-profit compensation disclosed on IRS Form 990. Figure [IA.6](#) shows that loan size inflation by non-profits is increasing and highly related to the primary flags. See the Internet Appendix for details on this analysis.

2.6 FinTech Differences?

Table [III](#) summarizes the percentage of loans with each of the four flags separately for FinTech and traditional lenders. The table also summarizes the percent of all loans with at least one flag and with two or more flags. For each individual measure, the denominator is the loans that could have such flag (i.e., only corporate and LLC loans for the business registry flag, only loans in industry-CBSA pairs with average compensation below \$33,333.33 for high implied compensation, and only loans with a matched EIDL Advance loan for the comparison to EIDL). For the overall flag measures, the denominator is all loans in the sample, which understates the incidence of suspicious loans since most of the flags are only applicable to a minority of the loans. Differences between FinTech and traditional flag percentages are reported in column (3). For all four individual measures, FinTech lenders have flag rates that are 2.35 to 5.19 times as high as traditional lenders, with particularly large differences for the high implied compensation and EIDL > PPP jobs flags. Overall, 23.6% of FinTech loans have at least one of the flags, compared to 7.4% for traditional loans. These differences are all highly significant with standard errors double-clustered by zip code and lender to conservatively allow for potential geographic and within-lender correlations.

To account for potential compositional differences between FinTech and traditional lenders, column (4) reports adjusted differences that control for geography, business type, and industry based on regressions controlling for loan size and number of jobs with zip code, business type and industry \times CBSA fixed effects.^{[23](#)} After accounting for these effects, the adjusted difference between FinTech and traditional flag rates is 3.5 percentage points (ppt) for the business registry flag (which is 81% of the rate for traditional loans), 1.1 ppt (106%) for the multiple loan flag, 9.4 ppt (99%) for the high implied compensation flag, and 6.5 ppt (134%)

²³Corresponding regressions results with and without the control variables and fixed effects are reported in Table [IA.II](#).

for the $EIDL > PPP$ jobs flag. These results indicate that even though loan composition explains part of the difference between FinTech and traditional loans, flag rates remain much higher for FinTech loans even after controlling for all observable characteristics. To further control for potentially non-linear loan characteristic effects, we match FinTech loans with traditional loans based on loan size, industry, county, and business type in column (5) FinTech with similar results.²⁴ It remains possible that other omitted variables or unobserved loan characteristics could explain some of the difference between FinTech and traditional loans, but these effects would have to be large to explain the results. To control for any unobserved differences across households, Table IA.III considers a restricted sample of residential addresses with loans from both traditional and FinTech lenders. Consistent with results in Table III, flag rates are elevated for FinTech loans across all of the potential misreporting measures, with highly statistically significant differences in all but one specification. Tests in the next section, including grouping around discontinuities and clustering, also help to address omitted variable concerns.

3 Suspicious Loans or Mistakes?

While they are suggestive of misreporting, the misreporting indicators in the previous section also have potentially innocent explanations. In this section, we develop and analyze five additional measures as external verification to assess the plausibility of alternative explanations. The additional measures involve discontinuities, rounded compensation levels, abnormal numbers of loans in industry-county pairs, clustering of loan features within lender-county pairs, and criminal records. We also explore differences between FinTech and traditional lenders.

3.1 Discontinuities at \$100,000 Compensation

PPP loan size is calculated as 2.5 times a borrower’s average monthly payroll, including up to \$100,000 in wages per employee.²⁵ This \$100,000 cutoff is a hard maximum for self-employment compensation. For other employees, payroll expenses also include employer insurance and retirement contributions and unemployment taxes, which can push included payroll expenses above \$100,000 per employee. Someone filling out a fraudulent PPP application and who may not have carefully read the PPP rules, might want to maximize their loan amount by submitting payroll expenses at or close to the \$100,000 per employee limit without the additional expenses that are eligible with proper payroll details.

²⁴Details on the matching process are provided in the Internet Appendix.

²⁵See the Internet Appendix for details on the SBA guidance for how to calculate loan size.

Figure 5 plots the distribution of implied compensation per employee and shows how it relates to the misreporting indicators from the previous section. The implied compensation distributions (up to \$130,000) for FinTech and traditional loans are plotted as orange and gray bars, respectively. Panel A shows that FinTech loans stand out as having more loans with high implied compensation right at and slightly under \$100,000, and traditional banks have more loans with implied compensation in the range of \$10,000 to \$75,000. The percent of loans with one of the four primary flags for FinTech and traditional loans are plotted as orange and gray dots along with third-degree polynomials and their associated 95% confidence intervals estimated separately above and below the \$100,000 compensation bin. As compensation increase from \$40,000 to \$100,000, the prevalence of the primary flags for FinTech loans increases from 6.1% to 43.1%. For traditional lenders, the increase is also present but much smaller. For FinTech loans with implied compensation above \$100,000, there is a sharp drop-off in the flag rate, which indicates that businesses that followed the detailed SBA guidelines for including non-wage payroll expenses for employees with wages above \$100,000 are less likely to have one of the primary misreporting flags.

Panel B repeat the same analysis separately for the business registry, multiple loan, high implied compensation, and EIDL > PPP Jobs flags. For each indicator there is a much steeper slope for FinTech loans than for traditional loans, and FinTech loans have a sharp discontinuity above \$100,000. The drop in the flags above \$100,000 indicate that borrowers with detailed insurance and tax expenses in excess of \$100,000 are also more likely to be reporting correctly than those at or below the threshold. In Table IA.IV, we formally test for discontinuities at \$100,000 of compensation after controlling for loan characteristics including number of employees, loan size, business type, zip code, and industry \times CBSA and find large and highly economically significant discontinuities for FinTech loans for all four measures. Differences are also statistically significant for traditional loans, though the effects are economically smaller. Overall, the increasing flag rates as compensation increases and the discontinuities around \$100,000 in annualized compensation are consistent with suspicious loans maximizing loan amounts.

Additionally, the SBA used a loan amount cutoff of \$150,000 for a more streamlined processing (fewer calculations and less documentation) of loan forgiveness applications.²⁶ Consistent with applicants or lenders being aware of the threshold and trying to avoid scrutiny, the percentage of flagged loans is high for loans up to \$150,000 and decreases after the threshold. This is true for both traditional and FinTech lenders, but much more pronounced for FinTech lenders (see Figure IA.7).

²⁶The shorter form and reduced requirements for loans of \$150,000 or below to receive forgiveness are outlined [here](#).

3.2 Rounded Loan Amounts

The PPP loan application instructs borrowers to enter their average monthly compensation and to calculate their loan amount as

$$\text{Loan Amount} = \text{Average Monthly Payroll} \times 2.5 + \text{EIDL Refinance Amount}.$$

Applicants are instructed to calculate average monthly payroll based on historical compensation (in 2019 in most cases) with detailed supporting documentation.²⁷ It is unlikely that actual monthly payroll would be a round number, especially after including unemployment insurance and employer insurance and retirement contributions. Rounded loan amounts suggest that the numbers are potentially fictitious as opposed to being based on actual documented data. If the flags we have previously identified reflect misreporting issues, then one might expect both a clustering of loans at round numbers and elevated flags at round numbers. However, if round numbers are simply a result of a borrower with valid documentation rounding numbers slightly downward to simplify calculations, then one would expect no elevated reporting issues at round numbers.

In Panel A of Figure 6, we first examine the distribution of the last four digits of loan amounts, excluding EIDL Refinancing, for FinTech and traditional loans. Loan amounts within 50 cents of a \$1,250 increment (which corresponds to \$500 of implied monthly payroll) are plotted as thicker and slightly darker bars with all other loans binned into \$1 wide bins plotted as the thinner, lighter bars. Loans with an implied compensation within \pm \$1,000 of \$100,000 are excluded to make sure these results are distinct from the maximum compensation result shown in Figure 5. Both FinTech and traditional loans exhibit rounding at \$1,250 increments, particularly at increments of \$2,500 (corresponding to \$500 and \$1,000 increments of implied monthly payroll). FinTech lenders have moderately more rounding with 10.2% of loans rounded to \$1,250 increments compared to 7.6% for traditional lenders.

The right axis of Panel A examines the prevalence of the primary misreporting flags. The percent of loans with a primary flag is plotted as solid dots at the \$1,250 loan increments and as hollow dots at other loan amounts (shown in \$250 wide bins). If rounded loans are more likely to be misreported, one would expect an elevated flag rates at round number thresholds. For FinTech loans, this is exactly what we observe. At rounded increments, the flag rates are consistently higher, by 2.41 ppt on average. This difference is highly significant, which can be seen by comparison to the dotted lines plotting a 95% confidence interval estimated with a third-degree polynomial estimated based on the non-rounded loans. For traditional lenders,

²⁷See the Internet Appendix for details on how the loan size was to be calculated and exclusions.

there is only small and weak evidence of elevated flags in some of the rounded bins. Thus, rounding appears to capture suspicious loans for FinTech lenders but less so for traditional lenders, which is also consistent with our findings in Figures 2, 3, 4, and Figure 5.

In Panel B, we consider each of the four primary misreporting indicators separately. The top left subpanel plots corporate and LLC loans, the top right subpanel plots all loans, the bottom left subpanel plots loans with an industry-CBSA pairs with average compensation of less than \$33,333.33, and the bottom right subpanel plots loans with matched EIDL Advances. The four plots show that rounded loans by FinTech lenders have elevated levels of all four primary misreporting flags. For traditional lenders, the business registry flag is slightly elevated at some round numbers, the multiple loan flag is slightly elevated levels at every other \$1,250 increment (but not at \$2,500 increments), and there is no evidence of elevation for the other flags. Overall, the fact that all of the loan flags are elevated at round loan amounts for FinTech loans provides additional validation for the suspicious behavior underlying these loans.

3.3 Loan Overrepresentation

If there is an organized effort to obtain funds for non-existent businesses, networks of illegitimate borrowers may fill out multiple applications in a similar manner and could cluster on characteristics such as industry and geography. Exhibit 2 shows examples from 4,300 \$20,000 first draw loans made by Cross River to businesses in the “Insurance Agencies and Brokerage Industry” in Illinois, mainly in the Chicago area, almost all of which have one employee. These are followed by examples from 938 \$20,000 first draw loans by Cross River to businesses engaged in “All Other Miscellaneous Crop Farming,” most of which have exactly one or eight employees.²⁸ Most of these loans are in urban areas of Chicago, frequently in apartment dwellings, where it is difficult to see how crop farming is performed. There are also another 3,068 \$20,000 first draw loans by Cross River in Illinois to borrowers in other industries (including 706 to business in “All Other Personal Services,” 351 to “General Freight Trucking, Local,” 337 to “Other Performing Arts Companies,” and 300 to “New Single-Family Housing Construction (except For-Sale Builders)”). In addition to having the same loan amount and similar industries, these \$20,000 loans were almost non-existent until the very end of round 2. Specifically, 40.8% were originated in late July and early August of 2020 during the final two weeks of round 2, and 56.8% were originated in round 3. Overall, 48.9% of Cross River’s Illinois loans between July 21, 2020 and August 8, 2020 and 17.6% of Cross River’s Illinois round 3 loans are for \$20,000, compared to 1.1% of Cross River’s Illinois

²⁸There are an additional 1,574 loans for amounts besides \$20,000 by Cross River in Illinois to business in “Insurance Agencies and Brokerage Industry” and 646 to the “All Other Miscellaneous Crop Farming” industry.

loans before July 21, 2020 and 2.1% of Cross River’s loans in other states. This pattern is particularly suspicious given that the US Census CBP reports 2,207 “Insurance Agencies and Brokerage Industry” establishments in Cook County, Illinois, which is about half the number of first draw loans made in this industry by Cross River alone (4,384 loans, of which 3,321 are for exactly \$20,000).²⁹ To systematically look for similar patterns throughout the PPP data, we compare PPP numbers to overall establishment counts in the 2019 US Census CBP database. Because the CBP data does not include self-employed and independent contractors as establishments, we exclude loans to these business types from our analysis.

Panel A of Figure 7 plots histograms of FinTech (red bars) and traditional (gray bars) lender loans by the ratio of first-draw PPP loans to census establishments in the loan’s industry and county. For FinTech lenders, 40.4% of loans exceed industry-county establishment counts, and this occurs 14.9% of the time for traditional lenders. For loans in the tails the differences are even more extreme with 33.7% of loans exceed industry-county establishment counts by a factor of more than two for FinTech lenders and 8.8% for traditional lenders. Even further in the right tail, 8.2% of FinTech loans exceed industry-county establishment counts by a factor of more than ten as compared to 0.9% of traditional loans.³⁰ It is possible that some excess PPP loans may be due to the missing establishments in the CBP data, industry misclassifications, or other errors in the data. Nonetheless, the large excess loan rate for FinTech lenders is difficult to explain, particularly since it is so much higher than traditional lenders.

Panel A of Figure 7 also plots, for FinTech and traditional lenders separately, the percentage of loans flagged by one of the four primary suspicious loan flags by the ratio between PPP first-draw loans and CBP establishments. The flag rate increases substantially as the loan-to-establishment ratio increases, particularly for FinTech lenders. Whereas 13.8% of FinTech and 5.8% of traditional loans in industry-county pairs with a loan-to-establishment ratio at or below one are flagged by at least one of the primary misreporting indicators, the flag rates are 43.1% for FinTech and 11.9% for traditional loans where loan-to-establishment ratios are above two.

Panel B of Figure 7 plots separate rates for each of the four suspicious loan flags, with consistent results for all measures. As one moves to ratios above one, indicating more PPP

²⁹Excluding Cross River’s loans, there are 1,700 first draw loans to Cook County businesses in this industry, which is already 77% of the establishment count provided by the CBP. Loan counts for “All Other Miscellaneous Crop Farming” also appear to be high, but the CBP data does not have a comparable establishment count for this industry because it does not include agricultural establishments.

³⁰Excess loan percentages are calculated by assigning a weight to each loan based on the inverse of its industry-county’s loan-to-establishment ratio. Specifically, let r be the loan-to-establishment ratio in the loan’s industry-county pair, the weight is 0 if $r \leq 1$ and $1 - 1/r$ if $r > 1$. The interval limits are changed to 2 (10) instead of 1 for the 33.7% (8.2%) and 8.8% (0.9%) figures.

loans in an industry-county than listed in the CBP, the number of suspicious loans flagged increases substantially for all of the suspicious loan measures. This is true for both FinTech and traditional lenders, but the increase is generally steeper for FinTech lenders, consistent with FinTech loans in industry-county pairs with high loan-to-establishment ratio being particularly suspicious.

3.4 Loan Clustering

In addition to exhibiting geographic and industry clustering, many of the examples discussed above also feature identical loan amounts and job numbers. If networks submitting fictitious loan applications repeat the same application information across multiple loans, lenders may have many loans in a geographic region with similar industries, loan amounts, or jobs reported. There will clearly be some loan similarities by chance and due to lender specialization, but it is instructive to quantify how frequently loans cluster. For each lender-county pair with at least 25 loans, we calculate concentration ratios for the industry, loan amount (rounded to \$100), and reported jobs (excluding one because it is common across all lenders and counties). The concentration ratios are based on the sum of squared shares of loans with a characteristic.³¹ Then, we rescale each of the concentration ratios to have a median of 1,000 and an interquartile range (IQR) of 300 so that the three concentration ratios have similar impacts on the overall concentration measure. Finally, we average the three concentration ratios for each lender-county pair.

The bars in Panel A of Figure 8 plot the distribution of scaled concentration ratios separately for FinTech and traditional loans. High concentration ratios are much more common for FinTech loans with 88.4% of FinTech loans in lender-county pairs with a scaled concentration ratio above 1,000, compared to 21.3% of loans for traditional banks. The dots in Panel A of Figure 8 plot how the incidence of the four primary suspicious loan flags changes with concentration ratio. When the scaled concentration ratio is below 1,000, 10.2% of FinTech loans and 6.8% of traditional loans have at least one flag. However, when the scaled concentration ratio is above 1,300, this grows to 36.2% for FinTech loans and 7.3% for traditional loans. Panel B of Figure 8 shows that similar patterns hold for each of the four suspicious loan flags individually. The overall pattern is similar to the previous measures: FinTech lenders have much higher loan concentration ratios, and high concentration ratios are highly related to the suspicious loan flags, particularly for FinTech loans. This pattern is exactly what one would expect if the indicators are picking up misreported FinTech loans

³¹For example, let $i = 1, 2, \dots, n$ represent the n industries in a given lender-county pair, then $Concentration_{industry} = \sum_{i=1}^n s_i^2$ where s_i is the percentage of loans in the lender-county that are in industry i times 100 (e.g., 6.2 for 6.2%). Note that this concentration ratio is the same as a Herfindahl-Hirschman Index (HHI), which is commonly used to measure market concentration.

and is difficult to explain with innocent mistakes or errors in the data.

3.5 Criminal Records

Recidivism statistics show that individuals with past criminal histories are more likely to commit crimes in the future (Alper et al. 2018). The PPP originally prohibited loans to businesses more than 20% owned by individuals currently subject to criminal charges, incarceration, probation, or parole or who had been convicted of a felony within the past five years. These restrictions were relaxed somewhat in June 2020 to permit loans to businesses owned by individuals facing misdemeanor charges and those with convictions, probation, or parole for most felonies more than a year in the past.³² To assess the prevalence of criminal records among PPP borrowers, we collect criminal histories for a random sample of 150,000 round 1 and 2 loans to individual names in the PPP data that can be matched to LexisNexis public records data.

Panel A of Figure 9 plots the percentage of borrowers with felony criminal records in 2000–2020 within the sample of 150,000 individual borrowers for whom we collected background information.³³ Felony criminal records are present for 4.9% of non-bank FinTech borrowers and 4.6% of online bank FinTech borrowers have criminal records compared to only 1.3% of traditional borrowers. There is also a strong relation between criminal records and both the primary and secondary indicators. The EIDL misreporting indicator seems to be capturing the highest percentage of criminals. We confirm that these relations are robust and statistically significant by regressing an indicator for having a criminal record on the other primary and secondary risk flags for loans originated by FinTech lenders.³⁴

Panel B of Figure 9 examines how criminal records vary across lenders with a clear positive relation between the percentage of a lender’s sampled borrowers with criminal records and the percentage of its overall loans with one of the primary suspicious loan flags. In particular, the four lenders with the highest criminal record percentages (MBE, Cross River, Fundbox, and Kabbage, all of which are FinTech) also have the highest primary flag rates.³⁵

³²The five-year criminal record prohibition was only retained for financial crimes such as fraud and embezzlement. As a result, many individuals with criminal records were legally eligible for PPP loans. Nonetheless, a criminal record is a potential risk factor.

³³Ninety-five percent confidence intervals based on standard errors clustered by zip code and lender are plotted on top of the bars. Panel A of Figure IA.8 replicates this figure using felonies from 2015-2020. While the percentage of borrowers with felonies is lower across the board, the relative results remain.

³⁴Results are reported in Table IA.V. The regressions control for loan size and number of jobs with business type, industry \times CBSA, and lender fixed effects. Standard errors are double-clustered by zip code and lender. In all cases, the coefficients are positive, statistically significant, and economically large for FinTech loans (Panel A) with almost no relation between the misreporting indicators and criminal records for traditional loans (Panel B).

³⁵Panel B of Figure IA.8 replicates this figure using felonies post-2005, post-2010, and post-2015. While the percentage of borrowers with felonies decreases as the time period is decreased, the relative results remain. Additionally, Panel C of Figure IA.8 replicates this figure using bankruptcy filings post-2015 and finds similar results.

3.6 Relation Between Primary and Secondary Flags

We have already seen that the primary flags are strongly predictive of one another, and the evidence in Figures 5–9 show strong relations between the primary and secondary flags. In Table IV, we more formally assess these relations with regression analysis controlling for loan size and number of jobs with zip code, business type, industry \times CBSA, and lender fixed effects. The dependent variable in the regressions is an indicator variable for the loan having at least one of the primary flags. Standard errors are double clustered by zip code and lender. The secondary flags are all interacted with an indicator variable for FinTech loans, so the direct coefficients represent effects for traditional loans. Four of these five effects are positive and significant with magnitudes ranging from 6.4% to 25.0% of mean misreporting rate. Further, all of the interactions between the secondary flags and the indicator for FinTech loans are large and positive, and almost all are significant. As a result, all of the secondary flags strongly relate to the primary flags for FinTech loans, with relations that are much stronger than for traditional loans. For compensation near \$100,000 and rounded compensation, the effects for FinTech loans are 5.75 and 3.77 times as high as the effects for traditional loans, respectively. For criminal records and high loan concentration, the effects for FinTech loans are over 2.41 and 2.13 times as large as the traditional loan effects, respectively. Lastly, for industry overrepresentation, there is a strong FinTech effect despite essentially no relation for traditional loans. We also examine relations between the primary and secondary flags at the lender level (see Figure IA.9) and find that except for monthly rounding, lenders with high levels of each secondary flags tend to be the same lenders that have high levels of the primary flags.

4 How Many PPP Loans Are Suspicious?

In this section, we quantify ranges of suspicious loans based on the primary and secondary flags developed in the previous two sections. Panel A of Figure 10 plots flag rates for each of the four primary flags along with overall suspicious lending rates. Our primary measure consists of loans that have at least one primary flag, plotted as the total height of the bars. By this measure, 1,515,887 loans representing 12.9% of the PPP and totaling \$68.9B are suspicious.³⁶

FinTech lenders are responsible for a disproportionate share of suspicious loans. Combined, non-bank FinTech and online bank FinTech originated 936,084 suspicious FinTech loans totaling \$21.6B. This means FinTech lenders originated 61.8% of flagged loans despite,

³⁶In addition, the EIDL > PPP flag also provides an indication of misreporting in the EIDL and EIDL Advance program. In particular, 207,595 EIDL Advances (10.3% of those matched to a PPP loan), totaling \$1.79B, have potential misreporting.

substantially outpacing their overall FinTech 33.7% market share of loans.³⁷ As a share of loans originated by each lender type, 7.4% of traditional loans have at least one of the primary suspicious loan flags compared to 25.2% for non-bank FinTech and 20.7% for online bank FinTech.

While some of the loans flagged as suspicious by the primary measures may be sincere mistakes or errors in the data, these four measures also surely miss many fraudulent loans. This is particularly true for the business registry and EIDL > PPP Jobs flags, which only apply to subsets of loans (corporate/LLC loans and loans with matched EIDL Advances, respectively).³⁸ Thus, despite having much higher rates for these flags within the relevant subsets of loans, these flags are relatively uncommon overall, especially for FinTech lenders. As a more lenient measure of suspicious lending that is less sensitive to these restrictions, Panel A of Figure IA.10 plots suspicious loan rates including all loans with any primary or secondary flag. By this measure, 5,826,006 loans totaling \$300B are suspicious with 2,679,848 suspicious FinTech loans (46.0% of flagged loans) totaling \$61.5B.

As a more conservative estimate, we consider loans that have at least one primary flag plus an additional primary or secondary flag. While this measure almost certainly misses considerable misreporting, it has the benefit of dropping sincere mistakes or errors in the data that are isolated to a single measure. Under this more conservative measure, 1,057,613 loans totaling \$35.4B are suspicious. Of these loans, 777,490 (\$17.5B) are FinTech. This is an even larger FinTech share than for the primary measure because 83.1% of FinTech loans with a primary flag are further confirmed by an additional flag while the corresponding figure is only 45.2% for traditional loans. The higher confirmation rate for FinTech loans is consistent with flagged FinTech loans being far more likely to be fraudulent as opposed to simply reflecting honest explanations or errors in the data.

The last three bars of Panel A plot suspicious lending rates by rounds of the program with the clear pattern that suspicious lending increased over time. In round 1, 6.4% are suspicious, compared to 7.9% in round 2 and 17.1% in round 3. The conservative measure with an additional confirmatory flag follows the same pattern.

In Panel B of Figure 10, we plot suspicious loan rates by lender. The total height of the bars plots the percent of loans with at least one primary flag, and the solid part of the bars plots the percent of loans with a primary flag that is confirmed with an additional primary or

³⁷FinTech represents a larger share of suspicious loans than suspicious loan dollar volume because FinTech loans tend to be smaller. The same pattern is reflected in FinTech overall market share, which is 33.7% of PPP loans and 12.4% of PPP dollar lending volume.

³⁸Corporation and LLC loans constitute 18.5% and 64.5% of FinTech and traditional loans, respectively, and 17.8% and 26.7% of FinTech and traditional loans have a matched EIDL Advance, respectively.

secondary flag. Average rates for the two measures are plotted as solid and dashed horizontal lines, respectively. Disparities across lenders are striking. Using the at least one primary flag measure, 13 out of 21 FinTech lenders have above average suspicious loan rates, and the 10 lenders with the most suspicious loans (eight of which are FinTech) all have at least a quarter of their loans implicated compared to the overall average of 12.9%. In the extreme, Lendistry, Capital Plus, and Prestamos have primary flag rates of 34.3%, 32.9%, and 30.7%, respectively. Even with the more conservative measure, requiring an additional primary or secondary flag, these three lenders have flag rates of 30.0%, 31.0%, and 29.2%, respectively. Prestamos is particularly striking because it is largest lender overall with 495,545 loans. Cross River (second largest FinTech lender and third largest overall lender with 479,869 loans) and Harvest (fourth largest FinTech lender and sixth largest overall lender with 433,305 loans) are also well above the average flag rate with primary flag rates of 20.2% and 28.0%, respectively. While most of the FinTech lenders cluster among the lenders with the most suspicious loans, there are a few exceptions. In particular, Square, Capital One, and Intuit have misreporting rates that are well under the average misreporting rates across all lenders.

4.1 Geography of Suspicious Lending

In addition to varying across lenders, suspicious lending also varies geographically. Panel A of Figure 11 plots the percent of loans with at least one primary flag in each county across the U.S. with considerable variation.³⁹ Areas with a particularly high percentage of flagged loans cluster near New Orleans, Atlanta, and surrounding areas in Louisiana, Mississippi, Georgia. Chicago and parts of South Carolina also exhibit elevated levels. Many counties in these areas have suspicious lending rates in excess of 25% whereas large parts of the country have suspicious loan rates under 10%. The geographic pattern is somewhat regional with elevated fraud rates in the Southeast, but there are elevated counties scattered across the country. There are also big differences across large cities. For example, Cook County, IL has a suspicious loan rate of 31.7% compared to suspicious loan rates of 8.8% in New York County and 6.1% in Los Angeles County.

In Panel B of Figure 11, we examine the relation between FinTech market share and suspicious loan rates across counties and zip codes. Each dot represents a zip code. The horizontal axis plots the percent of loans flagged at the county level, and the vertical axis plots the percent of loans flagged at the zip code level. There is significant variation across zip codes within counties, with flagged loan rates varying from 20% to 50% in many counties. Additionally, FinTech market share (represented by the color of the dots) is strongly related

³⁹Panel A Figure 11 shows geographic variation in FinTech market share.

to the percent of flagged loans not only across counties, but also across zip codes within counties; zip codes with the highest flagged loan rates consistently have the highest FinTech market share.⁴⁰

Is geographic variation in suspicious lending related to poverty, crime, or culture? Or does suspicious lending cluster in other ways? Table V further analyzes the geography of suspicious PPP lending by considering relations with demographic and cultural measures that are associated with other forms of financial misconduct (Grullon et al. 2010; Parsons et al. 2018; Griffin et al. 2019). The dependent variable is an indicator for whether a loan is flagged by at least one primary flag and the explanatory variables are county-level cultural and demographic measures.⁴¹ In column (1), public corruption convictions and religious affiliation, have a positive relation with the probability of a loan being flagged as suspicious and usage of a marital infidelity website (Ashley Madison) has a negative relation. The strongest relation is for the public corruption measure. A one standard deviation increase in per capita public corruption convictions is associated with a 1.23 ppt increase in the suspicious loan rate, which is 9.5% of the mean. The other two cultural variables are economically much less important. In column (2), we add county-level demographic control variables for population density, median income, percentage of the population that is non-white, percentage of adults who are college educated, and pre-pandemic unemployment. Suspicious lending rates decrease with population density, median income, and the percentage of college educated adults and increase with the percentage of the population that is non-white and pre-pandemic unemployment. Coefficients on all but the percentage of the population that is non-white are economically small.

In column (3), we add county-level FinTech market share. A one standard deviation increase (15.7 ppt) in FinTech market share in the county is associated with a 2.69 ppt increase in the suspicious loan rate, which is 20.9% of the mean misreporting rate. This is a much stronger relation than any of the other county variables, and coefficients for the cultural and demographic variables generally decrease or become statistically insignificant once FinTech market share is added to the regression. While this regression specification is not conducive to a causal interpretation, it indicates a strong FinTech association and indicates that cultural variables play a minor role.

⁴⁰Within a county, a 10 ppt rise in FinTech market share in a zip code is associated with a 3.21 ppt rise in suspicious lending (see Table IA.VII).

⁴¹The regressions are at the loan level to control for jobs reported, loan size, business type, and industry code \times state fixed effects. Standard errors are double clustered by zip code and lender. The independent variables are standardized to have a mean of 0 and a standard deviation of 1 at the county level. Thus, the coefficients can be interpreted as the change in the probability of a loan being flagged for a one standard deviation change in the variable. Table IA.VIII shows equivalent regressions at the county level.

Why does suspicious lending vary so much across geographies? Strong clustering in certain counties and zip codes suggests that suspicious borrowing is driven by more than just the idiosyncratic decisions of individual borrowers. One possibility is that referral fee programs, agent fees, kickback schemes, or local networks may arise in certain areas to systematically attract and facilitate suspicious lending.⁴² Because we do not observe the identity of agents directing or assisting the PPP borrowers, this possibility is difficult to directly test. Nonetheless, the geographic clustering of suspicious loans shown in Figure 11 is what one would expect from agents steering suspicious borrowers to FinTech lenders. If agents are utilizing more than one FinTech lender to originate suspicious loans, one might expect counties with many potentially misreported loans by one lender to have elevated levels of suspicious loans and FinTech lending more generally. Consistent with this premise, when one FinTech lender has a high flagged loan rate in a county, other FinTech lenders also tend to have elevated flagged loan rates (as shown in the lower triangle of Figure IA.12). Additionally, FinTech lenders with the highest overall flagged loan rates have a high correlation in their market shares at the county level, whereas market-share correlations between most other lenders are slightly negative (as shown in the upper triangular of Figure IA.12).

5 Why Does Suspicious Lending Concentrate in Fin-Tech?

FinTech lenders on average have much higher suspicious lending rates than traditional lenders, and Tables III, IA.II, and IA.VI show that their elevated suspicious lending is not explained by observable facets of loan composition. What could be driving the elevated flag rates for FinTech lenders?

5.1 FinTech Lender Background and Incentives

5.1.1 FinTech Lender Background

Differences across FinTech lenders give a first clue to this puzzle. While most FinTech lenders have high suspicious loan rates, Square and Intuit have among the lowest suspicious loan rates of all lenders. Online lending does not appear to be the problem in and of itself. One thing that sets Square and Intuit apart is that they have established relationships with customers based on a broad suite of payment, accounting, payroll, and other financial support services.

⁴²For example, Amur Equipment offered a referral fee program and explicitly stated that the referral program required “zero-touch and follow up on your end.” (see Tweet [here](#)). The PPP allowed lenders to pay agents 1% on loans up to \$350k, 0.5% on loans between \$350k and \$2M, and 0.25% on loans above \$2M (see instructions [here](#)). Further, some people filed PPP loans in return for upfront and backend fees or kickbacks, which was against SBA rules (e.g., see [here](#), [here](#), and [here](#)).

By contrast, the largest FinTech PPP lender, Prestamos, is a Community Development Financial Institution with locations in Arizona, Nevada, and New Mexico. The second largest FinTech lender, Cross River, is a small community bank in New Jersey that acts as a conduit for partner FinTechs. Similarly, Capital Plus Financial, the third largest FinTech PPP lender, is a small mortgage lender in Texas that traditionally focused on supporting Hispanic homeownership but now appears to be almost entirely focused on PPP lending. The number four FinTech lender, Harvest Small Business Finance, is also a small lender with limited history. Now that the PPP has ended, Harvest’s only current product appears to be SBA 7(a) commercial real estate loans. Benworth and Fountainhead, the other FinTech lenders in the top ten by number of PPP loans originated, follow a similar pattern with limited business outside of PPP lending. We systematically examine this relation and find that lenders who have fewer SBA loans pre-pandemic, have lent in SBA programs for fewer years, and for whom the PPP was their first experience with SBA lending (in particular new FinTechs) all have higher rates of flagged loans (as shown in Table IA.IX). FinTech lenders also relied more heavily on liquidity support from the Federal Reserve than traditional lenders.⁴³

The six largest FinTech lenders primarily originated loans that were sourced from other FinTech platforms. Cross River adopted this business model early in round 1 by partnering with other FinTechs such as Intuit and Kabbage to originate PPP loans (New York Times 2020). The other large FinTech lenders originated loans sourced by two marketing FinTechs that did not do any PPP lending until round 3, Womply and BlueAcorn (New York Times 2021a). Womply is a marketing technology firm with no lending history before participating in the PPP. It launched a platform called Fast Lane to facilitate PPP applications that were then originated by partner lenders including Harvest, Capital Plus, Benworth, and Fountainhead. BlueAcorn was founded in April 2020 exclusively to source PPP loans in partnership with Capital Plus and Prestamos. Both firms relied heavily on online advertising promoting easy access to PPP money.

5.1.2 FinTech Fluidity

To examine relationships between these lenders, Panel A, Figure 12 plots the network of relationships between lenders based on originating loans in the same draw to the same non-commercial address as identified by the multiple loan flag. The edges between lenders represent the number of addresses to which both lenders originated a loan within the same

⁴³Financing for FinTech PPP loans was in part provided with a credit facility, the Paycheck Protection Program Liquidity Facility (PPPLF), in which the Federal Reserve extended credit to lenders using PPP loans as collateral. Figure IA.13 plots flagged loan rates relative to PPPLF advance volume for the 100 largest PPP lenders. Whereas most traditional lenders did not use the PPPLF at all, it was a major source of funding for some of the large FinTech lenders.

draw. Node size is based on the number of loans at addresses flagged for having multiple loans. The thicker edges between FinTech lenders show that FinTech borrowers with multiple loans often received funds from more than one FinTech lender even within the same draw. Specifically, 61.2% of FinTech borrowers with multiple loans to the same address split their loans across multiple lenders. Shared FinTech lending to the same address is largely explained by FinTech portals sourcing loans for multiple lenders. For example, the plot shows a strong relationship between Prestamos and Capital Plus, the two lenders that partnered with BlueAcorn. This is likely from borrowers applying for multiple loans through BlueAcorn, some of which were originated by Prestamos while others were directed to Capital Plus. Similarly, there are strong relationships between Harvest, Benworth, Capital Plus, and Fountainhead, all of which are Womply partners. By contrast, for traditional lenders 81.4% percent of traditional borrowers who took out multiple loans received all their loans from the same lender. Some exceptions include Customers Bank, Amur Equipment, and Bank of America. Customers Bank has known FinTech affiliations even though it does not meet the formal FinTech criteria and Amur Equipment went from specializing in equipment financing to becoming one of the largest PPP lenders by advertising their “lighting fast portal” on social media.⁴⁴

As another way to examine relationships between lenders, we track borrowers switching lenders between their first and second PPP loan draws. If a borrower already received a first draw from the same lender, obtaining second draw only required refreshing the application with some additional information. This provided a strong incentive for borrowers to use the same lender. For borrowers with first draw loans flagged for potential misreporting that subsequently took out second draw loans, Panel B of Figure 12 shows the movement of these loans across rounds to different lenders.⁴⁵ The large movements and connections between FinTech lenders likely reflect online lending portals switching lenders. Overall, the graphs highlight the fluid nature of the FinTech space where online portals originating loans can easily originate their loans through different lenders, and suspicious borrowers can utilize several platforms or switch platforms. The lack of relationship banking within the FinTech space may be advantageous to expand access to capital (Erel and Liebersohn 2021), but it also appears to be expedient for dubious lending.

⁴⁴Customers Bank directly worked with multiple FinTech lenders, in particular Kabbage and Cross River (see press release [here](#)). See one of many posts on social media by Amur Equipment [here](#).

⁴⁵The thickness of the edges between lenders is proportional to the number of flagged loans that changed lenders between the first and second draws. The switches between the first and second draw are clockwise. Node size is based on the number of first draw loans (with a matching second draw) and second draw loans by each lender. Figure IA.14 replicates this figure using all second draw loans.

5.1.3 FinTech Revenue

PPP lending had the potential to be a profitable business for lenders. Lenders were initially compensated with processing fees of 5% for loans up to \$350,000, 3% for loans between \$350,000 and \$2,000,000, and 1% for loans of \$2,000,000 or more. For loans made in 2021, fees for small loans were increased to the lesser of 50% or \$2,500 for loans below \$50,000.⁴⁶ Based on this fee schedule, we estimate that PPP lending generated \$38.8B of lender processing fees, \$9.4B of which went to FinTech lenders (see Table [IA.X](#)). The top four FinTech lenders alone likely generated \$4.4B in processing fees, including \$1.18B to Prestamos, \$1.10 to Capital Plus, \$1.06B to Harvest, and \$1.04B to Cross River. The average processing fee for FinTech PPP loans was 18.3% of the loan balance, largely driven by the high processing fees for small loans in round 3.⁴⁷ We lack data on cost structure associated with PPP lending and do not observe how lender fees are shared with partner organizations used to source the loans such as Womply and BlueAcorn.⁴⁸

5.2 Did FinTech Lenders Improve Standards Over Time?

We consider two potential scenarios under which suspicious lending could arise:

- Scenario A: The lender does not want to facilitate fictitious loans but is not performing great due diligence. As it learns over time, the lender cracks down on the fraud.
- Scenario B: The lender is aware of the existence of or potential for fraud within its PPP loans but ignores this risk because there is little downside for the lender. This may be particularly true for lenders with little reputation or other business to protect.

Under scenario A, when lenders are new to PPP lending they may facilitate questionable loans, but over time as they experience more loans with improbable features, they should originate fewer of these loans. In this case, borrowers who wish to commit loan fraud would need to rotate among lenders. In scenario B, the amount of suspicious lending could grow through time as lenders develop a reputation for rapid and unquestioning approval.

Scenario A predicts:

1. Loan misreporting will decrease over time as lenders become more aware and develop

⁴⁶See fee schedule [here](#).

⁴⁷The average FinTech processing fee for round 1 and 2 was 4.96%. This dramatically increased to 21.7% in round 3.

⁴⁸Capital Plus (the second largest FinTech lender and fourth largest lender overall) received a PPP loan of \$376,800, reportedly to cover payroll for its 28 employees. The loan was approved in April 2020, potentially before their business opportunities as a PPP lender, most of which occurred in round 3, were apparent. Similarly, Benworth Capital Partners (fourth largest FinTech and eighth largest lender overall) received a PPP loan of \$100,600 for its 13 employees on April 5, 2020, DreamSpring received a PPP loan of \$757,753 for its 54 employees on April 27, 2020, and Amur Equipment was approved for a PPP loan of \$2,817,846 on May 2, 2020 but then repaid/canceled its loan 12 days later.

systems to screen out suspicious loans.

2. Suspicious borrowers will be less likely to receive a repeat loan from the same lender compared to other borrowers.
3. Regions with high misreporting in rounds 1 and 2 will face extra scrutiny from lenders, which will decrease round 3 misreporting.

Scenario B predicts:

1. Loan misreporting will grow over time as borrowers learn about the potential for fraud.
2. Borrowers with suspicious first draw loans in rounds 1 and 2 will be able to obtain second draw loans in round 3 from the same lenders.
3. Regions with high misreporting in rounds 1 and 2 will have the same or more misreporting in round 3.

Did lenders improve their loan screening over time? While we do not observe denied applications or specific lender practices, we can observe how loans that were approved and funded changed over time. We have already seen that the overall rate of suspicious lending grew over time from round 1 to round 3. Panel A of Figure 13 plots more granular suspicious loan rates on a weekly basis separately for non-bank FinTech, online bank FinTech, and traditional lenders. For the FinTech lenders, loans became more suspicious over time throughout rounds 1 and 2. The rate of suspicious lending dropped at the beginning of round 3, likely due to pent up demand for second draw loans from legitimate borrowers. Most round 3 FinTech lending occurred later in round 3 (see Figure 1), and as round 3 progressed, the suspicious loan rate rose dramatically to around 30% of loans flagged as suspicious during April and early May of 2021. PPP funds for most loans were exhausted on May 4, 2021 (New York Times 2021b). The suspicious loan rate fell after this date, but this could be due to loan composition since funding after May 4 was only available for prioritized community financial institutions and some loans that were already under review prior to May 4. Suspicious lending by traditional lenders also grew over time, but at a much lower rate. FinTech and traditional lenders both started the PPP with suspicious loan rates of around 10%, but by the end of the program the FinTech suspicious loan rate was close to 30%, more consistent with scenario B.⁴⁹

We also examine lending growth and changes in suspicious loan rates across rounds at the lender level. Panel A of Figure IA.16 shows that most lenders had higher suspicious lending rates in round 3 than in rounds 1 and 2. Additionally, many of the FinTech lenders with the

⁴⁹Panel B of Figure IA.10 shows similar trends for each primary flag individually.

highest suspicious loan rates in rounds 1 and 2 also had the most growth in lending and the most growth in suspicious loans in round 3. In Table VI, we ask whether lenders appear to be learning by regressing indicators for the four primary flags in round 3, individually and combined, on lenders' rounds 1 and 2 misreporting rates for the same flags. As in previous regressions, we control for loan size and jobs with zip code, business type, industry \times CBSA fixed effects. For FinTech lenders, we find highly economically and statistically significant positive relations across the board with largely insignificant and one negative relationship for high implied compensation for traditional lenders. For traditional banks there is overall no relationship between the lenders suspicious loan percentages in rounds 1 and 2 and their lending in Round 3. By contrast, FinTech lenders have persistent and increasing levels of suspicious loans through time, consistent with scenario B above.

To assess prediction 2, if lenders are taking steps to screen out questionable loans, then their borrowers with questionable first draw loans in rounds 1 and 2 may get rejected when they apply for a second draw loan in round 3. To examine this, we estimate regressions to determine whether a first draw borrower is more or less likely to receive a second draw loan from the same lender if its first draw loan is flagged by one of the primary misreporting indicators. Table VII shows that traditional loans which are flagged in the first two rounds have a statistically significant decrease in the probability of receiving a second draw loan from the same lender of 3.10 ppt (with t-stat of -11.94) and FinTechs have a smaller and statistically insignificant decrease of 0.98 ppt (with t-stat of -0.83).⁵⁰ This provides some indication that traditional banks were less likely to continue lending to borrowers with previous suspicious borrowing, but FinTechs do not seem to be screening or implementing procedures which make it less likely for questionable borrowers to continue receiving funds in the form of a second draw. Columns (3) and (4) of Table VII condition on the borrower receiving a second draw (either from the same or different lender) with similar results.⁵¹

To assess prediction 3, we examine whether areas with high misreporting in rounds 1 and 2 had higher or lower misreporting in round 3. Panel B Figure 13 plots the percentage of loans flagged in rounds 1 and 2 in each zip code on the horizontal axis, and the percentage of flagged loans in round 3 in the same zip code on the vertical axis. The left subpanel uses all loans, the middle uses FinTech loans, and the right uses traditional loans. Each dot represents a zip code, and the size of the dots corresponds to the number of loans in the

⁵⁰These results are based on $SameLender_i$ being set to 0 if the borrower did not get a second draw at all.

⁵¹In Figure IA.15, we show results separately for individual lenders with lender fixed effects and lender interactions. The inclusion of the lender fixed effects ensures that the reported coefficient is due solely to differences in the lender's behavior towards flagged and nonflagged loans rather than systematic changes in the lender's behavior. For most traditional lenders, borrowers with a flagged first-draw loan are less likely to receive a second draw loan from the same lender, but for several FinTech lenders, suspicious first-draw borrowers are slightly more likely to receive a second draw.

zip code. Purple to blue colors indicate that a zip code had fewer loans in round 3 than in rounds 1 and 2, yellow colors indicate that a zip code had about the same number of loans in round 3 compared to rounds 1 and 2, and orange to red colors correspond to an increase in the number of loans in the zip code.

The figure displays three interesting findings. First, most zip codes (85.4%) are above the 45-degree line in the left subpanel, indicating that misreporting rates increased in round 3 almost everywhere. In many zip codes (35.0%), the percentage of loans that are flagged as suspicious in round 3 is more than twice as high as in rounds 1 and 2. Second, the zip codes with the highest suspicious loan rates experienced the most growth in lending. Many zip codes with the highest level of flagged loans in round 3 have more than three times the number of loans in round 3 compared to rounds 1 and 2, suggesting that significant portions of zip code-level loan growth in round 3 may be due to suspicious lending practices. Third, the middle and right subpanels differentiate between FinTech and traditional lenders and show that lending growth and increased misreporting rates are almost entirely from FinTech lenders. Traditional lenders had only small increases in suspicious loan indicators, and their lending generally decreased. In contrast, FinTech lenders increased the number of loans they originated and increased their suspicious lending rates in almost all zip codes. Additionally, FinTech lending growth was highest in zip codes with the highest misreporting rates.⁵² We also test these results at the zip code-lender level in Table IA.XI with zip code and lender fixed effects and find that a 10 ppt increase in flagged loans in a zip code-lender pair in rounds 1 and 2 is associated with a 22.1 ppt increase in lending for a FinTech lender and an insignificant increase of 1.3 ppt for a traditional lender. There is also strong persistence of suspicious lending across rounds within zip code-lender pairs.

5.3 Repayments and Enforcement Actions

The economics of crime depends crucially on a crime’s expected penalty and probability of detection (Becker 1968). The US Department of Justice is pursuing criminal complaints alleging PPP fraud, and some borrowers have voluntarily repaid their loans without applying for loan forgiveness or had their loan canceled.⁵³ However, the magnitude of these enforcement actions is tiny. Compared to the 2.3 million loans we identify as suspicious, the DOJ

⁵²Results are similar at the county and state level. See Panels C and D of Figure IA.16. Panel B of Figure IA.11 also plots lending growth by county.

⁵³An earlier version of this paper based on the May 3, 2021 SBA data release was strongly criticized by several PPP lenders for including loans that were eventually canceled. Dropping the canceled loans led to only minuscule reductions in suspicious loans. The current version of the paper is entirely based on the June 30, 2021 SBA data release and thus does not include these canceled loans.

has publicized 162 criminal complaints regarding only 355 loans.⁵⁴ SBA data indicates that only 16,930 round 1 and 2 loans were repaid between December 1, 2020 and June 30, 2021. Repayment, enforcement action, and cancellation rates are all elevated for flagged loans (see Table [IA.XII](#)). While more enforcement actions may be forthcoming, there appears to be little penalty for most suspicious lending thus far.

6 Conclusion

We examine four primary and five secondary metrics related to potentially misreported loans. FinTech loans are highly suspicious at a rate of over six times that for traditional lenders. Eight of the ten lenders with the highest rates of suspicious loans are FinTech lenders and the remaining two traditional banks function much like FinTechs, including one with a “lighting fast portal.” We estimate the total amount of potential misreporting as 1.52 million loans with a balance of \$68.9 billion based on the four primary metrics, and \$35.4 billion (1.06 million loans) under a more conservative estimate requiring an additional indicator. The total amount of misreporting is likely larger than either estimate because many of our indicators are only available for a subset of loans, and a much larger set of loans has at least one of the nine indicators. In the early stages of the PPP, less than 10% of FinTech loans were potentially misreported, but the percentage of suspicious FinTech loans increased to more than 30% by April 2021. Extremely few of the suspicious loans have been prosecuted by authorities or repaid.

Our findings have important policy implications. First, the PPP did not include robust verification requirements, but traditional banks may have been more apt to follow standard lending practices. The lack of rigorous verification, seems to have led to substantial costs to taxpayers, especially in later rounds. Second, FinTech lending, though quite successful at adapting to new environments and quickly disbursing funds, seemingly needs to improve due diligence practices. Two established FinTech lenders persistently have low rates of misreporting, indicating that FinTech lending need not be substandard. Third, our evidence, along with evidence that the PPP saved relatively few jobs at a high cost ([Autor et al. 2020](#); [Chetty et al. 2020](#); [Granja et al. 2020](#)), provides growing evidence that the PPP may not have been an efficient source of capital allocation. Fourth, incentives in the PPP appear misaligned in that FinTech lenders with widespread indicators of misreporting made billions of dollars dispersing loans with lax oversight procedures.

⁵⁴See the Internet Appendix for additional details on the repayment and enforcement action data. Of the DOJ enforcement action loans with enough data to be matched to the PPP loan-level data, 153 loans were originated by FinTech lenders and 126 were originated by traditional banks. There are likely other cases that are still sealed, are in early stages of investigation, or are not included on the DOJ website for other reasons. We focus on loans from rounds 1 and 2 for this analysis to allow more time for repayments and enforcement actions.

Finally, the increasing scale of misreporting through time indicates that current penalty and enforcement systems are not effective. If the system is not changed, the most likely outcome is even more of the same. This paper is also an example of how forensic research ([Zitzewitz 2012](#)) can more fully investigate the rent-seeking dimension of finance ([Zingales 2015](#)). Government agencies can assist this transparency goal by making detailed data available to the public. We hope to see future research with additional forensic investigation of the PPP as well as other recent government and private lending programs.

References

- Alper, Mariel, Matthew R. Durose, and Joshua Markman, 2018, 2018 update on prisoner recidivism: A 9-year follow-up period (2005-2014), Special Report by Bureau of Justice Statistics: Office of Justice Programs.
- Amiram, Dan, and Daniel Rabetti, 2020, The relevance of relationship lending in times of crisis, Working Paper.
- Atkins, Rachel M. B., Lisa D. Cook, and Robert Seamans, 2021, Discrimination in lending? Evidence from the Paycheck Protection Program, Working Paper.
- Autor, David, David Choz, Leland D. Cranez, Mita Goldarx, Byron Lutzz, Joshua Montesz, William B. Petermanz, David Ratnerz, Daniel Villarz, and Ahu Yildirmazx, 2020, An evaluation of the Paycheck Protection Program using administrative payroll microdata, Working Paper.
- Avis, Eric, Claudio Ferraz, and Frederico Finan, 2018, Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians, *Journal of Political Economy* 126, 1912–1964.
- Bandiera, Oriana, Andrea Prat, and Tommaso Valletti, 2009, Active and passive waste in government spending: Evidence from a policy experiment, *American Economic Review* 99, 1278–1308.
- Bartik, Alexander W., Marianne Bertrand, Zoe Cullen, Edward L. Glaeser, Michael Luca, and Christopher Stanton, 2020, The impact of COVID-19 on small business outcomes and expectations, *Proceedings of the National Academy of Sciences* 117, 17656–17666.
- Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, 2021, Consumer-lending discrimination in the FinTech era, *Journal of Financial Economics*.
- Becker, Gary S., 1968, Crime and punishment: An economic approach, *Journal of Political Economy* 76, 169–217.
- Beggs, William, and Thuong Harvison, 2021, Fraud and abuse in the Paycheck Protection Program? Evidence from investment advisory firms, Working Paper.
- Begly, Taylor A., Umit G. Gurun, Amiyatosh Purnanandam, and Daniel Weagley, 2021, Disaster lending: “fair” prices, but “unfair” access, Working Paper.
- Ben-David, Itzhak, 2011, Financial constraints and inflated home prices during the real estate boom, *American Economic Journal: Applied Economics* 3, 55–87.
- Bloomberg Businessweek, 2020, PPP scammers made fintech companies their lenders of choice, Michelle F. Davis, October 7.
- Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.
- Chernenko, Sergey, and David S. Scharfstein, 2021, Racial disparities in the Paycheck Protection Program, Working Paper.

- Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team, 2020, The economic impacts of COVID-19: Evidence from a new public database built using private sector data, Working Paper.
- Deflo, Esther, 2017, The economist as plumber, *American Economic Review* 107, 1–26.
- Duchin, Ran, Xiumin Martin, Roni Michaely, and Ivy Wang, 2021, Concierge treatment from banks: Evidence from the Paycheck Protection Program, Working Paper.
- Erel, Isil, and Jack Liebersohn, 2021, Does fintech substitute for banks? Evidence from the Paycheck Protection Program, Working Paper.
- Faulkender, Michael W., Robert Jackman, and Stephen Miran, 2021, The job preservation effects of Paycheck Protection Program loans, Working Paper.
- Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, 2019, The role of technology in mortgage lending, *The Review of Financial Studies* 32, 1854–1899.
- Garmaise, Mark J., 2015, Borrower misreporting and loan performance, *The Journal of Finance* 70, 449–484.
- Glaeser, Edward L., and Claudia Goldin, 2006, *Corruption and Reform: Lessons from America's Economic History* (University of Chicago Press).
- Glancy, David, 2021, Bank relationships and the geography of PPP lending, Working Paper.
- Gopal, Manasa, and Philipp Schnabl, 2020, The rise of finance companies and fintech lenders in small business lending, Working Paper.
- Granja, João, Christos Makridis, Constantine Yannelis, and Eric Zwick, 2020, Did the Paycheck Protection Program hit the target?, Working Paper.
- Griffin, John M., Samuel Kruger, and Gonzalo Maturana, 2019, Personal infidelity and professional conduct in 4 settings, *Proceedings of the National Academy of Sciences* 116, 16268–16273.
- Griffin, John M., and Gonzalo Maturana, 2016, Who facilitated misreporting in securitized loans?, *The Review of Financial Studies* 20, 384–419.
- Grullon, Gustavo, George Kanatas, and James Weston, 2010, Religion and corporate (mis)behavior, Working Paper.
- Hanson, Samuel G., Jeremy Stein, Adi Sunderam, and Eric Zwick, 2020, Business credit programs in the pandemic era, *Brookings Papers on Economic Activity* 3–60.
- Hart, Oliver D., Andrei Shleifer, and Robert W. Vishny, 1997, The proper scope of government: Theory and an application to prisons, *Quarterly Journal of Economics* 112, 1127–1261.
- Howell, Sabrina T., Theresa Kuchler, David Snitkof, Johannes Stroebe, and Jun Wong, 2021, Racial disparities in access to small business credit: Evidence from the Paycheck Protection Program, Working Paper.

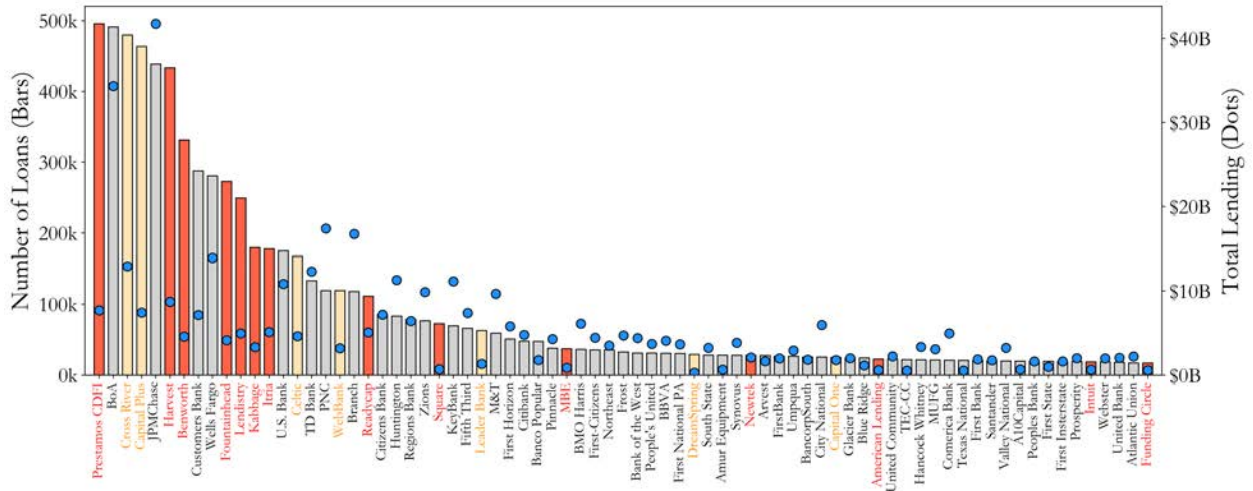
- Jagtiani, Julapa, and Catharine Lemieux, 2019, The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform, *Financial Management* 48, 1009–1029.
- Jiang, Wei, Ashlyn A. Nelson, and Edward Vytlačil, 2014, Liar’s loan? effects of origination channel and information falsification on mortgage delinquency, *The Review of Economics and Statistics* 96, 1–18.
- Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? Evidence from subprime loans, *The Quarterly Journal of Economics* 125, 307–362.
- Kruger, Samuel, and Gonzalo Maturana, 2020, Collateral misreporting in the residential mortgage-backed security market, *Management Science* 67, 2657–3320.
- Li, Lei, and Philip E. Strahan, 2021, Who supplies PPP loans (and does it matter)? Banks, relationships and the COVID crisis, *Journal of Financial and Quantitative Analysis, Forthcoming*.
- Meirer, Jean-Marie, and Jake Smith, 2021, The COVID-19 bailouts, Working Paper.
- Miami Herald, 2020, Copycat websites, newly minted companies, double dippers cashed in on PPP loans, Ben Wieder and Meghan Bobrowsky, September 11.
- Mian, Atif R., and Amir Sufi, 2017, Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005, *Review of Financial Studies* 30, 1832–1864.
- Neilson, Christopher, John Eric Humphries, and Gabriel Ulyssea, 2020, Information frictions and access to the Paycheck Protection Program, *Journal of Public Economics* 190.
- New York Times, 2020, The tiny bank that got pandemic aid to 100,000 small businesses, Stacy Cowley, June 23.
- New York Times, 2021a, How two start-ups reaped billions in fees on small business relief loans, Stacy Cowley and Ella Koeze, June 27.
- New York Times, 2021b, The Paycheck Protection Program is out of money., Stacy Cowley, May 4.
- Parsons, Christopher A., Johan Sulaeman, and Sheridan Titman, 2018, The geography of financial misconduct, *The Journal of Finance* 73, 2087–2137.
- Piskorski, Tomsaz, Amit Seru, and James Witkin, 2015, Asset quality misrepresentation by financial intermediaries: Evidence from the RMBS market, *The Journal of Finance* 70, 2635–2678.
- Project on Government Oversight, 2020, Lamborghinis, strip clubs, bogus companies, and lies: The first 56 Paycheck Protection Program fraud cases, Ryan Summers and Nick Schwellenbach, October 8.

- ProPublica, 2021, Hundreds of PPP loans went to fake farms in absurd places, Derek Willis and Lydia DePillis, May 18.
- Purnanandam, Amiyatosh, 2011, Originate-to-distribute model and the subprime mortgage crisis, *Review of Financial Studies* 24, 1181–1915.
- Shleifer, Andrei, and Robert W. Vishny, 1993, Corruption, *The Quarterly Journal of Economics* 108, 599–617.
- Svensson, Jakob, 2003, Who must pay bribes and how much? evidence from a cross section of firms, *Quarterly Journal of Economics* 118, 207–230.
- The Wall Street Journal, 2020, Evidence of PPP fraud mounts, officials say, Ryan Tracy, November 8.
- Zingales, Luigi, 2015, Presidential address: Does finance benefit society?, *The Journal of Finance* 70, 1327–63.
- Zitzewitz, Eric, 2012, Forensic economics, *Journal of Economic Literature* 50, 731–769.

Figure 1. Fintech Market Share

This figure shows the role that fintech lenders played in the PPP. Panel A shows the number of loans (bars) and dollar value of loans (dots) originated by the top 75 lenders (by number of loans). Panel B shows the percentage of loans originated by fintech lenders during each week of Round 1, 2, and 3 of the PPP on the left axis and the total number of loans originated each week on the right axis. In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders. Note that mid-August through December 2020 is not shown in Panel B since no PPP loans were originated during this period.

Panel A. Number of Loans and Dollar Value of Loans, by Lender (Top 75)



Panel B. Fintech Market Share, by Week

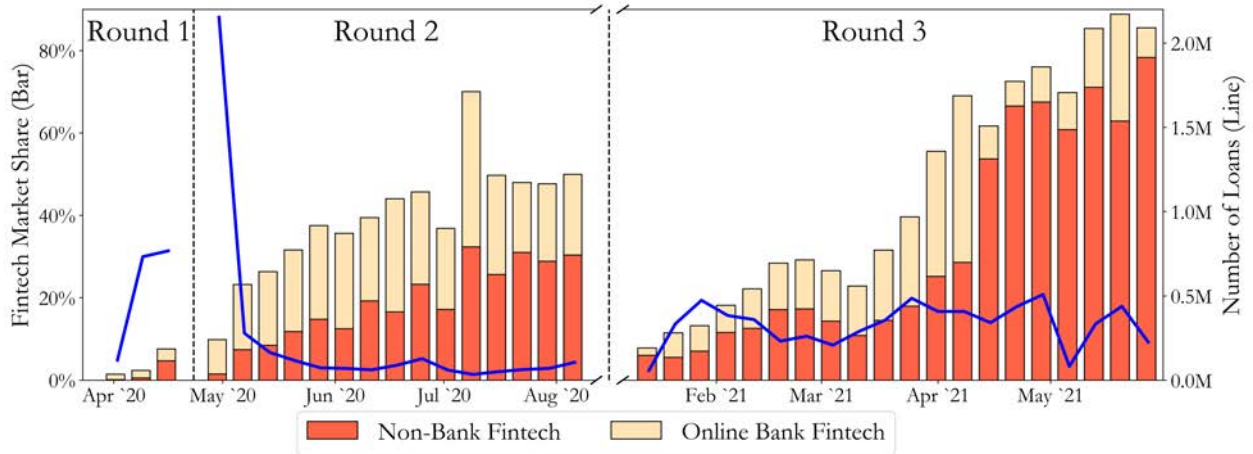
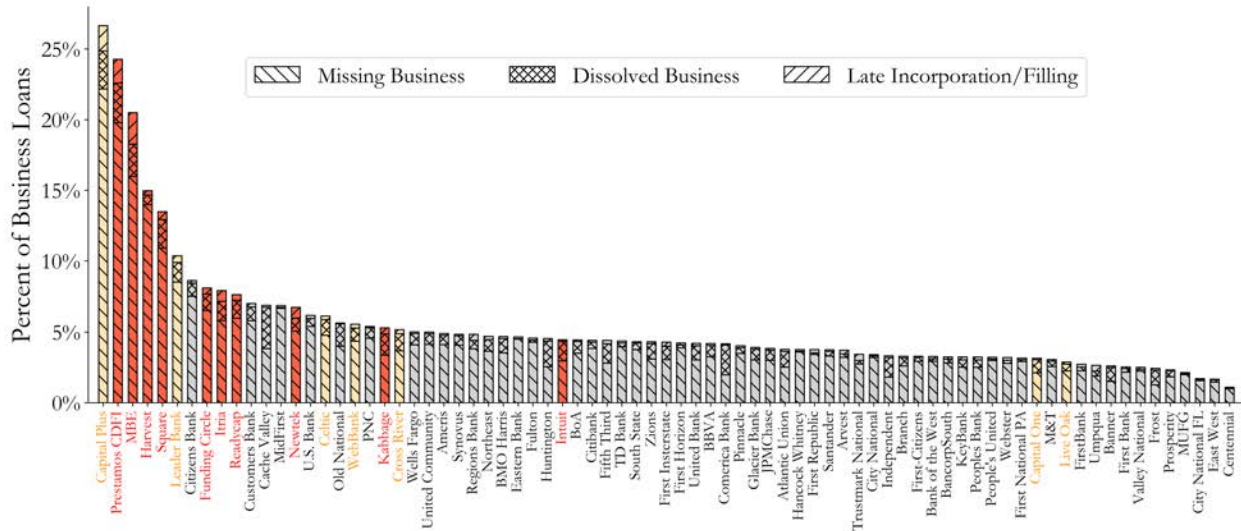


Figure 2. Business Registry and Multiple Loans Flags

This figure shows the prevalence of the business registry and multiple loans flagged loans by lender. Panel A shows the percentage of loans flagged for being incorporated after February 15, 2020 (“Late Incorporation/Filing”), being dissolved and inactive before approved for a PPP loan (“Dissolved Business”), or not being found in the business registry for its home state or in any other state while listing an address in its home state (“Missing Business”). Panel B shows the percentage of loans flagged as being located at a non-business, non-central (e.g., not an apartment or office building) address that received at least three loan within the given loan’s draw (i.e., the first or second draw). For Panel A, only loans to businesses organized as a corporation, subchapter S corporation, or LLC and not based in Illinois or a territory are considered; lenders originating at least 10,000 loans fitting these criteria are shown. For Panel B, all loans are considered and lenders originating at least 15,000 loans are shown. In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders.

Panel A. Business Registry Flagged Loans, by Lender



Panel B. Multiple Loans Flagged Loans, by Lender

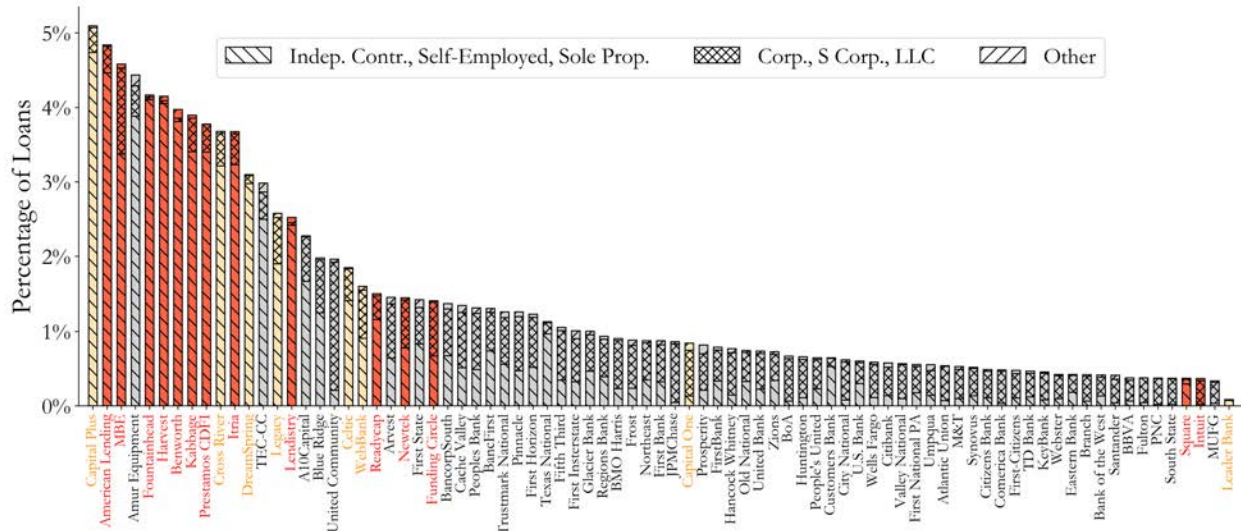
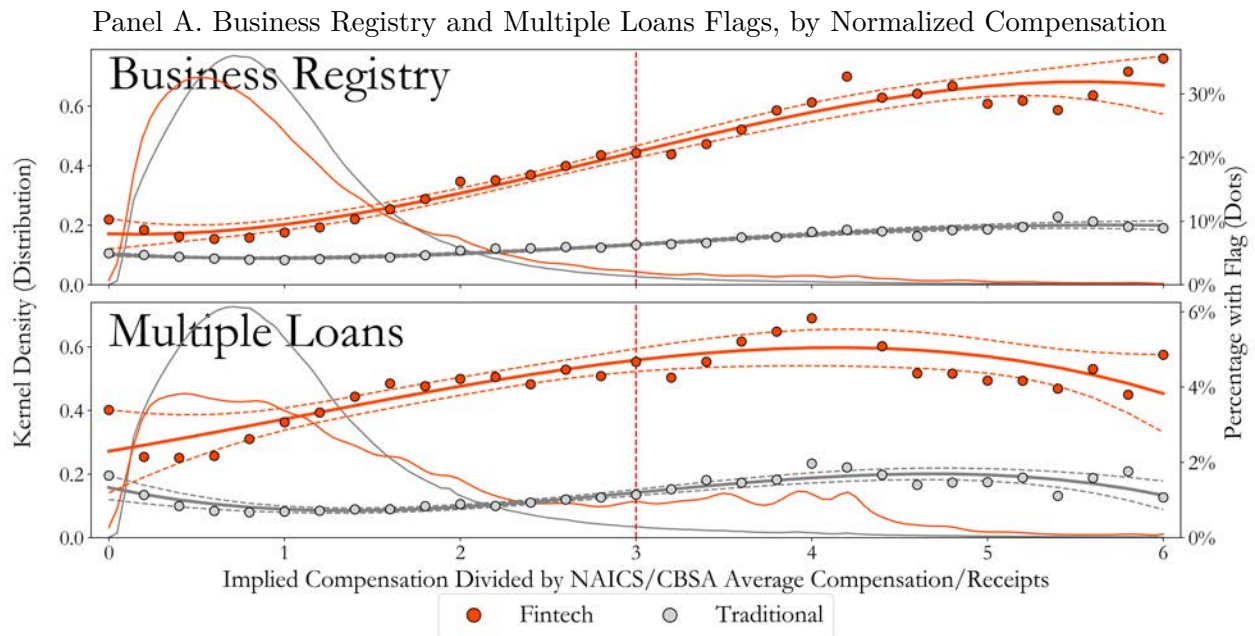


Figure 3. High Implied Compensation Flag

This figure shows the relationship between a loan's implied compensation per employee and the average compensation in the loan's industry (represented by NAICS [North American Industry Classification System] code) and region (represented by CBSA [core-based statistical area]). We define normalized compensation as the implied compensation of the loan divided by the average compensation in the loan's industry-CBSA. Panel A shows the relationship between normalized compensation and the business registry (top subpanel) and multiple loans flags (bottom subpanel). Panel B shows the percentage of loans with normalized compensation above 3 (i.e., implied compensation is more than three times the industry-CBSA average) by lender. For Panel A, the left axis shows the kernel density of loans (distribution) and the right axis shows the percentage of flagged loans in each bin (dots), where each bin is 0.2 units wide. The solid lines are third-degree polynomial fits for the percentage flagged and the dashed lines are 95% confidence intervals. For the business registry subpanel, only loans to businesses organized as a corporation, subchapter S corporation, or LLC and not based in Illinois are considered. For Panel B, only loans where the average compensation in the loan's industry-CBSA is less than \$33,333.33 are considered; lenders with at least 5,000 loans fitting this criterion are shown.



Panel B. High Implied Compensation, by Lender

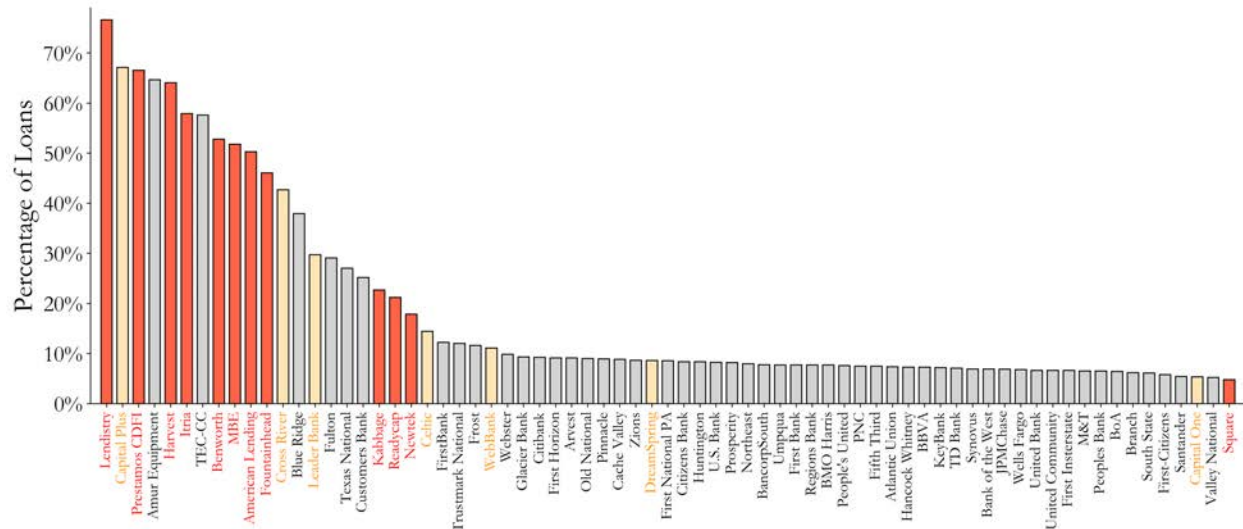
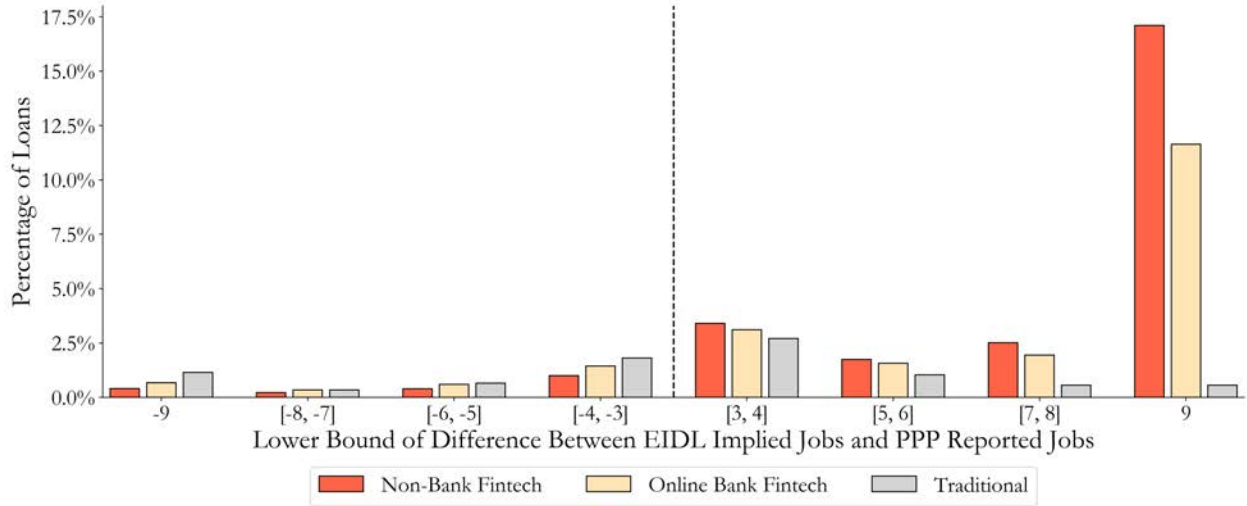


Figure 4. EIDL > PPP Jobs Flag

This figure shows the difference between the number of employees implied by a business's EIDL Advance amount ("EIDL Implied Jobs") and the number of jobs reported by the business on its PPP application ("PPP Reported Jobs"). Panel A shows the lower bound (in the absolute value sense) of the difference between the EIDL implied jobs and PPP reported jobs by lender type. Panel B shows the percentage of loans where the EIDL implied jobs is at least three more than the PPP reported jobs by lender. In both panels, only loans with a matched EIDL Advance are considered; lenders with at least 5,000 loans fitting this criterion are shown in Panel B. In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders.

Panel A. Difference Between EIDL Implied Jobs and PPP Reported Jobs



Panel B. EIDL > PPP Jobs, by Lender

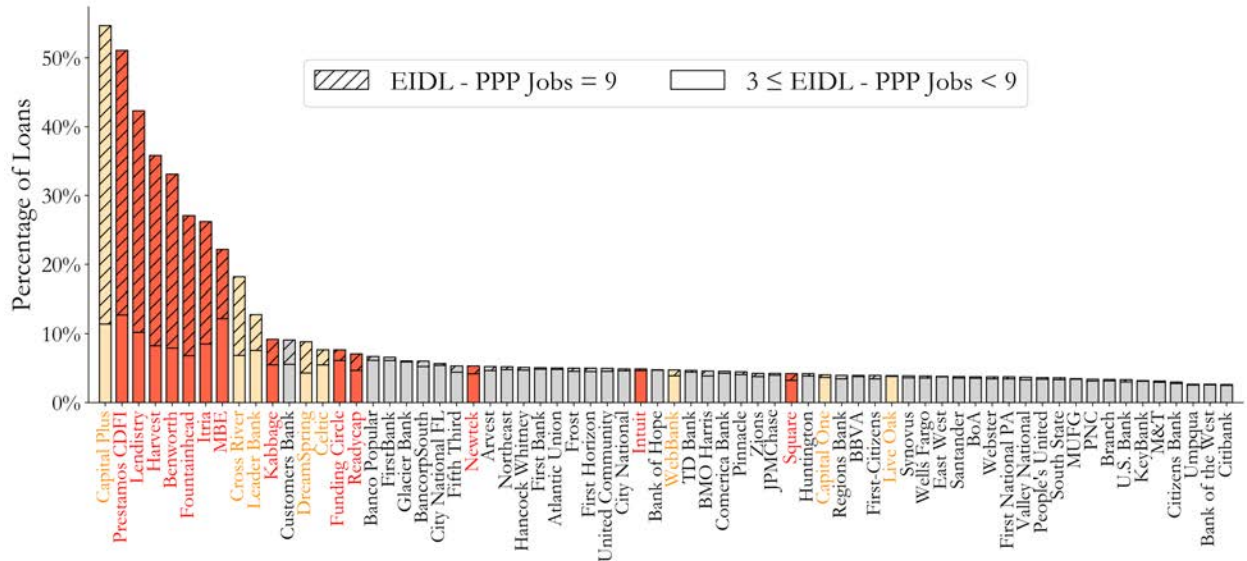
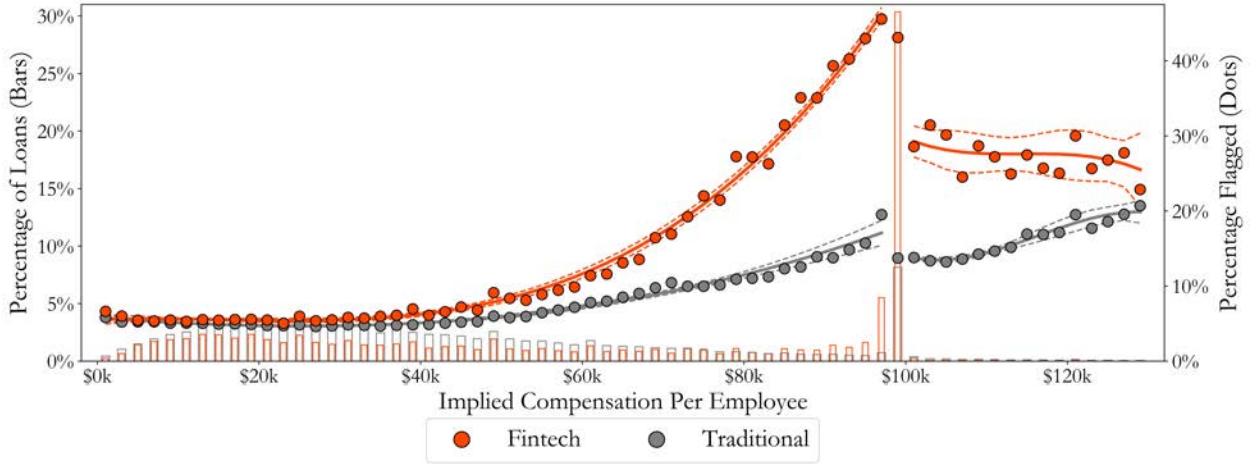


Figure 5. Discontinuities at \$100,000

This figure shows the prevalence of the primary flags by the implied compensation per employee. Panel A shows the relationship between implied compensation per employee and our four main flags combined together as at least one flag and Panel B shows the relationship for each primary flag separately. For both panels, loans are binned into \$2,000 wide bins (i.e., (\$0k, \$2k], ... , (\$98k, \$100k], ... , (\$128, \$130k]). For Panel A, the left axis shows the percentage of loans that in each bin (bars) and the right axis shows the percentage of the loans in the bin that are flagged by the given flag (dots). In Panel B, loans are filtered to corporation, S-corporation, and LLC loans for the Business Registry subpanel, loans for which we can determine industry-CBSA average compensation for the High Implied Compensation subpanel, and loans with a matched EIDL Advance for the EIDL > PPP subpanel. The solid lines are third-degree polynomial fits (weighted based on number of loans in the each bin), which are separately fitted for loans below \$98,000 and loans above \$100,000, and the dashed lines are 95% confidence intervals. Red represent fintech loans and grey represent traditional loans.

Panel A. Percentage Flagged, by Implied Compensation



Panel B. Individual Flags, by Implied Compensation

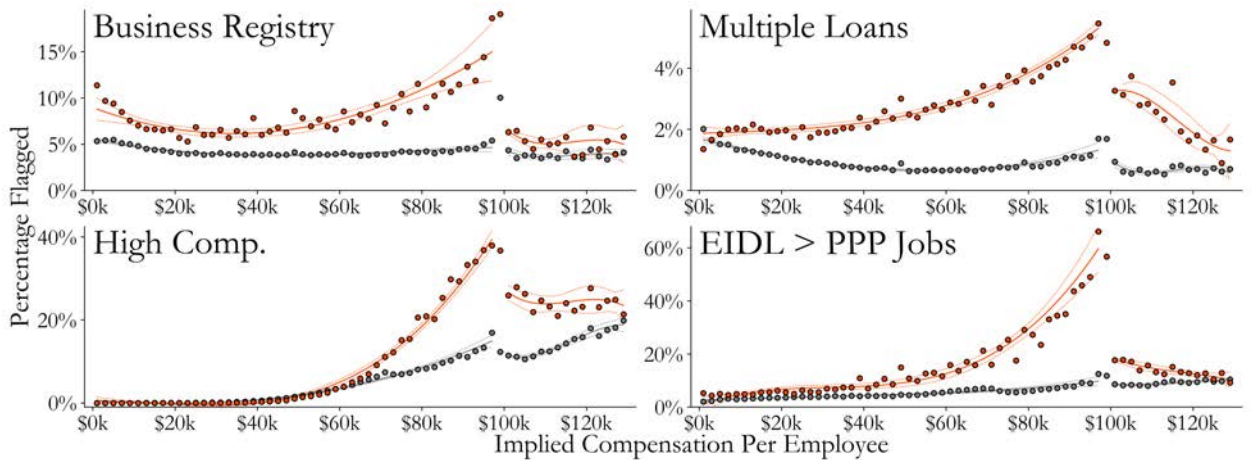
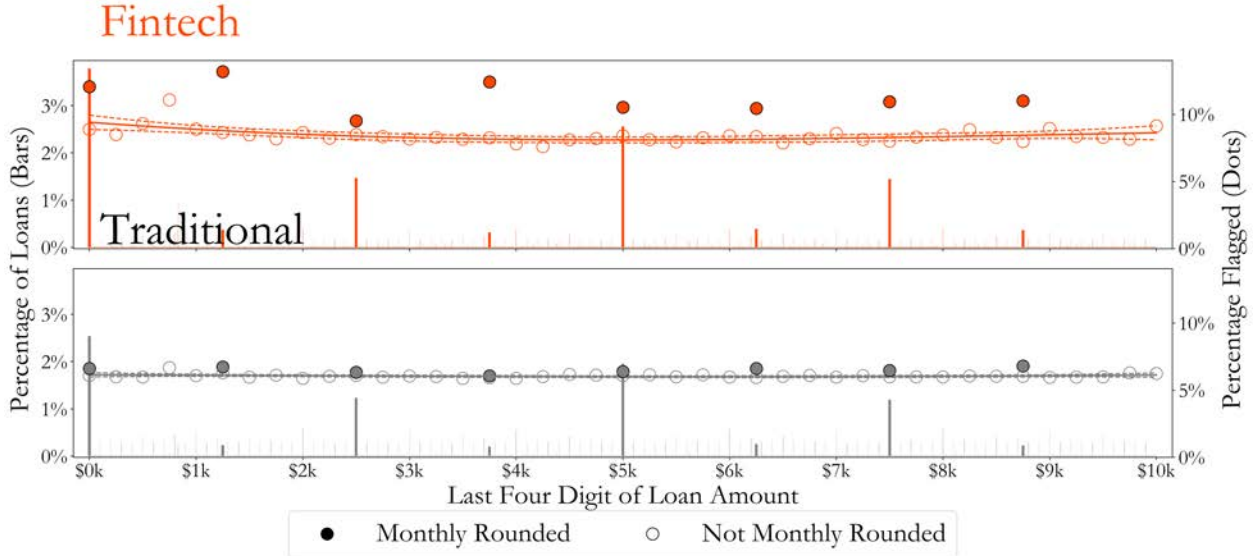


Figure 6. Rounded Loan Amounts

This figure shows the prevalence of the primary flags by whether the total monthly implied compensation of a loan is rounded to an interval of \$500 (i.e., loan amount is within ± 50 cents of an interval of \$1,250) and lender type. Panel A shows the relationship between implied compensation per employee and our four main flags combined together as at least one flag and Panel B shows the relationship for each primary flag separately. For both panels, the last four digits of the loan amount is considered (i.e., \$123,456.78 \rightarrow \$3,456.78). In Panel A, the top subpanel shows fintech loans and the bottom shows traditional loans. Further, the left axis shows the percentage of loans in each \$1 wide bin (bars for rounded compensation are thickened) and the right axis shows the percentage of loans that are flagged within each \$1 bins for monthly rounded (solid dots) and \$250 wide bins for non-rounded (hollow dots). In Panel B, loans are filtered to corporation, S-corporation, and LLC loans for the Business Registry subpanel, loans with industry-CBSA average compensation less than \$33,333.33 for the High Implied Compensation subpanel, and loans with a matched EIDL Advance for the EIDL > PPP Jobs subpanel. Additionally for both panels, loans with one job reported, loans with implied compensation within $\pm \$1,000$ of \$100,000, and second draw loans to hospitality businesses are excluded. The solid lines are third-degree polynomial fits for the percentage flagged in the non-rounded bins and the dashed lines are 95% confidence intervals.

Panel A. Percentage Flagged, by Lender Type and Rounding



Panel B. Individual Flags, by Lender Type and Rounding

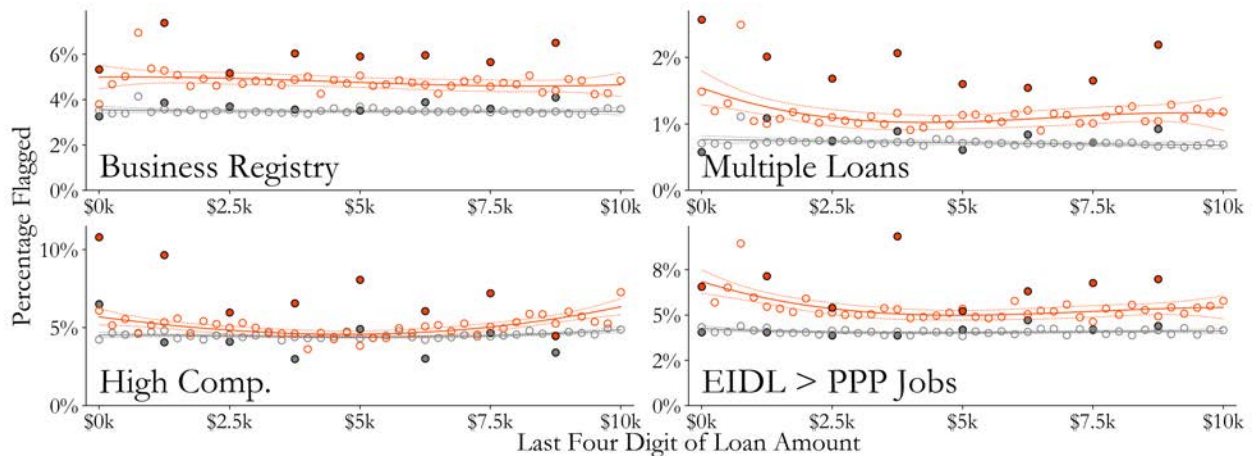
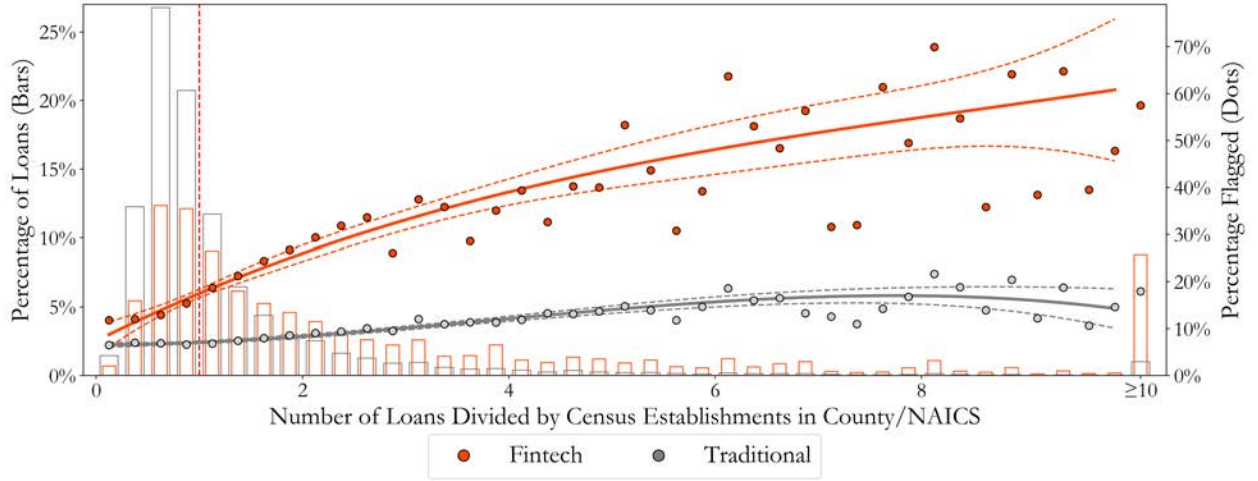


Figure 7. Overrepresentation of Industries in Counties

This figure shows overrepresentation of loans within industry-county pairs. We define the normalized number of loans as the number of first draw loans divided by the number of establishments (per the 2019 US Census County Business Patterns dataset) in an industry (represented by NAICS [North American Industry Classification System] code) and county pair. Panel A shows the relationship between normalized number of loans and our four main flags combined together as at least one flag and Panel B shows the relationship for each flag separately. Since the CBP does not include self-employed and independent contractors as establishments, we exclude loans to these business types. Note that 6.20% of fintech and 0.72% of traditional loans are in industry-county pairs with ratios of at least 10; these loans are represented in Panel A by the bars and dots at the far right labeled “ ≥ 10 ”. In both panels, loans are binned into 0.25 unit wide bins. The solid lines are third-degree polynomial fits for the percentage of flagged loans and the dashed lines are 95% confidence intervals.

Panel A. Percentage Flagged, by Normalized Number of Loans in Industry-County Pair



Panel B. Individual Flags, by Normalized Number of Loans in Industry-County Pair

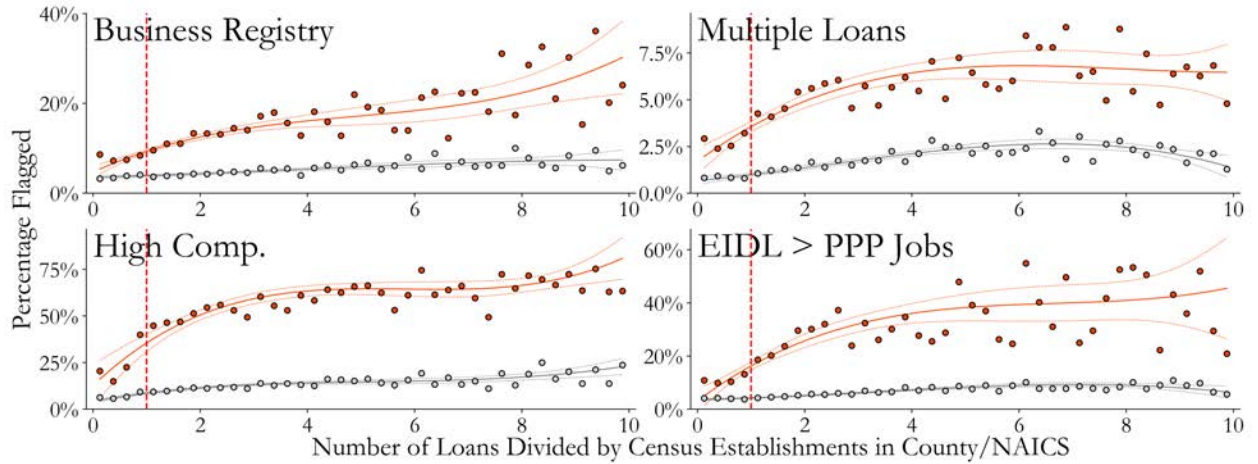
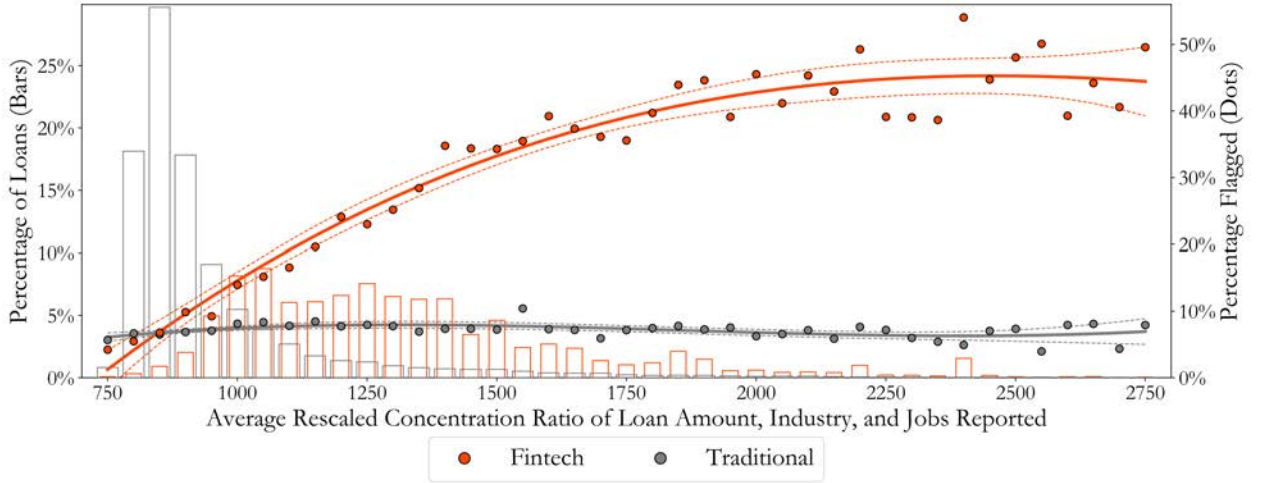


Figure 8. Clustering Within Lenders and Counties

This figure shows clustering of loans within lenders-county pairs. We calculate the concentration ratios of industries, loan amount (rounded to \$100), and jobs reported (excluding 1) for first draw loans in each lender-county pair, rescaled each concentration ratio to a median of 1,000 and IQR (interquartile range) of 300, and then take the average of the three rescaled concentration ratios. For example, let $i = 1, 2, \dots, n$ represent the n industries in a given lender-county pair, then $Concentration_{industry} = \sum_{i=1}^n s_i^2$ where s_i is the percentage of loans in the lender-county pair that are in industry i times 100 (e.g., 6.2 for 6.2%). Then, $Rescaled\ Concentration_{industry} = \frac{Concentration_{industry} - \text{Median}[Concentration_{industry}]}{75^{th}\text{Percentile}[Concentration_{industry}] - 25^{th}\text{Percentile}[Concentration_{industry}]} * 300 + 1000$. Panel A shows the relationship between the average rescaled concentration ratio and our four main flags combined together as at least one flag and Panel B shows the relationship for each flag separately. In both panels, only lender-county pairs with at least 25 loans are considered. Note that 2.4% of fintech loans and 0.5% traditional loans are outside the range of average rescaled concentration ratio shown in Panel A. In both panels, loans are binned into 50 unit wide bins; in Panel B, bins with fewer than 100 loans for which the given flag can be determined are excluded. The solid lines are third-degree polynomial fits for the percentage of flagged loans and the dashed lines are 95% confidence intervals.

Panel A. Percentage Flagged, by Average Rescaled Concentration Ratio in Lender-County Pair



Panel B. Individual Flags, by Average Rescaled Concentration Ratio in Lender-County Pair

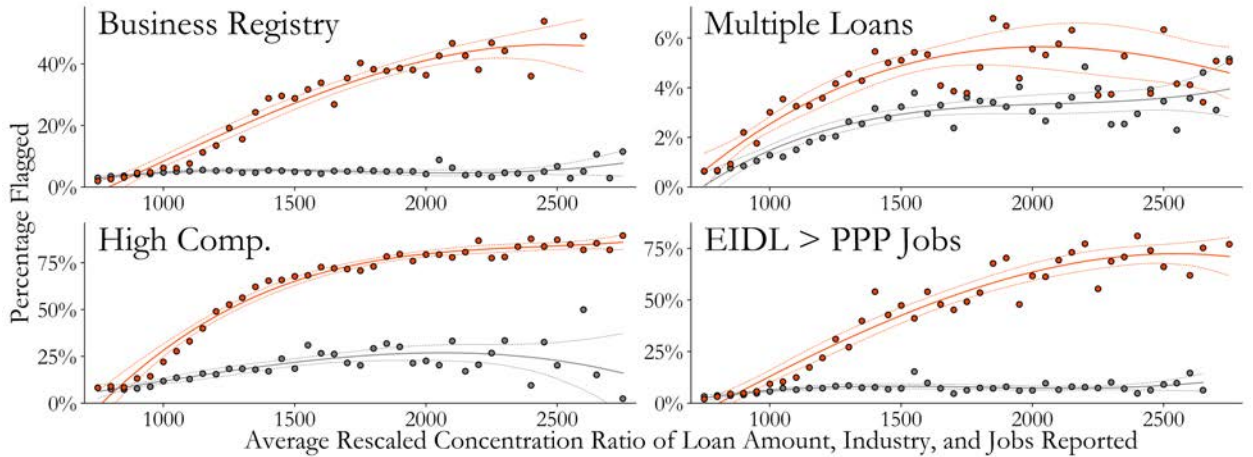


Figure 9. Criminal Records

This figure shows criminal records for a sample of 150,000 Round 1 and 2 loans to self-employed individuals, independent contractors, and sole-proprietors. Panel A shows the percentage of loans where the borrower has a felony from 2000-2020 on their record by lender type and the presence of misreporting indicators. The error bars denote 95% confidence intervals (based on standard errors double clustered by zip code and lender) for each percentage. Panel B shows the relationship between the percentage of loans in this sample that are flagged by at least one primary flags and the percentage of borrowers that have a felony from 2000-2020 on their record by lender. Lenders with at least 0.2% of the sample (300 loans) are shown. The dashed line is a linear fit and correlation is shown in the bottom left corner. In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders.

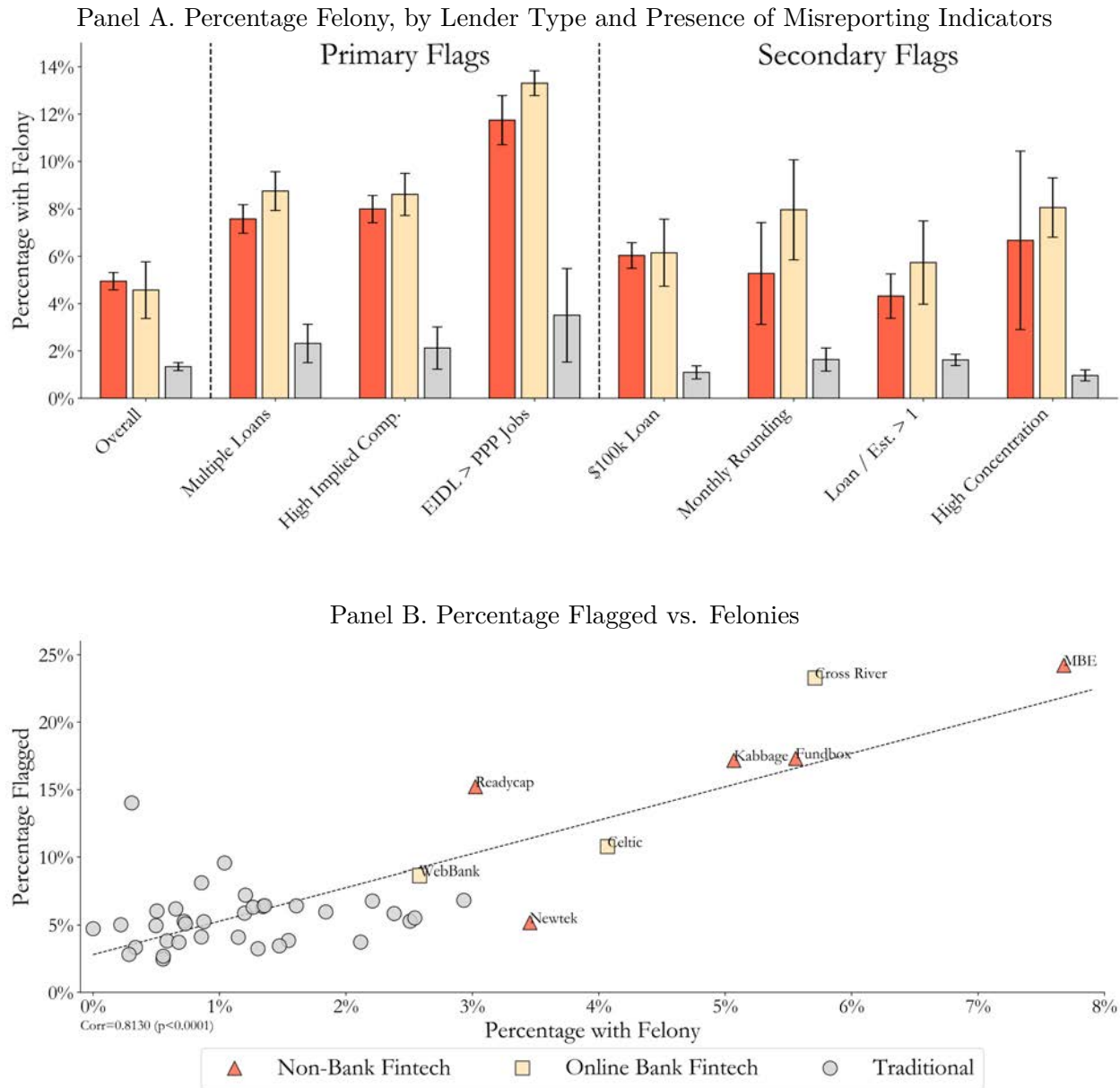
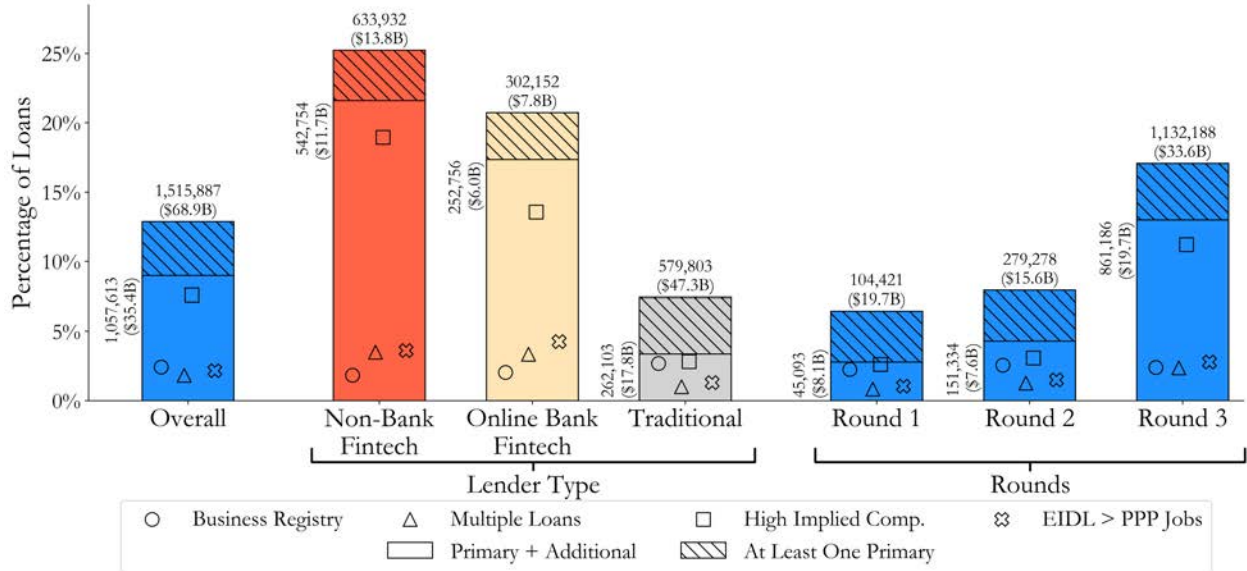


Figure 10. Overall Misreporting Flag Rates

This figure shows the variation in percentage of loans flagged. Panel A shows the percentage and dollar amounts of flagged loans overall, by lender type, and by round. Panel B shows the percentage of flagged loans by lender for the top 75 lenders (by number of loans). In both panels, the plain section of each bar represents the percentage of loans flagged by one primary flag and an additional flag (either another primary or a secondary) and the entire bar (plain and striped sections combined) represents loans flagged by at least one primary flag. In Panel A, the set of numbers to the left of each bar represent the number of loans and dollar value of loans flagged one primary flag and an additional flag and the set on top of each bar by at least one primary flag. The markers within each bar represent the percentage of loans flagged by each of the primary flags (unconditional of whether a flag can be determine for a given loan). In Panel B, the two horizontal lines represent the percentage of loans flagged by each measure across the entire sample (dashed for loans flagged by one primary flag and an additional flag and solid by at least one primary). In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders.

Panel A. Percentage of Loans Flagged, by Lender Type and Rounds



Panel B. Percentage of Loans Flagged, by Lender

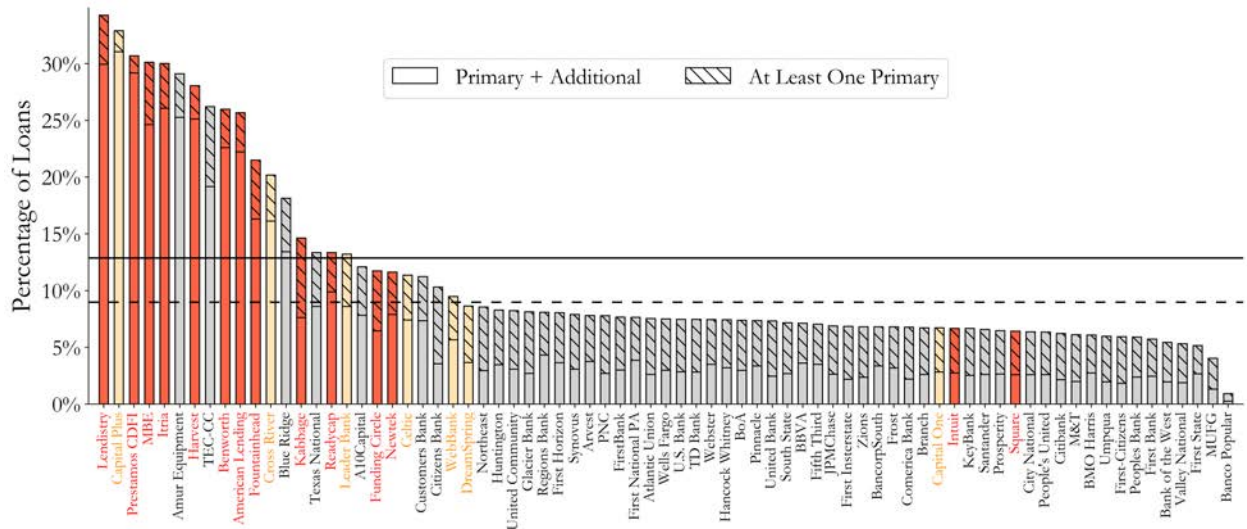
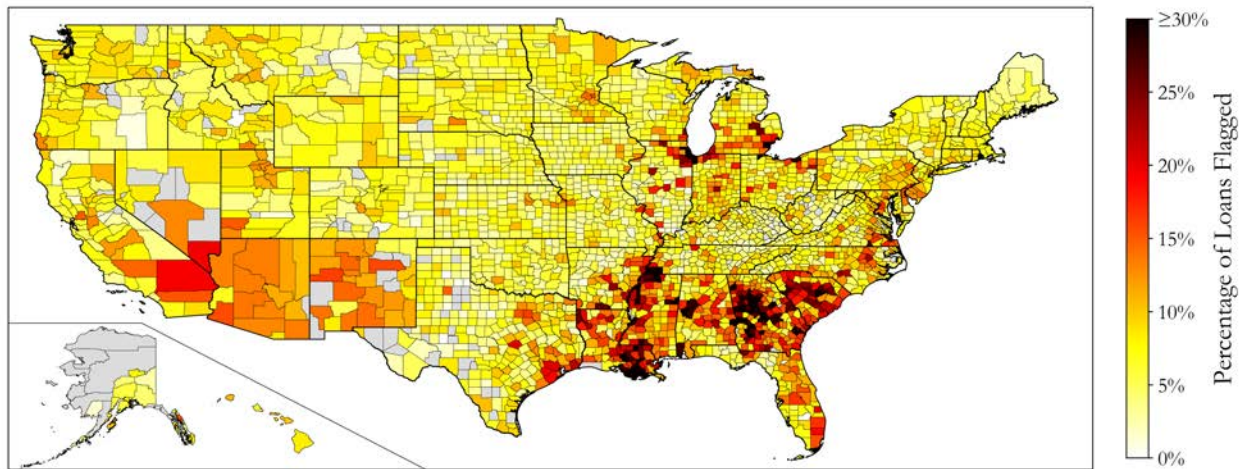


Figure 11. Geography of Flagged Loans

This figure shows geographic variation in the percentage of flagged loans. Panel A shows the percentage of flagged loans in each county and Panel B shows within county variation. In Panel A, counties are colored based on the color scheme shown by the bar to the right of the map and counties with fewer than 100 loans are colored grey. Panel B shows the percentage of flagged loans in each zip code on the vertical axis and the percentage of flagged loans in the corresponding county on the horizontal axis. Dots are colored by the percentage of fintech loans in each zip code and sized based on the number of loans in the zip code. Zip codes with at least 100 loans are shown. The dashed line is a linear fit and the correlation is shown in at the bottom left corner.

Panel A. Percentage of Flagged Loans, by County



Panel B. Within County Variation

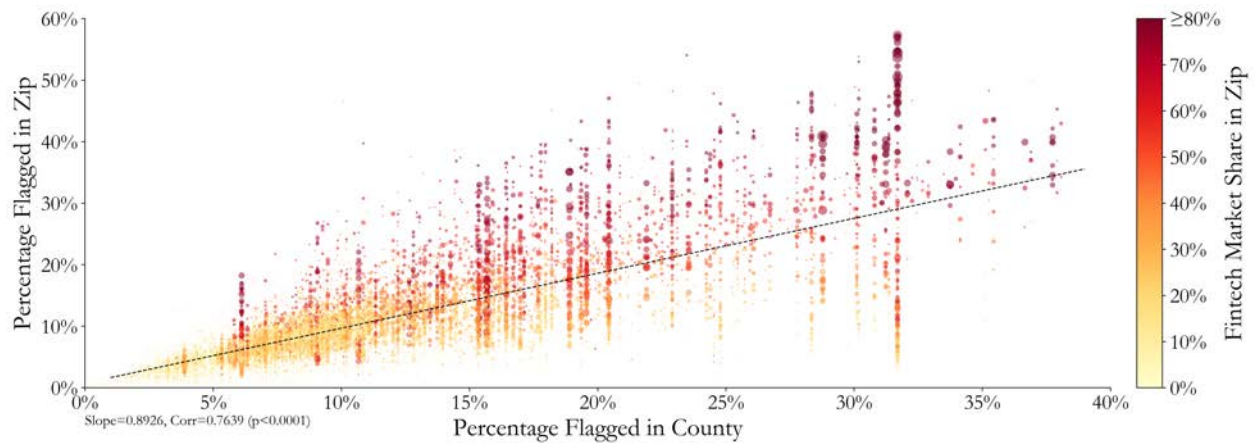
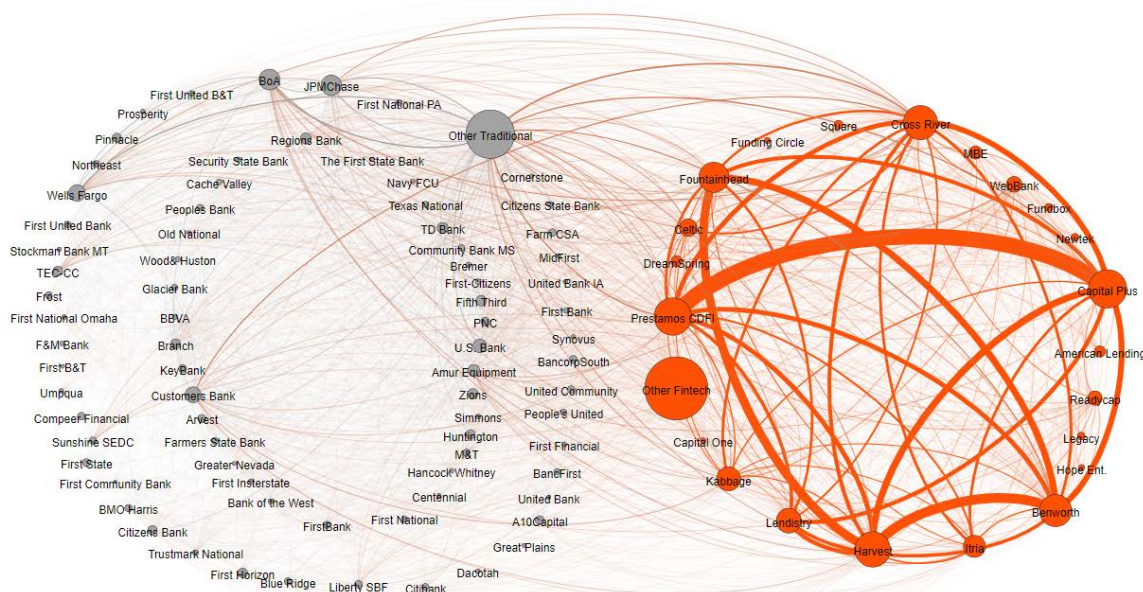


Figure 12. Lender Network

This figure shows connections between lenders. Panel A shows connections between lenders that were used by borrowers at the same address within the same draw. Panel B shows connections between lenders used by the flagged first draw borrowers across draws. In Panel A, node size is proportional to the number of loans with the multiple loans flag originated by each lender, edges are not directed, and edge width is proportional to the number of addresses that used both lenders. In Panel B, node size is proportional to the number of first draw loans (that also got a second draw loan from the same or different lender) and second draws originated by the lender, edges are directed, edge width is proportional to the number of flagged first draw borrowers moving clockwise from the first draw lender to the second draw lender. In both panels, red nodes are fintech lenders and grey nodes are traditional lenders. Top 100 lenders (by the same measure that node size is based on) are shown and the remainder are combined into the “Other” nodes (one for other fintech lenders and one for other traditional lenders).

Panel A. Multiple Lenders at Same Address



Panel B. Changes Between Draws

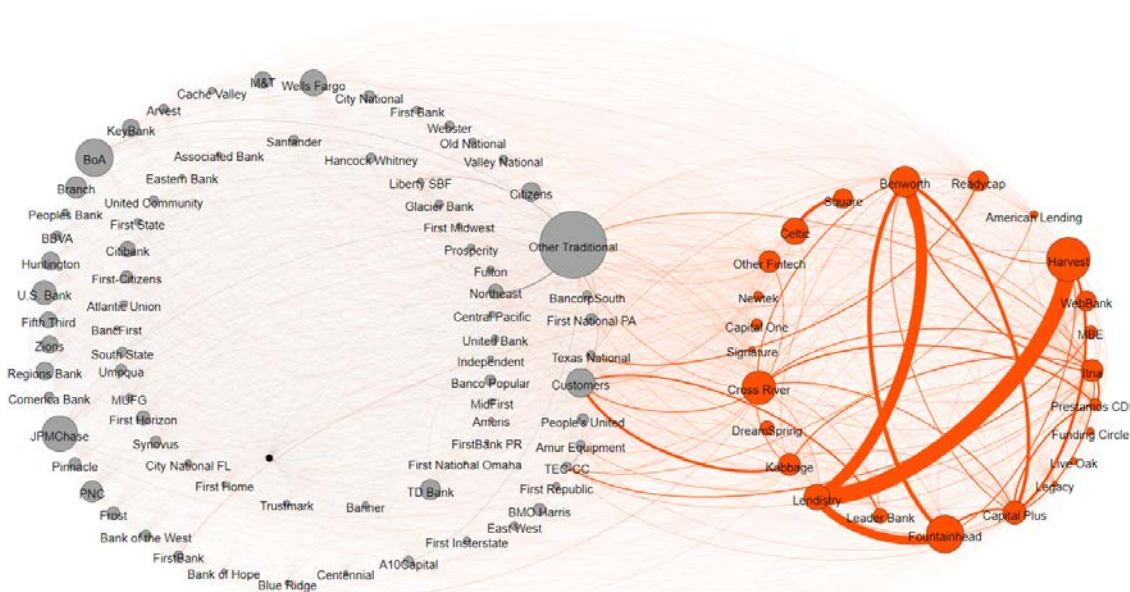


Figure 13. Persistence and Growth Across Rounds

This figure shows the persistence and growth of flagged loans across lending rounds. Panel A shows this by lender type and Panel B by zip code. In Panel A, each subpanel shows a lender type and each series is the percentage of loans flagged by the given measure across time. In Panels B, the percentage of loans flagged in rounds 1 and 2 are shown on the horizontal axis and in round 3 on the vertical axis. For Panel A, the vertical dotted lines split each subpanel into the three lending rounds. The solid lines are loans flagged by at least one primary flag and the dashed line is loans flagged by at least one primary flag and an additional flag (either another primary or a secondary). For Panel B, the left subpanel uses all loans, the middle uses fintech loans, and the right uses traditional loans. Zip codes with at least 100 loans and, for the fintech and traditional subpanels, 25 loans by the given lender type are shown. The black line is a 45-degree line and the correlation is presented in the bottom of each panel. The circle size corresponds to the number of loans in the zip code by the given lender type and color corresponds the growth/decline in lending in the zip code by the given lender type.

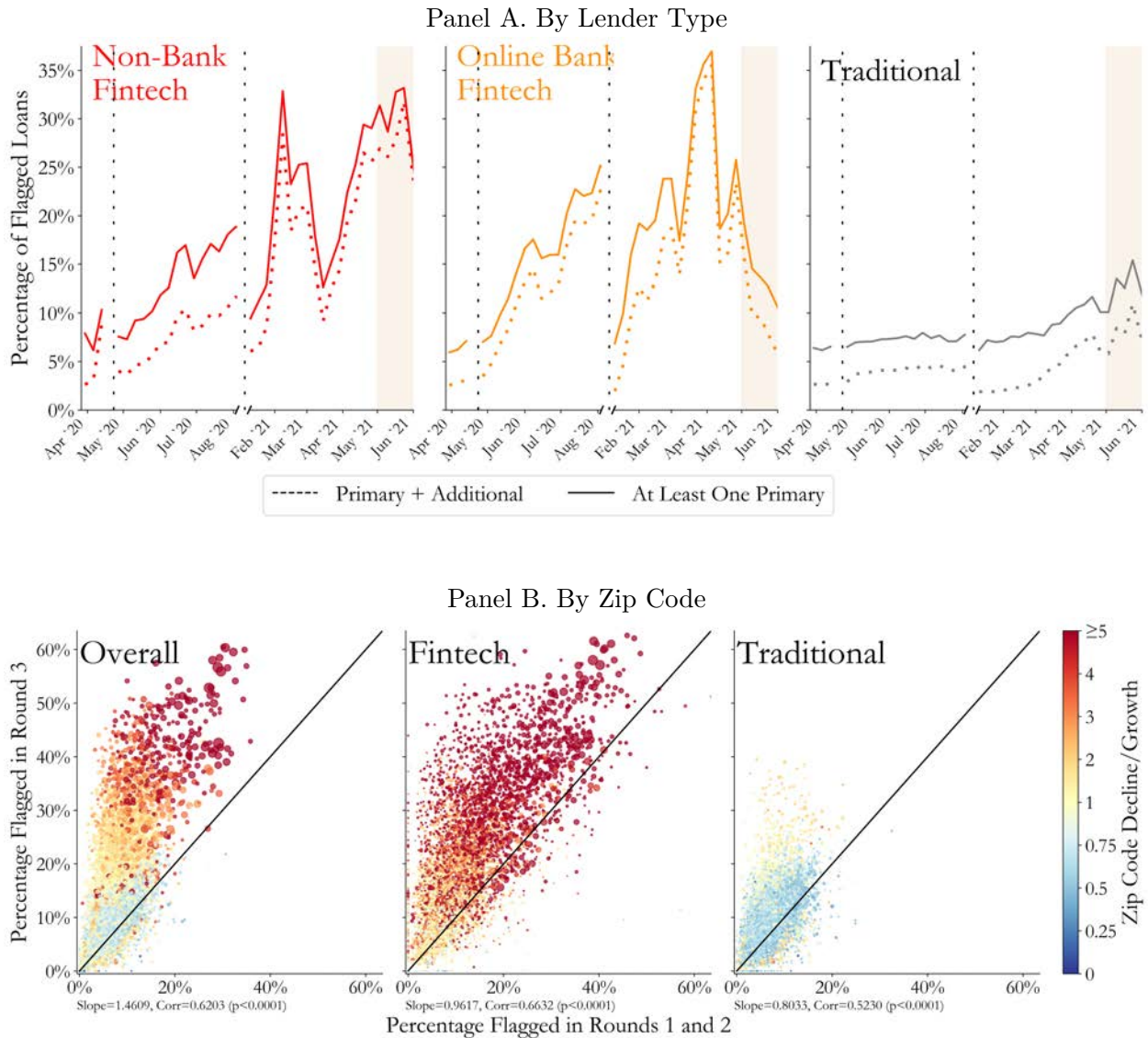


Exhibit 1. Examples of Suspicious Loans

This exhibit shows some examples of suspicious loans.

Panel A. 14 Loans to The Same Address, 13 Incorporated Late

Business Name (Redacted)	Date Incorporated	Date Approved	Lender	Industry	Loan Amount	Jobs Reported
FDML	4/2/2018	5/13/2020	Celtic	Indep. Artists, Writers, Performers	\$62,083	10
DREL	7/7/2020	7/8/2020	Kabbage	Musical Groups & Artists	\$53,125	10
LTTBTL	7/15/2020	7/15/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
KYHUL	7/15/2020	7/16/2020	Kabbage	Misc. Schools & Instruction	\$91,770	10
STWL	7/16/2020	7/17/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
STWL	7/16/2020	7/17/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
ATYL	7/19/2020	7/21/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
BLNL	7/19/2020	7/21/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
CAYL	7/21/2020	7/22/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
CTWIL	7/23/2020	7/23/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
DYJNL	7/22/2020	7/26/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
FML	7/27/2020	7/30/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
EIEL	7/22/2020	7/30/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
GGITL	7/30/2020	8/1/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10
JTBCL	8/5/2020	8/5/2020	Kabbage	Misc. Schools & Instruction	\$53,229	10

Panel B. Four Loans to Same Household

Individual Name (Redacted)	Age	Date Approved	Lender	Industry	Loan Amount	Jobs Reported
O. P.	21	7/10/2020	Kabbage	Lawn/Garden Equipment Manuf.	\$20,833	1
T. P.	49	7/10/2020	Kabbage	Lawn/Garden Equipment Manuf.	\$20,833	1
A. P.	20	7/15/2020	Kabbage	Nail Salons	\$20,833	1
G. P.	46	7/15/2020	Kabbage	Other Automotive Repair	\$20,833	1

Exhibit 2. Cross River Case Study

This exhibit shows examples of \$20,000 first draw loans by Cross River Bank in Illinois.

Individual Name (Redacted)	Date Approved	Lender	Industry	Loan Amount	Jobs Reported
J. C.	7/29/2020	Cross River	Insurance Agencies and Brokerages	\$20,000	1
R. J.	7/29/2020	Cross River	Insurance Agencies and Brokerages	\$20,000	1
4,300 \$20k Loans by Cross River in Illinois (Mostly in Chicago Area) to Individuals/Businesses (98% with 1 Employee) in "Insurance Agencies and Brokerages" Industry.					
C.C.	7/29/2020	Cross River	All Other Miscellaneous Crop Farming	\$20,000	8
M. A.	7/29/2020	Cross River	All Other Miscellaneous Crop Farming	\$20,000	8
938 \$20k Loans by Cross River in Illinois (Mostly in Chicago Area) to Individuals/Businesses (56% with 8 Employee and 22% with 1 Employee) in "All Other Miscellaneous Crop Farming" Industry.					
K. K.	7/30/2020	Cross River	All Other Miscellaneous Manufacturing	\$20,000	9
C.M.	8/7/2020	Cross River	Other General Government Support	\$20,000	50
3,068 \$20k Loans by Cross River in Illinois (Mostly in Chicago Area) to Individuals/Businesses in Various Industries (Including 706 to "All Other Personal Services," 351 to "General Freight Trucking, Local," 337 to "Other Performing Arts Companies," and 300 to "New Single-Family Housing Construction.")					

Table I. Summary Statistics

This table presents summary statistics for our sample. The sample includes all PPP loans approved from the start of the program (March 2020) through most of Round 3 (April 2021) that have not been repaid as of May 3, 2021. Fintech lenders are determined following Erel and Liebersohn (2021). *Loan Amount* is the initial approved amount minus any portion used to refinance an EIDL loan. *Implied Comp.* is determined following the guidelines in place when the loan was approved and is based on loan amount and jobs reported. *CBSA/NAICS Avg. Comp.* is average compensation in the loan's industry-CBSA based on the US Census CBP data, *CBSA/NAICS Avg. Receipts* is average receipts (for business types that were able to use gross income to calculate loan size) to nonemployer businesses in the loan's industry-CBSA based on the US Census NES data, and *Normalized Comp.* is the ratio of *Implied Comp.* and either *CBSA/NAICS Avg. Comp.* or, if the business was able to use gross income to calculate their loan amount, the larger of *CBSA/NAICS Avg. Comp.* and *CBSA/NAICS Avg. Receipts*. *Loans (Within Draw) at Address* is the number of loans (within the loan's draw) at the same residential address. *Frac. Corp, S Corp, LLC* is the percentage of loans to these business types, *Frac. Second Draw* is the percentage of Round 3 loans that are the borrower's second draw from the PPP, and *Frac. Matched EIDL Advance* is the percentage of loans with a matching EIDL Advance. *Frac. Fintech (Either Type)*, *Frac. Non-bank Fintech*, and *Frac. Online Bank Fintech* are the percentages of loans that are originated by the given type of lender.

	Fintech			Traditional		
	Mean	SD	Median	Mean	SD	Median
Num. Loans [Pct. Loans]	3,969,845 [33.73%]			7,798,844 [66.37%]		
Loan Amount	25,112	99,717	19,052	90,123	305,553	20,833
Jobs Reported	2.448	9.895	1.000	10.331	29.299	3.000
Implied Comp.	64,217	39,474	70,666	47,206	66,709	38,932
CBSA/NAICS Avg. Comp	46,911	38,966	36,837	49,657	37,302	42,698
CBSA/NAICS Avg. Receipts	41,917	30,714	30,437	48,908	33,596	39,172
Normalized Comp.	1.494	1.452	1.066	1.043	1.903	0.854
Num. Loans (Within Draw) at Address	1.335	0.843	1.000	1.197	0.792	1.000
Frac. Corp, S Corp, LLC	0.185			0.645		
Frac. Second Draw (Round 3 Loans)	0.268			0.594		
Frac. Matched EIDL Advance	0.178			0.267		

	Round 1	Round 2	Round 3
Num. Loans [Pct. Loans]	1,619,201 [13.8%]	3,518,979 [29.9%]	6,630,509 [56.3%]
Loan Amount	198,864	57,849	41,773
Jobs Reported	20.468	7.914	4.418
Implied Comp.	48,023	43,838	58,978
CBSA/NAICS Avg. Comp	49,339	51,832	46,795
CBSA/NAICS Avg. Receipts	-	-	43,523
Normalized Comp.	1.178	1.055	1.293
Num. Loans (Within Draw) at Address	1.277	1.215	1.271
Frac. Fintech (Either Type)	0.0479	0.204	0.479
Frac. Non-bank Fintech	0.0249	0.0748	0.333
Frac. Online Bank Fintech	0.230	0.130	0.145
Frac. Corp, S Corp, LLC	0.829	0.657	0.319
Frac. Second Draw	-	-	0.438
Frac. Matched EIDL Advance	0.292	0.301	0.190

Table II. Odds Ratios

In this table, we present the odds ratios between each of our four main indicators. Panel A shows the odds ratios for fintech and traditional loans combined. Panel B shows the odds ratios for fintech loans only in the lower triangular and traditional loans only in the upper triangular. Note that odds ratios are symmetric, which is why only values for the lower triangular are provided. Robust standard errors are double clustered by zip code and lender.

Panel A. Fintech and Traditional Loans Combined				
	Business Registry	Multiple Loans	High Implied Comp.	EIDL > PPP Jobs
Business Registry	-			
Multiple Loans	1.904*** (7.49)	-		
High Implied Comp.	2.965*** (9.34)	3.241*** (16.17)	-	
EIDL > PPP Jobs	1.496*** (5.95)	4.531*** (19.27)	14.746*** (19.90)	-
Panel B. Fintech Loans (Lower Triangular) and Traditional Loans (Upper Triangular)				
	Business Registry	Multiple Loans	High Implied Comp.	EIDL > PPP Jobs
Business Registry	-	1.338*** (8.10)	1.898*** (12.06)	1.184*** (6.13)
Multiple Loans	2.351*** (17.55)	-	1.655*** (4.40)	2.136*** (11.59)
High Implied Comp.	3.655*** (13.78)	2.085*** (6.28)	-	5.830*** (22.53)
EIDL > PPP Jobs	2.177*** (6.03)	3.093*** (11.07)	16.007*** (38.45)	-

z-statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Table III. Prevalence of Flags by Lender Type

This table presents the percentage of loans flagged by the four main flags, at least one of four flags, and at least two of the four flags. In Panel A, column (1) shows the percentage of fintech loans with the given flag, column (2) shows the percentage of traditional loans with the given flag, and column (3) shows the difference between the fintech and traditional percentages, column (4) shows the adjusted differences with zip code, business type, and NAICS \times CBSA fixed effects and controlling for jobs and loan size, and column (5) shows the differences between matched pairs of fintech and traditional loans. The N values show is the number of loans for which the flag can be determined and robust standard errors are double clustered by zip code and lender. For the matched differences, robust standard errors are four-way clustered by the zip code and lender of both matched loans. The full regression results for the unadjusted differences and adjusted differences are reported in Panel A and Panel B, respectively, of Table [IA.II](#).

	(1)	(2)	(3)	(4)	(5)
	Fintech	Traditional	Unadjusted Difference	Adjusted Difference	Matched Difference
Business Registry	0.103 N = 739,657	0.0434 N = 4,825,093	0.0587*** (3.09)	0.0350*** (3.55)	0.0268*** (3.12)
Multiple Loans	0.0344 N = 3,969,845	0.0100 N = 7,798,844	0.0244*** (8.71)	0.0106*** (5.19)	0.0170*** (3.76)
High Implied Comp.	0.489 N = 1,600,474	0.0942 N = 2,133,022	0.395*** (7.83)	0.0935*** (7.34)	0.0818*** (3.91)
EIDL > PPP Jobs	0.216 N = 705,837	0.0487 N = 2,083,055	0.168*** (4.34)	0.0651*** (4.13)	0.0794*** (4.62)
At Least One Flag	0.236 N = 3,969,845	0.0743 N = 7,798,844	0.161*** (8.12)	0.0560*** (8.81)	0.0535*** (6.71)
At Least Two Flags	0.0246 N = 3,969,845	0.00362 N = 7,798,844	0.0210*** (7.39)	0.00686*** (7.00)	0.0107*** (5.99)

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IV. Secondary Flags

In this table, we examine the relationship between our four main flags, which we combine to form *At Least One Flag*, and the secondary flags. We estimate OLS regressions with *At Least One Flag* as the dependent variable and the five secondary flags as independent variables. Each specification also include an interaction between the secondary flag and an indicator for whether the loan was originated by a fintech lender. *\$100k Implied Comp./Receipts* is a dummy variable equal 1 if the implied compensation/receipts per job is within \pm \$1,000 of \$100,000. *Monthly Rounding* is a dummy variable equal 1 if the loan amount is within \pm 50 cents of an interval of \$1,250. *Overrep. in County/NAICS* is a dummy variable equal 1 if the number of first draw loans to businesses not listed as self-employed and independent contractors in a loan's industry-county pair exceeds the number of establishments in the industry-county pair according to the US Census CBP data. *High Concentration* is a dummy variable equal 1 if the average rescaled concentration ratio in the loan's lender-county pair is above the 75th percentile. *Felony Post-2000* is a dummy variable equal 1 if the borrower has a felony on their criminal record from 2000 or after. *1(Fintech)* is a dummy variable equal 1 if the loan was originated by a fintech lender. For all specifications, loans are filtered to the sets for which we can determine the secondary flag. Further, for specification (2), one job loans and loans where $1(\$100k \text{ Implied Comp.}) = 1$ are excluded. Fixed effects are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Dep. Variable: At Least One Primary Flag					
	(1)	(2)	(3)	(4)	(5)
\$100k Implied Comp.	0.0238*** (2.65)				
Monthly Rounding		0.00426*** (5.84)			
Overrep. in County/NAICS			-0.00739** (-2.19)		
High Concentration				0.0141*** (4.23)	
Felony Post-2000					0.0265*** (2.71)

$\times 1(\text{Fintech})$	0.113*** (14.30)	0.0118** (2.40)	0.0604*** (8.87)	0.0159*** (2.77)	0.0374*** (3.01)
ln(Jobs Reported)	Yes	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes	No
Business Type FE	Yes	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes
Observations	11,085,989	5,120,313	6,215,638	7,615,688	123,670
Num. Lenders	4,849	4,768	4,798	4,221	2,578
R^2	0.324	0.103	0.347	0.346	0.383
Mean of Dep. Variable	0.134	0.0671	0.145	0.143	0.106

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table V. County Cultural Features

In this table, we examine the relationship between our four main flags, which we combine to form *At Least One Flag*, and cultural/regional features. We estimate OLS regressions with *At Least One Flag* as the dependent variable and the cultural/regional features as independent variables. All independent variables are rescaled at the county level to have a mean of 0 and a standard deviation of 1. *Political Corruption* is the number of public corruption convictions in 2004-2013 per million residents (as reported by the DOJ). *Religious Adherence* is the percent of the county's population with a religious affiliation as of 2010 (as reported by the Association of Religious Data Archives). *Ashley Madison Usage* is the paid Ashley Madison usage rate in the county (as reported by [Griffin et al. \(2019\)](#)). *Population Density* is the population per square mile as of 2019, *Median Income* is the median household income as of 2019, *Pct. Non-White* is the percentage of the population that is non-white as of 2019, *College Educated* is the percentage adults with a bachelor's degree or higher as of 2015-19, and *2019 Unemployment* is the unemployment rate as of 2019 (all from the Economic Research Service of the U.S. Department of Agriculture). *Pct. Fintech* is the percentage of PPP loans in the county originated by a fintech lender. Fixed effects are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Dep. Variable: At Least One Primary Flag			
	(1)	(2)	(3)
Public Corruption	0.0123*** (7.08)	0.00863*** (5.78)	0.00447*** (3.82)
Religious Affiliation	0.00122 (1.54)	-0.00136* (-1.66)	-0.0000799 (-0.10)
Ashley Madison Usage	-0.00304*** (-5.61)	0.00747*** (4.79)	0.000160 (0.13)
Population Density		-0.00112*** (-7.01)	-0.000861*** (-4.40)
Median Income		-0.00704*** (-7.27)	-0.00383*** (-4.19)
Pct. Non-White		0.0193*** (13.57)	-0.00316** (-2.27)
College Educated		-0.00446*** (-3.52)	0.000720 (0.84)
2019 Unemployment		0.00441*** (3.57)	0.00522*** (4.53)
Pct. Fintech			0.0269*** (9.94)
ln(Jobs Reported)	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes
NAICS \times State FE	Yes	Yes	Yes
Observations	11,651,511	11,651,511	11,651,511
Num. Lenders	4,874	4,874	4,874
R^2	0.248	0.250	0.252
Mean of Dep. Variable	0.129	0.129	0.129

t-statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Table VI. Persistence of Lender Behavior Across Rounds

In this table, we examine the persistence of lender behavior across rounds. We estimate OLS regressions with dummies for whether each Round 3 loan is flagged by our four main flags individually (specifications (1) through (4)) and at least one of them (specification (5)) as the dependent variables and the percentage of the lender's loans were flagged by the same flag in rounds 1 and 2 as the independent variable. Interactions with whether the loan was originated by a fintech or traditional lender are included in all specifications. For specifications (1) through (4), loans are filtered to the sets for which we can determine the flag (same as in Figures 2- 4). Further, to ensure we have accurate measures of past behavior, we require that each lender have at least 100 loans in Round 1 and 2 (combined) for which we can determine the given flag. Fixed effects are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Implied Comp.	(4) EIDL > PPP Jobs	(5) At Least One Flag
Past Pct. This Flag					
× 1(Fintech)	0.930*** (4.78)	0.640*** (3.45)	0.417*** (3.00)	0.872*** (2.94)	0.516*** (5.79)
× 1(Traditional)	0.201 (0.92)	0.225* (1.94)	-0.303** (-2.23)	-0.125 (-0.91)	-0.00962 (-0.09)
ln(Jobs Reported)	Yes	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes	Yes
NAICS × CBSA FE	Yes	Yes	Yes	Yes	Yes
Observations	1,853,605	5,503,222	1,649,381	1,029,259	5,503,222
Num. Lenders	2,475	3,133	1,574	1,439	3,133
R^2	0.142	0.052	0.638	0.350	0.377
Mean of Dep. Variable	0.0720	0.0220	0.351	0.139	0.171

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table VII. Likelihood of Receiving a Second Draw Loan

In this table, we examine whether lenders were more/less likely to provide a second draw loan to a borrower who's first draw loan is flagged by at least one of our primary flags. We estimate OLS regressions with a dummy for whether the same lender provided the first and second draw loans as the dependent variable and a dummy for whether the first draw loan was flagged by at least one of the primary flags as the independent variable. In specifications (1) and (2), if a borrower did not receive a second draw loan, the dependent variable is set to 0, and in specifications (3) and (4), only borrowers that received both a first and second draw loans are included in the sample. In the even specification, an interaction with whether the first draw loan was originated by a fintech lender is included. Fixed effects are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Dep. Variable: 1(First and Second Draw by Same Lender)				
	(1)	(2)	(3)	(4)
	Unconditional of Receiving Second Draw		Conditional on Receiving Second Draw	
First Draw Flagged	-0.0260*** (-6.64)	-0.0310*** (-11.94)	-0.00972** (-2.06)	-0.0161*** (-5.00)
× 1(Fintech)		0.0212* (1.66)		0.0254 (1.08)
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS × CBSA FE	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes
Observations	4,849,427	4,849,427	1,584,875	1,584,875
Num. Lenders	4,727	4,727	4,518	4,518
R^2	0.122	0.122	0.420	0.420
Mean of Dep. Var.	0.278	0.278	0.834	0.834

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Internet Appendix

Classifying Lenders as FinTechs

We use Erel and Liebersohn’s (2021) classifications for lenders that were active in rounds 1 and 2 (the sample period for Erel and Liebersohn (2021)), and we use the same methodology for classifying round 3 lenders that were not active enough to be classified in the earlier rounds. The method used by Erel and Liebersohn (2021) is summarized in their paper as, “We match this loan-level data to bank identifiers from the Federal Financial Institutions Examination Council (FFIEC) using the lender names provided. Most of the names are matched using automated name matching. Lenders which we are not able to match automatically are a combination of non-bank lenders, banks that have duplicate names, and banks that have idiosyncratic names. We therefore hand-match all PPP lenders who originate over 500 PPP loans, classifying separately non-bank lenders and banks which do not have a unique match in the FFIEC database... We identified Fintech lenders as any unregulated nonbank lender that participated in the program as well as any regulated online direct bank. Specifically, nonbank lenders are non-depository financial institutions, like Kabbage, that generally rely on FinTech in their lending... Online banks, however, are regulated deposit-taking banks but with only one administrative branch.”

Loan Size Calculation

Average monthly payroll expenses are to be based on the full 2019 calendar year for most loans. Exceptions include seasonal businesses which may use average monthly payroll for any twelve-week period between February 15, 2019 and February 14, 2020, and new businesses may use average monthly payroll over the period from January 1, 2020 to February 29, 2020. Additionally, second draw loans by businesses in the hospitality industry (NAICS starting with 72, representing 2% of all loans) were allowed to receive 3.5 times their average monthly payroll. We account for this when computing implied compensation, and we drop second draw loans to the hospitality industry from the rounding analysis because their different compensation multiplier could lead to different loan rounding. In determining the high implied compensation flag and for the \$100k discontinuity and rounding analyses, we subtract any portion of the loan amount used to refinance an EIDL. See SBA guidance entitled “How to Calculate First Draw PPP Loan Amounts,” which was updated over time. A list of all versions can be found [here]<https://www.sba.gov/document/support-how-calculate-first-draw-ppp-loan-amounts>.

Matching Analysis

The matching analysis reported in Table III is based on a combination of propensity score matching and exact matching. First, we estimate a propensity score for whether a loan is originated by a FinTech lender using a logistic regression and the following variables: 4-digit NAICS code (industry), CBSA, business type, loan amount, jobs reported, lending draw, lending round, week of loan approval, and whether the borrower received an EIDL Advance. Second, for each loan originated by a FinTech lender, we identify loans made by traditional lenders such that the loans were made to borrowers in the same industry, CBSA, business type, and year (either 2020 or 2021) and either both received EIDL Advances or did not. Exactly matching on these characteristics ensures that we can determine each of our flags for both loans. Finally, among the loans that

match exactly on these features, we match the loans that have the smallest absolute difference in propensity scores. In total, 3,614,103 of 3,969,845 FinTech loans are matched. Note that a traditional loan may be matched to more than one FinTech loan.

Non-profit Analysis

For this analysis, we compare the PPP loan amounts that non-profits received to estimated loan amounts that they are eligible for based on compensation reported in their latest Form 990 (most commonly from 2019). The main fields of interest from the Form 990 are lines 5-10 of Part IX (“Statement of Functional Expenses”) with the sum of these lines being our measure of the non-profit’s total compensation. Note that this measure of total compensation is likely higher than eligible compensation for the PPP because we do not cap compensation per employee at \$100,000, and the Form 990 compensation includes federal payroll taxes and employee benefit costs that were not covered by the PPP. We then use this total compensation measure to calculate the implied loan amount that the non-profit should be eligible following the SBA’s guidelines for determining the maximum loan size (as outlined the “Loan Size Calculation” section above). Loan amount inflation is defined as $\frac{\text{Loan Amount} - \text{Form 990 Implied Loan Amount}}{\text{Loan Amount}}$.

Repayment, Enforcement Action, and Canceled Loans Data

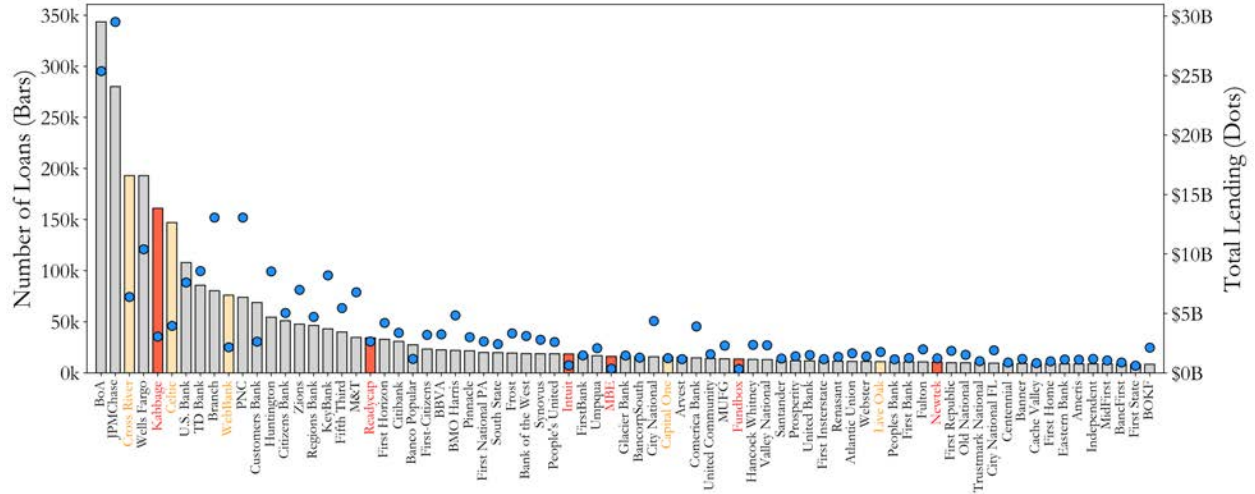
To examine round 1 and 2 loans that have been repaid, we use PPP loan-level data released by the SBA on December 2, 2020 and data from USASpending.gov as of June 30, 2021. The USASpending data provide monthly updates on the PPP loans, which allows us to observe which loans have been repaid by the borrower based on a loan balance being zeroed out without an associated U.S. Treasury account to fund forgiveness for the loan. It is also possible that some of these loans may have been canceled. The USASpending data do not include all of the loan details included in the SBA data, and after being repaid or canceled, the loans are removed from the SBA loan-level PPP data. Fortunately, the December 2020 version of the PPP loan-level data provides the same details as the main PPP data and covers all loans that had not been repaid as of December 2020. Many borrowers requested loans from multiple lenders to increase the odds of receiving funds; we exclude such loans by filtering out businesses with multiple loan requests where all but one is repaid.

To examine enforcement actions by the government, we collect information from Department of Justice criminal complaints against PPP borrowers that purportedly committed fraud based on <https://www.justice.gov/criminal-fraud/cares-act-fraud> and <https://www.arnoldporter.com/en/general/cares-act-fraud-tracker>. In total, we collect data on 162 complaints involving 355 loans. Of these 355 loans, 279 include enough information to be matched to the December 2020 version of the PPP loan-level data. Most of the unmatched loans were repaid before December 2020 and thus are not in the loan-level data. To examine canceled loans, we use loan-level data released by the SBA on May 3, 2021. This data includes all PPP loans that were approved as of May 3, 2021. We compare this data to our main loan-level data that was released on June 30, 2021 to determine which loans were canceled between May 3 and June 30, 2021.

Figure IA.1. Fintech Market Share

This figure shows the role of fintech lenders during the PPP (expanding on Figure 1). Panel A replicates Figure 1, Panel A based on loans from only rounds 1 and 2. Panel B shows the number of loans originated by lender type during each week of the PPP. In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders. Note that mid-August through December 2020 is not shown in Panel B since no PPP loans were originated during this period.

Panel A. Rounds 1 and 2 Lenders (Top 75)



Panel B. Lender Composition, by Week

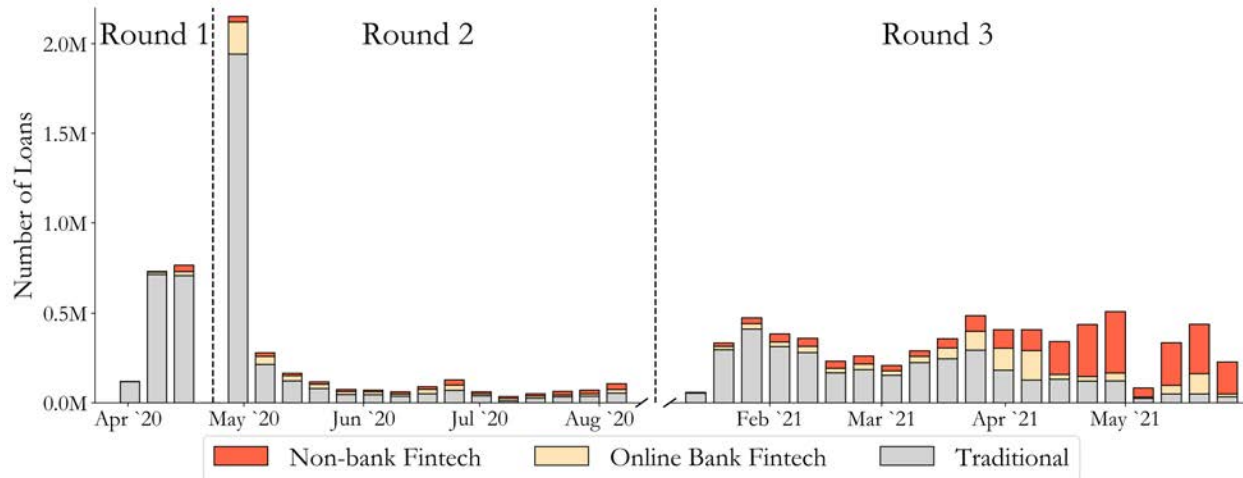
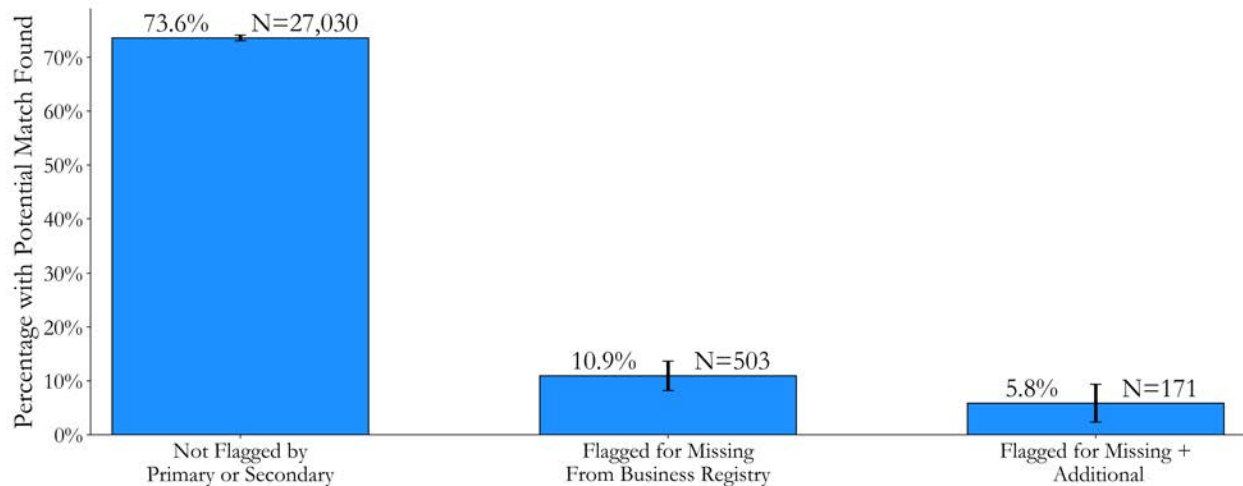


Figure IA.2. Florida Restaurant License Analysis

This figure shows the percentage of loans to Florida restaurants (defined as NAICS code starting with 722) structured as corporations, S-corporations, and LLC that can be matched the Florida Department of Business and Professional Regulation's (DBPR) restaurant/hotel license data (available at <http://www.myfloridalicense.com/DBPR/hotels-restaurants/public-records/#1506342906681-c46dd821-30bf>). Panel A shows the percentage of loans with potential matches for three different subsample of loans to Florida restaurants: those that are not flagged by any of our primary or secondary flags, those flagged as missing from the business registry, and those flagged as missing and also flagged by an additional primary or secondary flag. In this panel, a loan is considered to have a potential match if there exists a license with a licensee or location name that has at least a similarity ratio of 90% with the borrower name listed on the PPP loan and the loan and license listed the same location (zip, city, or county). The error bars show the 95% confidence intervals. In Panel B, we show the sensitivity of the match rate for the subsample flagged as being missing from the business registry. The horizontal axis shows the similarity ratio threshold used to determine if the loan and license matches. The blue area is the percentage of loans with a potential match if the loan and license are required to list the same location (zip, city, or county) and the light brown region shows the additional matches made if this location requirement is relaxed.

Panel A. Percentage with Potential Matches



Panel B. Sensitivity of Match Rate

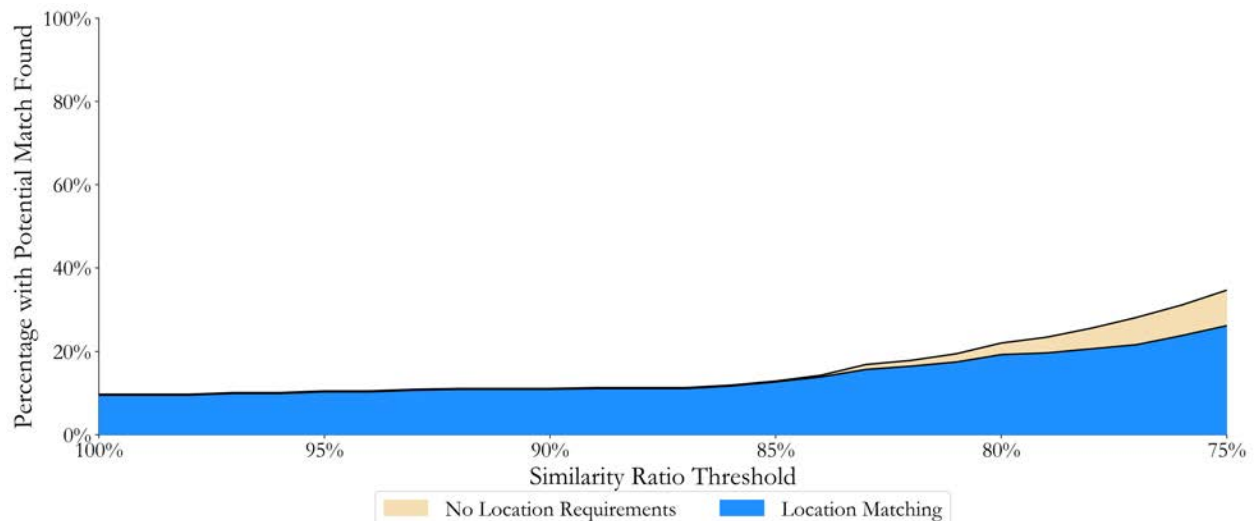
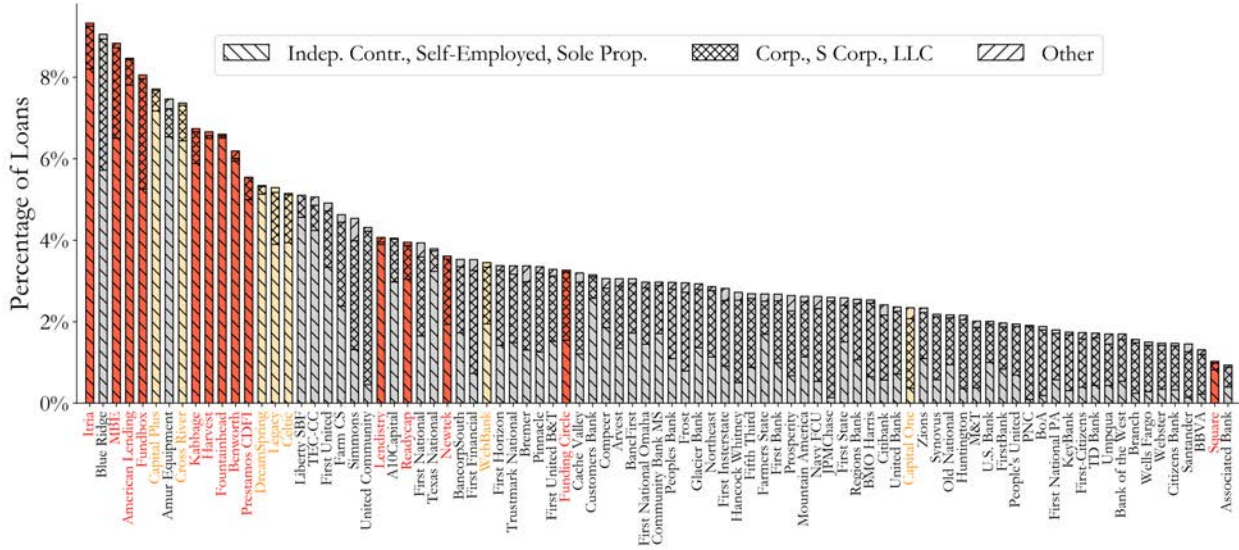


Figure IA.3. Multiple Loans

This figure expands on Figure 2, Panel B. Panel A is based on loans at residential addresses and Panel B decreases the number of loans at the same address in the same draw necessary for the loan to be flagged from three to two.

Panel A. Restricted to Loans Listing Residential Addresses



Panel B. More Than 2 Loans at Same Address in Given Loan Draw

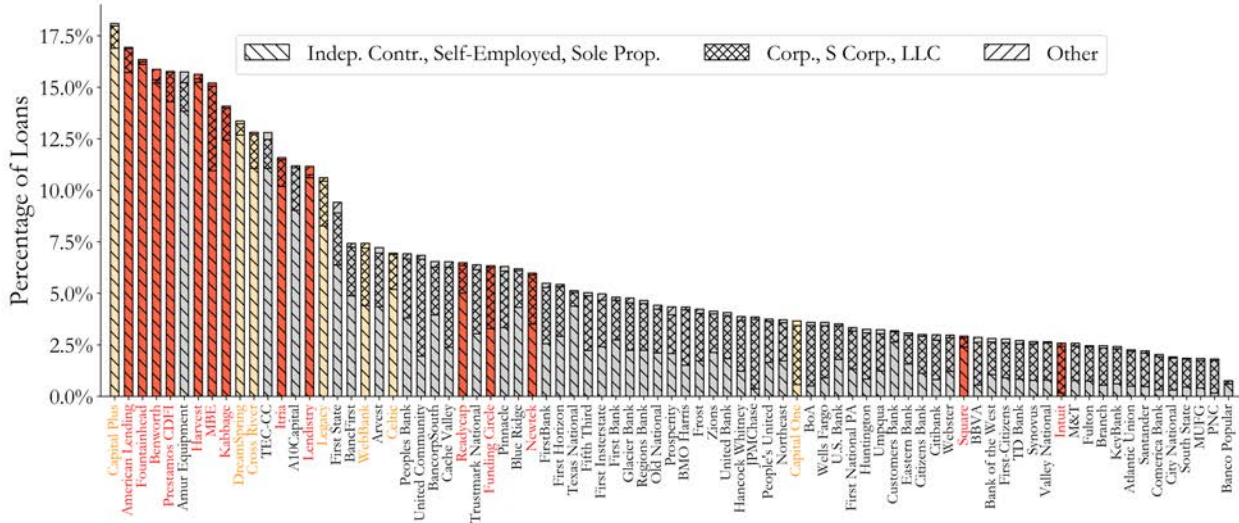


Figure IA.4. Implied Compensation

This figure compares the implied compensation of loans to the average compensation in its industry (NAICS) and CBSA. In Panel A, loans are split by round (rounds 1 and 2 in the left subpanel and round 3 in the right subpanel) and lender type. In Panel B, we focus on Round 3 loans to sole proprietorship, self-employed, independent contractors, and single member LLCs and split into loans approved before March 3, 2021 in the left subpanel and on or after March 3, 2021 in the right subpanel. In both Panels, loans are sorted and binned based on their industry-CBSA average compensation/receipts. The median annualized compensation and industry-CBSA average compensation/receipts of each bin is shown. The solid line is a 45-degree line.

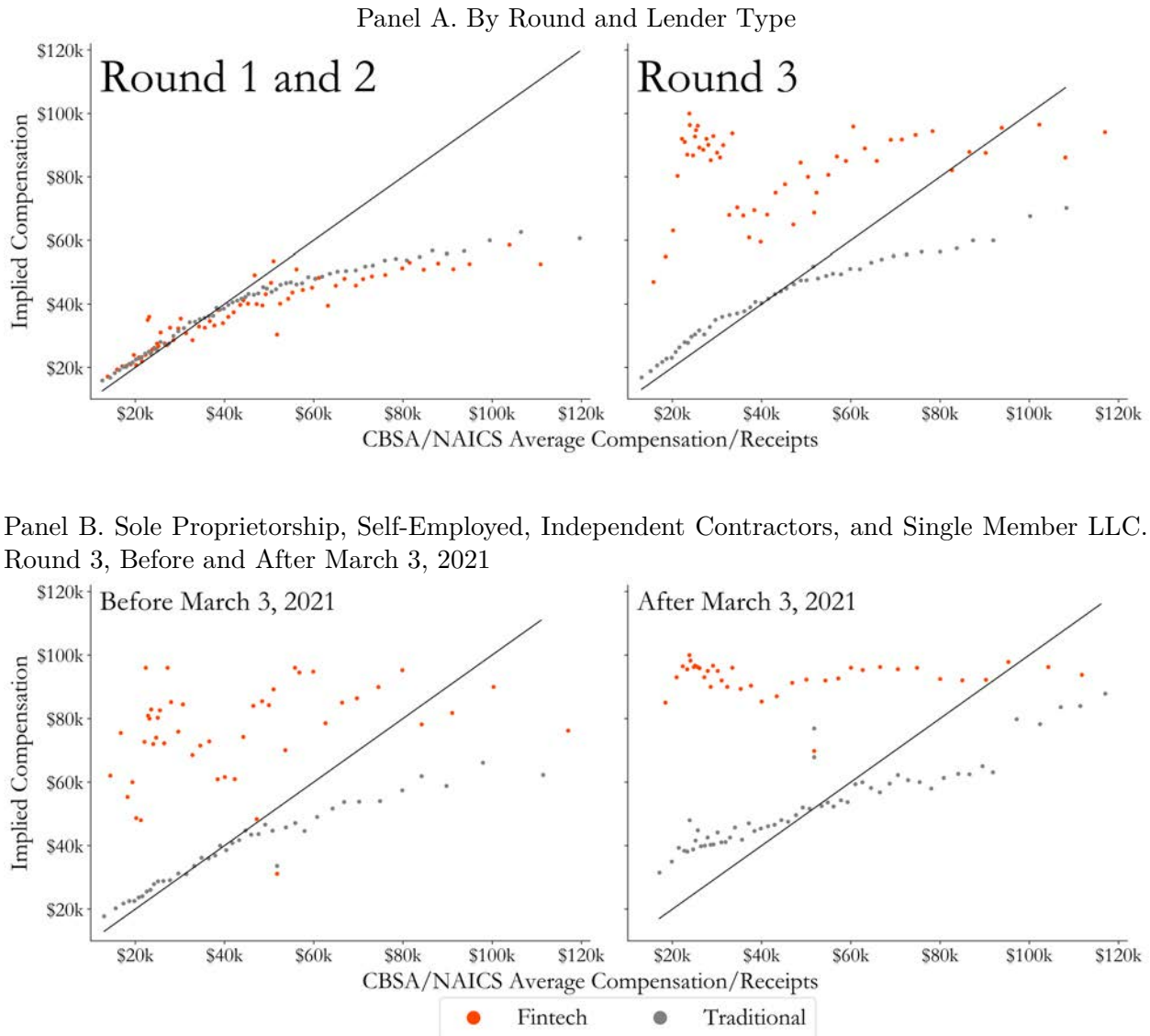


Figure IA.5. Relation Between Flags

This figure shows the relationship between the primary flags (business registry, multiple loans, high implied compensation, or EIDL > PPP jobs flags) at the lender level. Each subpanel is a scatterplot with the percentage of loans flag by one of the flags on each axis. Loans are filtered to the sets for which we can determine each flag (same as in Figures 2-4) for each axis separately (i.e., we do not require that both flags be able to be determined for a given loan). Lenders with at least 5,000 loans are shown; for the subpanels with the EIDL > PPP jobs flag, we additionally require that the lender have at least 1,000 loans with a matched EIDL Advance. The dashed line is a linear fit and the correlation is shown in the bottom left corner of each subpanel. Red triangles represent non-bank fintech lenders, cream squares represent online bank fintech lenders, and grey circles represent traditional lenders.

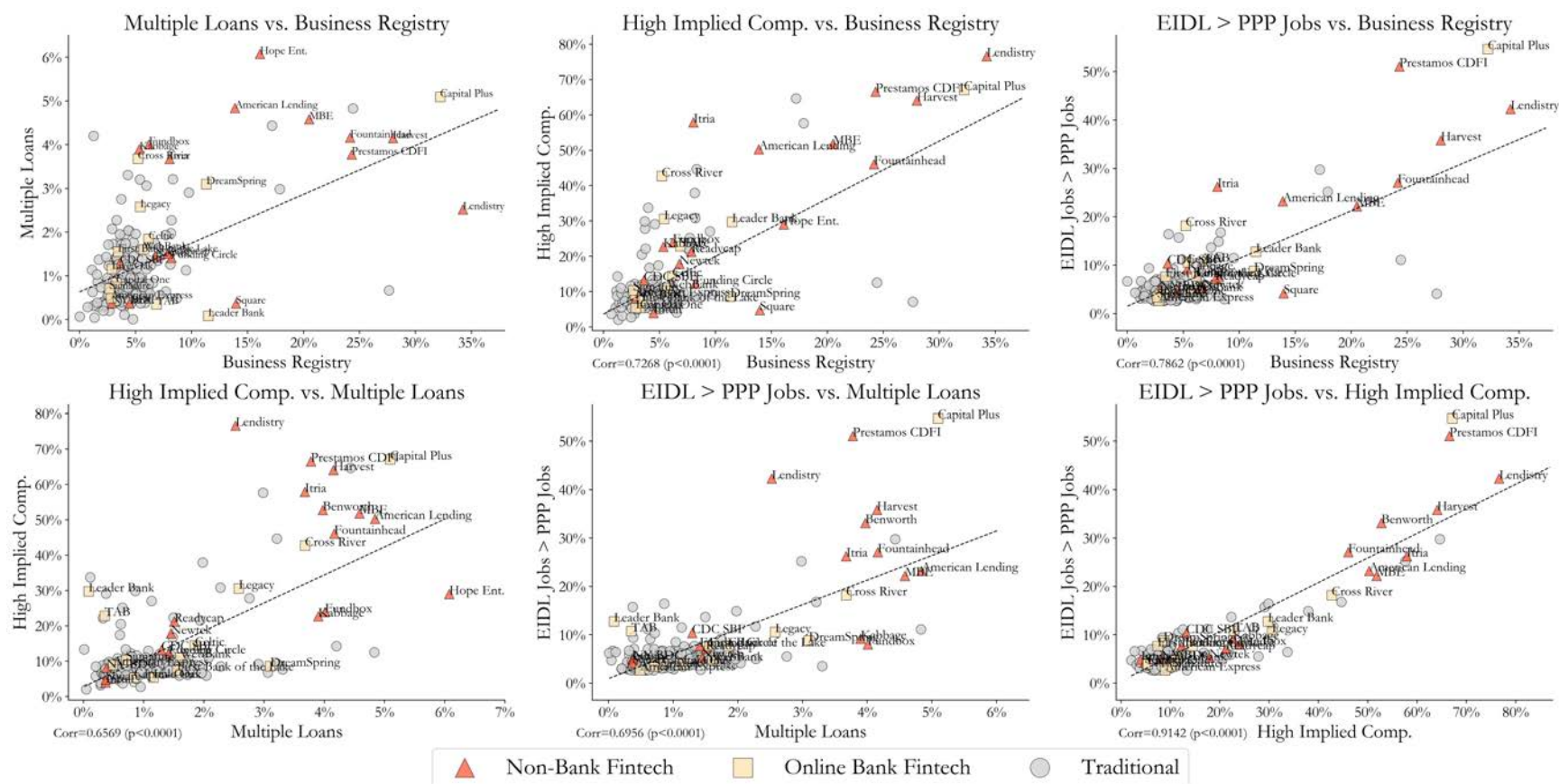


Figure IA.6. Loan Size Inflation by Non-Profits

This figure shows the relationship between non-profits inflating their loan size and being flagged by our primary measures. Loan amount inflation is defined $\frac{\text{Loan Amount} - \text{Form 990 Implied Loan Amount}}{\text{Form 990 Implied Loan Amount}}$ where Form 990 Implied Loan Amount is calculated based on the non-profit's 2019 Form 990 submission to the IRS. The left axis shows the percentage of loans with at least the given percentage of loan amount inflation that are flagged by our primary measures. The right axis shows the complementary cumulative distribution (i.e., the percentage of loans that have at least the given percentage of loan amount inflation). This analysis is based on 119,397 loans.

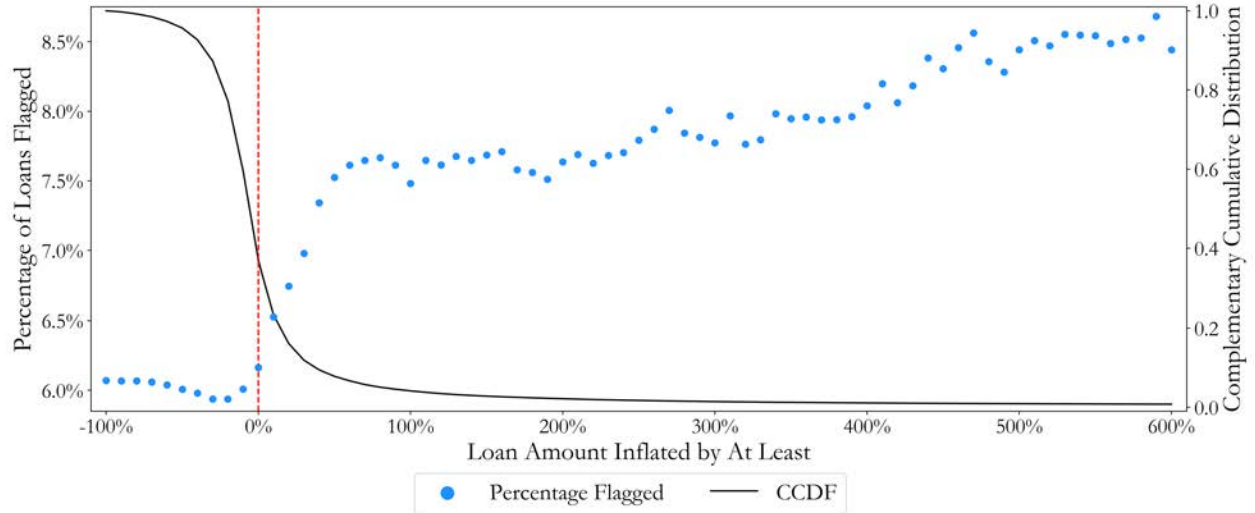


Figure IA.7. Discontinuities Around \$150,000 Loan Amount

This figure shows the percentage of loans that are flagged by loan amount. Panel A shows this across all loan amounts upto \$400,000, while Panel B focused on around \$150,000. In Panel A, loans are binned into \$1,000 wide bins. The left axis shows the percentage of flagged loans in each bin and the right axis shows the number of loans in each bin (on a log scale). In Panel B, loans are binned into \$250 wide bins, solid lines are linear fits (separately for bins below and above \$150,000), and the dashed lines are 95% confidence intervals. Fintech and traditional loans are shown separately. In both panels, the purple vertical line denotes the \$150,000 loan amount maximum threshold for streamline loan forgiveness

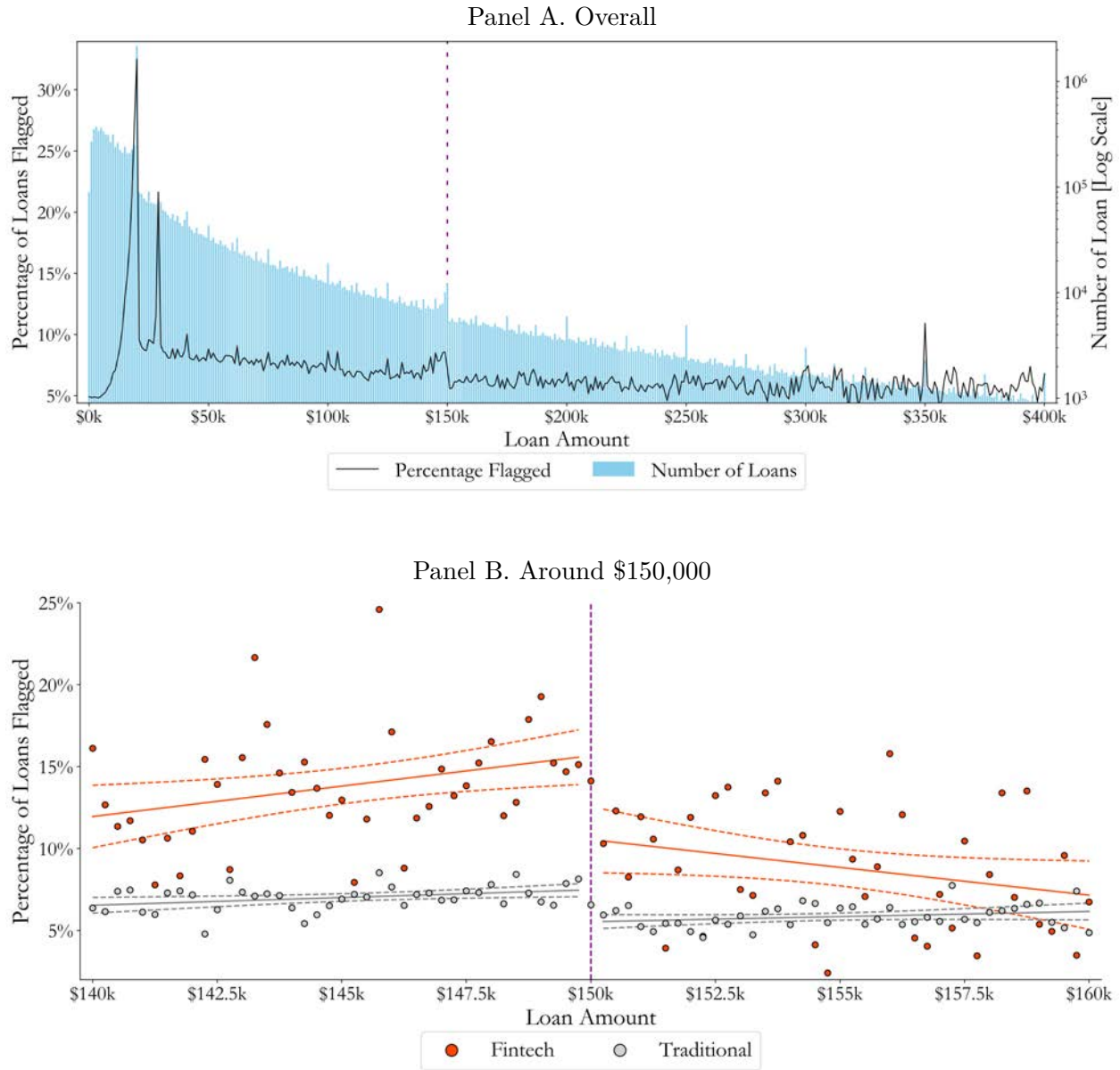
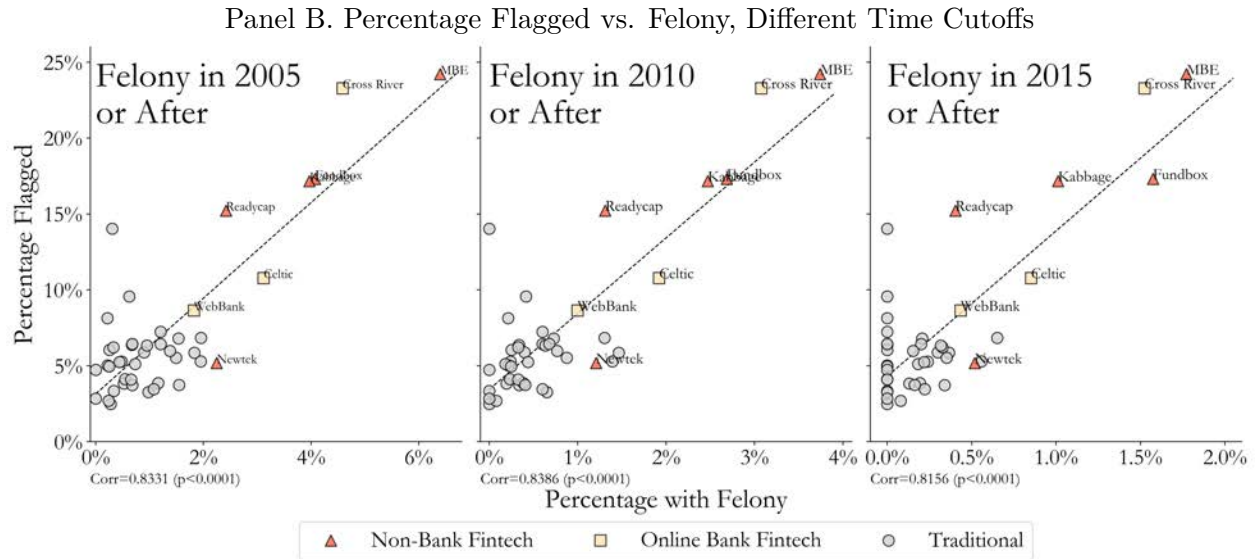
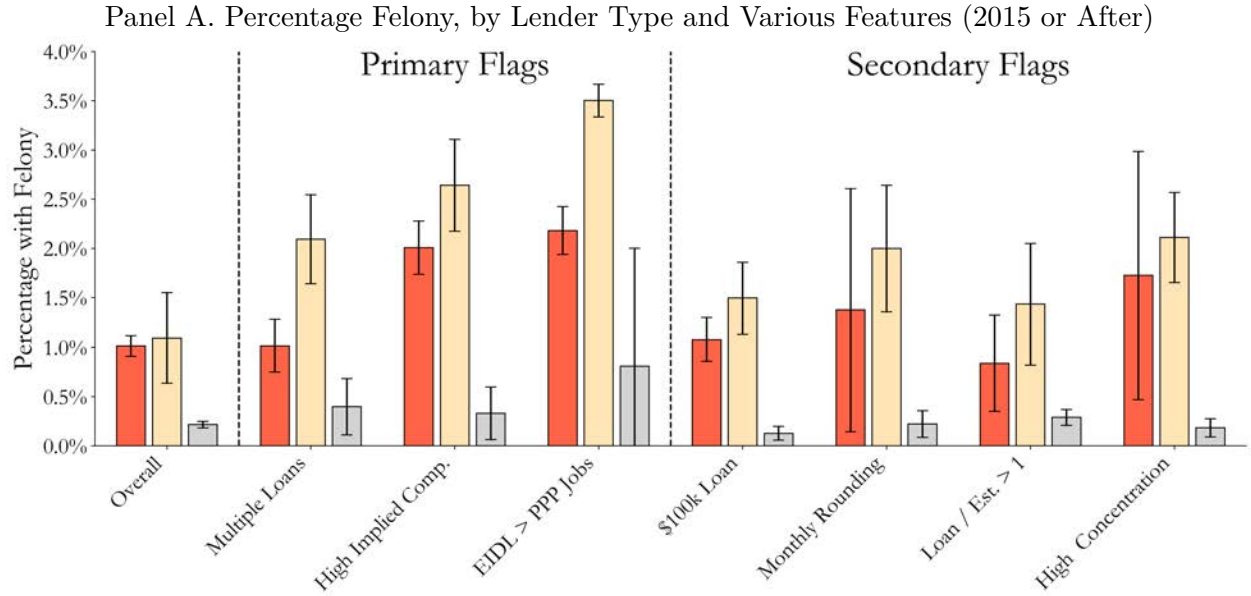
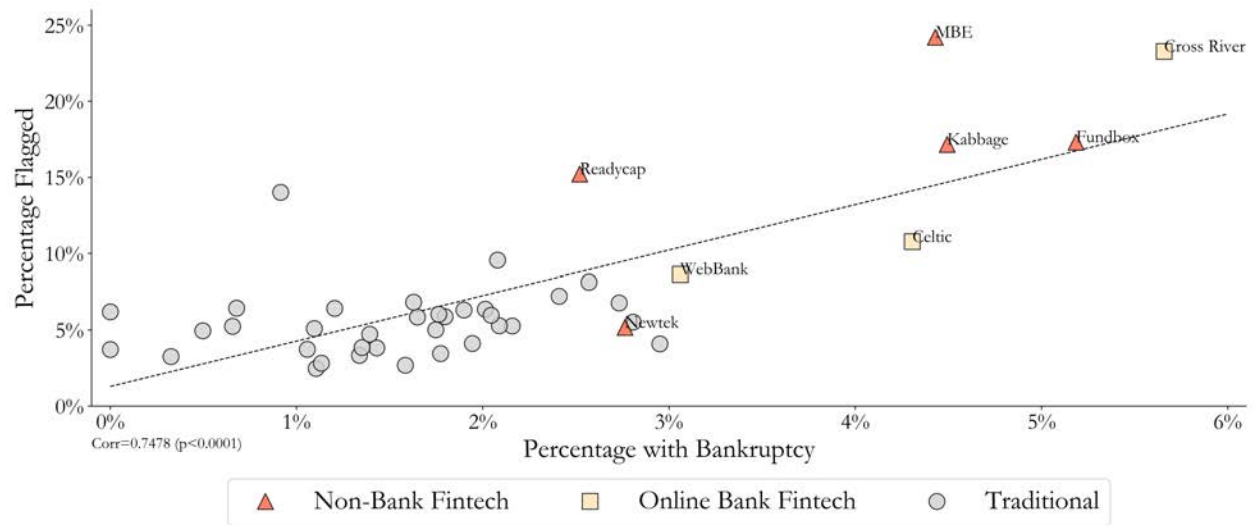


Figure IA.8. Criminal Records

This figure shows additional features (expanding on Figure 9) of our sample of criminal records for 150,000 Round 1 and 2 loans. Panel A replicates Figure 9, Panel A using felonies from 2015-2020, Panel B replicates Figure 9, Panel B using various time cutoffs, Panel C replicates Figure 9, Panel B using bankruptcies (2015-2020), and Panel D shows the percentage of felonies (2000-2020) by lender type and across implied compensation. In Panels B and C, lenders with at least 0.2% of the loans in the sample (300 loans) are shown. The dashed lines are linear fits and correlations are in the bottom corner. In Panel D, loans are binned into \$4,000 wide bins, solid lines are third-degree polynomial fits, and the dashed lines are 95% confidence intervals.



Panel C. Percentage Flagged vs. Bankruptcies



Panel D. Percentage with Felony, by Implied Compensation

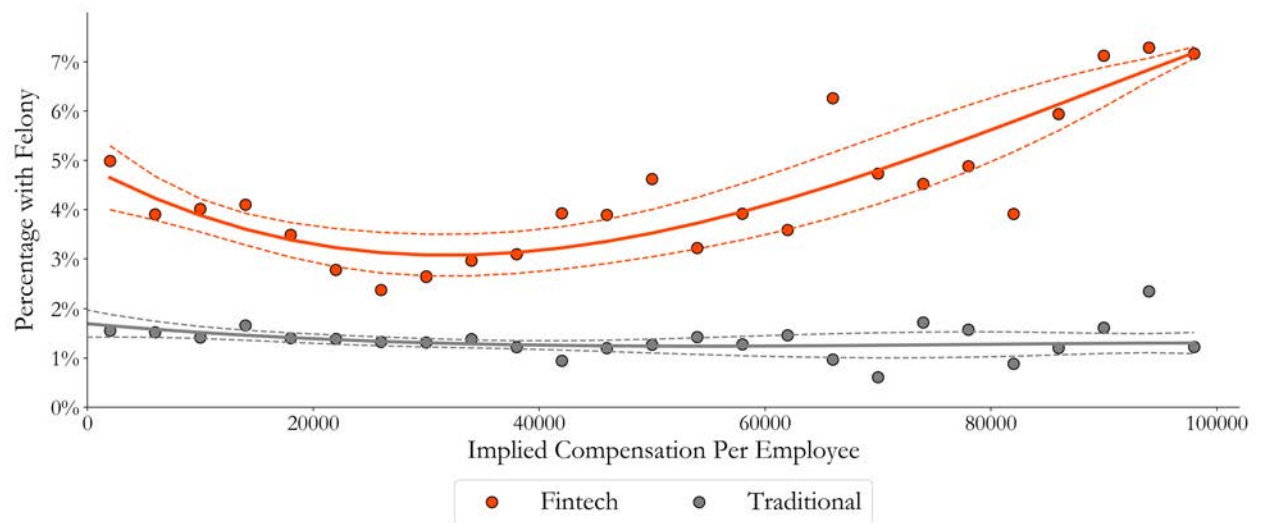


Figure IA.9. Relationship Between Primary and Secondary Flags

This figure shows the relationship between the primary flags (summarized as whether the loan is flagged by at least one of them) and the secondary flags by lender. The corresponding figure for criminal records is shown as Figure 9, Panel B. For monthly rounding, loans with one job reported are excluded; for overrepresentation, loans to self-employed and independent contractors, second draw loans, and loans in a industry-county pair not in the CBP data are excluded; for high concentration, second draw loans and loans in a lender-county pair with fewer than 25 first draw loans are excluded. No loans are excluded for percentage flagged by at least one primary flag. Lenders with at least 5,000 loans are shown. The dashed line is a linear fit and the correlation is in the bottom corner of each panel. Red triangles represent non-bank fintech lenders, cream squares represent online bank fintech lenders, and grey circles represent traditional lenders.

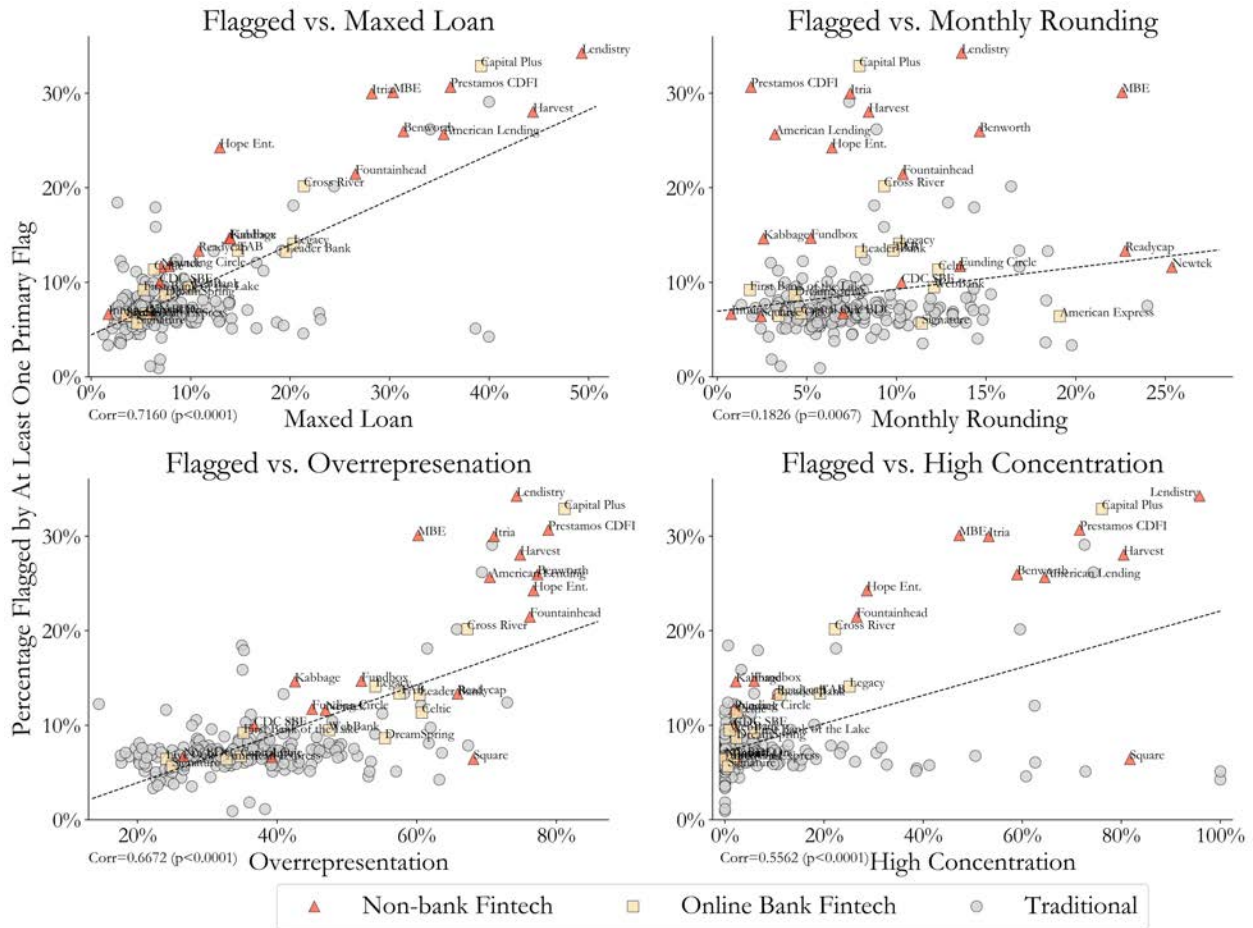
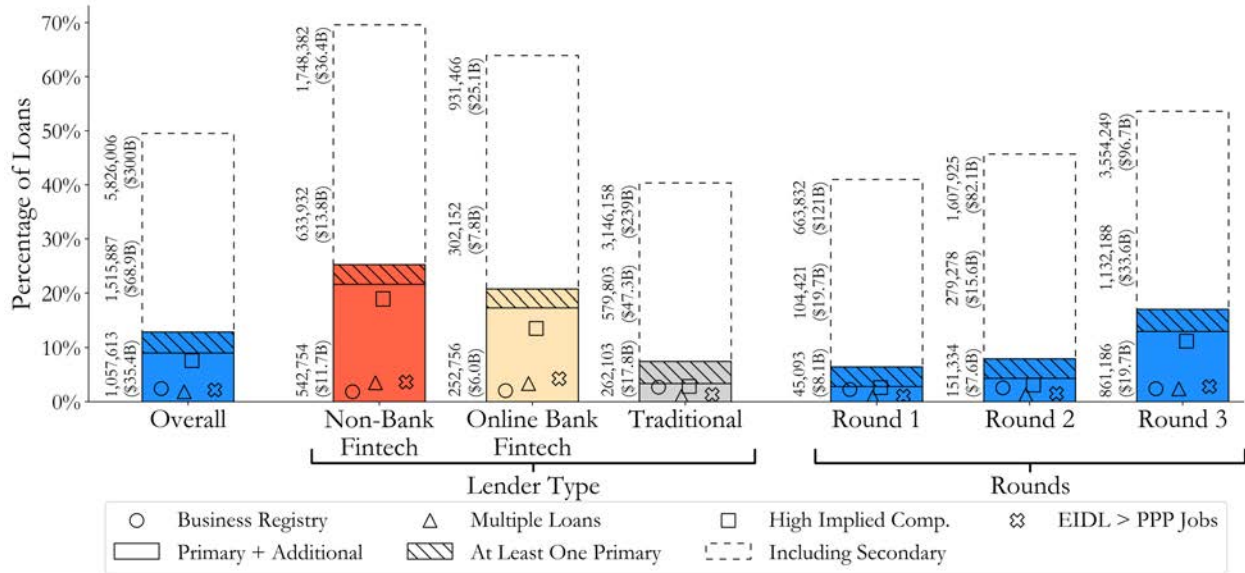


Figure IA.10. Overall Misreporting Flag Rates – Additional Details

This figure shows additional variation in percentage of loans flagged. Panel A replicates Figure 10, Panel A and adds loans flagged by at least one primary or secondary flag as the dashed, non-shaded portions of the bar. In Panel B, each subpanel shows a lender type and each series is the percentage of loans flagged by the given flag across time. The vertical dotted lines split each subpanel into the three PPP lending rounds. The loans used to calculate each series are filtered to the sets for which we can determine each flag (same as in Figures 2-4). Note that mid-August through December 2020 is not shown in Panel B since no PPP loans were originated during this period. In both panels, red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders.

Panel A. Percentage of Loans Flagged with Broader Flags, by Lender



Panel B. Percentage of Loans Flagged Over Time, by Lender Type

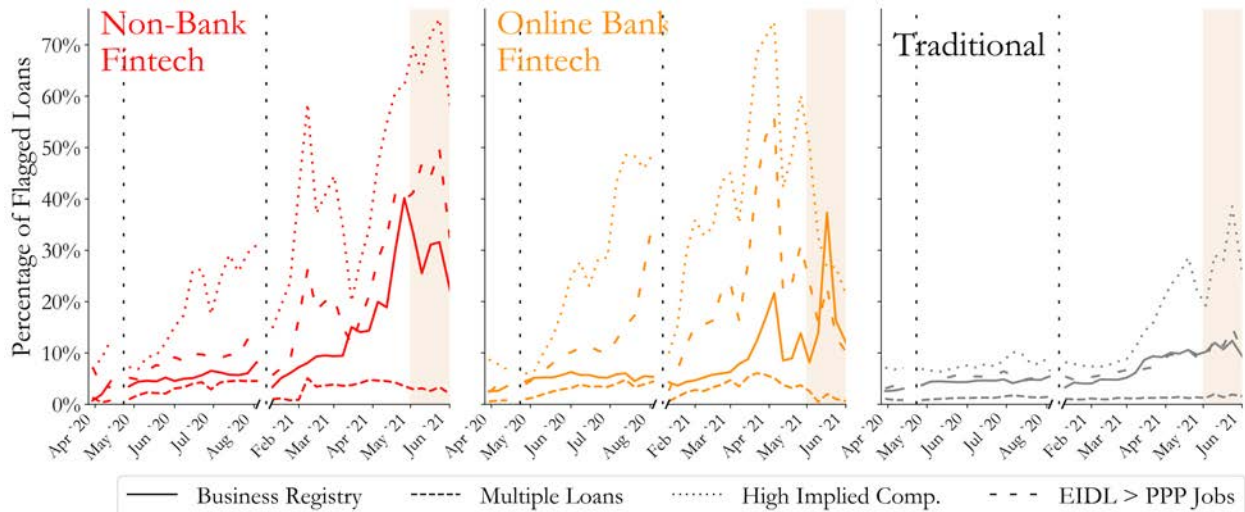
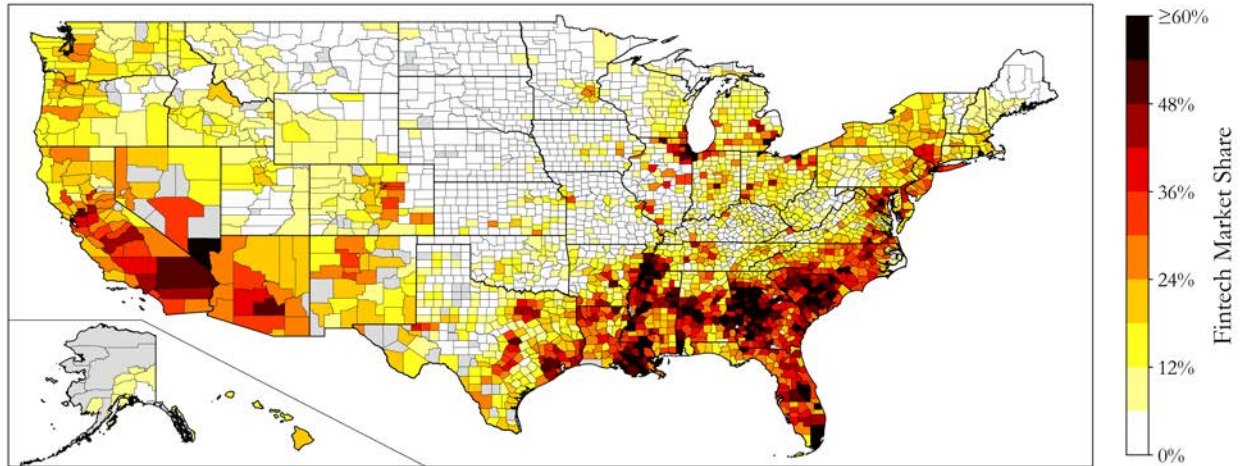


Figure IA.11. Geography

This figure shows additional geographic variation (extending Figure 11). Panel A shows the fintech market share in each county. Panel B shows the growth in lending between rounds 1-2 and round 3 in each county. In both panels, counties are colored based on the color scheme shown in the bar to the right of the maps and counties with fewer than 100 loans are colored grey.

Panel A. Fintech Market Share, by County



Panel B. Lending Growth, by County

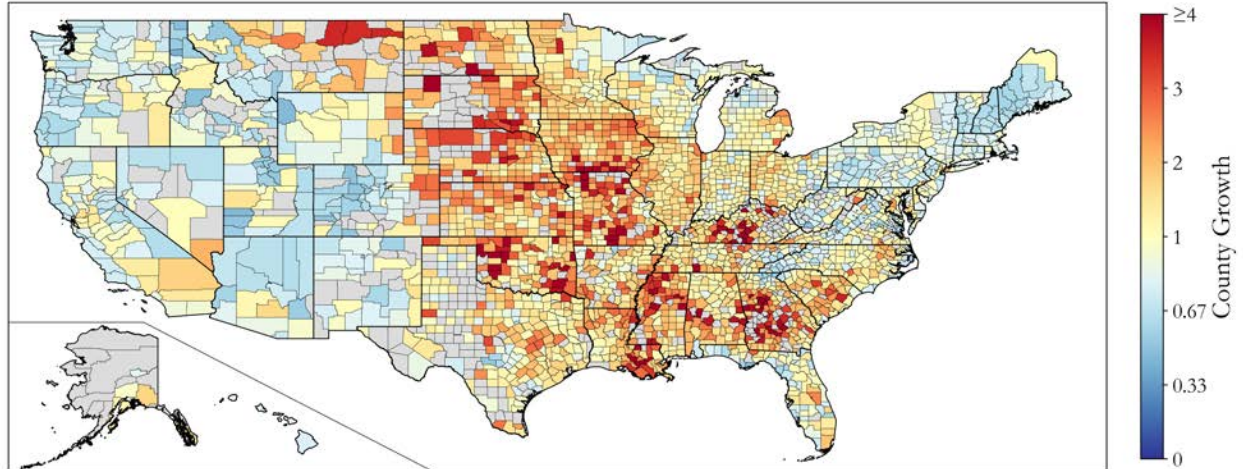


Figure IA.12. Pairwise Lender Correlations

This figure shows pairwise lender correlations using data at the county-lender level. The lower triangular shows correlations between the percentage of loans flagged by at least one primary flag and the upper triangular shows correlations between the lenders' market shares across counties. Lenders are ordered such that those with the highest percentage of loans flagged by at least one primary flag (across the entire sample) are at the top on the vertical axis and the left on the horizontal axis. Labels are colored red for non-bank fintechs, orange for online bank fintechs, and black for traditional lenders. Coloring of each square in the matrix is based pairwise correlation and the coloring scheme is shown at the bottom with darker red representing higher positive correlation and darker blue representing higher negative correlation. For each pairwise correlation, counties are filtered to the set that have 25 loans by both lenders. The top 75 lenders (by number of loans) are shown.

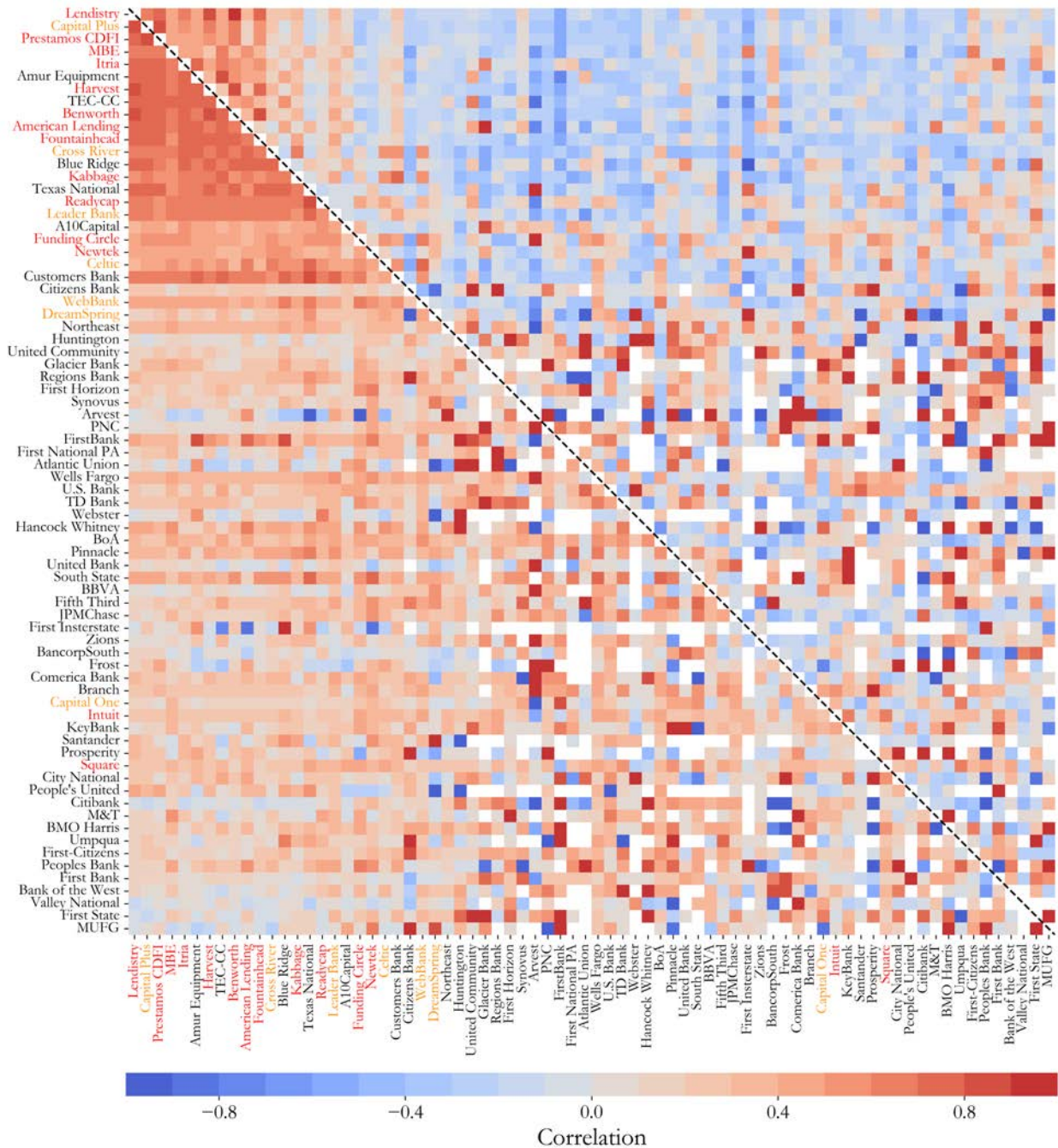


Figure IA.13. PPP Liquidity Facility

This figure shows the the total amount of PPP Liquidity Facility (PPPLF) advance received by each of the top 100 PPP lenders (by number of loans). The horizontal axis shows the total original outstanding advance amount for each lender and the vertical axis shows the percentage of PPP loans flagged. Lenders that did not receive a PPPLF advance are included with their PPPLF total advance amount set to zero. The circle size represents the total dollars lent by each lender. The dotted line is a line of best fit and the correlation is shown in the bottom left corner. Red represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders. Lenders that received at least \$500 million in PPPLF advances are labeled.

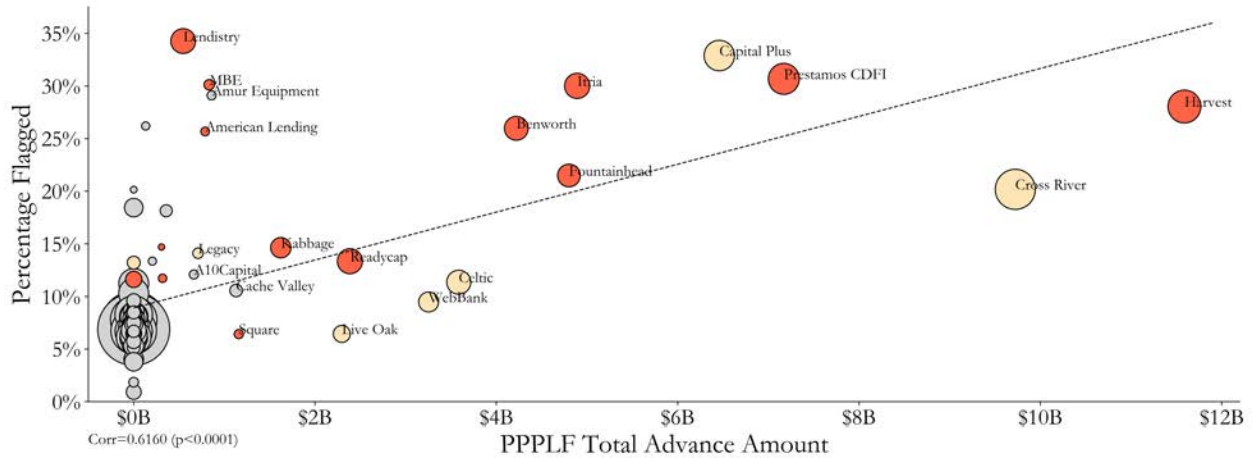


Figure IA.14. Lender Network

This figure replicates Panel B of Figure 12 using all loans where the borrower received a first and second draw loan. Node size is proportional to the number of first draw loans (which also received either a second draw from the same or different lender) and second draw originated by the lender. Edges are directed and have a width proportional to the number of loans moving clockwise from the first draw lender to the second draw lender. Red nodes are fintech lenders and grey nodes are traditional lenders. Pure red edges are between two fintech lenders, pure grey edges are between two traditional lenders, and darker red edges are between a fintech and traditional lender. Top 100 lenders (by the same measure used for node size) are shown and the remainder are combined into the “Other” nodes (one for other fintech lenders and one for other traditional lenders).

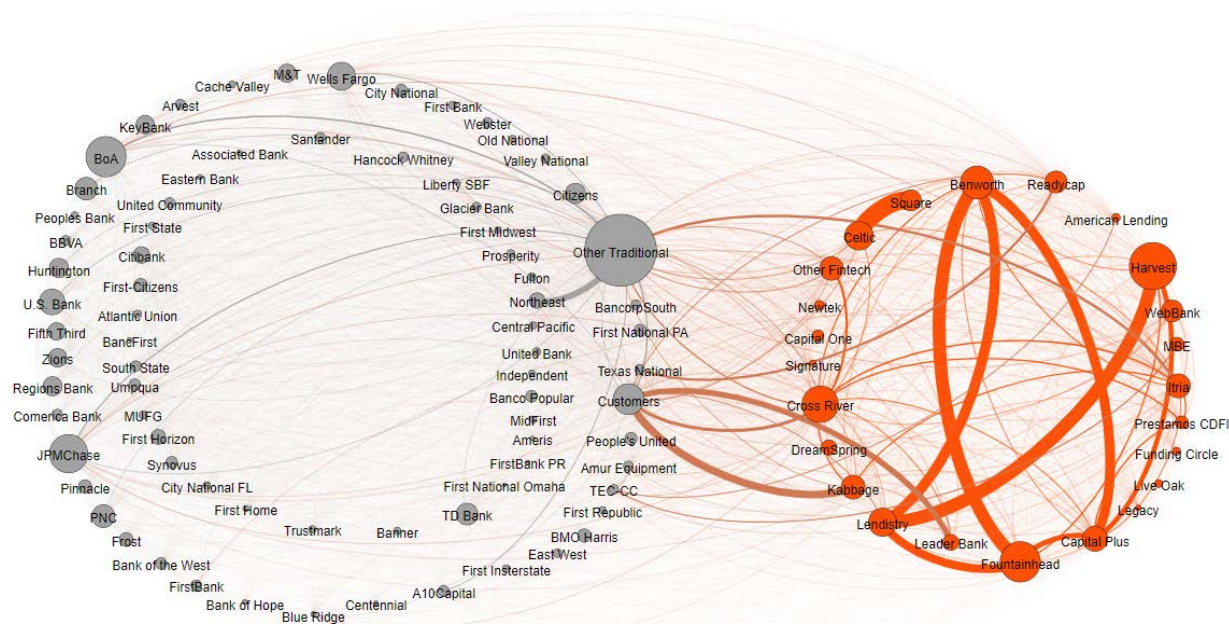
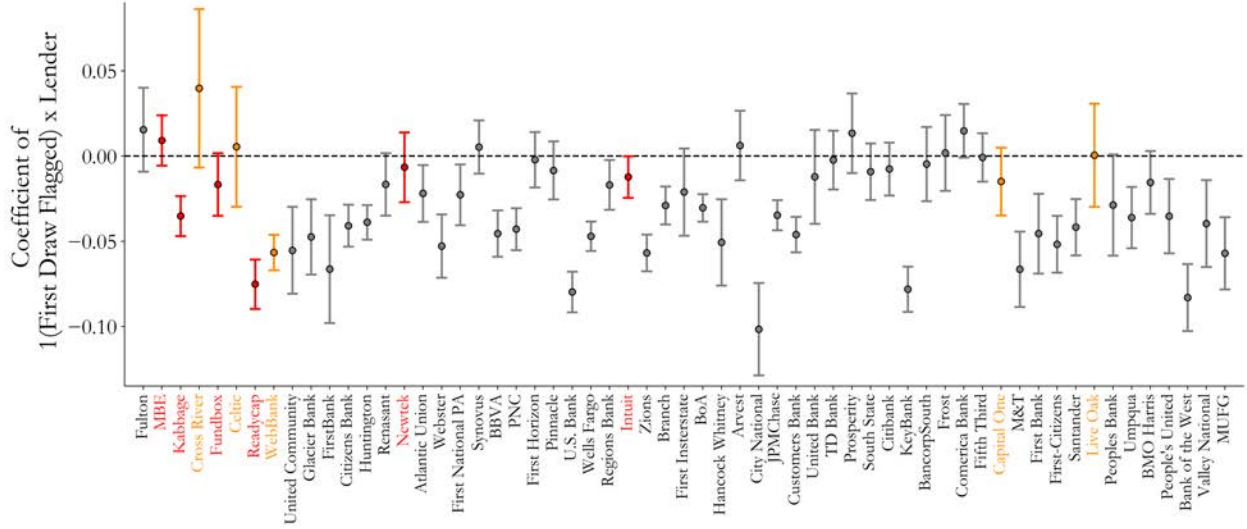


Figure IA.15. Likelihood of Receiving a Second Draw Loan

This figure examines whether lenders were more/less likely to provide a second draw loan to a borrower who's first draw loan is flagged by at least one of our primary flags. Similar to Table VII, we estimate a OLS regression with a dummy for whether the same lender provided the first and second draw loans as the dependent variable and a dummy for whether the first draw loan was flagged by at least one of the primary flags interacted with an indicator for each lender as the independent variables. The regression control for loan size and jobs and include zip code, business type, and NAICS \times CBSA fixed effects. For Panel A, if a borrower did not receive a second draw loan, the dependent variable is set to 0, and in Panel B, only borrowers that received both a first and second draw loans are included in the sample. In both panels, the hollow dots show the point estimates from the regression and the error bars show 95% confidence intervals corrected for multiple comparisons using a Bonferroni correction. Lenders with at least 10,000 loans in rounds 1-2 are shown and are sorted such that those with the highest percentage of flagged loans (in rounds 1-2) are on the left. Red error bars and labels represents non-bank fintech lenders, cream represents online bank fintech lenders, and grey represents traditional lenders.

Panel A. Unconditional of Receiving Second Draw



Panel B. Conditional on Receiving Second Draw

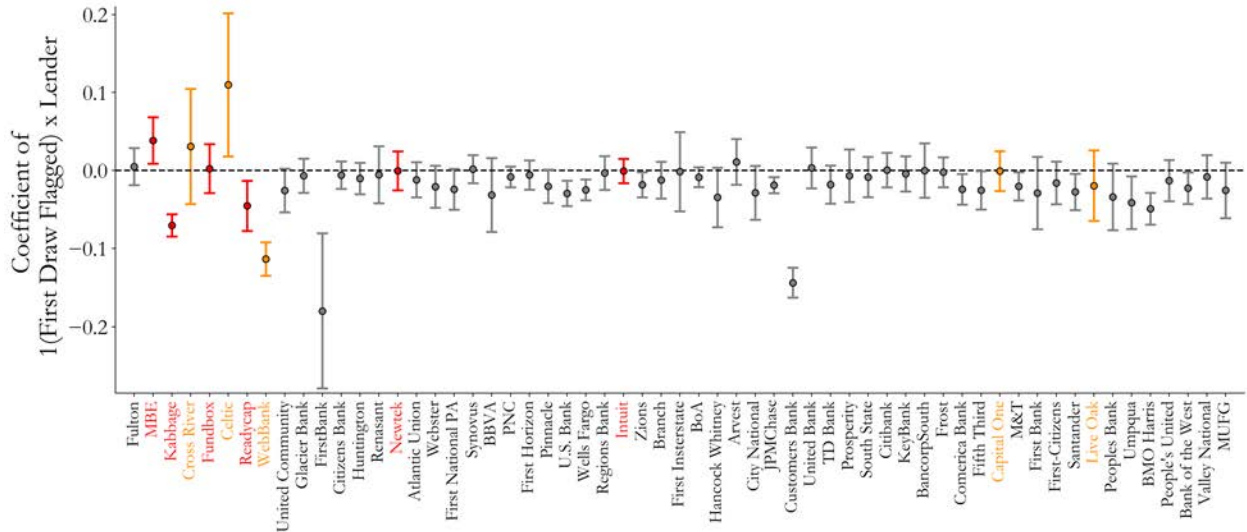
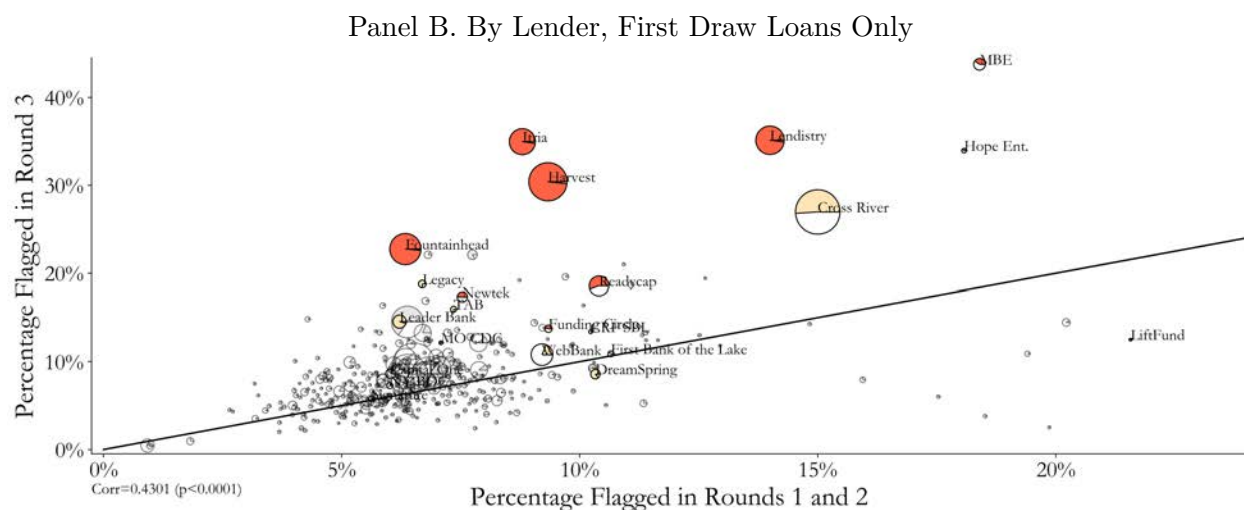
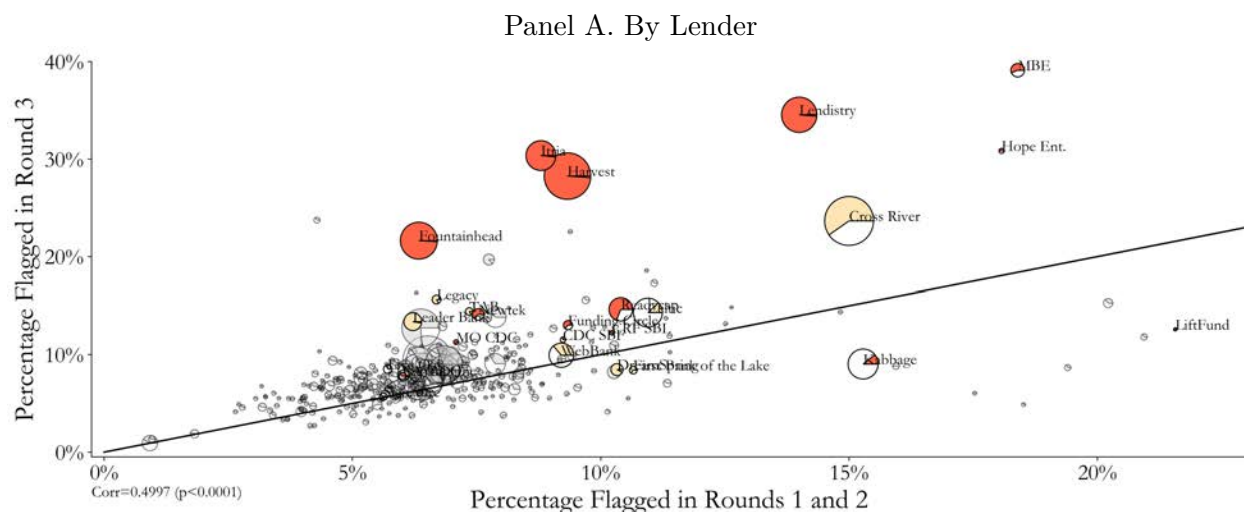
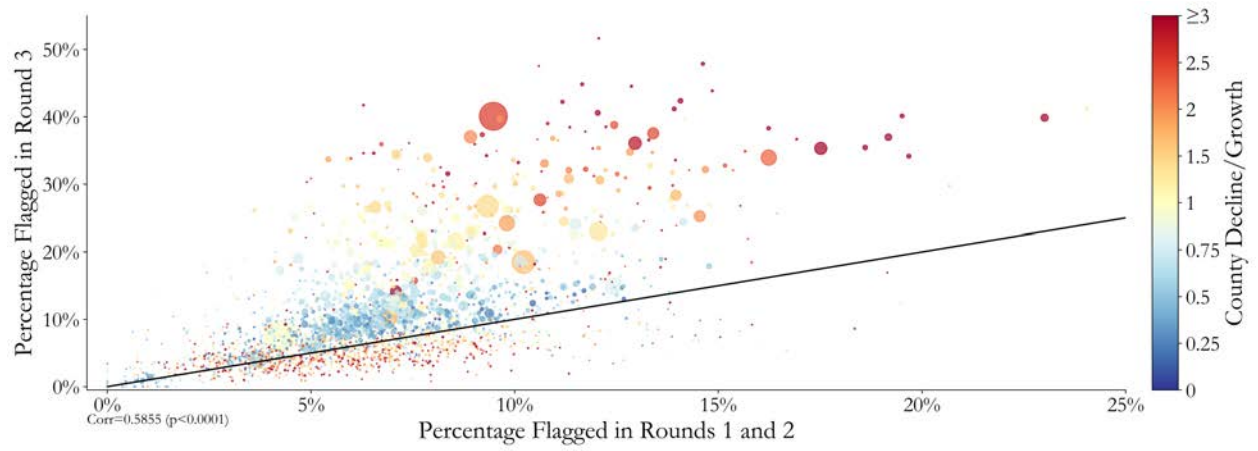


Figure IA.16. Persistence and Growth Across Rounds

This figure shows the persistence and growth of flagged loans across lending rounds. Panel A shows this by lender using all loans and Panel B by lender using first draw loans only. Panel C and D shows this by county and state (by zip code is shown as Figure 13, Panel B). For all panels, the percentage of loans flagged in rounds 1 and 2 are shown on the horizontal axis and round 3 on the vertical axis. For Panel A, lenders with at least 1,000 loans in rounds 1 and 2 combined and in round 3 are shown. For Panel B, lenders with at least 1,000 loans in rounds 1 and 2 combined and 250 first draw loans in round 3 are shown. In Panel C, counties with at least 100 loans in round 1 and 2 combined and in round 3 are shown. In all panels, the circle size corresponds to their total number of loans (across all rounds), the black line is a 45-degree line, and the correlation is presented in the bottom of each panel. For Panels A and B, the percentage of the circle that is shaded represents the proportion of loans that each lender provided in round 3 relative to in round 1 and 2. For Panel C and D, the circles are colored based lending growth in the county/state with the color scheme shown in the bar to the right of each panel.



Panel C. By County



Panel D. By State

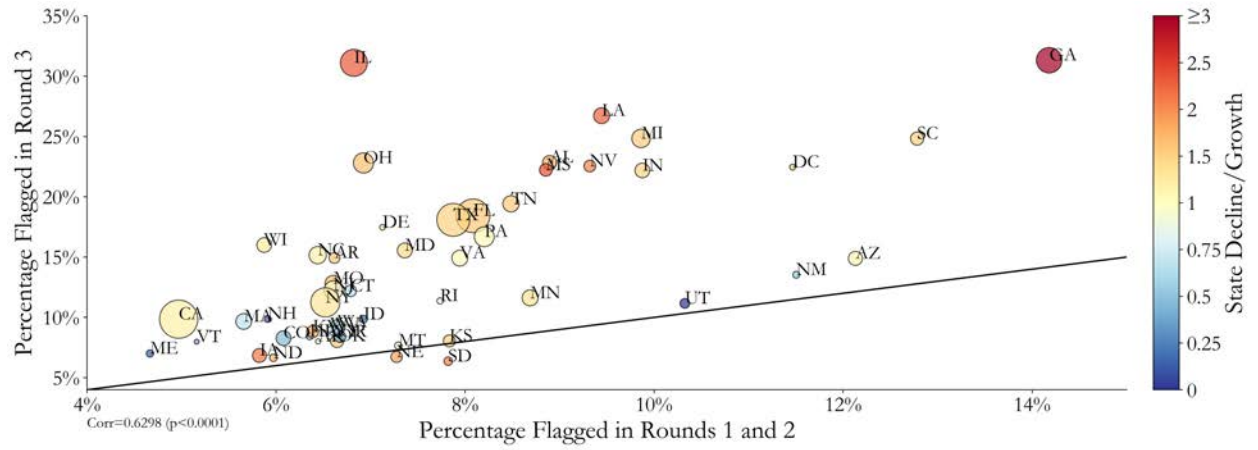
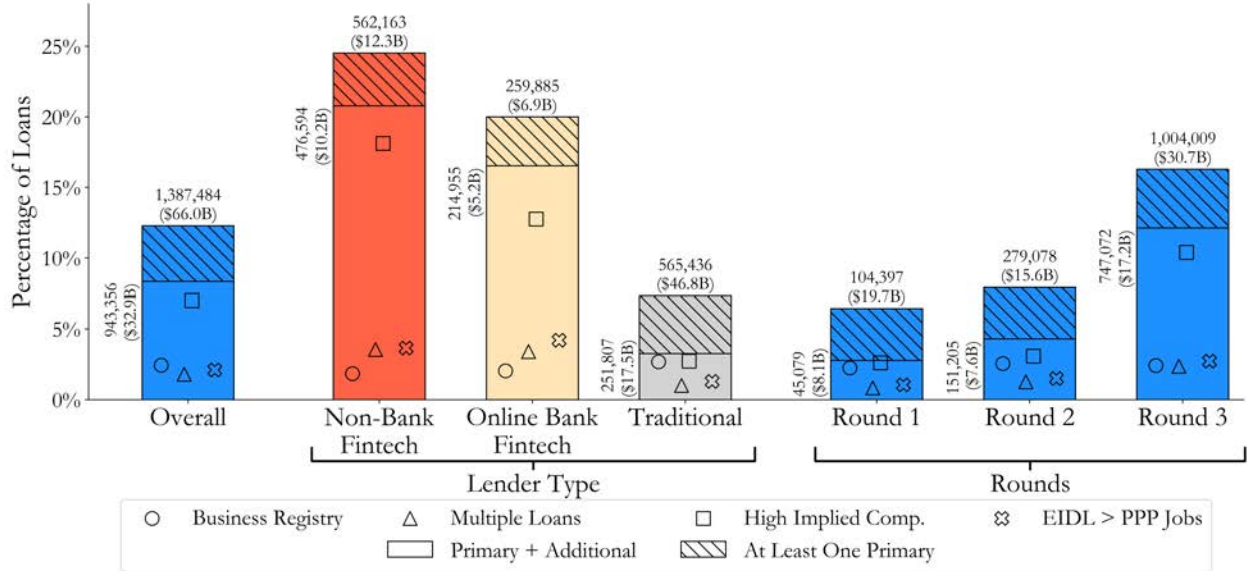


Figure IA.17. Lender Level, Excluding Undisbursed Loans

This figure replicates Figure 10 after excluding undisbursed loans. Undisbursed loans are those listed as “Active Un-Disbursed” in the loan-level data.

Panel A. Percentage of Loans Flagged, by Lender Type and Rounds



Panel B. Percentage of Loans Flagged, by Lender

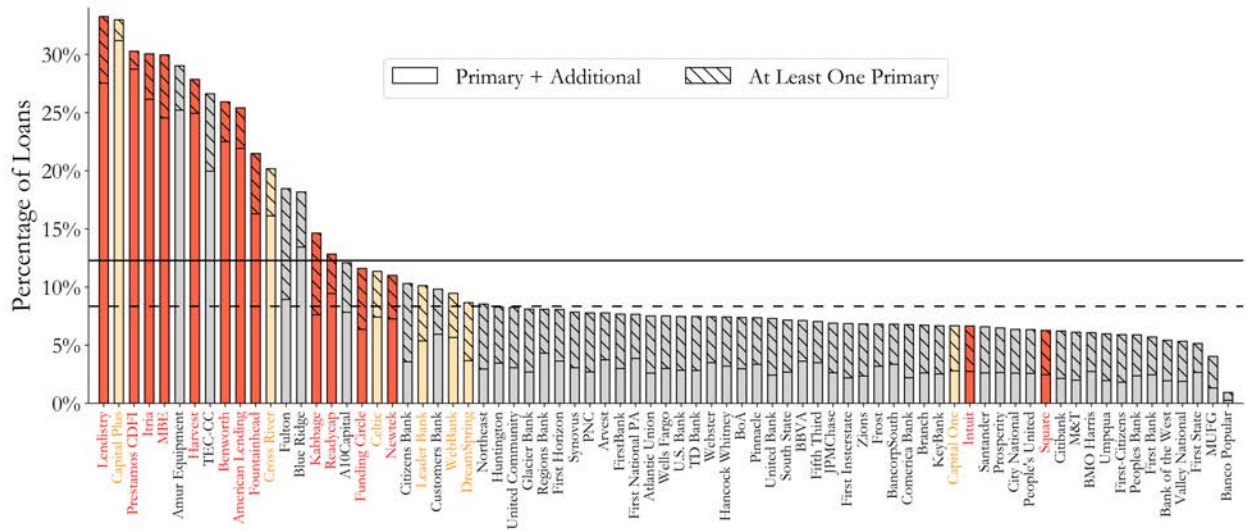


Figure IA.18. Repayment, Enforcement Actions, and Canceled Loans

This figure shows the percentage of loans that have been repaid between December 1, 2020 and June 30, 2021, part of a DOJ enforcement action, and canceled between May 3, 2021 and June 30, 2021. Loans are filtered based on the criteria at the bottom of bar. In total, 16,930 loans that were repaid between December 1, 2020 and June 30, 2021, 279 loans that are part of DOJ enforcement actions, and 95,526 loans have been canceled between May 3, 2021 and June 30, 2021. The percentage of repayment and enforcement actions is based on the loan-level data released by the SBA on December 1, 2020, and the percentage of canceled loans is based on the loan-level data released by the SBA on May 3, 2021 and .

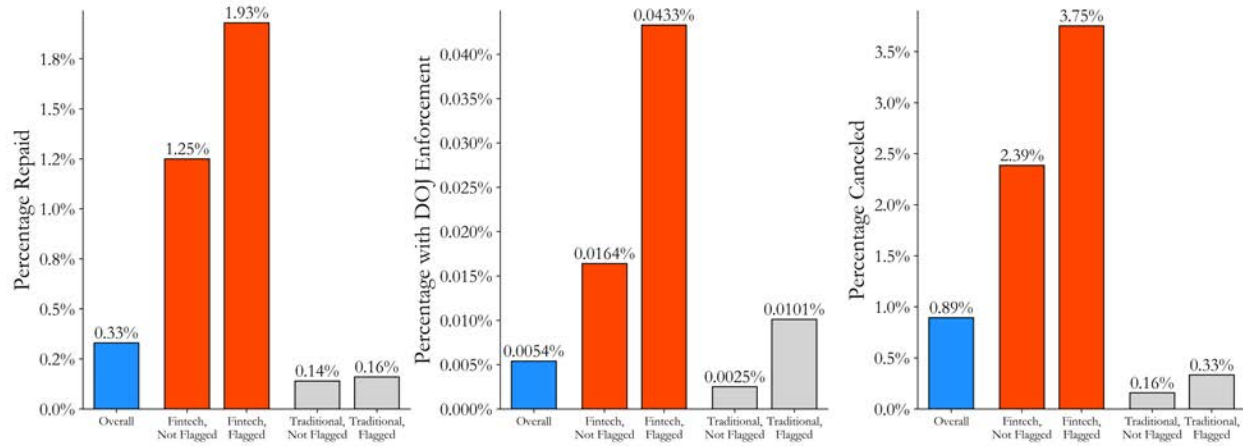


Table IA.I. Cross Verification of Flags

In this table, we examine the relationships between our four main flags. We estimate OLS regressions with dummies for each of our main flags as dependent variables and dummies for the other three flags as independent variables. Panel A shows the relationships without lender fixed effects and Panel B shows the relationship within lenders by adding lender fixed effects. In both panels, specification (1) uses business registry as the dependent variable, (2) uses multiple loans, (3) uses high implied compensation, and (4) uses EIDL > PPP jobs. Loans are filtered to the sets for which we can determine the flag being used as the dependent variable (same as in Figures!2-4). Fixed effects and control variables are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Panel A. Without Lender Fixed Effect				
Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Implied Comp.	(4) EIDL > PPP Jobs
Business Registry		0.00396*** (5.79)	0.0259*** (8.53)	0.00327** (2.19)
Multiple Loans	0.0226*** (7.57)		0.0365*** (6.09)	0.0616*** (13.39)
High Implied Comp.	0.0415*** (7.21)	0.00724*** (6.28)		0.194*** (28.62)
EIDL > PPP Jobs	-0.00698*** (-4.37)	0.0122*** (14.40)	0.0654*** (13.42)	
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes
1(EIDL Adv. Matched)	Yes	Yes	Yes	No
Lender FE	No	No	No	No
Observations	5,305,966	11,086,014	3,410,487	2,653,048
Num. Lenders	4,722	4,873	4,765	4,732
R^2	0.100	0.047	0.592	0.282
Mean of Dep. Variable	0.0513	0.0182	0.252	0.0928

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Panel B. With Lender Fixed Effect

Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Implied Comp.	(4) EIDL > PPP Jobs
Business Registry		0.00320*** (5.96)	0.0131*** (5.09)	-0.00158 (-0.79)
Multiple Loans	0.0140*** (6.96)		0.0293*** (7.45)	0.0462*** (11.08)
High Implied Comp.	0.0294*** (9.25)	0.00579*** (7.73)		0.164*** (23.30)
EIDL > PPP Jobs	-0.00706*** (-4.57)	0.0108*** (12.56)	0.0580*** (12.55)	
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip Code FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes
1(EIDL Adv. Matched)	Yes	Yes	Yes	No
Lender FE	Yes	Yes	Yes	Yes
Observations	5,305,887	11,085,989	3,410,400	2,652,927
Num. Lenders	4,648	4,849	4,679	4,628
R^2	0.116	0.051	0.608	0.311
Mean of Dep. Variable	0.0513	0.0182	0.252	0.0928

t-statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Table IA.II. Prevalence of Flags by Lender Types

Panel A shows the full results for the unadjusted and adjusted differences presented in Table III. Fixed effects and control variables are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Panel A. Unadjusted Percentages						
Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Comp.	(4) EIDL > PPP Jobs	(5) At Least One	(6) At Least Two
Fintech	0.0587*** (3.09)	0.0244*** (8.71)	0.395*** (7.83)	0.168*** (4.34)	0.161*** (8.12)	0.0210*** (7.39)
Observations	5,564,750	11,768,676	3,416,620	2,788,892	11,768,676	11,768,676
Num. Lenders	4,760	4,890	4,765	4,830	4,890	4,890
R^2	0.008	0.007	0.198	0.064	0.052	0.009
Mean of Dep. Var.	0.0512	0.0183	0.252	0.0911	0.129	0.0107

Panel B. Adjusted Percentages						
Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Comp.	(4) EIDL > PPP Jobs	(5) At Least One	(6) At Least Two
Fintech	0.0350*** (3.55)	0.0106*** (5.19)	0.0935*** (7.34)	0.0651*** (4.13)	0.0560*** (8.81)	0.00686*** (7.00)
ln(Jobs Reported)	Yes	Yes	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,305,966	11,086,014	3,410,487	2,653,048	11,086,014	11,086,015
Num. Lenders	4,722	4,873	4,765	4,732	4,873	4,783
R^2	0.097	0.048	0.597	0.276	0.310	0.081
Mean of Dep. Var.	0.0513	0.0182	0.252	0.0928	0.134	0.0113

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.III. Prevalence of Flags by Lender Type, with Address Fixed Effect

In this table, we examine the prevalence of our primary flags by lender type while including an address \times draw fixed effect. Both Panels include address \times draw fixed effects. Panel B also includes other controls and fixed effects as indicated at bottom of each column. Robust standard errors are triple clustered by zip code, lender, and address \times draw.

Panel A. Only Address Fixed Effect				
Dep. Variable:	(1) Business Registry	(2) High Implied Comp.	(3) EIDL > PPP Jobs	(4) At Least One of Three
Fintech	0.0277*** (3.18)	0.151*** (7.34)	0.0400*** (4.76)	0.0387*** (6.02)
Address \times Draw FE	Yes	Yes	Yes	Yes
Observations	135,383	167,961	73,374	934,794
Num. Lenders	3,555	2,832	2,780	4,564
R^2	0.572	0.801	0.837	0.632
Mean of Dep. Var.	0.0676	0.494	0.206	0.187
Panel B. Other Fixed Effects and Controls Added				
Dep. Variable:	(1) Business Registry	(2) High Implied Comp.	(3) EIDL > PPP Jobs	(4) At Least One of Three
Fintech	0.0187*** (2.80)	0.0233*** (2.98)	0.00988 (1.50)	0.0233*** (5.59)
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes
Address \times Draw FE	Yes	Yes	Yes	Yes
Observations	105,211	160,990	57,636	815,136
Num. Lenders	2,984	2,621	2,023	4,369
R^2	0.657	0.884	0.877	0.773
Mean of Dep. Var.	0.0690	0.505	0.231	0.208

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.IV. Discontinuity at \$100k

In this table, we examine the relationship between our four main flags and implied compensation. We estimate OLS regressions with dummies for each of our main flags as dependent variables and dummies for \$5k wide bins (i.e., (\$0k, \$5k], ..., (\$95k, \$100k], ..., (\$125k, \$130k]) of implied compensation as independent variables. The dummy variable for the (\$0k, \$5k] bin is used as the baseline. Panel A shows the results for fintech loans and Panel B for traditional loans. In both panels, specification (1) uses business registry as the dependent variable, (2) uses multiple loans, (3) uses high implied compensation, and (4) uses EIDL > PPP jobs. Loans are filtered to corporation, S-corporation, and LLC loans for specification (1), all loans for (2), loans for which we can determine industry-CBSA average compensation for (3), and loans with a matched EIDL Advance for (4). Fixed effects and control variables are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Panel A. Fintech Loans				
Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Implied Comp.	(4) EIDL > PPP Jobs
(\$0k, \$5k]	Used as Baseline			
(\$5k, \$10k]	-0.00903** (-2.45)	0.00812*** (3.63)	-0.0191** (-2.26)	-0.0123*** (-2.99)
(\$10k, \$15k]	-0.0185*** (-3.11)	0.0122*** (3.43)	-0.351** (-2.58)	-0.0173*** (-3.38)
(\$15k, \$20k]	-0.0155** (-2.40)	0.0136*** (3.23)	-0.0376** (-2.13)	-0.00893 (-1.39)
(\$20k, \$25k]	-0.0201** (-2.40)	0.0162*** (3.41)	-0.0249 (-1.24)	-0.00576 (-0.82)
(\$25k, \$30k]	-0.0143 (-1.54)	0.0170*** (3.35)	-0.0191 (-0.87)	-0.000893 (-0.11)
(\$30k, \$35k]	-0.0151 (-1.60)	0.0185*** (3.39)	-0.0161 (-0.68)	0.0102 (1.28)
(\$35k, \$40k]	-0.0132 (-1.31)	0.0209*** (3.55)	-0.0153 (-0.61)	0.0236** (2.43)
(\$40k, \$45k]	-0.00996 (-0.96)	0.0220*** (3.70)	-0.00535 (-0.20)	0.0210** (2.18)
(\$45k, \$50k]	-0.00674 (-0.61)	0.0270*** (4.22)	-0.0168 (-0.58)	0.0501*** (3.85)
(\$50k, \$55k]	-0.00296 (-0.26)	0.0255*** (4.02)	-0.00221 (-0.07)	0.0414*** (3.64)
(\$55k, \$60k]	-0.00612 (-0.52)	0.0277*** (3.99)	0.00916 (0.29)	0.0537*** (4.33)
(\$60k, \$65k]	-0.00380 (-0.32)	0.0294*** (4.25)	0.0168 (0.50)	0.0696*** (4.64)
(\$65k, \$70k]	0.0000156 (0.00)	0.0310*** (4.28)	0.0427 (1.26)	0.0877*** (5.51)
(\$70k, \$75k]	-0.00275 (-0.24)	0.0306*** (4.19)	0.0953*** (2.72)	0.0938*** (5.45)
(\$75k, \$80k]	0.0122 (1.07)	0.0352*** (4.78)	0.134*** (3.88)	0.127*** (5.63)
(\$80k, \$85k]	0.0124 (1.09)	0.0336*** (4.87)	0.179*** (5.09)	0.141*** (6.37)
(\$85k, \$90k]	0.00848 (0.72)	0.0365*** (4.83)	0.223*** (6.41)	0.174*** (7.49)
(\$90k, \$95k]	0.0170 (1.37)	0.0398*** (5.41)	0.267*** (7.43)	0.243*** (10.28)
(\$95k, \$100k]	0.0394** (2.37)	0.0400*** (5.73)	0.296*** (8.26)	0.299*** (17.54)
(\$100k, \$105k]	-0.0100 (-0.59)	0.0407*** (3.46)	0.254*** (6.78)	0.100*** (5.38)
(\$105k, \$110k]	-0.0136 (-0.94)	0.0400*** (3.84)	0.242*** (5.85)	0.106*** (5.17)
(\$110k, \$115k]	-0.0142 (-0.91)	0.0403*** (4.44)	0.239*** (5.81)	0.0879*** (4.89)
(\$115k, \$120k]	-0.0197 (-1.31)	0.0356*** (3.57)	0.260*** (6.46)	0.0949*** (5.59)
(\$120k, \$125k]	-0.0126 (-0.76)	0.0310*** (3.22)	0.281*** (6.91)	0.0724*** (3.91)
(\$125k, \$130k]	-0.0132 (-0.67)	0.0305*** (3.58)	0.258*** (5.92)	0.0625** (2.56)
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes
Observations	693,826	3,865,035	3,710,771	673,469
Num. Lenders	77	77	77	77
R^2	0.237	0.050	0.639	0.448
Mean of Dep. Var.	0.101	0.0348	0.181	0.220

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Panel B. Traditional Loans

Dep. Variable:	(1) Business Registry	(2) Multiple Loans	(3) High Implied Comp.	(4) EIDL > PPP Jobs
Used as Baseline				
(\$0k, \$5k]				
(\$5k, \$10k]	-0.00812*** (-4.58)	0.000627 (0.68)	-0.00744** (-2.25)	-0.0107** (-2.25)
(\$10k, \$15k]	-0.0125*** (-5.26)	0.00139 (1.10)	-0.00706 (-1.60)	-0.0116** (-2.08)
(\$15k, \$20k]	-0.0144*** (-5.14)	0.00147 (0.96)	-0.00109 (-0.23)	-0.00649 (-1.22)
(\$20k, \$25k]	-0.0161*** (-5.11)	0.00116 (0.66)	0.00754 (1.53)	-0.000960 (-0.17)
(\$25k, \$30k]	-0.0158*** (-4.39)	0.00108 (0.56)	0.0166*** (3.18)	0.00454 (0.78)
(\$30k, \$35k]	-0.0152*** (-3.98)	0.00102 (0.50)	0.0266*** (4.85)	0.0101 (1.61)
(\$35k, \$40k]	-0.0162*** (-3.91)	0.00766 (0.35)	0.0356*** (6.22)	0.0133** (2.16)
(\$40k, \$45k]	-0.0153*** (-3.54)	0.0000999 (0.43)	0.0463*** (7.68)	0.0187*** (2.80)
(\$45k, \$50k]	-0.0145*** (-3.24)	0.00169 (0.69)	0.0568*** (8.97)	0.0206*** (3.01)
(\$50k, \$55k]	-0.0141*** (-2.94)	0.00118 (0.46)	0.0690*** (10.33)	0.0251*** (3.40)
(\$55k, \$60k]	-0.0129*** (-2.69)	0.00146 (0.56)	0.0827*** (11.79)	0.0323*** (4.04)
(\$60k, \$65k]	-0.0138*** (-2.73)	0.00186 (0.71)	0.0958*** (12.93)	0.0381*** (4.04)
(\$65k, \$70k]	-0.0136*** (-2.63)	0.00184 (0.66)	0.114*** (14.40)	0.0460*** (4.94)
(\$70k, \$75k]	-0.0119** (-2.24)	0.00239 (0.83)	0.128*** (16.08)	0.0405*** (5.00)
(\$75k, \$80k]	-0.0110** (-2.02)	0.00339 (1.15)	0.136*** (17.48)	0.0347*** (4.26)
(\$80k, \$85k]	-0.0114** (-2.07)	0.00307 (1.03)	0.148*** (17.52)	0.0403*** (4.77)
(\$85k, \$90k]	-0.0106* (-1.91)	0.00405 (1.33)	0.166*** (17.68)	0.0474*** (5.54)
(\$90k, \$95k]	-0.00824 (-1.49)	0.00521 (1.69)	0.178*** (16.83)	0.0555*** (5.05)
(\$95k, \$100k]	0.0150** (2.52)	0.00475 (1.50)	0.190*** (13.96)	0.0764*** (4.38)
(\$100k, \$105k]	-0.0141** (-2.07)	0.00345 (1.11)	0.180*** (16.45)	0.0568*** (5.04)
(\$105k, \$110k]	-0.0162*** (-2.66)	0.00304 (0.93)	0.186*** (18.15)	0.0557*** (6.08)
(\$110k, \$115k]	-0.0141** (-2.23)	0.00310 (0.97)	0.203*** (19.56)	0.0654*** (6.87)
(\$115k, \$120k]	-0.0166*** (-2.59)	0.00429 (1.21)	0.223*** (22.51)	0.0683*** (6.45)
(\$120k, \$125k]	-0.0161** (-2.45)	0.00322 (0.98)	0.240*** (19.72)	0.0645*** (6.52)
(\$125k, \$130k]	-0.0170** (-2.46)	0.00402 (1.17)	0.258*** (24.81)	0.0698*** (6.94)
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes
Observations	4,512,085	7,079,739	6,052,314	1,931,998
Num. Lenders	4,643	4,794	4,787	4,655
R^2	0.074	0.058	0.260	0.108
Mean of Dep. Var.	0.0433	0.00923	0.0358	0.0486

t-statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Table IA.V. Criminal Records

In this table, we examine the relationship between each of our flags and criminal records for loans within a subsample of 150,000 rounds 1-2 loans. We estimate OLS regressions with a dummy for whether the borrower has a felony from 2000 or after on their record as the dependent variable and dummies for whether the loan is flagged by each of our flags individually as the independent variable. Panel A shows the relationship for fintech loans and Panel B for traditional loans. Loans are filtered to the sets for which we can determine the flag (same as in Figures 2-8). Note that the business registry flag is not included since we can only determine criminal records for loans to individuals while the business registry flag can only be determined for corporations and LLCs. Fixed effects and control variables are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Panel A. Fintech Loans							
Dep. Variable: Felony Post-2000	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Multiple Loans at Address	0.0141*** (2.77)						
High Implied Comp.		0.0256*** (7.22)					
EIDL Jobs > PPP Jobs			0.0762*** (11.90)				
\$100k Implied. Comp.				0.0128*** (4.56)			
Monthly Rounding					0.00899* (1.83)		
Overrep. in County/NAICS						0.0173*** (3.32)	
High Concentration							0.00263* (0.63)
ln(Jobs Reported)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	No	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes	Yes	No	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,365	17,511	10,625	54,376	54,376	28,176	52,490
Num. Lenders	69	45	47	69	59	69	61
R^2	0.127	0.113	0.183	0.126	0.126	0.008	0.122
Mean of Dep. Var.	0.0487	0.0590	0.0494	0.0481	0.0481	0.0490	0.0483

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Panel B. Traditional Loans

Dep. Variable: Felony Post-2000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Multiple Loans at Address	0.00851 (1.65)						
High Implied Comp.		0.00797 (1.50)					
EIDL Jobs > PPP Jobs			0.0378* (1.87)				
\$100k Implied Comp.				-0.00162 (-1.19)			
Monthly Rounding					0.00215 (0.73)		
Overrep. in County/NAICS						0.00311** (2.64)	
High Concentration							0.00524 (1.27)
ln(Jobs Reported)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	No	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes	Yes	No	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	64,251	11,298	7,067	64,251	64,251	44,551	57,913
Num. Lenders	2,466	1,036	596	2,466	2,466	2,163	2,195
R^2	0.225	0.305	0.379	0.225	0.225	0.053	0.219
Mean of Dep. Var.	0.0169	0.0171	0.0150	0.0136	0.0136	0.0133	0.0133

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.VI. Prevalence of Secondary Flags by Lender Types

In this table, we examine the prevalence of the secondary flags by lender type. We estimate OLS regressions with each of the secondary flags as the dependent variable and a dummy for whether the lender is a fintech as the independent variable. Loans are filtered to the sets for which we can determine each flag (same as in Figures 5-9). Fixed effects and control variables are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Panel A. Unadjusted Differences					
Dep. Variable:	(1) \$100k Comp.	(2) Monthly Rounding	(3) Overrep. in County/NAICS	(4) High Concentration	(5) Felony
Fintech	0.199*** (6.09)	0.0263* (1.84)	0.319*** (10.80)	0.404*** (4.73)	0.0337*** (8.56)
Observations	11,768,676	5,458,913	6,414,028	8,040,494	150,000
Num. Lenders	4,890	4,872	4,875	4,405	3,656
R^2	0.069	0.001	0.087	0.216	0.010
Mean of Dep. Var.	0.150	0.0787	0.467	0.225	0.0274
Panel B. Adjusted Differences					
Dep. Variable:	(1) \$100k Comp.	(2) Monthly Rounding	(3) Overrep. in County/NAICS	(4) High Concentration	(5) Felony
Fintech	0.0658*** (4.55)	0.0168 (1.23)	0.122*** (9.87)	0.279*** (4.62)	0.0130*** (7.51)
ln(Jobs Reported)	Yes	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes	No
Business Type FE	Yes	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	No	Yes	Yes
Observations	11,086,014	5,166,522	6,408,461	7,615,707	119,799
Num. Lenders	4,873	4,830	4,874	4,240	3,045
R^2	0.313	0.040	0.228	0.527	0.209
Mean of Dep. Var.	0.148	0.0800	0.467	0.214	0.0291

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.VII. Within State/County Variation

This table examines the degree to which geographic variation in flagged loans can be explained by fintech market share. We estimate OLS regressions with the percentages of flagged loans in each zip code as the dependent variable and the fintech market share in each zip code as the independent variable. Specification (1) examines the relationship across all zip codes, (2) examines the relationship within states, and (3) examines the relationship within counties. Zip codes with at least 100 loans are considered. Fixed effects are indicated at the bottom of each column. Robust standard errors are clustered at the county level.

Dep. Variable: Percentage Flagged in Zip Code			
	(1)	(2)	(3)
Fintech Market Share	0.281*** (18.14)	0.321*** (23.81)	0.321*** (16.52)
State FE	No	Yes	No
County FE	No	No	Yes
Observations	15,835	15,835	15,011
Num. Counties	2,903	2,903	2,079
R^2	0.554	0.685	0.824
Mean of Dep. Var.	0.104	0.104	0.106

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.VIII. County Cultural Features

This table replicates Table V at the county-level rather than the loan-level. The variables are as defined in Table V. All variables (independent and dependent) are rescaled to have a mean of 0 and a standard deviation of 1. Counties with at least 100 loans are considered. Fixed effects are indicated at the bottom of each column. Robust standard errors are clustered at the state level.

Dep. Variable: Percentage Flagged			
	(1)	(2)	(3)
Public Corruption	0.150** (2.28)	0.116* (1.95)	0.0291 (1.06)
Religious Affiliation	-0.0539* (-1.93)	-0.0502** (-2.18)	0.00475 (0.27)
Ashley Madison Usage	0.229*** (8.27)	0.152*** (4.75)	0.00704 (0.27)
Population Density		-0.0177 (-1.37)	-0.0528*** (-5.31)
Median Income		-0.0537 (-1.22)	-0.0696** (-2.60)
Pct. Non-White		0.447*** (8.48)	0.0316 (0.81)
College Educated		0.133*** (5.45)	0.140*** (8.43)
2019 Unemployment		-0.0652 (-1.17)	-0.0861** (-2.50)
Pct. Fintech			0.811*** (10.47)
State FE	Yes	Yes	Yes
Observations	3,013	3,012	3,012
R^2	0.417	0.531	0.693

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.IX. Previous SBA Lending

In this table, we examine lender-level relationship between suspicious lending in the PPP and previous SBA lending. $\ln(\text{Num. 7(a) Loans Pre-2020} + 1)$ is the natural log of the number of SBA 7(a) loans originated by the lender pre-2020. *Num. Year Since First 7(a) Loan* is the number of years between when the lender originated its first SBA 7(a) loan and 2020 (we have data going back to 1990, so this variable can take a maximum value of 30). *New Lender* is a dummy that takes one a value of 1 if the lender had not originated any SBA 7(a) loans pre-2020. $1(\text{Fintech})$ and $1(\text{Traditional})$ are indicator functions for whether the lenders is a fintech or traditional lender, respectively. Lenders with at least 1,000 PPP loans are considered. Robust standard errors are used.

Dep. Variable: Percentage Flagged by at Least One Primary Flag				
	(1)	(2)	(3)	(4)
$\ln(\text{Num. 7(a) Loans Pre-2020} + 1)$	-0.00220*** (-3.22)			
Num. Year Since First 7(a) Loan		-0.00131*** (-6.59)		
New Lender			0.0341*** (3.82)	
$\times 1(\text{Fintech})$				0.0378 (1.44)
$\times 1(\text{Traditional})$				0.0224*** (2.71)
$1(\text{Fintech})$				0.0484*** (4.92)
Observations	1,141	1,141	1,141	1,141
R^2	0.018	0.111	0.055	0.161
Mean of Dep. Var.	0.0736	0.0736	0.0736	0.0736

t-statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Table IA.X. Fees

In this table, we show the dollars value of loans originated, number of loans originated, and the estimated fees received by the top 75 PPP lenders (by number of loans). The list is sorted in descending order by estimated fees. The first forty lender are on this page and the remaining 35 are on the next page.

Lender	Lender Type	Dollars Lent	Number of Loans	Estimated Fees
JPMChase	Traditional	\$41,724,832,566	438,571	\$1,692,343,319
BoA	Traditional	\$34,318,969,946	491,037	\$1,471,270,588
Prestamos CDFI	Fintech	\$7,706,291,810	495,547	\$1,176,524,235
Capital Plus	Fintech	\$7,421,361,604	463,598	\$1,101,706,143
Harvest	Fintech	\$8,701,102,355	433,306	\$1,064,286,118
Cross River	Fintech	\$12,927,230,933	479,871	\$1,037,484,722
Benworth	Fintech	\$4,567,637,099	331,317	\$745,155,411
Wells Fargo	Traditional	\$13,894,207,311	280,694	\$691,472,470
Customers Bank	Traditional	\$7,137,984,962	287,470	\$651,576,422
Fountainhead	Fintech	\$4,150,617,279	272,705	\$642,366,086
Lendistry	Fintech	\$4,944,781,183	249,321	\$630,539,114
PNC	Traditional	\$17,390,351,788	119,354	\$601,990,855
Branch	Traditional	\$16,727,409,331	117,952	\$590,928,473
U.S. Bank	Traditional	\$10,834,749,543	174,825	\$506,152,576
TD Bank	Traditional	\$12,291,334,013	133,136	\$495,335,337
Itria	Fintech	\$5,102,367,975	177,790	\$488,823,558
Huntington	Traditional	\$11,284,620,628	83,188	\$402,221,521
KeyBank	Traditional	\$11,135,483,364	69,558	\$376,725,537
Zions	Traditional	\$9,873,297,016	76,534	\$356,275,679
M&T	Traditional	\$9,661,643,581	59,094	\$335,143,327
Readycap	Fintech	\$5,061,982,022	111,228	\$324,957,172
Citizens Bank	Traditional	\$7,185,615,514	85,678	\$294,273,948
Fifth Third	Traditional	\$7,397,570,873	65,997	\$270,545,233
Regions Bank	Traditional	\$6,423,446,492	78,544	\$266,969,904
Celtic	Fintech	\$4,627,792,089	167,203	\$235,744,519
First Horizon	Traditional	\$5,797,703,273	50,976	\$218,555,438
WebBank	Fintech	\$3,190,673,848	119,292	\$214,816,890
Citibank	Traditional	\$4,786,384,820	47,768	\$191,158,479
Kabbage	Fintech	\$3,317,781,390	179,820	\$187,787,604
BMO Harris	Traditional	\$6,117,701,222	36,237	\$186,505,159
City National	Traditional	\$5,954,598,740	25,543	\$185,730,854
First-Citizens	Traditional	\$4,432,975,119	35,578	\$177,150,216
Frost	Traditional	\$4,702,944,223	32,518	\$171,420,904
Pinnacle	Traditional	\$4,282,505,226	37,943	\$170,829,371
Northeast	Traditional	\$3,513,827,150	35,469	\$164,925,053
Bank of the West	Traditional	\$4,398,310,337	31,111	\$164,558,480
Leader Bank	Fintech	\$1,336,378,802	62,757	\$155,573,115
Square	Fintech	\$681,853,014	72,570	\$145,565,241
Synovus	Traditional	\$3,842,389,064	27,810	\$142,270,309
Comerica Bank	Traditional	\$4,959,450,533	21,027	\$142,259,791

Lender	Lender Type	Dollars Lent	Number of Loans	Estimated Fees
BBVA	Traditional	\$4,080,037,532	30,695	\$140,674,018
First National PA	Traditional	\$3,667,129,883	30,401	\$137,338,347
People's United	Traditional	\$3,697,583,298	30,952	\$135,503,885
South State	Traditional	\$3,238,850,035	28,037	\$131,941,339
Hancock Whitney	Traditional	\$3,337,831,677	21,601	\$117,588,402
Umpqua	Traditional	\$2,912,581,701	26,421	\$111,921,925
Valley National	Traditional	\$3,239,909,935	19,772	\$111,182,058
MUFG	Traditional	\$3,074,469,662	21,231	\$109,724,419
Banco Popular	Traditional	\$1,775,879,079	47,457	\$99,635,470
Newtek	Fintech	\$2,107,995,767	27,758	\$96,761,658
United Community	Traditional	\$2,220,819,108	22,239	\$90,727,018
Glacier Bank	Traditional	\$2,015,613,877	24,553	\$88,176,938
FirstBank	Traditional	\$1,984,886,943	26,429	\$87,215,605
BancorpSouth	Traditional	\$1,822,888,745	25,555	\$83,880,069
First Bank	Traditional	\$1,866,103,529	20,153	\$81,071,969
Atlantic Union	Traditional	\$2,213,653,002	16,923	\$80,697,177
Prosperity	Traditional	\$2,009,570,006	18,719	\$80,548,892
Arvest	Traditional	\$1,654,425,266	27,526	\$79,719,868
United Bank	Traditional	\$2,054,702,073	17,953	\$79,217,758
Capital One	Fintech	\$1,782,975,418	24,940	\$78,476,632
Blue Ridge	Traditional	\$1,163,973,161	24,171	\$77,397,714
Webster	Traditional	\$1,987,579,454	18,483	\$77,283,640
Santander	Traditional	\$1,768,965,307	19,963	\$73,347,246
Amur Equipment	Traditional	\$651,553,095	28,005	\$72,596,432
MBE	Fintech	\$881,524,566	37,195	\$72,475,542
Peoples Bank	Traditional	\$1,636,678,317	19,321	\$72,368,778
First Interstate	Traditional	\$1,637,638,148	18,854	\$71,495,727
American Lending	Fintech	\$608,503,695	22,542	\$59,858,806
A10Capital	Traditional	\$661,264,536	19,617	\$57,613,469
TEC-CC	Traditional	\$535,671,262	21,624	\$56,921,187
First State	Traditional	\$992,702,386	19,288	\$55,618,703
DreamSpring	Fintech	\$296,007,066	29,147	\$55,483,834
Texas National	Traditional	\$533,598,748	20,611	\$53,464,431
Funding Circle	Fintech	\$573,628,956	16,858	\$43,498,786
Intuit	Fintech	\$638,052,437	18,560	\$30,626,291

Table IA.XI. Persistence and Growth by Lender-Region Pairs

In this table, we examine persistence and growth of suspicious lending at the lender-region level. Panel A is based on data at the lender-zip code level and Panel B on lender-county level. We estimate OLS regressions with the percentage of flagged loans during rounds 1-2 within each lender-region pair interacted with whether the lender is a fintech or traditional lender as the independent variables. In specification (1), the dependent variable is whether the lender increased its lending within the region (as a percentage of its overall lending) between rounds 1-2 and round 3; in specification (2), the dependent variable is the percentage change in the percentage of the lender's loans that are in the region between rounds 1-2 and round 3; in specification (3), the dependent variable is the percentage of flagged loans during round 3 within each lender-region pair. In both panels, lender-region pairs with at least 25 loans in rounds 1-2 (combined) are considered. Fixed effects are as indicated at bottom of each column. Robust standard errors are double clustered by region (zip code in Panel A and county in Panel B) and lender.

Panel A. By Lender-Zip Code			
Dep. Variable:	(1) 1(Lending Growth)	(2) Lending Pct. Change	(3) Pct. Flagged in Round 3
Pct. Flagged in Rounds 1-2			
× 1(Fintech)	1.419*** (6.10)	2.209*** (6.36)	0.516*** (13.80)
× 1(Traditional)	0.00999 (0.11)	0.0128 (0.10)	0.114*** (4.06)
Zip Code FE	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes
Observations	37,271	37,271	37,271
Num. Lenders	1,635	1,635	1,635
R^2	0.410	0.575	0.320
Mean of Dep. Var.	0.349	-0.0732	0.0925

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Panel B. By Lender-County

Dep. Variable:	(1) 1(Lending Growth)	(2) Lending Pct. Change	(3) Pct. Flagged in Round 3
Pct. Flagged in Rounds 1-2			
× 1(Fintech)	1.275*** (3.59)	2.106*** (3.15)	0.466*** (4.15)
× 1(Traditional)	0.116 (1.11)	0.0739 (0.33)	0.173*** (4.80)
County FE	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes
Observations	20,047	20,047	20,047
Num. Lenders	2,347	2,347	2,347
R^2	0.372	0.530	0.356
Mean of Dep. Var.	0.427	0.0976	0.0947

t-statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Table IA.XII. Repayments, Enforcement Actions, and Canceled Loans

In this table, we examine the relationship between repaid loans, enforcement actions, and canceled loans and whether the loan is flagged. We estimate OLS regressions with a dummy for whether the loan is flagged by at least one primary flag as the independent variable and a dummy for whether the loan is repaid (Panel A), part of a DOJ enforcement action (Panel B), or canceled (Panel C) as the dependent variable. Even columns include lender fixed effects and odd columns do not. Panel A is based on the loan-level data released by the SBA on May 3, 2021, and Panel B and C are based on the loan-level data released by the SBA on December 1, 2020. Since loans originated by MBE Capital Partners makes up 42% of the repaid loans, Panel A includes specifications across all loans and excluding loans by MBE Capital Partners. Fixed effects and control variables are as indicated at bottom of each column. Robust standard errors are double clustered by zip code and lender.

Panel A. Repayment				
Dep. Variable: 1(Repaid)				
	(1)	(2)	(3)	(4)
	All Loans		Ex. MBE Capital Partners	
Flagged	0.00195 (1.41)	0.000250 (0.36)	0.000574** (2.00)	0.000907*** (3.91)
ln(Jobs Reported)	Yes	Yes	Yes	Yes
ln(Loan Amount)	Yes	Yes	Yes	Yes
Zip FE	Yes	Yes	Yes	Yes
Business Type FE	Yes	Yes	Yes	Yes
NAICS \times CBSA FE	Yes	Yes	Yes	Yes
Lender FE	No	Yes	No	Yes
Observations	4,866,890	4,866,836	4,843,296	4,843,296
Num. Lenders	4,823	4,769	4,822	4,768
R^2	0.031	0.159	0.027	0.049
Mean of Dep. Variable	0.00328	0.00328	0.00192	0.00192

Panel B. DOJ Enforcement		
Dep. Variable: 1(DOJ Enforcement Action)		
	(1)	(2)
Flagged	0.0000929** (2.35)	0.0000900** (2.26)
ln(Jobs Reported)	Yes	Yes
ln(Loan Amount)	Yes	Yes
Zip FE	Yes	Yes
Business Type FE	Yes	Yes
NAICS \times CBSA FE	Yes	Yes
Lender FE	No	Yes
Observations	4,866,890	4,866,836
Num. Lenders	4,823	4,769
R^2	0.025	0.026
Mean of Dep. Variable	0.0000541	0.0000541

Panel C. Canceled Loans

Dep. Variable: 1(Canceled)		
	(1)	(2)
Flagged	0.00478*** (3.49)	0.00131* (1.93)
ln(Jobs Reported)	Yes	Yes
ln(Loan Amount)	Yes	Yes
Zip FE	Yes	Yes
Business Type FE	Yes	Yes
NAICS \times CBSA FE	Yes	Yes
Lender FE	No	Yes
Observations	10,043,880	10,043,853
Num. Lenders	4,885	4,860
R^2	0.030	0.070
Mean of Dep. Variable	0.00893	0.00893

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table IA.XIII. Undisbursed Loans

In this table, we examine the impact of excluding undisbursed loans. Undisbursed loans are those listed as “Active Un-Disbursed” in the loan-level data. Panel A is based on loans originated by any lender, Panel B by fintech lenders, and Panel C by traditional lenders. Column (1) shows the percentage of flagged loans in our entire main loan-level data, Column (2) shows the percentage of flagged loans after excluding the 465,769 undisbursed loans, and Column (3) shows the percentage of flagged loans in 465,769 undisbursed loans. For rows based on individual flags, loans are filtered to the sets for which we can determine each flag; for rows based on sets of flags, all loans are included.

Panel A. Overall			
	(1) All Loans	(2) Ex. Undisbursed	(3) Undisbursed
Business Registry	0.0512	0.0499	0.182
Multiple Loans	0.0183	0.0180	0.0249
High Implied Comp.	0.252	0.233	0.641
EIDL > PPP Jobs	0.0911	0.0868	0.343
At Least One Flag	0.129	0.123	0.276
Primary + Additional	0.0881	0.0817	0.243
Pct. Fintech	0.337	0.318	0.809
Panel B. Fintech			
	(1) All Loans	(2) Ex. Undisbursed	(3) Undisbursed
Business Registry	0.102	0.0965	0.241
Multiple Loans	0.0344	0.0350	0.0286
High Implied Comp.	0.489	0.467	0.687
EIDL > PPP Jobs	0.216	0.206	0.420
At Least One Flag	0.236	0.229	0.303
Primary + Additional	0.198	0.190	0.274
Panel C. Traditional			
	(1) All Loans	(2) Ex. Undisbursed	(3) Undisbursed
Business Registry	0.0434	0.0430	0.111
Multiple Loans	0.0100	0.0100	0.00903
High Implied Comp.	0.0942	0.0907	0.388
EIDL > PPP Jobs	0.0487	0.0480	0.151
At Least One Flag	0.0743	0.0733	0.161
Primary + Additional	0.0321	0.03126	0.111

Table IA.XIV. Prevalence of Flags by Lender Type, Excluding Undisbursed Loans

This table replicates Table III after excluding undisbursed loans. Undisbursed loans are those listed as “Active Un-Disbursed” in the loan-level data.

	(1) Fintech	(2) Traditional	(3) Unadjusted Difference	(4) Adjusted Difference	(5) Matched Difference
Business Registry	0.0965 N = 711,086	0.0430 N = 4,801,246	0.0535*** (3.04)	0.0330*** (3.47)	0.0246*** (3.15)
Multiple Loans	0.0350 N = 3,593,267	0.0100 N = 7,709,653	0.0249*** (8.93)	0.0110*** (5.88)	0.0157*** (3.97)
High Implied Comp.	0.467 N = 1,232,410	0.0907 N = 2,026,091	0.377*** (7.65)	0.0925*** (7.41)	0.0847*** (3.96)
EIDL > PPP Jobs	0.206 N = 671,895	0.0480 N = 2,069,311	0.158*** (4.43)	0.0629*** (4.13)	0.0783*** (4.54)
At Least One Flag	0.229 N = 3,593,267	0.0733 N = 7,709,653	0.155*** (8.08)	0.0562*** (8.54)	0.0547*** (6.55)
At Least Two Flags	0.0244 N = 3,593,267	0.00353 N = 7,709,653	0.0209*** (7.38)	0.00709*** (7.07)	0.00981*** (5.20)

t-statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$