Journal of Financial Economics

Did pandemic relief fraud inflate house prices?☆

John M. Griffin, Samuel Kruger *, Prateek Mahajan

McCombs School of Business, University of Texas at Austin, 2110 Speedway, Austin, TX 78705, USA

ARTICLE INFO

JEL classification:
G21
G23
G28
H12
R21

Keywords:
Fraud
House price growth
Government spending

ABSTRACT

Pandemic fraud is geographically concentrated and stimulated local purchases, with effects on prices. Recipients of fraudulent Paycheck Protection Program (PPP) funds significantly increased their home purchasing rate compared to recipients of non-fraudulent PPP funds, and house prices in high-fraud ZIP codes increased 5.8 percentage points more than in low-fraud ZIP codes within the same county. In a horse race, pandemic fraud is one of the largest and most robust factors explaining house price appreciation during COVID. ZIP codes with fraud also experienced heightened vehicle purchases and other consumer spending in 2020-21, with a return to normal in 2022.

1. Introduction

Are the costs of financial fraud largely confined to stolen funds, or does fraud have other, perhaps unanticipated, distortive effects? (Akerlof and Romer, 1993) propose that fraud can have large unintended externalities that are often greater than the direct costs, including in the form of distorted asset prices. In this paper, we seek to test this theory by examining the use of fraudulent funds stolen from pandemic relief programs. Are the proceeds of the fraud used, at least in part, to purchase assets, and do they thereby put upward pressure on asset prices?

In the traditional (Becker, 1968) crime model, the optimal amount of resources to devote to the prosecution of crime and the nature of the punishment depend on both the direct and indirect costs of the crime. For financial fraud, direct cost estimates typically range between three and nine percent of GDP (Gee and Button, 2019). Indirect costs are even harder to detect and quantify. However, federal COVID relief spending of over $4.5 trillion may provide a setting to examine potential spillovers and externalities of fraud. Determining the extent to which this spending may have influenced the price of goods is difficult since most spending programs were mainly designed to offset income lost due to the pandemic and thus are cross-sectionally proportional to population and income. One potential source of regional variation in government spending is fraudulent transfers. There is growing evidence that a sizable portion of the funds distributed by the $793 billion Paycheck Protection Program (PPP), $384 billion Economic Injury Disaster Loan (EIDL) program, and $872 billion unemployment insurance programs may have been fraudulent (Griffin et al., 2023; Podkul, 2021; U.S. Small Business Administration, Office of Inspector General, 2023; Lardner et al., 2023; Khetan et al., 2024). From the ground level, one former U.S. Attorney described it as: “nothing like this has ever happened before… it is the biggest fraud in a generation” Dilanian and Strickler (2022).

☆ Philipp Schnabl was the editor for this article. We thank the editor, an anonymous referee, Jim Conklin, Mike Faulkender, William Fuchs, Andra Ghent, Caitlin Gorback, Christoph Herpfer, Elena Loutskina, Vrinda Mittal, Lin Peng, Alex Priest, Marius Ring, Zhaogang Song, Laura Starks, Constantine Yannelis, Eric Zwick, industry participants, and others for helpful comments. We are also thankful for comments from seminar and conference participants at the American Real Estate and Urban Economics Association (AREUEA) national conference, Baruch College, the Federal Reserve Bank of Dallas, the Financial Intermediation Research Society (FIRS) conference, NBER Summer Institute (Household Finance), the Office of the Comptroller of the Currency (OCC) Emerging Risks Symposium, Pre-WFA Summer Real Estate Research Symposium, the Society of Financial Studies (SFS) Cavalcade, the University of Kentucky Finance Conference, the University of North Carolina, the University of Pennsylvania, and the University of Texas at Austin. We also thank the University of Texas at Austin and the Integra Research Group for research support. Griffin is an owner of Integra Research Group and Integra FEC, which engage in research, financial consulting, and recovery on a variety of issues related to the investigation of financial fraud.

Received 28 May 2024; Received in revised form 12 February 2026; Accepted 1 March 2026

Available online 21 March 2026
0304-405X/

In this paper, we examine the effects of fraud in the PPP and other pandemic relief programs on local housing markets and consumer spending. This research builds on two previous papers that also examined pandemic relief fraud. Griffin et al. (2023) find that fraud in the PPP was widespread and accelerated quickly, in part due to lax standards by some FinTech lenders. Additionally, Griffin et al. (2025) find that many ZIP codes have almost no PPP fraud, while others have in excess of 40% of their loans flagged as suspicious. They examine numerous explanations for the variation in suspicious lending rates across regions and find that a measure of social connections between ZIP codes (based on Facebook data) is the strongest predictor of the spread of pandemic relief fraud. Information about how to obtain fraudulent loans circulated rapidly on social media. The spread of fraud over time can be explained by social connections between ZIP codes, resulting in concentrated pockets of fraud and substantial variation in fraud rates even within the same county. This paper assesses how recipients spent at least a subset of the ill-gotten funds and, more importantly, whether their spending potentially distorted local markets, particularly in the housing market. This is important for understanding externalities of fraud and is quantitatively one of the most important channels for explaining the post-COVID rise in house prices.

Pandemic relief programs were mainly designed to offset lost income to keep the economy from substantially shrinking. There are two reasons that fraudulent funds might result in different spending patterns than non-fraudulent funds. First, when relief programs operated as intended, recipients experienced no income shock and spending practices should not have shifted. However, if an individual committed PPP fraud and suddenly received a financial windfall, this could represent a large increase in their income and overall wealth. Second, individuals willing to engage in fraud may be more likely to spend these ill-gotten gains quickly rather than saving for the future. Anecdotal evidence suggests that many recipients of fraudulent PPP loans used the funds they received to purchase expensive houses, cars, and luxury items.[1] We examine how pandemic relief fraud may have had unintended consequences by creating excess demand for an important immovable regional good—housing.

Before examining house prices, we first investigate whether individuals who received pandemic relief payments through fraudulent means were more likely to purchase homes. We match a random sample of 250,000 individual PPP recipients to property ownership records from PropertyRadar and also utilize data on house purchases from LexisNexis and consumer address history from Verisk. The first two sources indicate a sizable upward shift in house purchase probability for recipients of flagged PPP loans compared to non-flagged recipients, and the third indicates a sizable upward shift in moving propensity compared to non-recipients. In particular, using a stacked difference-in-differences framework, we find that the probability that an individual purchased a house increased by 17% compared to non-flagged recipients and moving propensity increased by 22% compared to non-recipients in the eighteen months after receiving a suspicious loan.

Given the higher rate of house purchases by recipients of fraudulent PPP loans and the concentration of pandemic relief fraud, it is possible that these purchases, funded by ill-gotten funds from the PPP and other government programs, distorted local house prices. We primarily use information on PPP loans for our analysis due to the availability of more detailed information. However, results from Griffin et al. (2024) suggest that the geographic areas which received fraudulent funds through the PPP also have higher levels of fraudulent EIDL loans, EIDL Advances, and unemployment insurance claims. To control for macro factors that may have influenced regional house price growth, our analysis focuses on the ZIP code level while including county fixed effects. We find that ZIP codes in the top decile of suspicious lending per capita experienced house price growth that is 5.8 percentage points (ppt) higher than ZIP codes in the lowest decile of suspicious lending per capita. This difference represents a sizable 22.5% of the 25.9 ppt average increase in house prices during 2020 and 2021. Additionally, the effect of pandemic fraud is consistently positive across the initial period; the effect first becomes significant in April 2020 and persists until June 2022. After June 2022, areas with higher PPP fraud have lower house price growth, suggesting that the inflationary impact of pandemic fraud on house prices was temporary. Non-fraudulent PPP lending has no effect on house prices, consistent with these funds offsetting legitimate expenses and lost revenues as opposed to providing extra stimulus. Additionally, since individuals may purchase homes in nearby areas rather than only in their current ZIP code, we also calculate the average amount of fraud (both simple and distance-weighted) among ZIP codes within a five-mile radius of the focal ZIP code and find similarly strong house price effects.

During the period when fraudulent funds were received and spent, we expect to see an increase in demand in high-fraud areas due to the surge in buyers. In turn, this would lead to the heightened price growth discussed above. However, when the influx of fraudulent funds ceased, one would expect to see slightly less demand due to the elevated prices, leading to a price reversal. We examine proxies for net demand and find that beginning in June/July 2020, areas with high fraud had a larger percentage of listings sold within two weeks, a larger percentage of purchases made above listing price, more views from potential buyers per property, and lower inventory compared to prepandemic levels. These patterns weaken later in 2021 and reverse during the 2022–23 period: areas with high fraud had fewer listings sold quickly, fewer purchases above listing price, fewer views per property, and more inventory than prepandemic levels. Consistent with this, high-fraud ZIP codes experienced lower house price growth from mid-2022 to December 2023, leading to a reversal that amounts to 35% of the initial effect.

Other factors have also been proposed to explain house price growth during the COVID period. To assess pandemic fraud in relation to these competing explanations, we consider all proposed factors together within a common framework following the methodology used by Griffin et al. (2020). This framework is useful not just for assessing pandemic fraud, but also for understanding factors that affected house price growth from 2020 to 2021 more generally. Using both Bayesian Model Averaging and variable selection based on the Bayesian Information Criterion, PPP fraud consistently emerges as one of the strongest predictors of house price growth, alongside land unavailability. The models also consistently select teleworkability, previous (2018–2019) housing price growth, net migration, and remote work as predictors, although with smaller economic magnitudes.

A potential concern with our baseline analysis is that PPP fraud is not randomly assigned. It is possible that ZIP codes with high PPP fraud had pre-existing house price momentum (Guren, 2018) or omitted characteristics related to house price appreciation during the COVID period. We use five additional strategies to gauge the sensitivity of our results to these possibilities. First, we use a synthetic control method to create a control group with nearly identical house price trends during 2018 and 2019. These price trends remain identical during the first months of 2020 and only diverge in the summer of 2020, which is when pandemic fraud is most likely to have started affecting house prices. Second, since Griffin et al. (2025) find that fraud spread through social networks, we use social connections to fraud in distant parts of the country as an instrument for fraud in a given ZIP code. The instrumental variable (IV) estimates are consistent with, and even somewhat larger than, our baseline estimates. Restricting the instrument to distant connections, as far as 500 miles away, reduces concerns about migration, local omitted variables, or regional shocks that could violate the exclusion restriction. Further, overidentification tests using connections in different distance bands imply that any effect that social proximity to fraud has on house price growth directly, or indirectly through omitted variables, must be the same over different

[1] For example, here and here.

distances. Third, we refine this IV approach by focusing only on house price growth in 2021 and instrumenting for a ZIP code's 2021 fraud rate using its social proximity to 2020 suspicious lending. This helps alleviate reverse causality concerns because the variation in fraud is driven entirely by events that occurred before the outcome. To strengthen this argument, we also control for the ZIP code's own 2020 fraud rate and 2020 house price growth in addition to the baseline controls. The effect is 83% of the baseline IV result, consistent with much of the effect of fraud on house prices being in 2021 when the fraud was most prevalent. Fourth, in addition to controlling for demographic characteristics, we consider interactions between PPP fraud intensity and demographic characteristics. Specifically, we estimate the effect of PPP fraud in ZIP codes with above- and below-median values of income, poverty rate, population density, minority population share, educational attainment, and prepandemic unemployment. The effect of PPP fraud on house prices is similar across all demographic splits. However, we find that the effects are over 30% stronger in areas of less elastic housing supply, consistent with the demand shock from the influx of fraudulent funds being intermediated by local housing supply. Finally, following the logic of Oster (2019), we find that the effect of fraud on house prices is similar with or without extensive control variables. This mitigates omitted variable concerns if the included control variables are at least partially informative about the effects of unobservable variables.

To examine whether fraud affected consumer spending and inflation more generally, we analyze four additional sources of data. First, using ZIP code × month-level data on automobile title registrations for six large states, we find that ZIP codes with one standard deviation higher PPP fraud per capita have a highly statistically significant 2.38% increase in automobile title registrations from March 2020 to December 2021. Second, using census tract × year-level data on consumer spending from Mastercard, we find that census tracts with one standard deviation higher PPP fraud per capita experienced a highly statistically significant 0.595 percentile rank increase in spending per capita in 2020 and 2021 compared to 2019, with spending returning to normal in 2022. Third, using census tract × week-level data on consumer mobility, we find that census tracts with higher PPP fraud have more visits to a broad set of different types of shopping locations, including auto dealerships, grocery stores, furniture stores, restaurants, and financial institutions. For all three of these analyses, we also perform numerous cross-sectional splits by different demographic variables and find similar effects across these splits, indicating that the results are not driven by a subset of ZIP codes. Further, all three of these analyses are robust to controlling for detailed demographics, controlling for overall PPP loan take-up, and including county × time fixed effects to control for potential time-varying differences in county-level COVID policies. Fourth, using regional CPI (consumer price indices) from the Bureau of Labor Statistics (BLS), we find that CBSAs (Core Based Statistical Areas) with high PPP fraud per capita experienced elevated inflation starting in late 2021 and persisting through April 2023.[2]

Our paper contributes to four main literatures. First, our findings highlight the importance of understanding the potential externalities of fraud. Akerlof and Romer (1993) describe how insolvent banks can gamble through fraudulent lending, leading to boom and bust cycles in asset prices, as observed in commercial real estate during the 1980s Savings and Loan Crisis. Similarly, rampant non-agency mortgage fraud from 2003–06 distorted house prices and heavily contributed to the 2003–06 boom and the 2007–11 bust in house prices.[3] We focus on price distortions with an emphasis on house prices. Financial fraud can also have large externalities through decreased participation in the financial system (Guiso et al., 2008; Gurun et al., 2017). Kedia and Philippon (2007) show that fraudulent accounting practices can distort the allocation of economic resources across firms. To our knowledge, we are the first paper documenting economic spillovers from fraud in government programs.

Second, an emerging literature seeks to understand the forces that drove home price appreciation during the COVID period. Gupta et al. (2022) find that house prices and rents increased in areas farther from city centers, with stronger effects in MSAs with more remote workers. Other factors proposed to explain house price growth during the COVID period include suburban areas (Ramani and Bloom, 2022), higher proportions of stay-at-home residents (Gamber et al., 2023), the percentage of individuals in a CBSA able to telework (Dingel and Neiman, 2020), remote worker share (Mondragon and Wieland, 2025; Davis et al., 2024), lower population density (Liu and Su, 2021), and higher economic impact payments (Lin, 2025).[4] We extend this literature in two ways. First, we propose a new and strong channel for housing price growth during the COVID period that is independent of any of the other channels proposed in the literature. Second, we estimate detailed within-county, ZIP code-level horse races among carefully constructed proxies from the literature. We are the first to synthesize and rigorously compare these alternative explanations. We find that land unavailability and suspicious PPP lending have the strongest relationships with house price growth. Previous levels of remote work, teleworkability, migration during 2020 and 2021, and prior housing price growth are also related to house price growth during this period but have quantitatively smaller effects, particularly when considered in multivariate regressions alongside proxies for other potential channels.

Third, we contribute to further understanding the general impact of COVID relief spending, as well as its relation to potential inflation. Chetty et al. (2023) find that the cost of each job saved by the PPP was $377,000, Autor et al. (2022) find a cost of $169,000 to $258,000 per job saved, Dalton (2023) finds that only about 24% of PPP money went towards wage retention, and Cole (2024) finds a cost of $270,000 per job-year saved at small firms. Similarly, Granja et al. (2022) find small effects of the PPP on employment.[5] Diamond et al. (2025) model and quantify the effects of fiscal and monetary stimulus on inflation and house prices, concluding that both led to inflation and large increases in house prices. de Soyres et al. (2023) and Jorda and Nechio (2023) examine differences in fiscal spending across countries during 2020 and 2021 and show that countries with higher COVID-related fiscal spending, such as the U.S., are experiencing higher rates of post-COVID inflation. We provide a detailed cross-sectional analysis within the U.S. on the relationship between geographically concentrated PPP fraud and local house price growth and consumer spending.

Fourth, our paper also relates to a broader literature on the marginal propensity to consume (MPC) in response to wealth shocks. Evidence of

[2] Data on regional inflation is limited to just 23 CBSAs, but results are statistically significant even within this limited sample. Elevated inflation is primarily due to housing costs, but there is also some evidence of smaller inflationary effects on non-housing prices, including vehicles.

[3] Griffin et al. (2020) perform comparisons of measures proposed in the literature and find that excess credit from mortgage fraud (Griffin and Maturana, 2016) and excess subprime credit (Mian and Sufi, 2009, 2018) were the largest forces behind the 2003 to 2011 house price boom and bust cycle.

[4] Cherry et al. (2021) find that mortgage forbearance allowed borrowers to enter forbearance and miss $86 billion in mortgage payments, which led to lower delinquency rates for home mortgages, auto loans, student loans, and other consumer debt. Fuster et al. (2025) find that even though interest rates went to historic lows, the spread between mortgage rates and treasuries was at historic highs during the pandemic, partially due to labor and operational frictions. More generally, our paper adds to a literature showing that housing demand shocks can have large effects on house prices (e.g., Favilukis et al., 2012; Badarinza and Ramadorai, 2018; Hartman-Glaser et al., 2023; Aiello et al., 2025; Gorback and Keys, 2025).

[5] In contrast, Faulkender et al. (2021) find that the program was much more effective with an average cost per job saved of $34,000 to $37,000. A broader literature has also examined various aspects of differential access to the PPP (e.g., Denes et al., 2021; Cong and Rabetti, 2023; Bartik et al., 2020; Neilson et al., 2020).

violations of consumption smoothing is nearly as longstanding as Friedman's (1957) permanent income hypothesis itself. The empirical literature generally finds average MPCs for non-durable goods and services in the range of 15% to 25%, with large heterogeneities across households (Kaplan and Violante, 2022).[6] The literature generally finds higher MPC estimates for low-income and liquidity-constrained households (Johnson et al., 2006; Baker, 2018). MPC estimates are typically larger for smaller wealth shocks (Kaplan and Violante, 2014; Beraja and Zorzi, 2026), during economic crises (Gross et al., 2020), and during periods of low house price growth (Cho et al., 2024). However, there is also evidence of high MPCs even for households with high incomes and significant liquid-asset holdings (Kueng, 2018), and survey responses indicate that MPCs can sometimes be larger for larger wealth shocks after accounting for durable goods and the extensive margin Fuster et al. (2020). Much of this literature has used government payments as shocks (Souleles, 1999; Johnson et al., 2006; Agarwal et al., 2007; Parker et al., 2013; Kueng, 2018; Aydin, 2022; Baker et al., 2023). While the MPC literature primarily focuses on non-durable spending, several studies find that durable spending also responds to wealth shocks and is often a larger part of the overall consumption response as wealth shocks become larger (Fuster et al., 2020; Beraja and Zorzi, 2026). For example, Parker et al. (2013) find large durable spending increases as a result of the 2008 economic stimulus checks, mostly in the form of new vehicle purchases. Tauber and Van Zandweghe (2021) find that about half of the increase in durable goods spending in 2020 was due to fiscal stimulus. Fuster et al. (2020) find vastly heterogeneous responses across consumers.

Our paper contributes to the MPC literature by examining a new source of heterogeneity—spending responses of fraudulent fund recipients. These responses could differ from other wealth shocks for a number of reasons. For instance, individuals who commit fraud may have different preferences, liquidity constraints, or motivations to spend what they perceive as a windfall. To our knowledge, no paper in the large literature on MPCs has examined the spending responses of fraudulent fund recipients. Nevertheless, our primary empirical analysis focuses on the externalities of fraud proceeds on prices, particularly in the housing market. Although a consumption response is a necessary predicate for this externality, our setting is not ideal for directly estimating MPCs. Our spending data is limited, especially at the individual level. Furthermore, our measure of pandemic fraud — based solely on the PPP — likely understates the total amount of pandemic fraud at both the individual and ZIP code levels, as fraud was highly correlated across multiple government programs, including EIDL and unemployment insurance (Griffin et al., 2024). Additionally, the unique circumstances of the COVID period, such as heightened uncertainty, may mean our results are not generalizable to other time periods. Our primary contribution lies in understanding the potential externalities of fraud, identifying the underlying forces driving house price growth during the COVID period, and assessing the effectiveness and externalities of COVID-relief programs as discussed above.

Overall, our findings suggest that fraud in government programs can have unanticipated effects that extend beyond and are potentially even more costly than the fraudulent transfers themselves. These effects may have helped people who already owned houses but hurt individuals who bought houses at inflated prices. Consistent with this concern, our findings indicate that areas with large amounts of fraud are experiencing lower house price returns after June 2022 than they would have otherwise. To the extent that fraud can have large unanticipated effects, more resources should be spent on designing government programs to prevent fraud and on related law enforcement.

[6] Jappelli and Pistaferri (2010) and Havranek and Sokolova (2020) survey the empirical side of this large literature. Havranek and Sokolova (2020) find an average MPC of 0.37, but conclude that after correcting for methodological issues, selective reporting of estimates, and liquidity constraints, estimated MPCs are close to zero.

2. Data sources

We use loan-level indicators of suspicious PPP loans developed by Griffin et al. (2023) aggregated to various geographic levels. These indicators are based on loan-level PPP data released on January 2, 2022, 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 or canceled as of January 2, 2022. The data includes business name, address, business type (e.g., corporation, LLC, and self-employed), 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 11,469,801 loans originated by 4809 different lenders with a total value of $793 billion. The primary suspicious loan indicators are: loans to non-registered businesses, multiple loans at the same residential address, abnormally high implied compensation relative to industry averages at the CBSA level, and large inconsistencies (as large as tenfold) between the jobs reported by borrowers on their PPP applications and jobs reported to the contemporaneous EIDL Advance program, which had a different incentive structure. In addition to these loan-level primary measures, we consider the extent to which PPP lending at the industry-county level exceeds the number of establishments listed for that industry-county pair in U.S. Census data. See Internet Appendix Section A for additional details regarding these measures. Griffin et al. (2023) extensively validate these measures, including with four secondary measures of fraud and three independent external measures. The findings of Griffin et al. (2023) are also validated by a detailed Congressional investigation of PPP fraud that focused on many of the same FinTech lenders flagged by Griffin et al. (2023) (see Congressional report here).

We use the Zillow House Value Index (ZHVI) at both the ZIP code and county levels, which estimates the typical value for homes in the 35th to 65th percentile and is adjusted for seasonality and compositional changes in sales over time.[7] After selecting ZIP codes for which ZHVI data, the social proximity instrument, and baseline control variables are available, our sample is composed of 18,761 ZIP codes, which collectively include 93% of the U.S. population.[8] Data on additional ZIP code-level housing market metrics is from Redfin and Realtor.com. Data on home purchases for a sample of individual PPP recipients is from LexisNexis and PropertyRadar. Data on address histories is from Verisk (formerly Infutor) and Melissa Data.

Additionally, we use data from a number of sources to replicate proposed house price drivers, including: the percent of individuals who worked remotely in 2015–2019 and population density from the U.S. Census American Community Survey (ACS); teleworkability based on Dingel and Neiman (2020) and U.S. Census LODES; net migration in 2020–2021 from a FOIA request made by Ramani and Bloom (2022) to the USPS; distance to the central business district from Ramani and

[7] As of January 2023, Zillow started using neural networks to generate Zestimates and the ZHVI. Zillow provides ZHVI data back to 2012 based on their new neural network methodology. This version of our paper uses the current Zillow data throughout. A previous version (December 4, 2023) used ZHVI data based on Zillow's previous model-based valuation methodology and found similar results. Figure IA.1 compares monthly house price growth based on the neural network-based and pre-2023 legacy ZHVI.

[8] We start with the 23,325 ZIP codes for which ZHVI data is available during our period of interest. From this set, we can construct the social proximity instrument for 19,464 ZIP codes. Finally, we drop 142 of these ZIP codes due to missing values for a control variable and an additional 561 because they are singletons within their respective counties and thus are dropped when county fixed effects are included. Requiring availability of the social proximity instrument does not have a meaningful impact on our results (Table IA.9) nor does dropping ZIP codes with missing controls (specification 1 of Table IA.9).

Bloom (2022); house price growth in 2018–2019 from Zillow; and land unavailability from Lutz and Sands (2023). We also use county-level data on employment, spending, and small business revenue during the pandemic from the Economic Tracker by Opportunity Insights (described in Chetty et al. (2023)). Monthly vehicle title registration data for five states from January 2018 to December 2022 is from Cross-Sell, with additional data for Washington directly from the state. Annual data on vehicles per household is from the U.S. Census ACS. Annual consumer spending data at the census tract level is from Mastercard's Center for Inclusive Growth. Consumer mobility data is from SafeGraph. Bimonthly regional CPI data is from the Bureau of Labor Statistics. Demographic data at the ZIP code and county levels is from the U.S. Census ACS and IRS Statistics of Income.

3. Background

3.1. Geographic summary

In addition to being widespread, pandemic relief fraud was also highly concentrated geographically. For example, Griffin et al. (2025) find that whereas over 30% of loans in Cook County, IL (Chicago) are suspicious, New York County and Los Angeles County both have suspicious loan rates under 10%. Suspicious loan rates are even more varied at the ZIP code level, often ranging from close to 0% to over 50%, even within the same county.[9] While the forthcoming version of Griffin et al. (2025) focuses on the PPP since the available data for it is the richest, the working paper version finds strong geographic correlations between PPP fraud and suspicious activity in other pandemic relief programs, such as the EIDL and unemployment insurance (Griffin et al., 2024). As a result, economic stimulus from pandemic fraud was highly concentrated geographically and had the potential to have distortionary effects on local house prices and purchases of other assets. Additionally, the large degree of variation in suspicious lending within counties allows for within-county analysis with county fixed effects to account for regional trends, leading to more powerful tests. Griffin et al. (2025) also examine a wide array of cross-sectional variables that could be related to the fraud, including culture, crime rates, past identity theft, demographics, and various measures of social capital, and find that most of these factors have limited associations with fraud rates. In contrast, social proximity to fraud in other regions of the U.S. has an incremental adjusted R^2 that is ten times larger than that of the next-most-impactful variable and is greater than the incremental adjusted R^2 from jointly adding all of the control variables.

3.2. Magnitude of pandemic relief fraud

PPP fraud is clearly geographically concentrated. Is it large enough to meaningfully impact local spending and the prices of housing and other goods? (Griffin et al., 2023) estimate that PPP fraud totaled approximately $117.3 billion, which is 14.8% of the program, and this estimate may be conservative given that the paper only uses publicly available data. Moreover, Griffin et al. (2024) find that fraud in other programs (specifically the EIDL program and unemployment insurance) is highly geographically correlated with PPP fraud. As a result, areas with high PPP fraud also had an influx of fraudulent funds from other pandemic relief programs. For example, EIDL support to small businesses totaled more than $384 billion in the form of EIDL loans and EIDL Advance grants. A June 27, 2023 report by the Office of the Inspector General (OIG) of the SBA indicates that it has identified over $136 billion in loans provided through the COVID-19 EIDL program as potentially fraudulent, which represents 33% of total disbursed funds.[10] Expanded unemployment benefits during COVID amounted to $872.5 billion, and an audit of the UI programs in four large states by the OIG of the Department of Labor found that 20% of Pandemic Unemployment Assistance (PUA) funds were lost to fraud.[11] We may never know the precise magnitudes of pandemic fraud, but if a similar fraud rate of 20% holds for other programs, total pandemic relief fraud could be over $900 billion. Even if the rate is half as much, the total would still be economically substantial.

To put pandemic relief fraud in context relative to the U.S. housing market, the total value of all homes sold in the U.S. was approximately $2.2 trillion in 2020 and $2.8 trillion in 2021, but the median down payment was only 17% for repeat buyers and 7% for new buyers.[12] At a 15% down payment rate, this would amount to approximately $750 billion in down payments on housing in 2020 and 2021. Thus, if even a small percentage of pandemic fraud proceeds was used to purchase houses, it could have an economically substantial impact on the housing market.

Another way to gauge the magnitude of the housing demand shock is that a one standard deviation shock (6.8 fraudulent loans per thousand people) amounts to approximately $5 million in additional PPP fraud proceeds for a typical ZIP code.[13] For context, the value of house purchases in a typical ZIP code in 2020 and 2021 totaled around $120 million.[14] If average down payments are around 15%, this amounts to aggregate down payments of $18 million in a typical ZIP code. Thus, even a small share of the $5 million in incremental PPP fraud proceeds could have a significant impact. Further, total pandemic relief fraud in these ZIP codes is likely even higher given that EIDL and unemployment insurance fraud are highly correlated with PPP fraud.

4. Did fraudulent PPP recipients buy houses?

Determining the effect of pandemic relief programs on asset prices is challenging because most programs were designed to offset lost income, and, if implemented correctly, would not provide an income shock to the borrower. For example, PPP loans were designed to cover business expenditures, such as employee payrolls, for businesses that were struggling due to lost pandemic revenue. PPP loans that offset revenue declines brought on by the pandemic would not generate excess cash for the business owner but instead would (perhaps only partially) make up for business expenses.[15] In contrast, individuals who

[9] Figure IA.2 replicates Figure 1 from Griffin et al. (2025) and shows the strong geographic clustering both across counties (Panel A) and across ZIP codes within counties (Panel B).

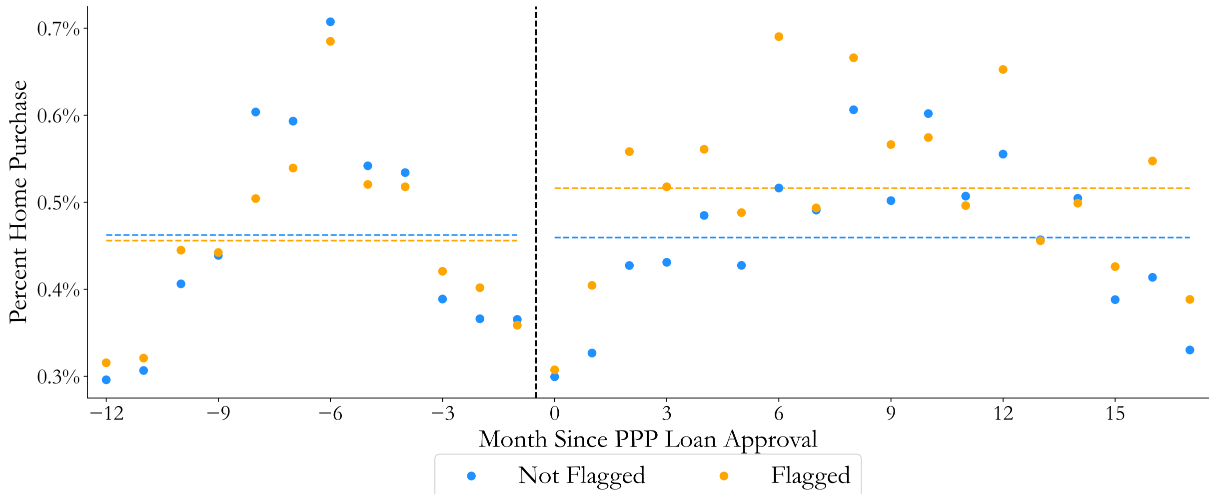
[10] The types of fraud indicators used by the OIG include $55.7 billion based on common or suspicious IP addresses; $34.2 billion due to various hold codes placed on loans by the SBA due to the loan being flagged; $20.5 billion based on duplicated or invalid Employer Identification Numbers (EINs); $31.7 billion based on bank accounts receiving multiple loans or individuals changing their bank account from the one listed on their application; $5 billion to sole proprietors or independent contractors without EINs; and the rest due to other indicators such as hotline complaints and suspicious phone numbers, physical addresses, and email addresses. The OIG report also separately identified $64 billion in potentially fraudulent PPP loans, which represents 8% of total disbursed funds. See report here.

[11] See report here. An analysis by the U.S. Government Accountability Office (GAO) estimates the amount of fraud in pandemic UI programs at between $100 billion and $135 billion (see report here).

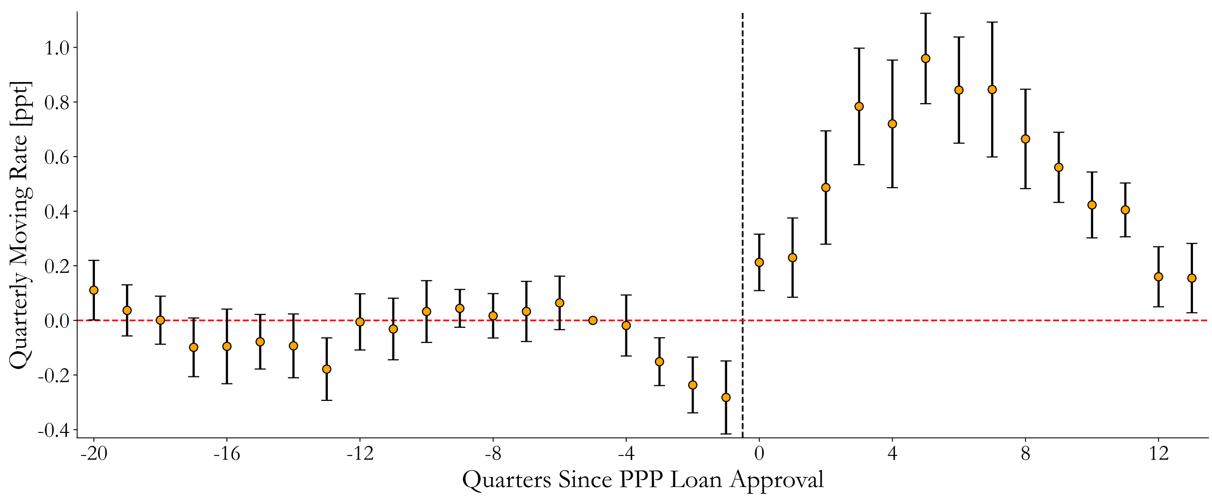
[12] Source and source.

[13] The average flagged loan is for $45,293 and the population in an average ZIP code is 15,881 people. Thus, $0.0068 \times 15,881 \times$ $45,293 = $4.9 million.

[14] The average ZIP code has 4474 single-family housing units that were worth $269,114 in December 2019. Thus, the housing stock in the average ZIP code was worth $1.2 billion. A turnover of 5% per year implies that the flow of house purchases during 2020 and 2021 was $120 million in the average ZIP code. The estimated turnover of 5% is based on annual U.S. home sales of approximately 6.4 million units (e.g., see here) and a housing stock of approximately 141 million units (e.g., see here).



(a) Purchases, Comparing Recipients of Flagged to Recipients of Non-Flagged PPP Loans



(b) Moving, Comparing Recipients of Flagged PPP Loans to Non-Recipients

Fig. 1. Effect on Housing Purchases and Moving Propensity. This figure shows the relationship between individuals receiving a flagged PPP loan and their likelihood of housing purchase (Panel A) and moving (Panel B). In Panel A, data from PropertyRadar for a sample of 250,000 loans is used to determine home purchases from one year before the individual received their PPP loan to 18 months after. The sample of non-flagged loans is reweighted to match the distribution of the timing of PPP loan approval in the sample of flagged loans. The horizontal lines are the average monthly likelihood of an individual in each group buying a house during the pre-/post-period. In Panel B, address history data from Verisk is used to track the moving propensity of recipients of PPP loans flagged by at least one primary flag and a comparison set of individuals who likely did not receive a PPP loan and were living in the same census block group as the recipient when the PPP loan was received. For each PPP recipient and their corresponding comparison set, we track moving rates from twenty quarters before the PPP loan was received to thirteen quarters afterward. To ensure correct weighting, a simple average of movement rates for each quarter across all addresses in each loan's comparison set is taken. The quarter that is five periods before the PPP loan was received is the omitted period, and the regression includes loan × 1(Suspicious PPP) and loan × relative quarter fixed effects. The error bars represent 95% confidence intervals based on standard errors that are double clustered by county and week of PPP approval. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

received funds through fraudulent PPP loans, EIDL advances and loans, or unemployment insurance claims potentially gained a windfall of new wealth that they could either spend or save. We first examine house purchases using property records from PropertyRadar for a random sample of 250,000 individual PPP borrowers. The sample was collected in February 2023 and consists of individual borrowers who received PPP loans during all three rounds of the PPP with data on house purchases through the end of 2022. Round 3 of the PPP ended in June 2021, so we observe at least 18 months of post-PPP house purchase activity for all individuals in the sample. We match the names of individuals purchasing houses in the PropertyRadar data to the names of PPP borrowers, limiting the sample to names that are unique.[16]

Fig. 1, Panel A plots monthly house purchase rates for PPP loan recipients before and after receiving a PPP loan in event time relative to the date that the PPP loan was approved. Flagged and non-flagged PPP borrowers follow parallel trends before receiving PPP loans, with an average monthly home purchase rate of 0.46% for both flagged and non-flagged PPP recipients. After receiving PPP loans, this purchase rate

[15] Loans and grants made by the EIDL program were also designed to replace lost revenue, and unemployment insurance is designed to cover at least part of the income lost due to unemployment.

[16] We determine unique names by using voter rolls from nine states that collectively represent 27% of the U.S. population; a name is unique if it occurs only once across all nine states. While this methodology may add noise relative to LexisNexis individual-level data (described below), estimates are similar and 85.3% of individuals in both samples have the same house purchase data.

Table 1
Effect on housing purchases and moving propensity.

Panel A: Purchases, comparing recipients of flagged to recipients of non-flagged PPP loans

	1(Housing purchase during month) × 12			
	PropertyRadar		LexisNexis	
	(1)	(2)	(3)	(4)
1(Post) × 1(Flagged)	0.00863***	0.00865***	0.0109***	0.00790***
	(5.22)	(5.59)	(4.44)	(4.87)
1(Post)	0.00686**		−0.00517	
	(2.16)		(−1.24)	
Loan FE	Yes	Yes	Yes	Yes
Month of Year FE	Yes	Yes	Yes	Yes
1(Post) × ln(Loan amount)	No	Yes	No	Yes
1(Post) × County FE	No	Yes	No	Yes
1(Post) × Business type FE	No	Yes	No	Yes
1(Post) × Week approved FE	No	Yes	No	Yes
Observations	19,500,000	19,500,000	9,600,000	9,600,000
R^2	0.0279	0.0279	0.0202	0.0206

Panel B: Moving, Comparing recipients of flagged PPP loans to non-recipients

	1(Moved during quarter)		
	Primary (1)	Primary + Additional (2)	Two Primary (3)
1(Post) × 1(Suspicious PPP)	0.00579***	0.00644***	0.00848***
	(9.43)	(10.95)	(8.21)
Loan × 1(Suspicious PPP) FE	Yes	Yes	Yes
Loan × Relative Quarter FE	Yes	Yes	Yes
Observations	26,964,652	19,272,016	2,669,000
R^2	0.534	0.534	0.534

This table examines whether individuals purchased homes (Panel A) and moved (Panel B) after receiving a flagged PPP loan. In Panel A, data on home purchases for a sample of 250,000 loans from PropertyRadar is used in columns (1) and (2) and for a sample of 150,000 loans from LexisNexis in columns (3) and (4). For each individual, we include monthly observations for the five years before they received their PPP loan to 18 (4) months after for PropertyRadar (LexisNexis). $1(Flagged)$ takes a value of 1 if the PPP loan is flagged by at least one primary flag. In Panel B, address history data from Verisk is used to track the moving propensity of recipients of flagged PPP loans and a comparison set of individuals who likely did not receive a PPP loan and were living in the same census block group as the recipient when the PPP loan was received. For each PPP recipient and their corresponding comparison set, we include quarterly observations from twenty quarters before the PPP loan was received to thirteen quarters afterward. To ensure correct weighting, a simple average of movement rates for each quarter across all addresses in each loan's comparison set is taken. Column (1) is based on loans flagged by at least one primary flag, column (2) on loans flagged by at least one primary and an additional (either primary or secondary) flag, and column (3) is based on loans flagged by at least two primary flags. Fixed effects and controls are indicated at the bottom of each column. t-statistics based on robust standard errors that are double clustered by PPP loan and month in Panel A and by county and week of PPP approval in Panel B are in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

increased by 6.3 bps (a 14% increase relative to the pre-period house purchase rate) for flagged recipients and remained largely unchanged for non-flagged recipients.

Table 1, Panel A estimates difference-in-differences regressions. Specifically, we estimate the following stacked difference-in-differences model:

$$Purchase_{i,\tau} = \beta 1(Post)_\tau 1(Flagged)_i + \gamma 1(Post)_\tau + LoanFE + MonthOfYearFE + \epsilon_{i,\tau} \quad (1)$$

The regression uses monthly observations for each PPP recipient in the sample, and the dependent variable is an indicator variable for whether the PPP recipient purchased a house in that month. The sample starts five years before each PPP loan was approved and ends 18 months after. $1(Post)_\tau$ is an indicator that takes the value of one starting in the month that the PPP recipient's loan was approved, which can be as early as April 2020 or as late as June 2021, depending on the recipient. $1(Flagged)_i$ is an indicator that takes the value of one for loans that are flagged by at least one primary flag.

In column (1), the regression controls for loan fixed effects and month-of-year fixed effects. The main coefficient of interest is β, which estimates the differential purchasing rate for recipients of flagged loans compared to recipients of non-flagged loans. All coefficients are multiplied by twelve to represent annual effects. Consistent with Panel A of Fig. 1, the house purchase probability for flagged PPP recipients increases by 0.86 ppt more than for non-flagged PPP recipients. Non-flagged PPP recipients experience an increase in house purchase probability as indicated by the value of γ at 0.69 ppt on *Post*. The value of β at 0.86 ppt, which is large relative to the average annual house purchase rate of 5.01%, represents a 17% relative increase in house purchases. Column (2) adds *Post* × county fixed effects, *Post* × business type fixed effects, *Post* × week approved effects, and log loan amount interacted with *Post*. The value of β remains significant with an unchanged estimate of 0.86 ppt. Consistent with Griffin et al.'s (2023) finding that PPP fraud was concentrated in FinTech loans, we find that the effects are concentrated in FinTech loans. Non-flagged FinTech loans are also associated with increased home purchase rates (see Table IA.1), consistent with these loans potentially exhibiting high rates of undetected fraud.[17]

As an additional examination of house purchases with a different sample, we examine house purchases using detailed property records from LexisNexis for a random sample of 150,000 individual PPP borrowers used by Griffin et al. (2023) for their felony analysis.[18] Columns

[17] Griffin et al. (2023) find that suspicious loan flags are a stronger signal of potential fraud for FinTech loans and, through a variety of secondary indicators, find that even non-flagged FinTech loans contain high rates of fraud.

(3) and (4) report the results, which are essentially the same as the baseline estimates in columns (1) and (2). Monthly home purchase results are also similar for the LexisNexis sample (see Figure IA.3).[19]

The above analyses compare recipients of flagged loans to recipients of non-flagged loans. In order to compare recipients of flagged loans to the typical household, we make use of a dataset from Verisk (formerly Infutor) that allows us to track the movement of a substantial portion of the U.S. population.[20] Once we have matched an individual to the Verisk data, we can track them across their ten most recent addresses. To form a comparison set for each recipient of a PPP loan, we consider all individuals living in the same census block group (CBG) as the recipient at the time they received their PPP loan but who did not receive a PPP loan themselves.[21] For each recipient of a PPP loan and for the individuals in the corresponding comparison set, we track their movements over the period spanning from twenty quarters before the PPP loan was received to thirteen quarters afterward (which corresponds to the end of 2024 for the last PPP loans approved in May/June 2021).

Having determined quarterly moving rates for PPP loan recipients and for their corresponding comparison sets, we perform a stacked difference-in-differences analysis. Specifically, we estimate the following dynamic stacked difference-in-differences model:

$$Move_{i,\tau} = \sum_{\tau \neq -5} \beta_\tau 1(Period = \tau) 1(SuspiciousPPP)_i + LoanBySuspiciousFE + LoanByRelativeQuarterFE + \epsilon_{i,\tau} \quad (2)$$

The main coefficient of interest is β_τ, which estimates the differential moving rate over time for recipients of flagged loans compared to non-recipients.[22] Fig. 1, Panel B plots these dynamics. From twenty quarters before to four quarters before the PPP loan is received, we find little to no differences in moving rates between recipients of suspicious PPP loans and their comparison set. During the three quarters immediately before the PPP loan is received, the recipients of suspicious PPP loans have lower moving rates than their comparison set.[23] After receiving their PPP loan, recipients of suspicious PPP loans immediately have higher moving rates compared to their comparison set. The difference in moving rates peaks five quarters after the PPP loan is received at 0.96 ppt and is, on average, 0.53 ppt per quarter over the fourteen-quarter period after the PPP loan was received. Column (1) of Table 1, Panel B collapses β_τ from Eq. (2) into a single coefficient for the interaction between $1(SuspiciousPPP)_i$ and an indicator variable for quarters after the PPP loan has been received. Columns (2) and (3) of Table 1, Panel B show that recipients of loans flagged by more stringent loan-level measures of fraud have even larger differential moving rates.[24] Table IA.7 shows that the moves by recipients of suspicious loans result in larger increases in neighborhood quality than moves by recipients of non-flagged loans.[25] A natural question is whether this differential moving propensity translates into differential housing purchases. While we cannot directly examine this using the Verisk data, Internet Appendix Section B outlines why heightened moving propensity likely implies heightened purchasing rates for recipients of suspicious loans compared to the typical household, as shown directly for recipients of suspicious loans compared to recipients of non-flagged loans previously using the data from PropertyRadar and LexisNexis. Overall, the cumulative evidence from our analysis based on data from three different sources indicates that fraudulent loans stimulated house purchases.

5. Does PPP fraud predict house price growth?

Because of the geographic clustering of PPP fraud and geographically correlated fraud in other pandemic relief programs, house purchases by recipients of fraudulent pandemic relief funds have the potential to distort local house prices. While fraudulent funds could have been used in many ways, the level of fraud combined with its high geographic concentration suggests that fraud may have distorted local house prices. We examine whether PPP fraud levels had a discernible effect on ZIP code-level house prices after controlling for other factors.

In this section, we examine the relation between ZIP code-level house prices and PPP fraud, as well as other potential factors. To control for macro factors that may influence regional house prices, our analysis focuses on the ZIP code level with county fixed effects. County fixed effects are used in most of our analyses since counties differed dramatically in their COVID policies, and these fixed effects also capture broader regional trends. First, we regress ZIP code-level house price growth on several measures of potentially fraudulent lending while controlling for demographic variables. Second, we use a

[18] The sample consists of individual borrowers who received PPP loans in rounds 1 and 2. The LexisNexis data was collected in March 2021 and includes data on house purchases through the end of 2020. Rounds 1 and 2 of the PPP occurred from April to August 2020, with most loans occurring by the end of May. As a result, we observe at least four months of post-PPP house purchase activity for all individuals in the sample.

[19] The advantage of the baseline sample is that it includes a larger sample of PPP borrowers from all three rounds and a longer post-PPP period. The LexisNexis sample only includes borrowers with loans in rounds 1 and 2 (i.e., in 2020). This sample, however, has the advantage of being matched to LexisNexis based on both name and address and uses LexisNexis data on home purchases at the individual level.

[20] See Internet Appendix Section B for additional details regarding this analysis and the matching process between the Verisk and PPP loan-level data based on dates and standardized addresses (using the U.S. Census Geocoder and the USPS Address Standardization API). We are able to track movements for recipients of over 2.81 million PPP loans, including recipients of 396,539 loans flagged by at least one primary flag. Table IA.2 compares the loan characteristics and local demographics of the matched loans with those of the overall population of PPP loans. See Diamond et al. (2019) for an analysis of the representativeness of the Verisk data more generally.

[21] CBGs are small geographic areas, and individuals living in the same CBG tend to be very similar in terms of socioeconomic characteristics. There are 239,780 CBGs across the U.S. with an average population of 1382 and a median area of 0.52 square miles. The comparison set is composed of individuals living at an average of 255 other addresses in the same CBG as the recipient.

[22] $LoanBySuspiciousFE$ controls for time-invariant differences across individuals. $LoanByRelativeQuarterFE$ makes sure comparisons are being made between recipients and their corresponding comparison set of non-recipients during each period. The change in estimation strategy between our purchasing and moving analyses is due to the difference in the comparison being made, i.e., between recipients of suspicious and non-suspicious loans for purchasing and between recipients of suspicious loans and non-recipients for moving.

[23] This difference is driven by individuals who received their PPP loan in 2021; the movement rates of recipients of suspicious PPP loans in 2020 are similar to their comparison set during the entire pre-period (Figure IA.4). Excluding the four quarters before the PPP loan was received does not meaningfully change the results (Table IA.3).

[24] Table IA.4 finds heightened moving rates based on additional loan-level measures of fraud. Table IA.5 considers heterogeneity in the effects of suspicious lending on differential moving propensity by various demographics. Figure IA.5 shows that recipients of non-flagged PPP loans have higher moving rates than non-recipients, though the coefficients are approximately half of those of the recipients of suspicious PPP loans. The triple-differences between recipients of suspicious and non-flagged loans are highly significant (Table IA.6). Though less common than for flagged loans, Griffin et al. (2023) demonstrate that PPP fraud was widespread even among non-flagged loans.

[25] We also examined moving propensity based on change-of-address data from Melissa Data for the same sample of 150,000 used for the LexisNexis analysis described previously and find that recipients of suspicious PPP loans were more likely to move after receiving a PPP loan compared to recipients of non-flagged PPP loans (Figure IA.6 and Table IA.8).

Table 2
Effect on house price growth.

	House price growth from January 1, 2020 to December 31, 2021				
	(1)	(2)	(3)	(4)	(5)
Flagged Per Capita	0.0179*** (10.38)	0.0211*** (9.84)			
High Loan-to-Est. Per Capita			0.0222*** (11.97)		
High Similarity Per Capita				0.0224*** (11.06)	
Flagged Composite Per Capita					0.0268*** (13.34)
County FE	Yes	Yes	Yes	Yes	Yes
Past HP Growth	No	Yes	Yes	Yes	Yes
Loans Per Capita	No	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
Observations	18,761	18,761	18,761	18,761	18,761
Num. Counties	2215	2215	2215	2215	2215
R^2	0.781	0.828	0.824	0.828	0.827
Mean of Dep. Var.	0.259	0.259	0.259	0.259	0.259

This table examines the relationship between house price growth and various measures of suspicious PPP lending. *Flagged Per Capita* is a ratio of the number of flagged PPP loans in the ZIP code to the ZIP code's population. *High Loan-to-Est. Per Capita* is based on the number of PPP loans in county-industry pairs where there are more than two times as many PPP loans as establishments per the US Census CBP. *High Similarity Per Capita* is based on the number of PPP loans in county-lender pairs with high levels of similarity in terms of loan amount, jobs reported, and industry. *Flagged Composite Per Capita* is based on the number of loans that are either flagged, in industry-counties where there are more than two times as many PPP loans as establishments, or in lender-counties with high levels of similarity. See Griffin et al. (2023) for additional details about these measures. *Past HP Growth* and *Loans Per Capita* control for house price growth in 2018–19 and PPP lending intensity, respectively, using percentile fixed effects. The controls included are log population density, vacancy rate, log housing units, and log average household income. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the population of the ZIP code in 2019. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

synthetic control methodology to generate counterfactuals for ZIP codes with similar house price trends. Third, we examine differences in high- and low-fraud areas and undertake several analyses to examine whether these differences are driving our results. Fourth, we use fraud rates in distant parts of the country that are socially connected to a given ZIP code as an instrument for local PPP fraud rates. Fifth, we examine and control for variables associated with house price growth in previous research, including measures specific to potential channels influencing house price growth during the COVID pandemic. We use variable selection procedures to compare the relative importance of these variables. Sixth, we explore the demand mechanism by examining additional ZIP code-level housing metrics from Realtor.com and Redfin. Finally, we examine longer-term price movements to test an earlier out-of-sample reversal prediction with more recent house price data.

5.1. Weighted least squares regressions

To understand the potential relation between suspicious lending and house prices across ZIP codes within counties, we regress house price growth on flagged PPP loans per capita with county fixed effects. Specifically, we estimated regressions of the form:

$$HPGrowth2020\text{-}21_i = \beta FlaggedPerCapita_i + CountyFE + Controls_i + \epsilon_i \tag{3}$$

The dependent variable in the regression is house price growth in ZIP code i between January 2020 and December 2021. *Flagged Per Capita* is standardized such that one unit represents one standard deviation.[26] The coefficient of interest is β, which estimates the effect on house price growth associated with a one standard deviation increase in *Flagged Per Capita*. The regressions include county fixed effects to isolate differential changes within counties. In addition to county fixed effects, the regressions also control for house price growth between January 2018 and December 2019, PPP loans per capita, population density, housing vacancy rates, number of housing units, and average household income as indicated in the table. Previous house price growth and PPP loans per capita are controlled for non-parametrically using percentile fixed effects, which allow for non-linear relationships.[27] The results are also similar when non-flagged loans per capita is used as a control instead of total loans per capita (as shown in Table IA.12). The regressions are weighted by the ZIP code's 2019 population to ensure our estimates are nationally representative.[28] Table 2 shows the relation between various measures of suspicious lending and house price growth.

Column (1) estimates a regression of house price growth on flagged PPP loans per capita at the ZIP code level with county fixed effects and no other control variables. The resulting coefficient of 0.0179 indicates that, on average, ZIP codes with one standard deviation higher suspicious PPP lending per capita experienced house price growth in 2020 and 2021 that was 1.79 ppt higher. Column (2) reports our baseline estimate, with control variables added for past house price growth, overall PPP lending per capita, and ZIP code demographic variables. Including these control variables modestly increases the coefficient estimate. A one standard deviation increase in *Flagged Per Capita* is associated with a 2.11 ppt increase in house prices between January

[26] Standardization throughout the paper is based on standard deviations weighted by population.

[27] Tables IA.10 and IA.11 show that the results are also robust to controlling for previous house price growth and loans per capita linearly or with higher order polynomials.

[28] Results are similar when ordinary least squares (OLS) is used (Table IA.13).

Table 3
Effect on house price growth, suspicious lending in five-mile radius.

	House price growth from January 1, 2020 to December 31, 2021			
	Simple average		Distance weighted average	
	(1)	(2)	(3)	(4)
Flagged Per Capita	0.0208***		0.0243***	
	(10.74)		(12.08)	
Flagged Composite Per Capita		0.0262***		0.0316***
		(8.63)		(11.35)
County FE	Yes	Yes	Yes	Yes
Past HP Growth	Yes	Yes	Yes	Yes
Loans Per Capita	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	18,761	18,761	18,761	18,761
Num. Counties	2215	2215	2215	2215
R^2	0.824	0.824	0.826	0.826
Mean of Dep. Var.	0.259	0.259	0.259	0.259

This table examines the relationship between house price growth and average suspicious lending rates in a five-mile radius around a ZIP code. In columns (1) and (2), simple averages of suspicious lending rates among all ZIP codes in a five-mile radius of the focal ZIP code are used. In columns (3) and (4), distance-weighted averages (specifically, weighted by 1/(1 + distance)) of suspicious lending rates among all ZIP codes in a five-mile radius of the focal ZIP code are used. ***Past HP Growth*** and ***Loans Per Capita*** control for house price growth in 2018–19 and PPP lending intensity, respectively, using percentile fixed effects. The controls included are log population density, vacancy rate, log housing units, and log average household income. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the population of the ZIP code in 2019. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

2020 and December 2021. This is a sizable 8.2% of the 25.9 ppt average increase in house prices during this period.[29] The relation between flagged loans and house price growth is concentrated in FinTech loans, consistent with more house purchasing by this group in the previous section and Griffin et al.'s (2023) findings that fraud was more pronounced in FinTech loans and that suspicious loan flags are a stronger predictor of fraud for FinTech loans (see Table IA.14).

Griffin et al. (2023) show that excess PPP loans relative to establishments and highly similar loans (with nearly identical loan features) are common and strongly correlate with other suspicious loan indicators.[30] Results based on these alternative measures in columns (3) and (4) are slightly stronger than the baseline results shown in column (2). Finally, in column (5), we consider a composite measure that is based on a combination of *Flagged Per Capita*, *High Loan-to-Establishment Per Capita*, and *High Similarity Per Capita*. Specifically, it is the ratio of the number of PPP loans that fit the criteria for any of these three measures to population, which we call *Flagged Composite Per Capita*. This more comprehensive measure indicates an even larger impact of suspicious lending on house prices. A one standard deviation increase in *Flagged Composite Per Capita* is associated with a 2.68 ppt increase in house prices, which is 10.4% of the average increase in house prices.

The results in Table 2 are robust to propensity score weighting to control for previous house price growth (Table IA.15), alternative suspicious lending measures focused on the dollar value of suspicious loans and the percent of loans that are suspicious (Table IA.16), alternative fixed effects (Table IA.17), using within-CBSA variation instead of within-county (Table IA.18), controlling for prepandemic house price levels (Table IA.19), controlling for COVID mortgage forbearance (Table IA.20), alternative standard error clustering (Table IA.21), restrictions to different subsets of loans (Table IA.22), restrictions to different subsets of states (Figure IA.7), and alternative house price data including the pre-2023 non-neural-network version of the Zillow Home Value Index (Table IA.23) and house price data from Realtor.com (Table IA.24).[31] Controlling for economic impact payments (i.e., stimulus checks) does not affect results (Table IA.25). Results using rent growth are positive but smaller than house price growth (Table IA.26, Panel A). There is no evidence of heterogeneity based on political leanings or exposure to COVID (Table IA.27). In contrast to the predictiveness of suspicious PPP loans, overall PPP lending rates are, if anything, negatively related to house price growth (Table IA.16), which is consistent with legitimate PPP funds being used to replace lost business revenues rather than creating a windfall for owners.

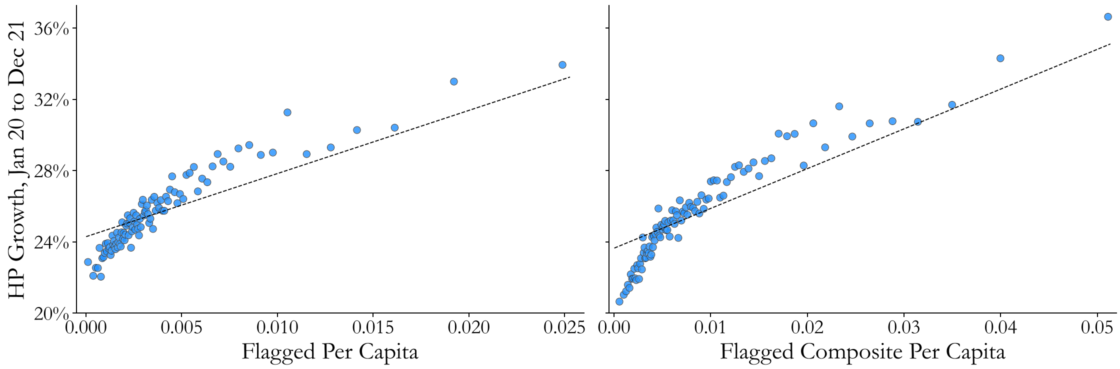
Because ZIP code house price indices may include some noise and house purchases need not occur in exactly the same ZIP code as the fraud recipient's current address and ZIP code, we also consider more aggregated versions of our house price regressions.[32] Columns (1) and (2) of Table 3 report results using the average fraud rate across all ZIP codes within a five-mile radius of the focal ZIP code. Results are nearly identical to the baseline results in Table 2. Columns (3) and (4) of

[29] The 2.11 ppt coefficient estimate is also large relative to the standard deviation of house price growth across ZIP codes during this period, which is 10.5 ppt.

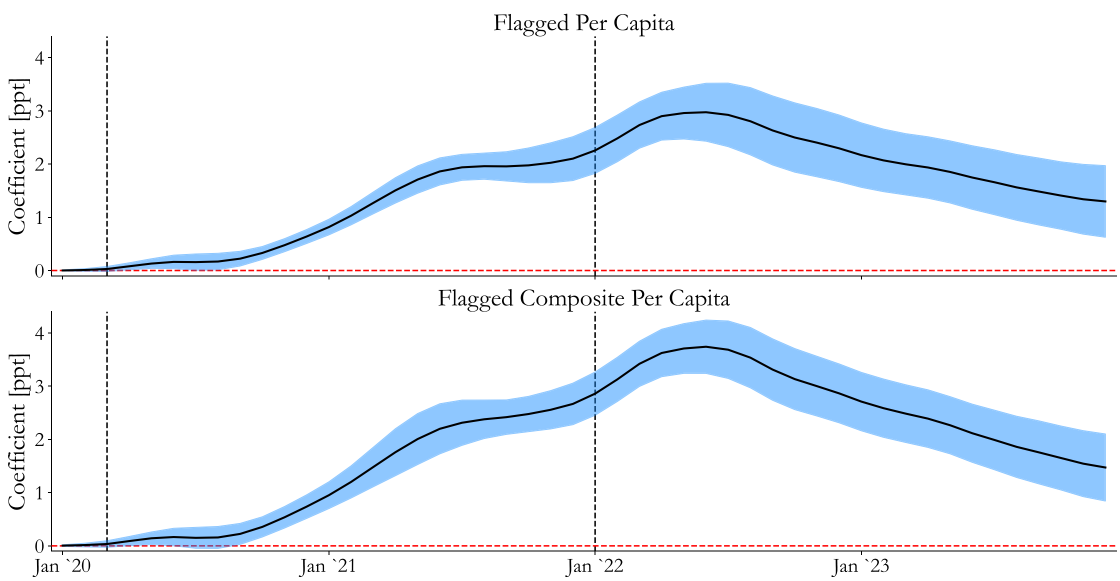
[30] The first measure is the ratio of the number of business PPP loans in the ZIP code that are in county-industry-pairs with excess loans, defined as county-industry-pairs with more than twice as many business PPP loans as establishments recorded in U.S. Census County Business Patterns (CBP) data, to population. Establishment counts in the CBP data are at the county-industry-pair level. This analysis is restricted to C-corporation, S-corporation, LLC, and sole-proprietorship loans because self-employed and independent contractors are not included in the CBP data. The second measure is the ratio of number of first-draw PPP loans in the ZIP code that are in county-lender pairs with high similarity, defined as average of the loan amount, number of jobs, and industry concentration indices of the county-lender being above the 75th percentile, to population. See Griffin et al. (2023) for more details on the construction of these measures.

[31] In Figure IA.8, we permute the suspicious lending measure between ZIP codes, both within the same county and across the nation, and show that the effect of suspicious lending is much smaller in all 10,000 permutations performed (the largest being less than 20% of the true coefficient); this implies the result we found is extremely unlikely to occur by chance.

[32] Based on address history data from Verisk, 30.9% of moves are within the same ZIP code and another 10.6% are between ZIP codes that are within five miles of each other.



(a) Binscatters, With County FEs and Controls



(b) Effect Over Time

Fig. 2. Effect on House Price Growth. This figure shows the relationship between suspicious lending and house price growth. Panel A shows the relationship using binscatters. Panel B shows the relationship over time. Panel C examines heterogeneity in the relationship across demographics. All three panels are estimated using ZIP code-level data, include county fixed effects, and control for house price growth in 2018 to 2019 and loans per capita using percentile fixed effects. Further, they control for log population density, vacancy rate, log housing units, and log average household income. The left subpanel of Panel A and the top subpanels of Panels B and C are based on *Flagged Per Capita*. The right subpanel of Panel A and the bottom subpanels of Panels B and C are based on *Flagged Composite Per Capita*. *Flagged Per Capita* is a ratio of the number of flagged PPP loans in the ZIP code to the ZIP code's population. *Flagged Composite Per Capita* is based on the number of loans that are either flagged, in county-industry pairs where there are more than two times as many PPP loans as establishments, or in county-lender pairs with high levels of similarity. See Griffin et al. (2023) for additional details about these measures. In Panels B and C, the measures of suspicious lending are standardized, so the coefficients represent the house price effect of a one standard deviation change in suspicious lending. The splits in Panel C are based on the median value of the demographic. To have a nationally representative estimate, all three panels use weighted least squares (WLS) regressions with the weight being the population of the ZIP code in 2019. The error bars in Panels B and C represent 95% confidence intervals based on standard errors clustered by county.

Table 3 repeat the same analysis with distance-weighted average fraud rates for ZIP codes within the same five-mile radius, again with nearly identical results. Table IA.28 considers an even broader aggregation of the data at the county level with CBSA fixed effects. Results are positive and significant with coefficients that are about half the size of our baseline estimates, consistent with much of the effect being at a more local neighborhood level within counties.

To graphically illustrate the relation between flagged loans per capita and house price growth between January 2020 and December 2021, the left (right) subpanel of Fig. 2, Panel A shows the relations between *Flagged Per Capita* (*Flagged Composite Per Capita*) and house price growth using binscatters. These binscatters include county fixed effects and the same controls used in the regressions reported in Table 2. The relation between suspicious lending and house price growth from January 2020 to December 2021 is close to linear, with a 5.8 ppt difference in house price growth between ZIP codes in the lowest decile of fraud rates and those in the highest decile.

To further understand the time-series dynamics of the relation between suspicious lending and house prices, we re-estimate specification (1) of Table 2 for the cumulative house price growth from January 2020 to each subsequent month. The coefficients on *Flagged Per Capita* for each month are reported, with a corresponding 95% confidence interval, in the top subpanel of Fig. 2, Panel B. The coefficient first becomes significantly positive in April 2020 and generally strengthens each month until it peaks in June 2022. After June 2022, the effect begins to decrease, and by the end of December 2023, the effect is half of its peak size. The bottom subpanel of Fig. 2, Panel B shows that the cumulative effect of pandemic fraud on house prices over time follows a similar pattern for the broader composite measure of suspicious lending, *Flagged Composite Per Capita*. The general trends in the coefficient match reasonably well with the timing of the PPP and other relief programs that started earlier or ended later, such as the EIDL program and expanded unemployment benefits.

To assess whether results are concentrated in ZIP codes with particular demographics, we add interactions between PPP fraud and indicator variables for whether the ZIP code is above or below the national median of different demographic characteristics (also controlling for the demographic indicator variables themselves). Panel C of Fig. 2 shows the results. The first column of Panel C plots separate PPP fraud coefficients for low- and high-income ZIP codes. Results are almost identical and remain large and highly significant for both subsets of ZIP codes. The same is true for ZIP code subsets based on poverty rates, population density, minority population share, educational attainment, and prepandemic employment. Effects are also strong across sample splits by specific race and ethnicity and when racially or ethnically homogeneous ZIP codes are excluded (see Table IA.29).[33] Even when coefficients differ a bit, the differences are not statistically significant. Consistent results across diverse ZIP codes point to a broad-based effect of PPP fraud as opposed to an effect that is concentrated in a particular demographic group. This is reassuring and may alleviate some concerns about omitted variables and non-random assignment of PPP fraud.

The influx of fraudulent funds represents a demand shock to housing, and the price effect is likely to be mediated by local housing supply elasticity, which is highly heterogeneous across locations. In particular, the effects of fraud should be more severe in locations with less elastic housing supply. Using the data from Baum-Snow and Han (2024), we split ZIP codes into terciles within counties based on their elasticities and examine whether the effect of pandemic fraud on house price growth is lower for the most elastic tercile compared to the two less elastic terciles. We find that a one standard deviation increase in pandemic fraud is associated with a 1.65 ppt increase in house price growth from January 2020 to December 2021 for elastic ZIP codes versus 2.41 ppt for the less elastic ZIP codes. This represents a difference of over 30% in the effect's magnitude, which is statistically significant at the 1% level (Figure IA.9 and Table IA.30).

5.2. Pre-trend analysis

One potential concern with our house price regressions is that PPP fraud is correlated, to some extent, with pre-existing house price trends.[34] We control for these trends in our baseline regressions with flexible percentile fixed effects for historical house price growth. In this section, we employ the synthetic control methodology to control for potential house price momentum.

We use synthetic controls based on Abadie et al. (2014) to construct a control group. The treated group consists of ZIP codes in the top quartile of the *Flagged Per Capita* measure within each county. The sample is limited to counties where the difference between the top and bottom quartile of *Flagged Per Capita* within the county is at least half the national standard deviation. This requirement is met by 283 counties, collectively representing approximately 25% of the U.S. population. For each treated ZIP code, we develop a synthetic control using ZIP codes in the same county that are in the bottom quartile of the *Flagged Per Capita* measure. This is achieved by finding weights that minimize the squared error in monthly house price growth from January 2018 to December 2019 between the synthetic control and the treated ZIP code.

In Fig. 3, we plot average house price growth for top quartile PPP fraud ZIP codes and synthetic control bottom quartile PPP fraud ZIP codes. The synthetic control methodology is designed to generate treatment and control groups with similar overall house price changes from January 2018 to December 2019, which the plot clearly demonstrates. After January 2020, the treatment and control groups continue to follow similar trends for several months. There is nothing mechanical about this result. The identical trends during these months indicate that the methodology successfully creates control groups with similar price trends. There is also no evidence of any differential impact of COVID during its early stages in March and April of 2020. By contrast, treatment and control groups start to diverge significantly in July 2020. This aligns with the expected timing for pandemic relief fraud to have an impact on house prices, as the PPP and other programs accelerated in April and it likely takes a few months to search for and purchase a house. ZIP codes in the top quartile of PPP fraud experienced an average house price growth of 28.3% in 2020 and 2021, compared to an average of 22.8% growth for the bottom quartile synthetic control ZIP codes. The difference of 5.5 ppt is economically significant and highly statistically significant, as shown by the 95% confidence intervals plotted as dotted lines. Figure IA.11 shows similar results based on the *Flagged Composite Per Capita* measure and several other measures and also shows that PPP lending in general did not affect house prices.

An alternative to the synthetic control methodology is to directly match high-fraud treated ZIP codes to low-fraud control ZIP codes with similar house price growth during 2018 and 2019. For this analysis, we match ZIP codes within CBSAs instead of counties because there are frequently not enough close matches within counties. We also restrict the sample to CBSAs with at least ten ZIP codes. In a previous version of this paper using legacy data from Zillow (before its adoption of a new neural network methodology), there was clear separation between matched high- and low-fraud ZIP codes, similar to the synthetic control results. After switching to the new Zillow data, there continues to be a positive difference between house price growth in matched high- and low-fraud ZIP codes, but the difference is much smaller than before (see Figure IA.12 for results with both the legacy data and the current Zillow data). All other results are virtually identical after updating the data (e.g., see Table IA.23), but we note this discrepancy.

5.3. Covariates

A second concern is that PPP fraud may correlate with other characteristics related to house price growth during 2020 and 2021. For example, if PPP fraud was concentrated in areas with strong economic growth during this period, the apparent relation between fraud and house price growth could be due to economic fundamentals rather than demand driven by fraud proceeds. Table 4, Panel A presents balance tests summarizing how ZIP code characteristics vary across areas with low, medium, and high fraud levels. High-fraud ZIP codes tend to have higher poverty rates, higher nonwhite population shares, and shorter distances to central business districts. Differences between low- and high-fraud areas are statistically significant for two-thirds of the variables. However, the economic magnitudes of these differences are modest when compared to the standard deviation of the underlying variables. The largest difference between low- and high-fraud ZIP codes is 0.36 of the standard deviation of the variable across ZIP codes. For five (nine) of the twelve variables examined, the absolute difference between low- and high-fraud areas is less than 0.1 (0.2) of the standard deviation of the variable across ZIP codes. As a result, the distribution of the variables across ZIP codes is highly similar across terciles of fraud (see Figure IA.13), which provides strong common support for the variables across ZIP codes with all levels of fraud.

[33] All demographic variables are from the U.S. Census American Community Survey. Poverty rate is the percentage of households with income below the poverty threshold, which varies based on family size and composition. Educational attainment is the percentage of adults with at least a bachelor's degree. Minority population share is the percentage of non-white individuals.

[34] To illustrate this potential concern, Figure IA.10 plots average house prices (indexed to January 2020) in ZIP codes in the top and bottom quartile of flagged PPP loans per capita for counties where the difference between the top and bottom quartile of fraud rates within the county is at least half the national standard deviation.

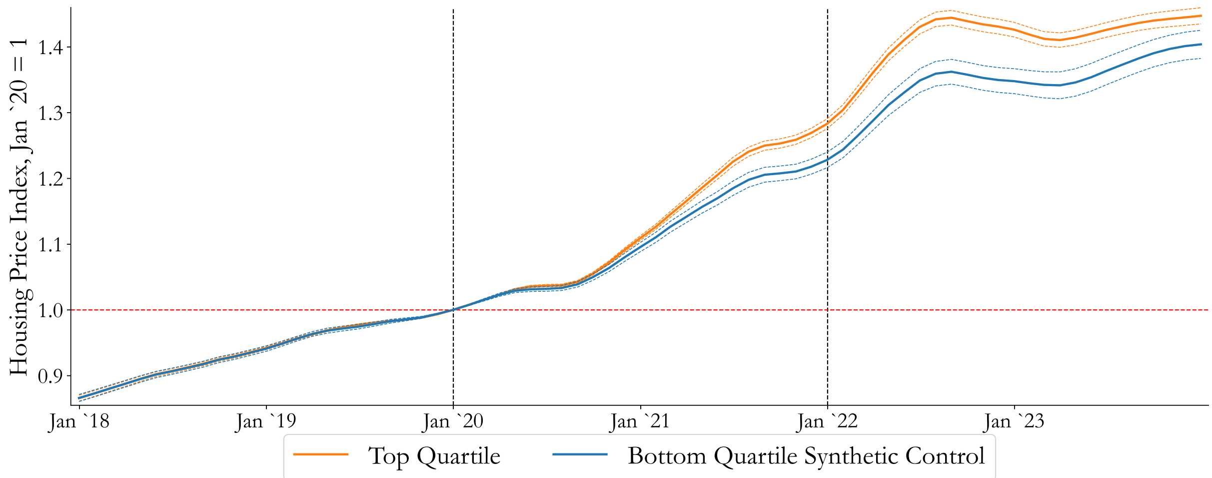


Fig. 3. Effect on House Price Growth, Synthetic Control. This figure shows the effect of suspicious lending on house price growth. A synthetic control method is used to create controls for each ZIP code in the top quartile of *Flagged Per Capita* using all ZIP codes in the same county that are in the bottom quartile of *Flagged Per Capita*. ZIP codes are split into quartiles within each county. Counties where the difference between the 75th and 25th percentiles of Flagged Per Capita within the county is at least half as large as the standard deviation across the entire nation are included. The dashed lines are 95% confidence intervals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4

Balance tables by within-county terciles of Flagged Per Capita and social proximity to suspicious lending.

Panel A: Within-county terciles of Flagged Per Capita

	Low fraud		Medium fraud		High fraud		High vs. Low		
	Mean	Median	Mean	Median	Mean	Median	Diff.	SE	Diff./SD
Log(Income)	11.117	11.042	11.107	11.022	11.076	10.979	−0.041***	(0.015)	−0.092
Pct. Poverty	0.112	0.095	0.125	0.109	0.143	0.125	0.031***	(0.002)	0.364
Log(Population density)	5.755	5.350	5.755	5.436	5.559	5.282	−0.197***	(0.045)	−0.097
Pct. Non-White	0.151	0.095	0.183	0.119	0.223	0.136	0.071***	(0.007)	0.346
Educational attainment	0.279	0.236	0.286	0.238	0.282	0.229	0.003	(0.005)	0.020
Pre-Pandemic unemployment	0.049	0.044	0.050	0.045	0.054	0.048	0.005***	(0.001)	0.172
Log(Distance to CBD)	2.803	2.872	2.630	2.719	2.478	2.666	−0.325***	(0.021)	−0.334
Land unavailability	0.256	0.250	0.254	0.238	0.258	0.237	0.002	(0.004)	0.011
Remote work 2015–19	0.052	0.046	0.051	0.045	0.053	0.044	0.001	(0.001)	0.026
Teleworkable	0.360	0.353	0.358	0.349	0.352	0.344	−0.008***	(0.001)	−0.173
Net migration 2020–21	−0.001	−0.002	−0.006	−0.006	−0.010	−0.011	−0.009***	(0.002)	−0.107
House price growth 2018–19	0.109	0.104	0.110	0.106	0.116	0.109	0.007***	(0.002)	0.098

Panel B: Within-county terciles of social proximity to suspicious lending

	Low SP		Medium SP		High SP		High vs. Low		
	Mean	Median	Mean	Median	Mean	Median	Diff.	SE	Diff./SD
Log(Income)	11.188	11.097	11.144	11.042	11.006	10.936	−0.182***	(0.014)	−0.417
Pct. Poverty	0.100	0.085	0.119	0.103	0.154	0.139	0.054***	(0.002)	0.627
Log(Population density)	5.449	5.027	5.693	5.383	5.783	5.570	0.334***	(0.048)	0.166
Pct. Non-White	0.125	0.076	0.167	0.111	0.250	0.168	0.126***	(0.006)	0.590
Educational attainment	0.291	0.246	0.296	0.243	0.268	0.222	−0.022***	(0.005)	−0.139
Pre-Pandemic unemployment	0.047	0.041	0.048	0.044	0.057	0.051	0.011***	(0.001)	0.332
Log(Distance to CBD)	2.940	2.977	2.641	2.723	2.372	2.518	−0.568***	(0.026)	−0.589
Land unavailability	0.256	0.248	0.255	0.239	0.257	0.237	0.002	(0.004)	0.010
Remote work 2015–19	0.058	0.051	0.053	0.046	0.048	0.041	−0.010***	(0.001)	−0.259
Teleworkable	0.363	0.355	0.360	0.350	0.349	0.342	−0.014***	(0.001)	−0.326
Net migration 2020–21	0.010	0.003	−0.003	−0.004	−0.018	−0.017	−0.028***	(0.002)	−0.355
House price growth 2018–19	0.104	0.100	0.107	0.104	0.122	0.114	0.018***	(0.002)	0.252

This table examines differences in various ZIP code-level socioeconomic variables and factors the literature has found to be associated with house price growth during 2020–21 by the level of suspicious lending in the ZIP code (Panel A) and the ZIP code's social proximity to suspicious lending (Panel B). ZIP codes in each county are split into terciles based on their ratio of flagged loans to population (Panel A) and on their social proximity to suspicious lending (Panel B). Mean and median values of each variable among ZIP codes in each tercile are provided. In addition, the difference in the mean value of the variable between the low- and high-fraud terciles, the standard error for the difference, and the standardized mean difference (difference divided by the standard deviation) are provided. Robust standard errors are clustered by county and are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Our baseline specifications address potential alternative drivers of house price growth in several ways. First, our analyses include county fixed effects, which should absorb most differences in economic growth and job opportunities across different areas. Second, the baseline regressions include extensive control variables: PPP loans per capita, past house price growth, population density, housing vacancy rates, number of housing units, and average household income. Results are also robust to different control variables and fixed effect specifications (see Table IA.17). The stability of the effect of fraud on house prices with and without control variables mitigates potential concerns about omitted variables under the logic of Oster (2019).[35] Finally, the lack of heterogeneity in the effect of fraud on house prices across numerous sample splits based on demographic and economic characteristics implies that any potential omitted variables would need to affect a diverse set of ZIP codes in a similar manner.

5.4. Instrumental variables regressions

To further address at least some of the potential omitted variable concerns, we instrument for PPP fraud using fraud in geographically distant but socially connected ZIP codes. This identification strategy builds on Griffin et al. (2025), which shows that the spread of fraud is strongly related to social connections and that a ZIP code's fraud rate is strongly predicted by fraud rates in other ZIP codes with which it has strong social connections. The specific instrument we consider is the average (weighted by social connection strength) *Flagged Per Capita* in other ZIP codes to which a ZIP code is connected, which Griffin et al. (2025) term "social proximity to fraud". Social connectedness between ZIP codes i and j is measured as the Facebook friendships between users in ZIP code i and ZIP code j scaled by the product of Facebook users in ZIP code i and ZIP code j, using data from Bailey et al. (2018b) and Bailey et al. (2020). Additional details on the construction of the instrument are discussed in Internet Appendix Section C. While social connections to other ZIP codes are endogenous, this identification strategy has the benefit of isolating variation related to distant social connections, as opposed to anything that might jointly influence housing markets and PPP fraud at the local ZIP code level. Balance tests across levels of social proximity to fraud are reported in Panel B of Table 4, and covariate distributions are plotted in Figure IA.14. The distributions are highly overlapping, albeit with larger differences than for the balance tests of fraud itself, which could be due to measurement error that the instrument is correcting.[36]

Table 5 shows the relationship between suspicious lending rates and house price growth based on the instrumental variable described above. Specifically, we estimate the following two-stage least squares regression:

$$SocialProximityToFraud_i = \frac{\sum_{CBSA(i)\neq CBSA(j)} SCI_{i,j} \times FlaggedPerCapita_j}{\sum_{CBSA(i)\neq CBSA(j)} SCI_{i,j}}$$

$$FlaggedPerCapita_i = \gamma SocialProximityToFraud_i + CountyFE + Controls_i + \epsilon_i \quad (4)$$

$$HPGrowth2020\text{-}21_i = \beta \widehat{FlaggedPerCapita}_i + CountyFE + Controls_i + \nu_i \quad (5)$$

As in our WLS estimates, the measures of suspicious lending are standardized to have a mean of zero and a standard deviation of one, so the coefficients represent the effect of a one standard deviation increase in suspicious lending rates. The coefficient of interest is β, which estimates the effect on house price growth associated with a one standard deviation increase in *Flagged Per Capita*. In addition to county fixed effects, both stages also control for house price growth in 2018 and 2019, PPP loans per capita, population density, housing vacancy rates, number of housing units, average household income, and the share of friends of Facebook users in the ZIP code who live within 50 and 150 miles of the ZIP code. Both previous house price growth and PPP loans per capita are controlled for non-parametrically using percentile fixed effects, which allow for non-linear relationships. The estimates are weighted by the ZIP code's 2019 population to ensure our estimates are nationally representative. In column (1), we instrument for a ZIP code's *Flagged Per Capita* using social connections outside of the CBSA where the ZIP code is located. A one standard deviation increase in instrumented *Flagged Per Capita* is associated with a 3.47 ppt increase in house prices from January 2020 to December 2021. This is a sizable 13.5% of the 25.7 ppt average increase in house prices during this period.

To the extent that distant social connections are less likely to affect house price growth through, for example, migration, local omitted variables, or regional shocks, social proximity to fraud based on only distant ZIP codes may be more likely to meet the exclusion restriction.[37] To examine this, columns (2), (3), and (4) of Table 5 instrument for a ZIP code's *Flagged Per Capita* using social connections to ZIP codes that are at least 100, 250, and 500 miles away, respectively. The results in all four specifications are extremely similar (between 3.36 and 3.52 ppt, with t-statistics above 5.4 and first-stage F-statistics of at least 28.5) and cannot be statistically distinguished from one another. In column (5), we include multiple instruments based on social connections to ZIP codes in non-overlapping rings (between 100 and 250 miles, between 250 and 500 miles, and over 500 miles). The estimate is again nearly identical, and including multiple instruments allows us to conduct a J-test of overidentifying restrictions. The J-statistic of 2.464 with a p-value of 0.292 indicates that estimates are consistent regardless of which subset of the instruments is used. This implies that any effect that social proximity to fraud has on house price growth directly, or indirectly through omitted variables, must be the same over different distances.[38]

Social connections have been shown to affect other economic outcomes, particularly for house price expectations. Bailey et al. (2018a,c) show that individuals whose friends experience house price increases have heightened house price expectations and are more likely to purchase a home. To examine the potential effects of house price expectations being transmitted through social connections, we construct a measure of social proximity to house price growth in a manner analogous to social proximity to suspicious lending. After controlling for social proximity to house price growth, social proximity to fraud

[35] In addition to estimating regressions with and without control variables, Table IA.17 also controls for additional potential channels that are discussed in the next subsection. Using the approach proposed by Oster (2019), if omitted variables have the same proportional impact on the flagged per capita coefficient as observed control variables (equal selection assumption) and could potentially increase the R^2 of the regression from 0.831 to 1.0, the coefficient on flagged per capita in column (4) would only decrease from 0.0175 to 0.0170 [$0.0170 = 0.0175-(1-0.831)\times(0.0177-0.0175)/(0.831-0.763)$]. Relaxing the equal selection assumption and maintaining the assumption that omitted variables could increase the R^2 to 1.0, other omitted variables would need to have approximately 35.2 times the proportional impact of the observed control variables to decrease the coefficient to zero. It is worth noting that these controls include drivers of house price growth that have already been documented in the literature, which should arguably have the largest impact and thus make it even more unlikely for unobserved variables to explain the effect.

[36] See Pancost and Schaller (2025) for a discussion of measurement error and instrumental variable regressions.

[37] This is similar to the logic provided by Bailey et al. (2018a,c) for using house price experiences of out-of-commuting-zone friends as an instrument for house price experiences of all friends. Hu (2021) uses floods in distant but socially connected counties to study flood insurance purchases since distant flooding is likely orthogonal to an individual's local flood risk.

[38] As one example of how the direct effects of social connections on house prices vary by distance, Table IA.31 shows that the effect of social connections on migration rates is decreasing with distance.

Table 5
Effect on house price growth, IV.

	House price growth in 2020–21					2021 only
	Outside CBSA (1)	≥100 Mi (2)	≥250 Mi (3)	≥500 Mi (4)	Concentric rings (5)	2020 Outside CBSA (6)
Flagged Per Capita	0.0347*** (6.35)	0.0353*** (6.51)	0.0336*** (6.17)	0.0340*** (5.43)	0.0367*** (6.41)	
2021 Flagged Per Capita						0.0287*** (2.76)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Past HP Growth	Yes	Yes	Yes	Yes	Yes	Yes
Loans Per Capita	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
2020 Flagged Per Capita	No	No	No	No	No	Yes
2020 House price growth	No	No	No	No	No	Yes
Observations	16,869	18,761	18,761	18,761	18,725	16,868
Num. Counties	1789	2215	2215	2215	2213	1789
Mean of Dep. Var.	0.257	0.259	0.259	0.259	0.259	0.148
First stage F-stat	34.33	38.04	34.27	28.48	14.38	14.81
Hansen's J-stat					2.465	
(p-value)					(0.292)	

This table examines the relationship between house price growth and suspicious PPP lending using instrumented variables based on social connectedness between ZIP codes. Column (1) is based on social connectedness between each ZIP code and ZIP codes that are outside the given ZIP code's CBSA. Columns (2), (3), and (4) are based on social connectedness between each ZIP code and ZIP codes that are at least 100, 250, and 500 miles away, respectively. Column (5) includes three instruments based on social connections to ZIP codes in non-overlapping rings (between 100 and 250 miles, between 250 and 500 miles, and over 500 miles) at the same time. The J-stat and p-value for an overidentification test are provided at the bottom of column (5). Column (6) is based on fraud rates in 2020 and social connectedness between each ZIP code and ZIP codes that are outside the given ZIP code's CBSA. Columns (1) to (5) use house price growth during the entire 2020–21 period as the dependent variable, whereas column (6) uses house price growth during only 2021 as the dependent variable. *Past HP Growth* and *Loans Per Capita* control for house price growth in 2018–19 and PPP lending intensity, respectively, using percentile fixed effects. The controls included are log population density, vacancy rate, log housing units, log average household income, and the share of Facebook friends within 50 and 150 miles. *2020 Flagged Per Capita* and *2020 House Price Growth* are the fraud rate and house price growth for each ZIP code during 2020. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. To have a nationally representative estimate, we use weighted least squares (WLS) regressions in both stages, with the weight being the population of the ZIP code in 2019. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

continues to have a strong effect on house price growth (see Table IA.32). This indicates that it is unlikely that social proximity to fraud is capturing house price expectations being transmitted through social connections. We also find that the effect of house price growth in socially proximate areas diminishes with distance, whereas the effect of social proximity to fraud is robust even when based on social connections over 500 miles away. This supports the aforementioned logic for using distant social connections and testing overidentifying restrictions.[39] We also re-examine our findings on heterogeneity by housing supply elasticity using the social proximity instrument and find that the effect of fraud on house price growth is 35% larger in less elastic areas.[40]

To address potential contemporaneous transmission of fraud and house price expectations and possible reverse causality concerns (e.g., house price growth leading to more pandemic fraud, perhaps because homebuyers in these areas commit fraud to afford a home), we also consider a lagged version of the instrument. In this approach, we instrument for 2021 fraud in a given ZIP code with its social proximity to 2020 fraud. This strategy uses variation in fraud rates that occurred entirely before the change in house price being studied (2020 fraud rates and 2021 house price growth). Using this instrument, we then regress house price growth in 2021 on fraud in 2021, instrumented using social proximity to 2020 fraud, while controlling for the ZIP code's own 2020 fraud rate, house price growth in 2020, and all other controls used in the baseline IV analysis. Column (6) of Table 5 reports the results. Consistent with much of the effect of fraud on house prices occurring in 2021 (Figure IA.15), the effect we find using this strategy is 83% [2.87 ppt/3.47 ppt] of the baseline IV result (which is based on the entire 2020–21 period).[41]

5.5. Other proposed factors affecting house prices

A growing body of literature proposes several factors that potentially affected house prices during the COVID period. We construct measures for these factors, adhering as closely as possible to the existing literature. The factors considered include: prior remote work from 2015 to 2019 (Mondragon and Wieland, 2025; Davis et al., 2024), the percent of teleworking individuals in the CBSA prior to COVID (Dingel and Neiman, 2020), population density (Liu and Su, 2021), net migration during 2020 and 2021 (Ramani and Bloom, 2022), distance to the central business district (Gupta et al., 2022), previous (2018–2019) house price growth, and land unavailability (Lutz and Sands, 2023).[42] All independent variables are standardized to have a standard deviation of one to allow for easier comparisons of the economic magnitude

[39] Additional IV analysis and robustness tests are discussed in Internet Appendix Section C.

[40] Using the version of the instrument constructed based on social connections more than 500 miles away, an instrumented increase of one standard deviation in pandemic fraud is associated with a 2.56 ppt increase in house price growth from January 2020 to December 2021 for elastic ZIP codes versus 4.01 ppt for less elastic ZIP codes (Table IA.33). This is a 35% difference in the effect's magnitude, which is statistically significant at the 1% level.

[41] For context, the house price growth in 2021 was 57% [14.8 ppt/25.7 ppt] of the total growth over the entire 2020–21 period. Table IA.34 shows additional variations of these results.

[42] Lin (2025) finds that MSAs with higher economic impact payments (i.e., stimulus payments) and child tax credits per capita experienced higher house price growth. We do not include this channel in our baseline analysis because we do not find support for a positive relationship between economic impact payments and house price growth at the ZIP code level using the standardized regression framework described below (Table IA.25). Controlling for economic impact payments does not affect results for PPP fraud.

Table 6
Effect on house price growth, univariate and multivariate.

	House price growth from January 1, 2020 to December 31, 2021								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Flagged Per Capita	0.0221*** (8.99)								0.0175*** (7.41)
Log of Dist. to CBD		0.0161*** (4.47)							0.00880*** (3.01)
Log of Pop. Density			−0.0189*** (−4.03)						−0.000170 (−0.07)
Land unavailability				0.0208*** (8.36)					0.0157*** (7.72)
Remote Work 2015–19					0.00771*** (4.52)				0.00569*** (3.44)
Teleworkable						−0.0187*** (−5.66)			−0.0120*** (−4.27)
Net migration 2020–21							0.0129*** (6.55)		0.00674*** (5.12)
HP Growth 2018–19								0.0260*** (6.00)	0.0206*** (6.61)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Past HP Growth	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Loans Per Capita	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	12,305	12,305	12,305	12,305	12,305	12,305	12,305	12,305	12,305
Num. Counties	1013	1013	1013	1013	1013	1013	1013	1013	1013
R^2	0.814	0.811	0.807	0.810	0.803	0.806	0.810	0.797	0.828
Mean of Dep, Var.	0.257	0.257	0.257	0.257	0.257	0.257	0.257	0.257	0.257

This table examines the univariate and multivariate relationships between various proposed variables and house price growth. *Past HP Growth* and *Loans Per Capita* control for house price growth in 2018–19 and PPP lending intensity, respectively, using percentile fixed effects. The controls included are vacancy rate, log housing units, and log average household income. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the population of the ZIP code in 2019. Only ZIP codes for which all variables can be determined are used. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

of the coefficients. We include county fixed effects, past house price growth, PPP loans per capita, vacancy rate, housing units, and average household income as control variables in all regressions. We are able to estimate the specification for a large cross-section of 12,305 ZIP codes for which we have data for all of the proposed factors.

The univariate regressions for each of the proposed factors are shown in Table 6. All of the factors are statistically significant, but with varying economic magnitudes.[43] In the multivariate regression (column (9)), the effects of most of the factors remain statistically significant but are attenuated relative to the univariate regressions to varying degrees. The coefficient on *Flagged Per Capita* remains statistically and economically significant, with a magnitude that is close to the univariate coefficient. The coefficient on population density becomes insignificant. To better visualize these relationships, Panel A of Fig. 4 shows the univariate and multivariate coefficients with 95% confidence intervals. *Flagged Per Capita*, land unavailability, and house price growth in 2018–19 have the largest coefficients, at slightly over 2 ppt of house price growth per standard deviation. Teleworking, 2020–2021 migration, prior remote working, and distance to the central business district are all also statistically significant in the multivariate specifications, though with considerably smaller coefficients.[44]

We consider alternative specifications, including averaging each of the factors over a 5-mile radius, only including county fixed effects, and using OLS instead of WLS (see Figure IA.16). We also examine heterogeneity in the effects across splits based on land unavailability, previous house price growth, pre-COVID house prices, COVID mortgage forbearances, ZIP code-level beta with national house prices during 2000–2019, and ZIP code-level house price volatility during 2000–2019 (see Figure IA.17). Across all of these variations, the effect of suspicious lending on house prices is similar to the baseline results. To understand the effects of each factor on house prices over time, we perform the same monthly analysis as in Panel B of Fig. 2 for each of the other proposed factors (see Figure IA.18).

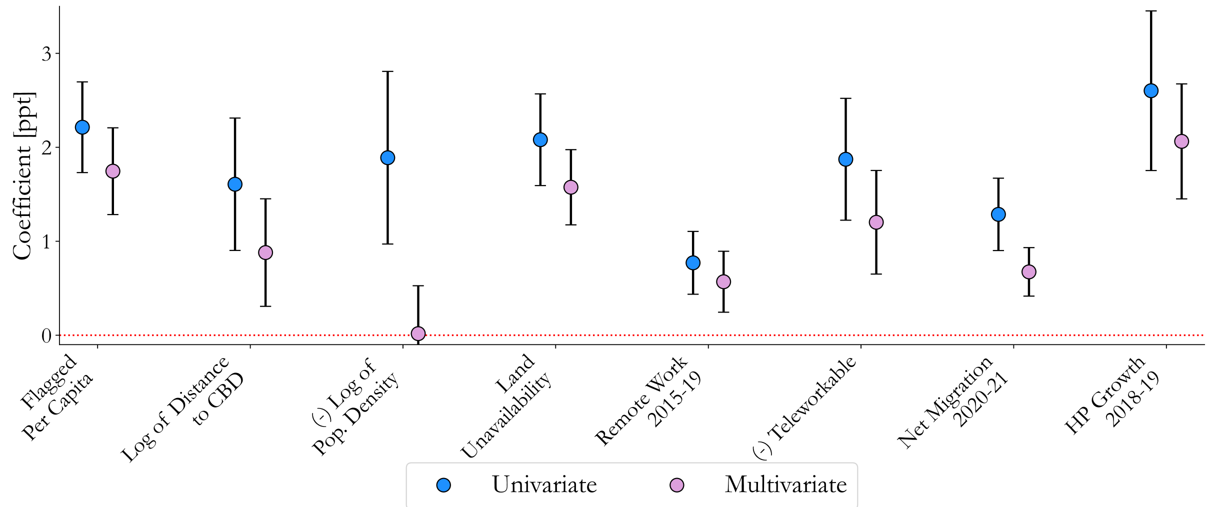
Correlation between the factors could complicate the multivariate estimates. To more formally assess which factors most robustly predict house price growth, we apply the Bayesian Model Averaging approach to model selection as suggested by Fernández et al. (2001) and Ley and Steel (2009), and following Griffin et al. (2020), which examined the determinants of 2003 to 2012 house price growth. Following the assumptions recommended by Ley and Steel (2009) for modeling and prior distributions, the procedure estimates posterior distributions for the probability that a variable is included in a model and the coefficient conditional on inclusion. Posterior coefficient distributions conditional on inclusion in the model are plotted in Panel B of Fig. 4. The probability of inclusion in the model is shown by the bars above each of the distributions.[45] The procedure always includes *Flagged Per Capita*, land unavailability, 2018–2019 housing price growth, teleworking, net migration, remote work, and distance to CBD. Population density is only selected in 0.9% of models. Conditional on inclusion, the coefficients are furthest from zero for house price growth in 2018–19, *Flagged Per Capita*, land unavailability, and teleworking, in that order.

Table 7 reports optimal model selection based on the Bayesian Information Criterion. The column number corresponds to the best model if the model is limited to that number of proposed factors (between one and eight). If the model is restricted to one factor, *Flagged*

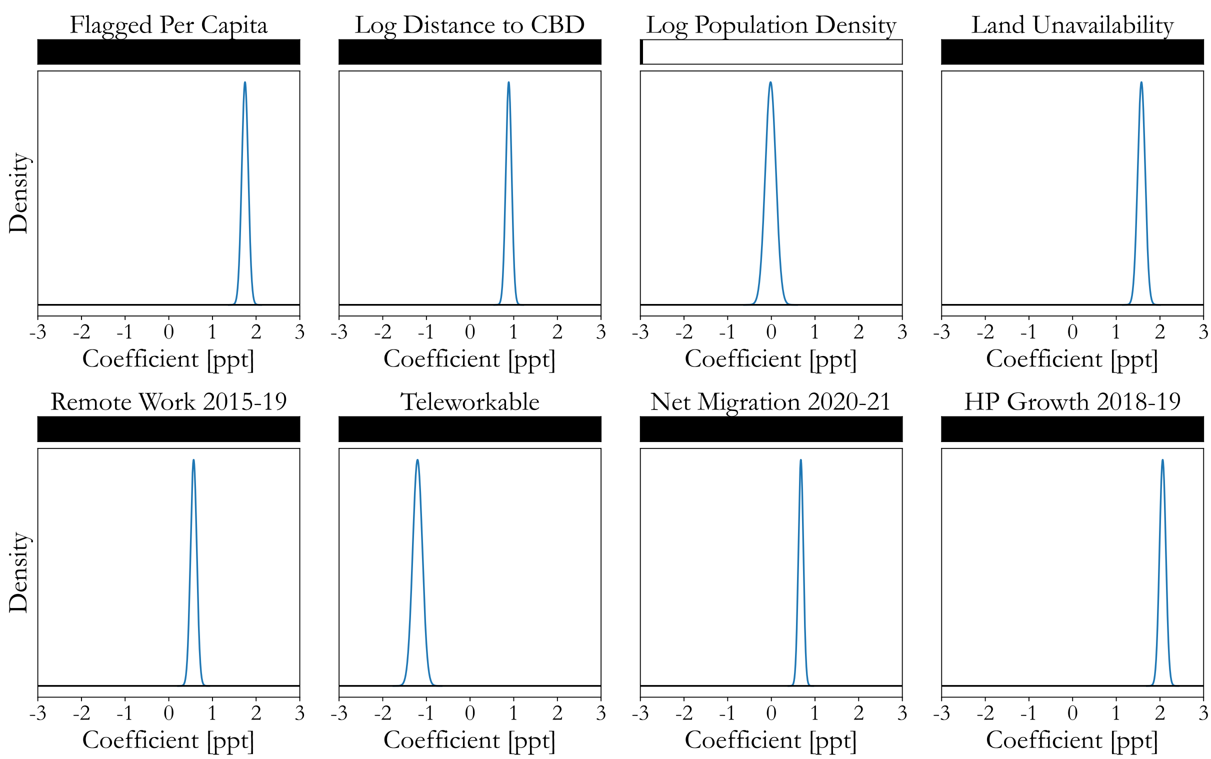
[43] The univariate relationships are similar to those documented in the existing literature.

[44] Table IA.35 shows that the effect of *Flagged Per Capita* is robust to controlling for alternative measures of migration, such as migration during 2020 and 2021 separately or decomposing net migration into inflows and outflows.

[45] Column 9 of Table 7 also reports the posterior inclusion probabilities for each proposed factor.



(a) Univariate and Multivariate Coefficients



(b) Bayesian Model Averaging

Fig. 4. Effect of other proposed variables on house price growth. This figure shows the effect of each proposed variable on house price growth. Panel A shows univariate and multivariate coefficients using WLS regressions. Panel B shows the posterior coefficient distribution, conditional on inclusion, from multivariate regressions using Bayesian model averaging. The black bars at the top of each distribution plot the posterior inclusion probability. All regressions control for vacancy rate, log housing units, log average household income, and for overall PPP loans per capita and house price growth in 2018–19 using percentile indicator variables to allow for non-linearity, and include county fixed effects. All proposed variables are standardized to have a mean of 0 and a standard deviation of 1. The weighted least squares (WLS) regressions are weighted by the population of the ZIP code in 2019. Only ZIP codes for which all proposed variables can be determined are used in both panels. The error bars in Panel A represent 95% confidence intervals based on standard errors clustered by county. The regressions corresponding to Panel A are shown in Table 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Per Capita is included. Furthermore, *Flagged Per Capita* is consistently included when the model is restricted to any number of factors. The optimal model, across any number of the proposed factors according to the Bayesian Information Criterion, includes seven of the eight factors (all but population density). Overall, the evidence supports many of the proposed factors and shows that regardless of which channels are considered, pandemic fraud continues to be a strong predictor of house price growth during 2020 and 2021. The relationship between fraud and house price growth has a magnitude that is at least as high as any other proposed factor and is robust to controlling for any combination of other factors.

Table 7
Effect on house price growth, variable selection.

	House price growth from January 1, 2020 to December 31, 2021								
	Regression coefficients and t-statistics								PIC Prob.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Flagged Per Capita	0.0269^{***} (11.52)	0.0223^{***} (9.79)	0.0202^{***} (8.02)	0.0186^{***} (7.65)	0.0176^{***} (7.41)	0.0176^{***} (7.57)	0.0175^{***} (7.42)	0.0175^{***} (7.41)	1.000
HP Growth 2018–19		0.0215^{***} (5.96)	0.0231^{***} (6.93)	0.0217^{***} (6.72)	0.0217^{***} (6.74)	0.0208^{***} (6.78)	0.0206^{***} (6.60)	0.0206^{***} (6.61)	1.000
Log of Dist. to CBD			0.0142^{***} (3.94)	0.0127^{***} (3.97)	0.00990^{***} (3.20)	0.00903^{***} (2.95)	0.00886^{***} (2.89)	0.00880^{***} (3.01)	1.000
Land unavailability				0.0163^{***} (8.16)	0.0161^{***} (7.77)	0.0159^{***} (7.44)	0.0158^{***} (7.59)	0.0157^{***} (7.72)	1.000
Net migration 2020–21					0.00825^{***} (6.20)	0.00740^{***} (5.74)	0.00676^{***} (5.42)	0.00674^{***} (5.12)	1.000
Teleworkable						-0.0108^{***} (−3.78)	-0.0120^{***} (−4.27)	-0.0120^{***} (−4.27)	1.000
Remote work 2015–19							0.00569^{***} (3.44)	0.00569^{***} (3.44)	1.000
Log of Pop. Density								−0.000170 (−0.07)	0.009
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loans Per Capita	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	12,305	12,305	12,305	12,305	12,305	12,305	12,305	12,305	12,305
Num. Counties	1013	1013	1013	1013	1013	1013	1013	1013	1013
R^2	0.798	0.811	0.818	0.823	0.826	0.827	0.828	0.828	
Mean of Dep. Var.	0.257	0.257	0.257	0.257	0.257	0.257	0.257	0.257	0.257

This table shows the results of a variable selection process in which the optimal model, with between one and eight of the potential independent variables, is selected based on the Bayesian Information Criterion. Column (9) reports posterior inclusion probabilities for each variable based on Bayesian model averaging. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the ZIP code's population as of 2019. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

5.6. Demand mechanisms and potential reversal

The mechanism driving higher house price growth in ZIP codes with higher levels of PPP fraud is that recipients of fraudulent funds respond to this wealth shock, at least in part, by purchasing houses. This increases demand and pushes up house prices in the short term. However, if housing prices were at an equilibrium prior to this shock, house prices should reverse once the influx of fraudulent funds ceases and demand returns to its prior levels. That is, in subsequent periods, the higher house prices fueled by pandemic fraud should lead some marginal homeowners to sell and some marginal potential buyers to decide against purchasing. Thus, in ZIP codes receiving substantial fraudulent funds, we expect to see an increase in net demand followed by the opposite after most fraudulent funds are spent. In an earlier version of this paper, we made the out-of-sample prediction that there could be a reversal in house prices in fraudulent ZIP codes.[46]

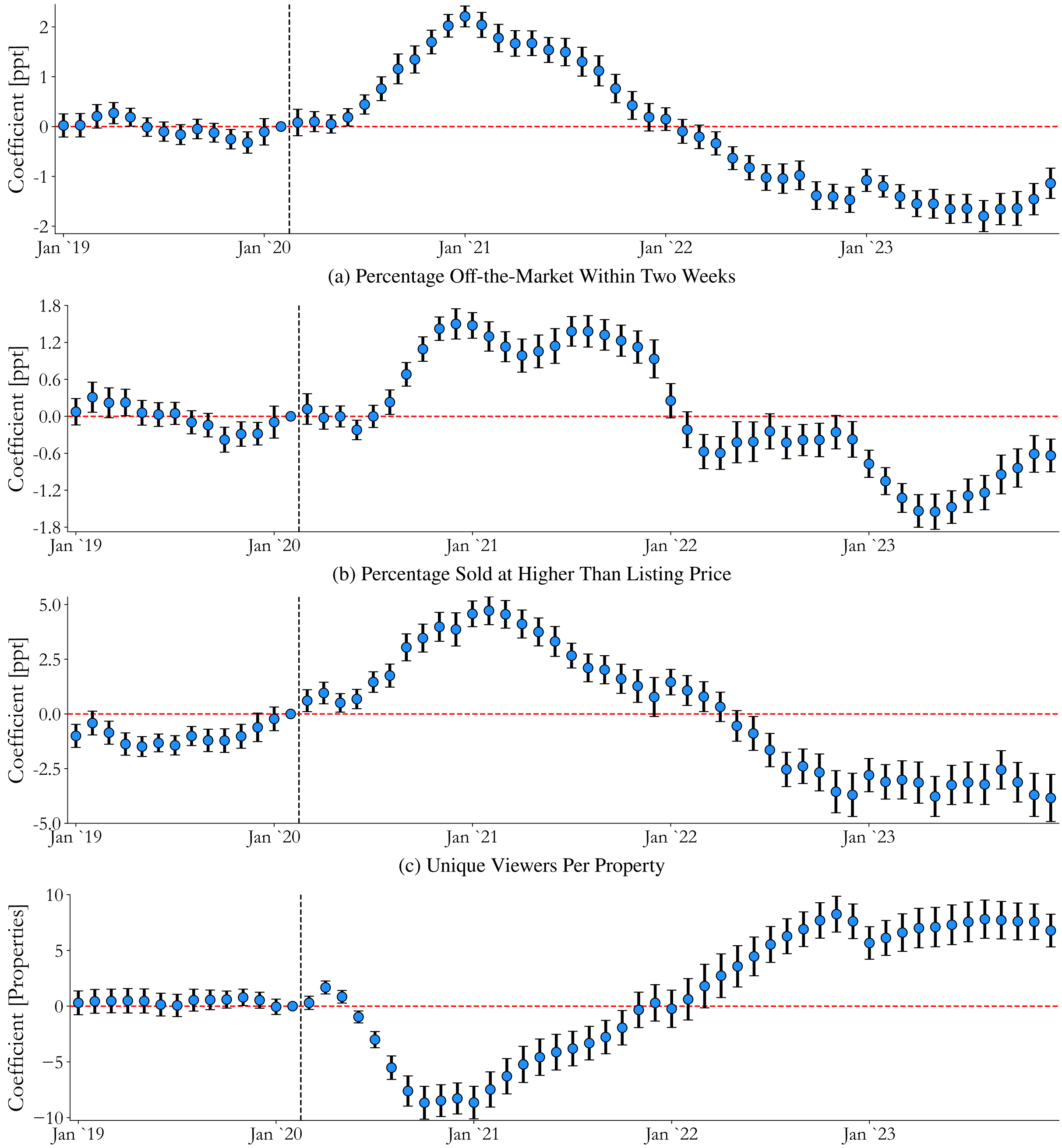
We test the intuition for this demand mechanism by examining additional ZIP code-level housing metrics from Realtor.com and Redfin. In particular, we examine the percentage of listings that are off the market within two weeks, the percentage of purchases made above the listing price, the number of views per property by potential buyers, and inventory. Fig. 5 shows the effects of fraud on these metrics using ZIP code-month-level difference-in-differences regressions based on data from January 2019 to December 2024. For all four variables, the coefficients are close to zero in 2019. The effect on the percentage of listings off the market within two weeks begins increasing in June 2020, the percentage of houses purchased at higher than list price begins increasing in August 2020, the views per property begin increasing in July 2020, and inventory begins decreasing in July 2020. The effects on most of these variables peak between January 2021 and mid-2021. The estimated effects then return to their baseline February 2020 levels by around January 2022 for the percentage sold in two weeks, February 2022 for the percentage purchased at higher than list price, May 2022 for views per property, and January 2022 for inventory. Most of the variables show a reversal that persists throughout 2023.[47] Overall, the measures show a temporary surge in proxies for net demand, followed by a reversal for much of 2022–23.

We next examine longer-term price movements to test our earlier out-of-sample reversal prediction with more recent house price data. Panel B of Fig. 2 and Fig. 3 already show evidence of a price reversal graphically. To test this pattern more formally, we regress house price growth in 2022 and 2023 on pandemic fraud rates following the same regression framework we previously used for house price growth in 2020 and 2021. Table 8 shows that for a one standard deviation increase in flagged (composite) loans per capita, house price growth was 0.74 (1.09) ppt lower from January 2022 to December 2023. This reversal amounts to 35% (41%) of the initial effect of pandemic fraud on house price growth from January 2020 to December 2021 (Table 2).[48] A similar reversal is also found when local pandemic fraud rates are instrumented with the social proximity to fraud instrument (Table

[46] For instance, in a November 2022 draft, we wrote "Because pandemic fraud was a transitory and temporary demand shock, our findings point to an out-of-sample prediction that areas with excess fraud will experience more rapid home price depreciations, which will have further distortionary effects on COVID-era home buyers".

[47] The ratio of purchase price to listing price and a market demand index from Realtor.com also show consistent patterns (Figure IA.19).

[48] Table IA.36 finds a similar degree of reversal using the average fraud rate in a five-mile radius of the focal ZIP code. Rental price growth in high fraud areas did not experience a reversal (Table IA.26, Panel B).



(a) Percentage Off-the-Market Within Two Weeks

(b) Percentage Sold at Higher Than Listing Price

(c) Unique Viewers Per Property

(d) Inventory

Fig. 5. Effect on other housing market metrics. This figure shows the effect of a one standard deviation change in *Flagged Per Capita* on various housing market metrics. We estimate the following regression and plot β_t: $Metric_{i,t} = \sum_{t\neq Feb2020} \beta_t 1(Month = t) FlaggedPerCapita_i + ZIPCodeFE + CountyMonthFE + Month_t \times Controls_i + \epsilon_{i,t}$ where i is a ZIP code and t is a month. Panel A shows the effect on the percentage of properties off the market within two weeks, Panel B on the percentage of properties sold at higher than the listing price, Panel C on unique viewers per property (relative to a typical property across the US), and Panel D on inventory (the number of active listings). Panels A and B (C and D) are based on metrics from Redfin (Realtor.com). The dependent variable is the difference between the metric during the given month and its value in the same month of 2018 for 2019, and between the metric during the given month and its average value in 2018 and 2019 for 2020 onward. The controls are loans per capita using percentile fixed effects, log population density, vacancy rate, log housing units, and log average household income. The error bars correspond to 95% confidence intervals based on standard errors that are double clustered by ZIP code and month.

IA.37 and Figure IA.15). We also examine the effects of other drivers of house prices discussed in Section 5.5 on home prices from January 2022 to December 2023. Distance to the central business district, land unavailability, teleworkability, and 2018–19 housing price growth are related to housing price growth during 2022–23 in the same direction as in the 2020–21 period, though generally with smaller coefficients. Among all the factors examined, pandemic fraud is the only one whose effect on house prices consistently reverses direction between these two periods while remaining statistically significant (Table IA.38, Figure IA.18, and Figure IA.20).

Overall, the evidence suggests that the influx of fraudulent funds in certain areas created a short-term surge in demand that subsequently

Table 8
Effect on house price growth, 2022–23.

	House price growth from January 1, 2022 to December 31, 2023				
	(1)	(2)	(3)	(4)	(5)
Flagged Per Capita	−0.00606*** (−4.49)	−0.00744*** (−4.40)			
High Loan-to-Est. Per Capita			−0.0124*** (−8.13)		
High Similarity Per Capita				−0.00897*** (−6.09)	
Flagged Composite Per Capita					−0.0109*** (−6.02)
County FE	Yes	Yes	Yes	Yes	Yes
Past HP Growth	No	Yes	Yes	Yes	Yes
Loans Per Capita	No	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	Yes
Observations	18,761	18,761	18,761	18,761	18,761
Num. Counties	2215	2215	2215	2215	2215
R^2	0.810	0.823	0.824	0.823	0.823
Mean of Dep. Var.	0.120	0.120	0.120	0.120	0.120

This table examines the relationship between house price growth in 2022–23 and various measures of suspicious PPP lending. All variables, fixed effects, and controls are as defined in Table 2. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the population of the ZIP code in 2019. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

reversed. The surge in demand was accompanied by a rise in house prices, and the subsequent fall in demand led 35% of the initial effect on house prices to dissipate by December 2023.

6. Other spending and inflation effects

The stimulating effects of PPP fraud are not conceptually limited to the housing market. Anecdotal evidence suggests that recipients of fraudulent PPP loans also frequently purchased cars and luxury products.[49] In this section, we first examine the relationship between PPP fraud and vehicle purchases at the ZIP code level. Next, we examine the connection between PPP fraud and general consumer spending at the census tract level. Finally, we explore how PPP fraud relates to regional differences in inflation.

6.1. Vehicle purchases

To investigate vehicle purchases, we utilize monthly data from January 2018 to December 2022 on vehicle title registrations at the ZIP code level for five large states (California, Texas, Florida, Illinois, and Ohio) from Cross-Sell, supplemented by similar publicly available data for the state of Washington.[50] If fraudulent PPP loan proceeds were used, in part, to purchase vehicles, we would expect an increase in vehicle title registrations in high-fraud ZIP codes following the initiation of the PPP. To test this hypothesis, we estimate a regression of the form:

$$Log(VehicleRegistrations)_{it} = \sum_{t \neq Feb2020} \beta_t 1(Month = t) FlaggedPerCapita_i + ZIPCodeFE + CountyMonthFE + \epsilon_{it} \quad (6)$$

The dependent variable in the regression is the log of the number of vehicle title registrations in ZIP code i during month t. *Flagged Per Capita* is standardized such that one unit represents one standard deviation. The coefficients of interest are β_t, which estimate the effect on vehicle title registrations, relative to the February 2020 baseline, associated with a one standard deviation increase in *Flagged Per Capita*. The regressions include ZIP code fixed effects and county × month fixed effects to isolate differential changes within counties. Standard errors are double clustered by county and month.

The left plot in Panel A of Fig. 6 shows the effect of a one standard deviation increase in *Flagged Per Capita* on vehicle title registrations over time. During the pre-COVID period, the average effect is close to zero (−0.19%), although some months show a positive effect and others a negative one.[51] After the onset of COVID, a one standard deviation increase in *Flagged Per Capita* is associated with a 1.43% increase in vehicle title registrations over the March 2020 to December 2022 time period. This effect is largely concentrated from March 2020 to December 2021, during which a one standard deviation increase in *Flagged Per Capita* is associated with a 2.08% increase in vehicle title registrations. In contrast, during the January 2022 to December 2022 time period, a one standard deviation increase in *Flagged Per Capita* is associated with a 0.23% increase in vehicle title registrations. This pattern aligns with the expected short-term stimulus to vehicle purchases during the PPP in 2020 and 2021.[52]

The right plot in Panel B of Fig. 6 shows a binned scatter plot across ZIP codes of the percentage change in vehicle title registrations versus *Flagged Per Capita*, controlling for county fixed effects. The vertical axis represents the percentage change in the number of vehicle title registrations from the 2018–2019 period to the 2020–2021 period. Consistent with the regression results in the left panel, there is a positive relationship between the percentage change in vehicle title registrations and *Flagged Per Capita*.[53]

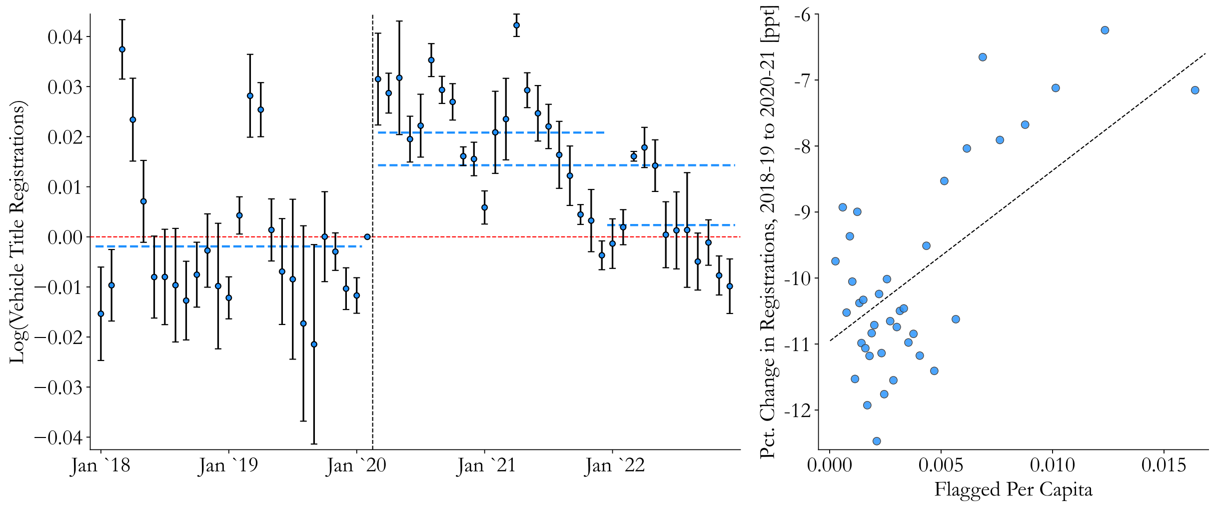
[49] For example, here and here.

[50] These six states collectively represent 36.6% of the U.S. population.

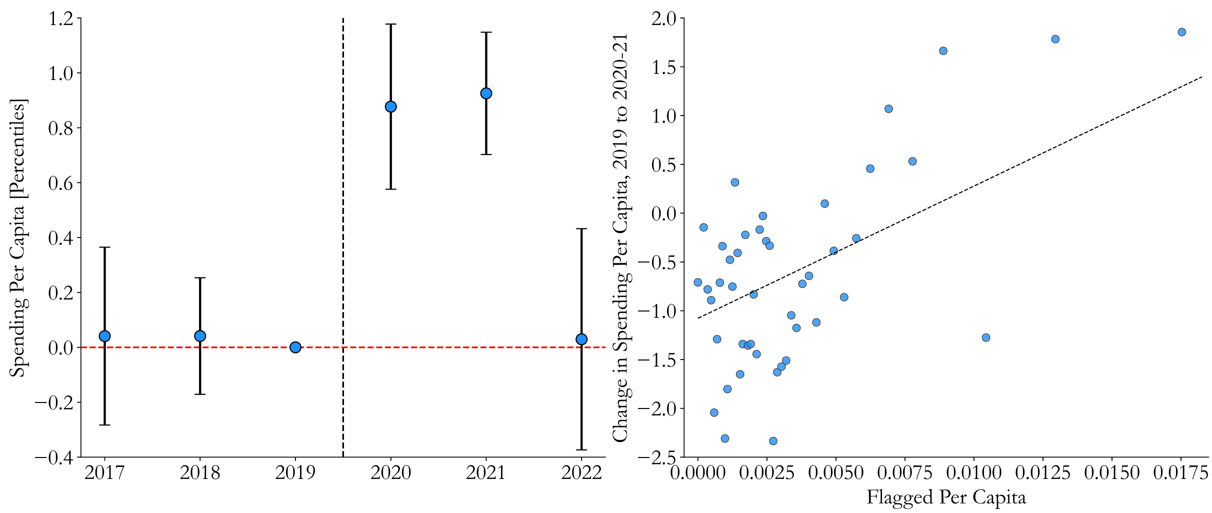
[51] We interpret the coefficients as percentage changes based on the log approximation.

[52] Figure IA.21, Panel A shows the relative vehicle title registration rates across terciles of *Flagged Per Capita*.

[53] Note that vehicle title registrations declined by an average of 9.8% between these periods. This relationship is highly statistically significant with a t-statistic of 6.35 based on standard errors clustered by county. Figure IA.21, Panel B shows similar results for percentage change in registrations from 2018–2019 to 2020 alone, as well as a longer post-COVID time period including all of 2020–2022.



(a) Vehicle Title Registrations From 6 States



(b) Consumer Spending

Fig. 6. Effect on Vehicle Purchases and Consumer Spending. This figure shows the effect of suspicious lending on vehicle purchases and consumer spending. Panel A uses vehicle title registration data from six states (California, Texas, Florida, Illinois, Ohio, and Washington) at the ZIP code-month level. Panel B uses annual data from 2017 to 2022 at the census tract level from Mastercard's Center for Inclusive Growth. Mastercard ranks each census tract's consumer spending per capita each year in the national distribution and releases the percentile rank of the tract. Data from 2017 to 2022 is used. The left subpanels examine the effects of a one standard deviation change in the number of flagged PPP loans per capita. The left subpanel of Panel A includes ZIP code and month × county fixed effects, and the left subpanel of Panel B includes census tract and year × county fixed effects. The error bars in the left subpanels represent 95% confidence intervals based on standard errors that are double clustered by county and month in Panel A and clustered by county in Panel B. In the left subpanel of Panel A, the horizontal blue lines represent the average coefficient over the period spanned by each line. The right subpanels examine the within-county effects of flagged PPP loans per capita on the percentage change in vehicle registrations between 2018–19 and 2020–21 in Panel A and the percentile change in spending per capita between 2019 and 2020–21 in Panel B. To have a nationally representative estimate, both panels use weighted least squares (WLS) regressions with the weight being the ZIP code's (census tract's) population as of 2019 in Panel A (Panel B). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 9 collapses the β_t coefficients in Eq. (6) into a single coefficient for the interaction between *Flagged Per Capita* and an indicator variable for the time period beginning in March 2020. Columns (1) to (3) use data from January 2018 to December 2021, with *Post* defined as 1 for March 2020 to December 2021 and 0 otherwise. Standard errors are again clustered by county and month, and ZIP code and county × month fixed effects are included. Column (1) estimates a regression similar to the regressions plotted in Fig. 6, Panel A. A one standard deviation increase in *Flagged Per Capita* is associated with a 2.29% increase in registrations. In column (2), we add a post-period indicator variable interacted with total PPP loans per capita. This specification distinguishes the stimulative effect of fraudulent PPP loans from any stimulative effect of PPP loans more generally. The coefficient on *Flagged Per Capita* increases to 3.10% and remains highly statistically significant. A potential concern is that PPP fraud could be correlated with other characteristics that predict vehicle title registration growth during this period. Column (3) addresses this concern by adding detailed demographic data interacted with post-period indicator variables. While adding these demographic variables somewhat decreases the *Flagged Per Capita* coefficient (from 3.10% to 2.39%), the result remains substantial and highly statistically significant (with a t-statistic of 9.94). Columns (4) to (6) replicate columns (1) to (3) using data from January 2018 to December 2022 and with *Post* defined as 1 for March 2020 to December 2022 and 0 otherwise. The results are similar but with smaller magnitudes, as expected based on Panel A of Fig. 6. Figure IA.21, Panel C examines heterogeneity in these

Table 9
Effect on vehicle purchases.

	Log(Vehicle title registrations)					
	March 2020 to December 2021			March 2020 to December 2022		
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flagged Per Capita	0.0229*** (5.42)	0.0310*** (3.76)	0.0239*** (9.94)	0.0163*** (4.32)	0.0201*** (2.98)	0.0163*** (9.56)
Post × Loans Per Capita		−0.0150** (−2.23)	−0.00587** (−2.36)		−0.00708 (−1.24)	−0.00136 (−0.56)
Post × Median income			0.0203** (2.36)			0.0275*** (3.40)
Post × Poverty			0.00267 (0.37)			0.00295 (0.46)
Post × Population density			0.0173*** (4.12)			0.0146*** (3.65)
Post × Pct. Non-White			−0.00786 (−1.57)			−0.00156 (−0.36)
Post × Educational attainment			−0.0285** (−2.47)			−0.0233** (−2.20)
Post × Pre-Pandemic unemployment			0.0104** (2.58)			0.00599 (1.67)
ZIP Code FE	Yes	Yes	Yes	Yes	Yes	Yes
Month × County FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	278,400	278,400	278,400	348,000	348,000	348,000
R^2	0.981	0.981	0.981	0.979	0.979	0.979

This table examines the effect of suspicious lending on vehicle purchases using vehicle title registration data from six states (California, Texas, Florida, Illinois, Ohio, and Washington) at the ZIP code-month level. *Post* is a dummy variable that takes a value of 1 from March 2020 to December 2021 (2022) in columns (1) to (3) (columns (4) to (6)) and a value of 0 from January 2018 to February 2020 in all columns. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the ZIP code's population as of 2019. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are double clustered by county and month are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

effects across demographic splits and finds that the effect is present and generally highly statistically significant across all splits.

To examine the effect of PPP fraud on vehicle purchases nationwide, we also analyze census tract-level data on vehicles per household from the American Community Survey (ACS).[54] During the pre-COVID years, no relation exists between PPP fraud and vehicles per household. However, a one standard deviation increase in *Flagged Per Capita* is associated with a 0.29% increase in vehicles per household during 2020, a 0.43% increase during 2021, a 0.51% increase during 2022, and a 0.66% increase during 2023 (see Figure IA.22).[55]

[54] The ACS is an annual survey conducted by the U.S. Census. We use annual data from 2015 to 2023 on the number of vehicles per household over the preceding five years by census tract. There are over three times as many census tracts as ZIP codes. We estimate a regression similar to Eq. (6) using this data, but with the dependent variable being the log of vehicles per household and including census tract and county × year fixed effects.

[55] ACS data at the census tract level is only released as five-year estimates. For example, the 2023 estimate is based on data collected during 2019 to 2023, which will cause any effects observed to be understated. Assuming there is no relation between PPP fraud and vehicles per household before 2020, the effects observed in 2020, 2021, 2022, and 2023 are one-fifth, two-fifths, three-fifths, and four-fifths of the true effect, respectively. The right plot in Figure IA.22 shows the within-county relationship between *Flagged Per Capita* and the percentage change in vehicles per household from 2019 to 2023. Consistent with the results in the left panel, there is a strong positive relation (t-stat=13.31). We also find evidence of increased auto debt in MSAs with high PPP fraud (Figure IA.23). The pattern is consistent with PPP fraud recipients initially paying down their debts and then eventually increasing their auto debts as they used PPP funds for down payments on vehicle purchases.

6.2. Consumer spending

Next, we examine the effects of PPP fraud on consumer spending more broadly. Consumer spending data at the census tract level is from Mastercard's Center for Inclusive Growth and is based on anonymized and aggregated transactions on the Mastercard network. For further privacy, Mastercard ranks each census tract's consumer spending per capita each year in the national distribution and releases only the percentile rank of the tract for each year. If fraudulent PPP loans stimulated consumer spending, we would expect elevated spending in census tracts with higher PPP fraud per capita. To examine this, we estimated regressions of the form:

$$SpendingPerCapita_{it} = \sum_{t\neq 2019} \beta_t 1(Year = t) FlaggedPerCapita_i + TractFE + CountyYearFE + \epsilon_{it} \tag{7}$$

The dependent variable in the regression is the tract's percentile rank of spending per capita. *Flagged Per Capita* is standardized so that one unit represents one standard deviation. The coefficients of interest are β_t, which estimate differences in percentile rank of spending per capita relative to the 2019 baseline that are associated with a one standard deviation increase in *Flagged Per Capita*.[56]

[56] The regressions include census tract fixed effects and county × year fixed effects to isolate differential changes within counties. In subsequent analysis, we add interactions with demographic characteristics and consider heterogeneous effects across census tracts with different demographic characteristics. Standard errors are clustered by county. Double clustering by county

Table 10
Effect on consumer spending.

	Spending Per Capita percentile rank			
	(1)	(2)	(3)	(4)
Post × Flagged Per Capita	0.874***	1.039***	0.541***	0.738***
	(8.59)	(10.05)	(4.08)	(7.03)
Post × Loans Per Capita		−0.232***	0.0569	−0.103
		(−2.92)	(0.52)	(−1.50)
Post × Median income			−0.0667	−0.525***
			(−0.34)	(−3.02)
Post × Poverty			0.0995	0.168
			(0.72)	(1.23)
Post × Population density			0.566*	0.758***
			(1.92)	(5.71)
Post × Pct. Non-White			0.523***	0.0882
			(3.12)	(0.64)
Post × Educational attainment			−1.831**	−1.593**
			(−1.99)	(−1.99)
Post × Pre-Pandemic unemployment			0.185*	0.452***
			(1.72)	(4.39)
Census tract FE	Yes	Yes	Yes	Yes
Year × County FE	Yes	Yes	Yes	No
Year FE	No	No	No	Yes
Observations	307,385	307,385	307,385	307,385
R^2	0.348	0.348	0.348	0.305

This table examines the effect of suspicious lending on consumer spending using annual data at the census tract level from Mastercard's Center for Inclusive Growth. Mastercard ranks each census tract's consumer spending per capita each year in the national distribution and only releases the percentile rank of the tract for each year. Data from 2017 to 2021 is used. *Post* is a dummy variable that takes a value of 1 if the year is 2020 or 2021 and 0 otherwise. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the census tract's population as of 2019. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are clustered by county are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

The left plot in Panel B of Fig. 6 shows the effect of a one standard deviation increase in *Flagged Per Capita* over time. During the pre-COVID period, *Flagged Per Capita* has no relationship to spending per capita. After the onset of the PPP, a one standard deviation increase in *Flagged Per Capita* is associated with a 0.88 and 0.93 percentile rank increase in spending per capita in 2020 and 2021, respectively, compared to 2019. Spending levels then normalize in 2022. This aligns with the expected short-term stimulus during the PPP in 2020 and 2021.[57]

The right plot in Panel B of Fig. 6 shows a binned scatter plot across census tracts of the change in spending per capita percentiles versus *Flagged Per Capita*, controlling for county fixed effects. The vertical axis plots the change in consumer spending per capita percentiles from 2019 to the average of 2020 and 2021. Consistent with the results in the left panel, there is a positive relationship between spending growth and *Flagged Per Capita*.[58]

Table 10 collapses the β_t coefficients in Eq. (7) into a single coefficient for the interaction between *Flagged Per Capita* and an indicator variable for the PPP years (2020 and 2021). The sample begins in 2017 and ends in 2021 to compare PPP years with preceding years. Column (1) estimates a regression similar to the regressions plotted in Panel B of Fig. 6. A one standard deviation increase in *Flagged Per Capita* is associated with a 0.874 percentile rank increase in 2020 and 2021 consumer spending relative to 2019, which is nearly identical to the effects estimated in Panel B of Fig. 6 for 2020 and 2021. In column (2), we add a post-period indicator variable interacted with total PPP loans per capita. This specification distinguishes the stimulative effect of fraudulent PPP loans from any stimulative effect of PPP loans more generally. The coefficient on *Flagged Per Capita* increases to 1.039 and remains highly statistically significant. A potential concern is that PPP fraud could be correlated with other characteristics that predict spending growth during this time period. Column (3) addresses this concern by adding detailed demographic data interacted with the post-period indicator variable. While adding these control variables somewhat decreases the *Flagged Per Capita* coefficient (from 1.039 to 0.541), the result remains substantial and highly statistically significant (with a t-statistic of 4.08). Column (4) replaces the year × county fixed effect with a year fixed effect and shows that the effect of PPP fraud on consumer spending is even larger in this case, indicating that PPP fraud also predicts consumer spending differences across counties.

To get a more granular view of consumer spending, we also examine consumer mobility data from SafeGraph, which uses cellphone GPS pings to track visits by consumers to various types of locations.[59] The SafeGraph data allows us to examine weekly visits from 2018 to 2021 by residents of a given census tract to specific types of locations—ranging from auto dealerships to financial institutions. To examine the effects of pandemic fraud on consumer mobility, and by extension consumer spending, we estimate regressions similar to Eq. (7) but with the dependent variable being visits per tracked device to a given type of location during each week.[60] Fig. 7 presents the results, with each

and year is not feasible due to having only six years of data. Purely cross-sectional results based on changes in spending (discussed below) are also highly significant.

[57] The effect of PPP fraud on consumer spending in 2020 and 2021 is similar for tracts with above and below median values of various demographics, which provides reassurance as to the consistency of the effect (Figure IA.24).

[58] This relationship has a t-statistic of 8.16 based on standard errors clustered by county.

[59] This dataset has been used as a proxy for consumer spending previously (e.g., Bizjak et al., 2022; Gurun et al., 2023; Noh et al., 2025).

[60] Specifically, the dependent variable is the $IHS(Visits)_{itl} - IHS(Devices)_{it}$ where $IHS(.)$ is the inverse hyperbolic sine function (to allow for zeros in

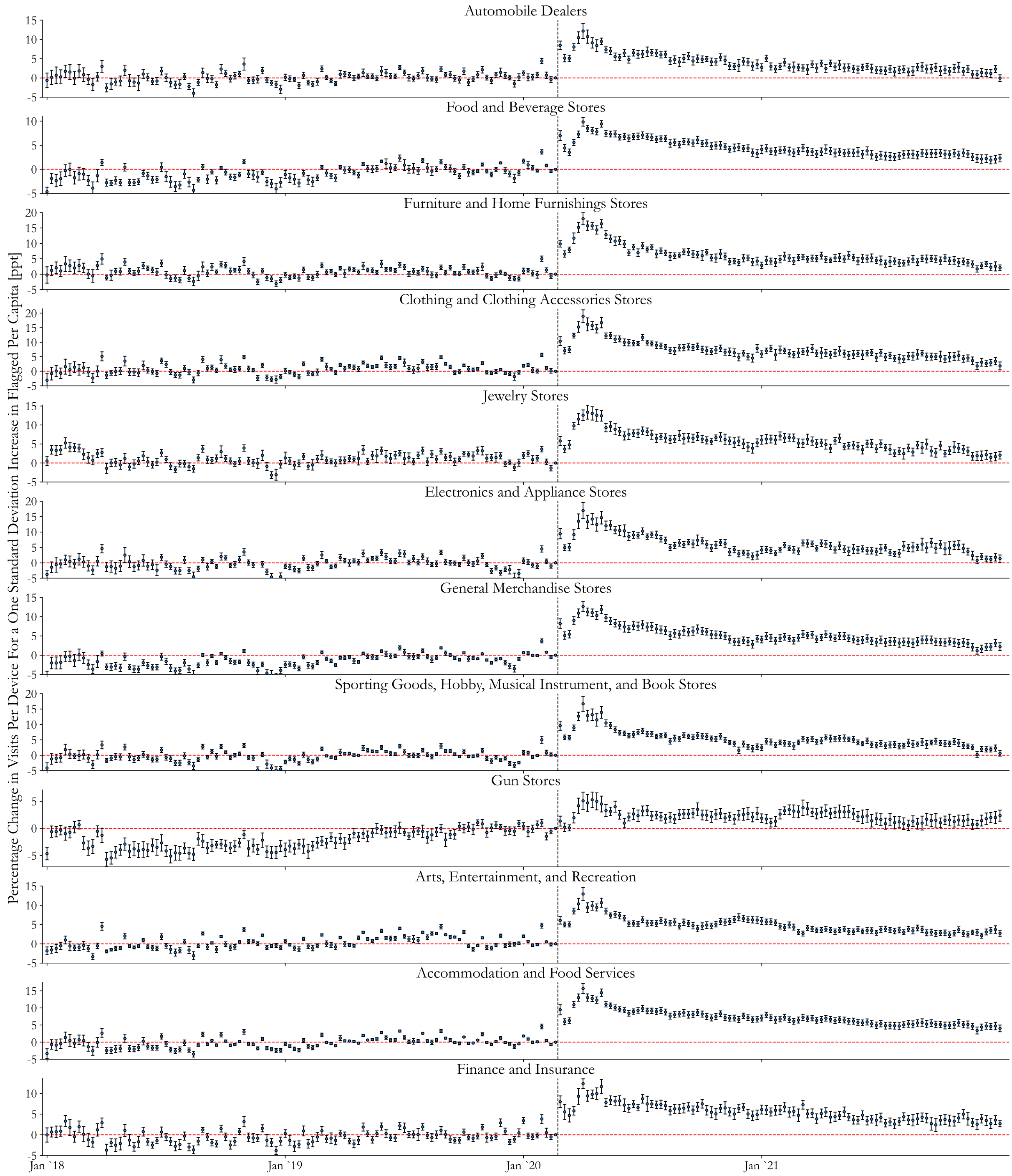


Fig. 7. Effect on consumer visits to various types of locations. This figure shows the effect of suspicious lending on consumer visits to various types of locations. Data on consumer visits is from SafeGraph and has been aggregated to the census tract × week × location type level. Location types are defined by NAICS codes; gun stores are identified using a combination of NAICS codes and a keyword search of location names. Each subpanel examines the effects of a one standard deviation change in the number of flagged PPP loans per capita on visits to the given type of location. The dependent variable is $IHS(Visits)_{itl} - IHS(Devices)_{it}$, where $IHS(.)$ is the inverse hyperbolic sine function, i is a census tract, t is a week, and l is a type of location. Note that the coefficients can be interpreted as a percentage change by the same logic as when the dependent variable is log transformed. The regressions include census tract fixed effects, week × county fixed effects, and control for post × loans per capita. Both the independent and dependent variables are winsorized at the 1% tails. The week of February 17th to 24th, 2020 is the omitted week. The error bars represent 95% confidence intervals based on standard errors that are double clustered by county and week. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the census tract's population as of 2019.

subpanel showing the effect of a one standard deviation increase in *Flagged Per Capita* on visits to a given type of location. During the pre-COVID period, *Flagged Per Capita* does not appear to have any systematic relation with consumer visits to most types of locations.[61] After the onset of the PPP, and other forms of stimulus more broadly, there is a meaningful and systematic relationship between *Flagged Per Capita* and visits across all the types of locations examined. The effect of *Flagged Per Capita* on visits declines over time but remains meaningfully elevated compared to prepandemic levels throughout 2021.[62] These results suggest that the effect of pandemic fraud on consumer spending was wide-ranging across a variety of different spending categories.

The results in Panel B of Fig. 6, Fig. 7, and Table 10 consistently point to elevated consumer spending and elevated visits to retail locations by residents of census tracts with higher levels of PPP fraud. As with the previous analysis of the housing and vehicle markets, PPP fraud is not randomly assigned, meaning we do not have a perfect shock for causal interpretation. Nevertheless, elevated spending and retail visits in 2020 and 2021, with a return to normal in 2022, aligns with the expected behavior of a short-term stimulus like PPP fraud and is consistent with the effects observed in the housing market and vehicle markets.

6.3. Inflation

Finally, we examine the effects of PPP fraud on regional inflation. The Bureau of Labor Statistics (BLS) releases bi-monthly 12-month regional consumer price indices (CPI) by Core Based Statistical Area (CBSA), though only for 23 CBSAs. Within this limited dataset, we examine how PPP fraud may have affected overall price levels at the CBSA level by estimating regressions of the form:

$$12\text{-}monthInflation_{it} = \sum_{t \neq JanFeb2020} \beta_t 1(Bi\text{-}month = t) FlaggedPerCapita_i + CBSAFE + Bi\text{-}monthFE + \epsilon_{it} \quad (8)$$

The dependent variable in the regression is the CBSA's 12-month inflation, calculated on a bi-monthly basis using regional CPI data from the BLS. The coefficients of interest are β_t, which estimate differences in inflation rates relative to the January/February 2020 baseline that are associated with a one standard deviation increase in *Flagged Per Capita*.[63]

Fig. 8 shows the results. During the pre-COVID period from 2010 to 2019, regional inflation rates had no relationship to *Flagged Per Capita*. There is also little relationship between PPP fraud and inflation growth from March to August 2020. Inflation then begins to accelerate in CBSAs with high PPP fraud in September/October 2020, with statistically significant differences by January/February 2021, and larger and significant effects throughout 2021 and most of 2022, peaking in July/August 2022. The effects are positive but smaller and insignificant from November/December 2022 to November/December 2023. These patterns correspond closely to overall inflation, which first increased above 2% in March 2021 and remained elevated through mid-2023, with peak annualized inflation of 9.06% in June 2022. We also separately consider the housing and non-housing components of CPI and find evidence of increased housing inflation and a smaller impact on non-housing inflation (see Figure IA.26, Panel A).[64]

Table 11 collapses the β_t coefficients in Eq. (8) into a single coefficient for the interaction between *Flagged Per Capita* and an indicator variable for the post-PPP time period, starting in March/April 2020. Column (1) reports results for overall inflation. Consistent with Fig. 8, a one standard deviation increase in *Flagged Per Capita* is associated with 0.43 ppt higher 12-month inflation during the post-PPP time period. This is a substantial increase and is highly statistically significant, with a t-statistic of 3.40.[65] In column (2), we add the interaction between a post-period indicator variable and overall PPP loans per capita to assess whether inflation is coming from PPP loans in general as opposed to fraudulent PPP loans. The coefficient for overall PPP loans per capita is close to zero, and the coefficient on *Flagged Per Capita* increases slightly after controlling for overall PPP loans per capita.

Column (3) of Table 11 repeats the same regressions for housing inflation and shows an even larger effect of 0.71 ppt. Controlling for overall PPP loans per capita in column (4) again slightly increases the result for flagged PPP loans. Columns (5) and (6) focus on non-housing inflation and show an insignificant effect in column (5) and a significant though smaller effect of 0.25 ppt in column (6).

Overall, the results in Fig. 8 and Table 11 indicate a strong relation between PPP fraud and inflation, even with relatively limited data for only 23 CBSAs. This inflation is primarily concentrated in housing but is apparent to a smaller extent in non-housing CPI components. Combined with the analyses of vehicle purchases and consumer spending in the previous subsections, this suggests that PPP fraud broadly stimulated consumer spending beyond the housing market. Differential price pressure across CBSAs is likely most pronounced for housing because it is an immovable good with distinct local markets.

7. Conclusion

The U.S. government responded to the COVID-19 pandemic with massive relief spending and minimal fraud safeguards. The result was hundreds of billions of dollars in pandemic relief fraud, much of which flowed to highly concentrated geographic areas. Our analysis highlights

---

$Visits_{itl}$), i is a census tract, t is a week, and l is a type of location. Normalizing by the number of tracked devices is important since the SafeGraph panel of tracked devices changes over time. The coefficients can be interpreted as percentage changes. The regressions include census tract fixed effects, week × county fixed effects, and control for post × loans per capita. Standard errors are double clustered by county and week. Both the independent and dependent variables are winsorized at the 1% tails. The week of February 17th to 24th, 2020 is the omitted period. Location types are defined by NAICS codes; gun stores are identified using a combination of NAICS and a keyword search of location names.

[61] The pre-trends for gun stores are meaningful and appear to be systematic; this may be due to only being able to track visits to 5634 gun stores.

[62] Because our SafeGraph data only extends to the end of 2021, we are unable to assess impacts on visits further in the future. Figure IA.25 examines heterogeneity in the effect of fraud on consumer foot traffic by demographics and finds broadly similar results across demographic splits. Table IA.39 collapses the β_t coefficients shown in Fig. 7 into a single coefficient for the interaction between *Flagged Per Capita* and an indicator variable for post-period and includes variations that control for week × demographics and variations using data for the entire 2018 to 2021 period or excluding 2021.

[63] 12-month inflation is determined as the current CPI divided by the CPI 12 months earlier. We consider inflation from January 2010 to December 2023. *Flagged Per Capita* is standardized so that one unit represents one standard deviation. Standard errors are double clustered by CBSA and bi-month.

[64] In Section 6.1, we find that recipients of fraudulent PPP loans were more likely to purchase vehicles from March 2020 to December 2021. Given that automobiles are a movable good and the automobile index depends on the composition of purchases (which was affected by supply issues in late 2020 and 2021), it is not clear if any effect would be detectable in regional car prices. In Panel B of Figure IA.26, we find that the vehicle component of regional CPI is somewhat elevated in high PPP fraud CBSAs in mid-2020 before returning to normal by the end of 2020.

[65] Figure IA.27 plots the CBSA-level data with particularly high fraud and elevated inflation in Atlanta. Regression results are robust to excluding any single CBSA. Excluding Atlanta results in a coefficient of 0.35 ppt with a t-statistic of 2.74. The t-statistics reported in parentheses are based on standard errors that are double clustered by CBSA and bi-month. Additionally, since the dependent variable is based on overlapping periods, we report t-statistics based on Newey–West standard errors with 6 lags in square brackets.

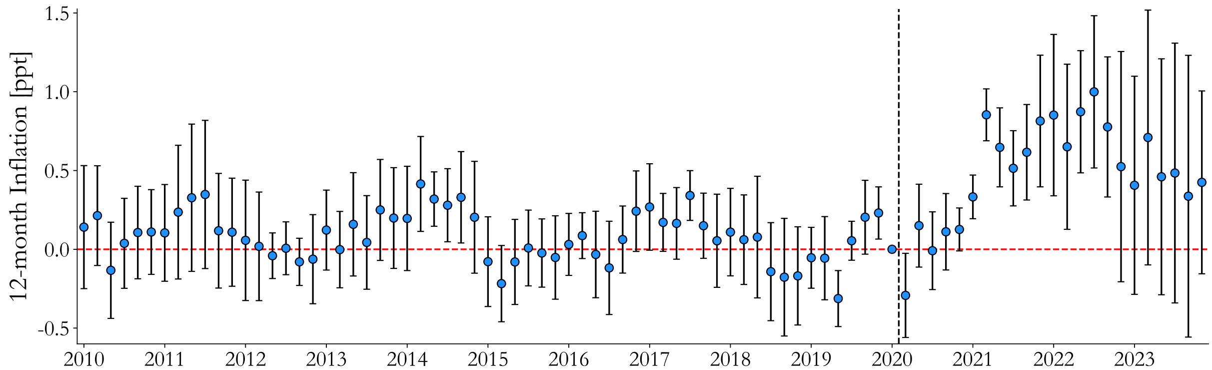


Fig. 8. Effect on Regional Inflation. This figure shows the effect of suspicious lending on regional inflation using regional all-items consumer price indices (CPI) from the BLS. Data for 23 CBSAs is released bi-monthly. 12-month inflation is determined by dividing the given month's CPI by the CPI 12 months earlier. *Flagged Per Capita* is standardized, so the coefficients represent the inflation effect of a one standard deviation change in suspicious lending. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the CBSA's population as of 2019. The error bars represent 95% confidence intervals based on standard errors that are double clustered by CBSA and bi-month.

Table 11
Effect on regional inflation.

	12-month inflation					
	Overall		Housing component		Excluding shelter	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Flagged Per Capita	0.00429*** (3.40) [3.45]	0.00613*** (3.13) [3.27]	0.00707*** (4.10) [3.82]	0.00964*** (3.69) [3.37]	0.00139 (1.48) [1.46]	0.00252** (2.28) [1.68]
Post × Loans Per Capita		−0.00259* (−1.91) [−1.49]		−0.00363 (−1.21) [−0.98]		−0.00159 (−1.43) [−1.15]
CBSA FE	Yes	Yes	Yes	Yes	Yes	Yes
Bimonthly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1317	1317	1317	1317	1317	1317
Num. CBSAs	23	23	23	23	23	23
R^2	0.895	0.896	0.749	0.751	0.924	0.924
Mean of Dep. Var.	0.0254	0.0254	0.0271	0.0271	0.0236	0.0236

This table examines the effect of suspicious lending on regional inflation using regional consumer price indices (CPI) from the BLS. Data for 23 CBSAs is released bi-monthly. 12-month inflation is determined by dividing the given month's CPI by the CPI 12 months earlier. Data from January 2010 to December 2022 is used. *Post* is a dummy variable that takes a value of 1 if the bi-month is on or after March 2020 and zero otherwise. Columns (1) and (2) are based on the all-items regional CPI, columns (3) and (4) are based on the housing component of the regional CPI, and columns (5) and (6) are based on the all-items-excluding-shelter regional CPI. To have a nationally representative estimate, we use weighted least squares (WLS) regressions with the weight being the CBSA's population as of 2019. All independent variables are standardized to have a mean of 0 and a standard deviation of 1. Fixed effects are as indicated at the bottom of each column. t-statistics based on robust standard errors that are double clustered by CBSA and bi-month are reported in parentheses. t-statistics based on Newey–West standard errors with 6 lags are reported in square brackets. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

that recipients of fraudulent funds were 17% more likely to purchase houses compared to non-flagged PPP recipients and 22% more likely to move compared to similar households that did not receive PPP funds in the eighteen months after receiving their PPP loan. Consequently, high-fraud ZIP codes experienced an increase in net demand, which resulted in a 5.8 percentage point larger increase in house price growth, indicating a sizable distortionary effect on home prices. This house price distortion is apparent in monthly data beginning in May 2020 and continues through June 2022. Synthetic control analysis shows that ZIP codes with high PPP fraud experienced divergent growth starting in mid-2020 compared to low-fraud ZIP codes within the same county that had similar house price trends pre-COVID. When compared to other variables examined in the literature, flagged loans per capita and land unavailability emerge as the economically largest home price predictors, although remote work, teleworking, migration, and past price growth are also statistically significant indicators. After the effect of fraudulent spending wears off, distortions to the housing market recede, and house price growth in high-fraud areas begins to reverse by mid-2022, continuing through at least the end of 2023. Additionally, vehicle purchases, consumer mobility patterns, and consumer spending are elevated in areas with high PPP fraud during 2020 and 2021, and pandemic fraud appears to have had a meaningful effect on overall inflation.

Our findings support the idea that unintended externalities of fraud, including distorted asset prices, can be broader and more costly than the direct costs of fraud (Akerlof and Romer, 1993). Our findings also indicate that rent-seeking in the financial system (Zingales, 2015) can have large spillovers into real markets. Given that our findings show that fraudulent transfers can be wealth shocks that generate economic distortions not created by normal transfers, future government program designs should take more proactive steps to prevent fraud on the front-end. Furthermore, since pandemic fraud appears to have overwhelmed government prosecutors, lawmakers and authorities should allocate more targeted resources for back-end auditing and prosecution. Additional research could explore other externalities of fraud, such as potentially encouraging future criminal behavior.

CRediT authorship contribution statement

John M. Griffin: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Samuel Kruger:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Prateek Mahajan:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Griffin is an owner/CEO of Integra Research Group LLC and Integra FEC LLC, both of which engage in litigation consulting on a variety of issues, including fraud discovery and recovery. Integra Research Group LLC provided research support and computational resources for this research project.

Kruger and Mahajan have not received financial support from any interested party while working on this research project, and neither is currently engaged in consulting or employment outside of the University of Texas. Mahajan was employed by Integra FEC LLC from May 2017 to August 2020, prior to the commencement of this research project.

References

Abadie, A., Diamond, A., Hainmueller, J., 2014. Comparative politics and the synthetic control method. Am. J. Political Sci. 59, 495–510.

Agarwal, S., Liu, C., Souleles, N.S., 2007. The reaction of consumer spending and debt to tax rebates—Evidence from consumer credit data. J. Political Econ. 115 (6), 986–1019.

Aiello, D., Baker, S.R., Balyuk, T., Di Maggio, M., Johnson, M.J., Kotter, J.D., 2025. The Effects of Cryptocurrency Wealth on Household Consumption and Investment. Working Paper, Brigham Young University.

Akerlof, G.A., Romer, P.M., 1993. Looting: The economic underworld of bankruptcy for profit. Brookings Pap. Econ. Act. 1993, 1–73.

Autor, D., Cho, D., Crane, L., Goldar, M., Lutz, B., Montes, J., Peterman, W., Ratner, D., Villar, D., Yildirmaz, A., 2022. The $800 billion Paycheck Protection Program: Where did the money go and why did it go there? J. Econ. Perspect. 36 (2), 55–80.

Aydin, D., 2022. Consumption response to credit expansions: Evidence from experimental assignment of 45,307 credit lines. Am. Econ. Rev. 112 (1), 1–40.

Badarinza, C., Ramadorai, T., 2018. Home away from home? Foreign demand and London house prices. J. Financ. Econ. 130 (3), 532–555.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., 2018a. The economic effects of social networks: Evidence from the housing market. J. Political Econ. 126 (6), 2224–2276.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., Wong, A., 2018b. Social connectedness: Measurement, determinants, and effects. J. Econ. Perspect. 32 (3), 259–280.

Bailey, M., Dávila, E., Kuchler, T., Stroebel, J., 2018c. House price beliefs and mortgage leverage choice. Rev. Econ. Stud. 86 (6), 2403–2452.

Bailey, M., Farrell, P., Kuchler, T., Stroebel, J., 2020. Social connectedness in urban areas. J. Urban Econ. 118, 103264.

Baker, S.R., 2018. Debt and the response to household income shocks: Validation and application of linked financial account data. J. Political Econ. 126 (4), 1504–1557.

Baker, S.R., Farrokhnia, R.A., Meyer, S., Pagel, M., Yannelis, C., 2023. Income, liquidity, and the consumption response to the 2020 economic stimulus payments. Rev. Financ. 27 (6), 2271–2304.

Bartik, A.W., Bertrand, M., Cullen, Z., Glaeser, E.L., Luca, M., Stanton, C., 2020. The impact of COVID-19 on small business outcomes and expectations. Proc. Acad. Sci. 117 (30), 17656–17666.

Baum-Snow, N., Han, L., 2024. The microgeography of housing supply. J. Political Econ. 132 (6), 1897–1946.

Becker, G.S., 1968. Crime and punishment: An economic approach. J. Political Econ. 76 (2), 169–217.

Beraja, M., Zorzi, N., 2026. Durables and the marginal propensity to spend. Am. Econ. Rev. Forthcoming.

Bizjak, J., Kalpathy, S., Mihov, V., Ren, J., 2022. CEO political leanings and store-level economic activity during the COVID-19 crisis: Effects on shareholder value and public health. J. Financ. 77 (5), 2949–2986.

Cherry, S., Jiang, E., Matvos, G., Piskorski, T., Seru, A., 2021. Government and private household debt relief during COVID-19. Brookings Pap. Econ. Act. Fall Edition, 141–199.

Chetty, R., Friedman, J.N., Stepner, M., Opportunity Insights Team, 2023. The economic impacts of COVID-19: Evidence from a new public database built using private sector data. Q. J. Econ. 139 (2), 829–889.

Cho, Y., Morley, J., Singh, A., 2024. Did marginal propensities to consume change with the housing boom and bust? J. Appl. Econometrics 39 (1), 174–199.

Cole, A., 2024. The Impact of the Paycheck Protection Program on (Really) Small Businesses. Working Paper, Arizona State University.

Cong, L.W., Rabetti, D., 2023. Firm Disclosure Under Relationship Lending: Theory and Evidence from Bailout Loans. Working Paper, Cornell University.

Dalton, M., 2023. Putting the Paycheck Protection Program into perspective: An analysis using administrative and survey data. Natl. Tax J. 76 (2), 393–437.

Davis, M.A., Ghent, A.C., Gregory, J., 2024. The work-from-home technology boon and its consequences. Rev. Econ. Stud. 91 (6), 3362–3401.

de Soyres, F., Santacreu, A.M., Young, H.L., 2023. Demand-supply imbalance during the COVID-19 pandemic: The role of fiscal policy. Fed. Reserv. Bank St. Louis Rev. First Quarter 2023, 21–50.

Denes, M., Lagaras, S., Tsoutsoura, M., 2021. First Come, First Served: The Timing of Government Support and Its Impact on Firms. Working Paper, Carnegie Mellon University.

Diamond, W., Landvoigt, T., Sánchez, G.S., 2025. Printing away the mortgages: Fiscal inflation and the post-covid boom. J. Financ. Econ. 171, 104072.

Diamond, R., McQuade, T., Qian, F., 2019. The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco. Am. Econ. Rev. 109 (9), 3365–3394.

Dilanian, K., Strickler, L., 2022. "Biggest fraud in a generation": The looting of the Covid relief plan known as PPP. https://www.nbcnews.com/politics/justice-department/biggest-fraud-generation-looting-covid-relief-program-known-ppp-n1279664.

Dingel, J.I., Neiman, B., 2020. How many jobs can be done at home? J. Public Econ. 189, 104235.

Faulkender, M.W., Jackman, R., Miran, S., 2021. The Job Preservation Effects of Paycheck Protection Program Loans. Working Paper, University of Maryland at College Park.

Favilukis, J., Kohn, D., Ludvigson, S.C., Van Nieuwerburgh, S., 2012. International Capital Flows and House Prices: Theory and Evidence. National Bureau of Economic Research, pp. 235–299.

Fernández, C., Ley, E., Steel, M.F., 2001. Benchmark priors for Bayesian model averaging. J. Econometrics 100, 381–427.

Friedman, M., 1957. The Permanent Income Hypothesis. Princeton University Press, pp. 20–37.

Fuster, A., Hizmo, A., Lambie-Hanson, L., Vickery, J., Wilen, P., 2025. How resilient is mortgage credit supply? Evidence from the COVID-19 pandemic. J. Finance Forthcoming.

Fuster, A., Kaplan, G., Zafar, B., 2020. What would you do with $500 Spending responses to gains, losses, news, and loans. Rev. Econ. Stud. 88 (4), 1760–1795.

Gamber, W., Graham, J., Yadav, A., 2023. Stuck at home: Housing demand during the COVID-19 pandemic. J. Hous. Econ. 59, 101908.

Gee, J., Button, M., 2019. The Financial Cost of Fraud 2019: The Latest Data from Around the World. Crowe UK, United Kingdom.

Gorback, C., Keys, B., 2025. Global capital and local assets: House prices, quantities, and elasticities. Rev. Financial Stud. Forthcoming.

Granja, J., Makridis, C., Yannelis, C., Zwick, E., 2022. Did the paycheck protection program hit the target? J. Financ. Econ. 145 (3), 725–761.

Griffin, J.M., Kruger, S., Mahajan, P., 2023. Did FinTech lenders facilitate PPP fraud? J. Financ. 78, 1777–1827.

Griffin, J.M., Kruger, S., Mahajan, P., 2024. Is Fraud Contagious? Social Connections and the Looting of COVID Relief Programs. Working Paper, University of Texas at Austin.

Griffin, J.M., Kruger, S., Mahajan, P., 2025. Is fraud contagious? Social connections and the looting of COVID relief programs. Rev. Financial Stud. Forthcoming.

Griffin, J.M., Kruger, S., Maturana, G., 2020. What drove the 2003–2006 house price boom and subsequent collapse? Disentangling competing explanations. J. Financ. Econ. 141, 1007–1035.
Griffin, J.M., Maturana, G., 2016. Did dubious mortgage origination practices distort house prices? Rev. Financ. Stud. 29, 1671–1708.
Gross, T., Notowidigdo, M.J., Wang, J., 2020. The marginal propensity to consume over the business cycle. Am. Econ. J.: Macroecon. 12 (2), 351–384.
Guiso, L., Sapienza, P., Zingales, L., 2008. Trusting the stock market. J. Financ. 63 (6), 2557–2600.
Gupta, A., Mittal, V., Peeters, J., Van Nieuwerburgh, S., 2022. Flattening the curve: Pandemic-induced revaluation of urban real estate. J. Financ. Econ. 146 (2), 594–636.
Guren, A.M., 2018. House price momentum and strategic complementarity. J. Political Econ. 126 (3), 1172–1218.
Gurun, U.G., Nickerson, J., Solomon, D.H., 2023. Measuring and improving stakeholder welfare is easier said than done. J. Financ. Quant. Anal. 58 (4), 1473–1507.
Gurun, U.G., Stoffman, N., Yonker, S.E., 2017. Trust busting: The effect of fraud on investor behavior. Rev. Financ. Stud. 31 (4), 1341–1376.
Hartman-Glaser, B., Thibodeau, M., Yoshida, J., 2023. Cash to spend: IPO wealth and house prices. Real Estate Econ. 51 (1), 68–102.
Havranek, T., Sokolova, A., 2020. Do consumers really follow a rule of thumb? Three thousand estimates from 144 studies say ''probably not''. Rev. Econ. Dyn. 35, 97–122.
Hu, Z., 2021. Social interactions and households' flood insurance decisions. J. Financ. Econ. 144, 414–432.
Jappelli, T., Pistaferri, L., 2010. The consumption response to income changes. Ann. Rev. Econ. 2 (Volume 2, 2010), 479–506.
Johnson, D.S., Parker, J.A., Souleles, N.S., 2006. Household expenditure and the income tax rebates of 2001. Am. Econ. Rev. 96 (5), 1589–1610.
Jorda, O., Nechio, F., 2023. Inflation and wage growth since the pandemic. Eur. Econ. Rev. 156, 104474.
Kaplan, G., Violante, G.L., 2014. A model of the consumption response to fiscal stimulus payments. Econometrica 82 (4), 1199–1239.
Kaplan, G., Violante, G.L., 2022. The marginal propensity to consume in heterogeneous agent models. Ann. Rev. Econ. 14 (Volume 14, 2022), 747–775.
Kedia, S., Philippon, T., 2007. The economics of fraudulent accounting. Rev. Financ. Stud. 22 (6), 2169–2199.
Khetan, U., Leder-Luis, J., Wang, J., Zhou, Y., 2024. Unemployment Insurance Fraud in the Debit Card Market. Working Paper, University of Iowa.
Kueng, L., 2018. Excess sensitivity of high-income consumers. Q. J. Econ. 133 (4), 1693–1751.
Lardner, R., McDermott, J., Kessler, A., 2023. The great grift: How billions in COVID-19 relief aid was stolen or wasted. https://apnews.com/article/pandemic-fraud-waste-billions-small-business-labor-fb1d9a9eb24857efbe4611344311ae78.
Ley, E., Steel, M.F., 2009. On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. J. Appl. Econometrics 24, 651–674.
Lin, L., 2025. Fiscal Stimulus Payments, Housing Demand, and House Price Inflation. Working Paper, University of Pittsburgh.
Liu, S., Su, Y., 2021. The impact of the COVID-19 pandemic on the demand for density: Evidence from the U.S. housing market. Econom. Lett. 207, 110010.
Lutz, C., Sands, B., 2023. Highly Disaggregated Land Unavailability. Working Paper, University of North Carolina at Charlotte.
Mian, A., Sufi, A., 2009. The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis. Q. J. Econ. 124, 1449–1496.
Mian, A., Sufi, A., 2018. Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. Rev. Financ. Stud. 30, 1832–1864.
Mondragon, J.A., Wieland, J., 2025. Housing Demand and Remote Work. Working Paper, Federal Reserve Bank of San Francisco.
Neilson, C., Humphries, J.E., Ulyssea, G., 2020. Information frictions and access to the paycheck protection program. J. Public Econ. 190.
Noh, S., So, E.C., Zhu, C., 2025. Financial reporting and consumer behavior. Account. Rev. 100 (1), 407–435.
Oster, E., 2019. Unobservable selection and coefficient stability: Theory and evidence. J. Bus. Econom. Statist. 37 (2), 187–204.
Pancost, A., Schaller, G., 2025. Investigating Instruments with Meta-Regressions. Working Paper, University of Texas at Austin.
Parker, J.A., Souleles, N.S., Johnson, D.S., McClelland, R., 2013. Consumer spending and the economic stimulus payments of 2008. Am. Econ. Rev. 103 (6), 2530–2553.
Podkul, C., 2021. How unemployment insurance fraud exploded during the pandemic. https://www.propublica.org/article/how-unemployment-insurance-fraud-exploded-during-the-pandemic.
Ramani, A., Bloom, N., 2022. The Donut Effect of Covid-19 on Cities. Working Paper, National Bureau of Economic Research.
Souleles, N.S., 1999. The response of household consumption to income tax refunds. Am. Econ. Rev. 89 (4), 947–958.
Tauber, C., Van Zandweghe, W., 2021. Why Has Durable Goods Spending Been So Strong During the COVID-19 Pandemic? Working Paper, Federal Reserve Bank of Cleveland.
U.S. Small Business Administration, Office of Inspector General, 2023. COVID-19 pandemic EIDL and PPP loan fraud landscape. https://www.sba.gov/document/report-23-09-covid-19-pandemic-eidl-ppp-loan-fraud-landscape, (Accessed 25 December 2025).
Zingales, L., 2015. Presidential address: Does finance benefit society? J. Financ. 70 (4), 1327–1363.