Which lenders had the highest minority share among their Payment Protection Program (PPP) loans?

Sabrina T. Howell, Theresa Kuchler, Johannes Stroebel*

This version: January 23, 2020

I Introduction

The economic crisis induced by the COVID-19 pandemic was devastating for small businesses in the United States, and minority-owned small businesses closed at much higher rates than White-owned businesses (Greenwood et al., 2020; Bartik et al., 2020; Fairlie, 2020; Alekseev et al., 2020). In March 2020, the U.S. Congress responded with provisions in the CARES Act that authorized $650 billion in Payment Protection Program (PPP) loans to small businesses. While guaranteed by the government, PPP loans were processed and disbursed by private lenders. Section 1102 of the CARES Act explicitly specified that the program should prioritize "small business concerns owned and controlled by socially and economically disadvantaged individuals."[1] However, the SBA did not issue specific guidance for distributing the loans, leaving private financial institutions administering the loans to independently determine which businesses to serve first or at all. Quickly, media reports raised concerns that minority-owned businesses struggled to access PPP loans (Zhou, 2020; Beer, 2020).

This short research note explores how PPP lenders differed in the share of their loans that went to minority-owned and especially to Black-owned businesses. Answering this question is challenging, since owner race was not systematically reported in PPP applications to the SBA. We employ data newly released by the SBA on December 1, 2020, which include the business names of all 5.16 million PPP borrowers.[2] The names permit us to build on a well-established literature that predicts race given a person's name and location (Imai and Khanna,

*All authors are at NYU Stern. Preliminary analysis, final numbers subject to change. Please contact Howell at Sabrina.howell@nyu.edu or 917-519-8052 with questions. We are grateful to Georgij Alekseev and Jun Wong for superb research assistance. For generous data and service support, we are also grateful to Brock Blake and Katherine Chandler at Lendio, as well as Kurt Ruppel and Kyle Mack at Middesk. Howell's work on this project is funded by the Kauffman Foundation.

[1] A similar desire for PPP loans to reach underserved communities was expressed in the December 2020 Framework for a *Bipartisan Emergency COVID Relief Act* (2020), a bi-partisan proposal for further COVID relief.

[2] The newly released SBA data contains 4,837 unique institutions who participated as lenders in the PPP program. Together, they gave out 5,156,850 PPP loans with SBA approval dates between April 3 and August 8, with a hiatus from April 17 until April 26. During that period the first round of funding ran out, and Congress had not yet appropriated the second round.

2016; Humphries et al., 2019; Tzioumis, 2018). Using a random forest algorithm, we enhance the accuracy of these predictions by training our model on a subset of PPP borrowers who self-reported race.

In analysis using 3.34 million of the PPP loans, we document substantial variation across types of financial institutions in the propensity to lend to minority-owned businesses, a finding that is robust to alternative proxies for which business is minority-owned. Our results suggest that fintech lenders appear to have played an important role in extending PPP loans to Black- and Hispanic-owned businesses.

Ongoing work with business checking account data, which provides the bank relationship before the PPP, as well as credit and debit card monthly revenue, will help us to shed light on the mechanisms. They will allow us to better understand (i) whether the differential approval rates by race are related to pre-existing banking relationships, and (ii) whether the timing of receiving a loan early relative to the late in the PPP matters for firm outcomes.

II PPP Loan Characteristics by Lender Type

We categorize the set of PPP lenders into one of 13 mutually exclusive groups. We include each of the top four banks by assets separately: JP Morgan Chase, Bank of America, Wells Fargo, and Citibank. The next category is large banks, which we define as those with more than $100 billion in assets, excluding the four largest banks. We then split the remaining banks by the median asset value to create a medium bank category with more than $2.2 billion in assets and a small bank category with less than $2.2 billion in assets. The remaining categories are credit unions, Business Development Corporations (BDCs), Minority Development Institutions (MDIs), Community Development Financial Institutions (CDFIs), nonprofits, and fintech lenders.[3]
Table 1 shows loan summary statistics by lender type. Small and medium banks accounted for roughly half of all loans with median loan values of $27,400 and $35,500, respectively. The larger banks made the majority of the remaining loans. Fintechs made 845,495 loans with a median value of $15,975. The institutions traditionally used by government to reach underserved communities are CDFIs, nonprofits and MDIs. CDFIs and nonprofits made 87,383 loans, while MDIs made 126,701 loans. Both groups had a median value of about

[3]The list of fintech lenders we include is in Table A.1. We include in the fintech category the lenders officially designated by the SBA as fintechs as well as online lenders that we know originate primarily for or via fintech partners or platforms, online lenders founded since 2005, and online lenders that received venture capital (VC) investment. It is important to note that the Lendio loans are only residual Lendio loans after removing loans that Lendio intermediated as a matching platform but were issued by another fintech. That is, the loans assigned to Lendio are those that Lendio sourced but were originated by a non-fintech lender. A second note is that we classify Customers Bank, Celtic and Cross River as fintech because the vast majority of their PPP loans were sourced and processed by a fintech, which partners with them to provide banking services. For example, Cross River and Customers are each the originator for about 20% of Kabbage's PPP loans. That is, the loans ascribed to Kabbage in this table represent only about 60% of Kabbage's total fintech loans; the remainder are among the loans ascribed to Cross River and Customers. Overall, the single largest fintech by loan volume is Cross River, a bank. However, they originated an overwhelming quantity of loans for fintech partners such as Kabbage, were founded in 2008, and they received VC.

Table 1: Loan Summary Statistics by Lender Type

	Number Lenders	Number Loans (100k)	Total Amt Lent (Bill)	Mean Loan Amt	Median Loan Amt	Share Loans < $22,880 (Median)
All	4,838	5,156.851	522.950	101,409	22,880	50%
Bank of America	1	337.885	25.155	74,449	20,833	53%
Citibank	1	30.216	3.334	110,354	30,300	42%
JPMC Bank	1	274.837	28.932	105,270	28,740	43%
Wells Fargo Bank	1	188.059	10.034	53,357	18,777	59%
Bus Dev Corp	3	6.572	0.455	69,305	20,800	55%
Large Banks (assets>$130b ex top 4)	16	543.785	78.625	144,589	32,600	40%
Medium Banks ($2b<assets<$130b)	506	1,272.885	186.366	146,412	35,500	39%
Small Banks (<$2b assets)	3,089	1,235.073	128.010	103,646	27,400	45%
Credit Union	927	207.960	9.837	47,300	16,456	64%
CDFI/Nonprofit	159	87.383	6.773	77,512	20,833	53%
Minority Dev Inst	110	126.701	10.210	80,583	20,800	55%
Fintech	25	845.495	35.218	41,653	15,975	73%

$20,800. Therefore, fintech lenders originated much smaller loans than other lenders, suggesting they served smaller firms on average.

III Methodology

To explore lending differences across types of financial institutions, we run the following regression at the level of the individual loan recipient for several racial categories, with a focus on Black-owned businesses:

$$Black_i = \beta LenderType_i + \alpha_{City} + \alpha_{State} + \alpha_{Industry} + \alpha_{LoanAmountPctile} + \varepsilon_i \tag{1}$$

Lender type includes indicators for each of the 13 lender type groups. With no controls, the estimates reflect the unconditional share of loans to Black-owned businesses by lender type. In some specifications, we also include dummy variables to account for the role of location (captured by the borrower's city), industry (captured by NAICS 3-digit industry classification, including categories such as "Health and Personal Care Stores," "Truck Transportation," and "Food Services and Drinking Places"), and loan size (captured by indicators for loan amount percentiles).

Whether a business is minority-owned is reported for a subset of about 536,000 PPP borrowers who choose to self-identify their race in the loan application, and for which the lender also chose to report this information to the SBA. For our primary measure, we therefore build on a well-established literature that predicts race given a person's name. We identify an individual owner or senior executive name in two ways. The first source is information on current firm officers drawn from Secretary of State registrations on 70% of PPP borrowers, which the data analytics firm Middesk provided to us.[4] Second, where this is not available but the business is a non-employer firm such as a sole proprietorship, we make use of the fact that the business name often corresponds to the name of the owner. In sum, we observe an owner

[4]We use the name of the first officer listed where there are multiple.

name for 4.1 million firms.

We create a random forest model to predict race based on the name. The first name is matched with a list of first names as in Tzioumis (2018) and the last name is matched with a list of surnames as in Imai and Khanna (2016). After dropping names that do not match these lists, we have 3.43 million names. For this sample, we input the race distributions for both names into the random forest model and train the model on the subset where race is self-identified. We enhance the prediction by adding Bayesian posterior probabilities from combining the first name with the business location (Imai and Khanna, 2016). Afterward, we validate the model on the remainder of these businesses before applying it to the full sample. The model provides us with a probability that each loan is to a minority-owned business.

While predicting race with a name produces a somewhat different signal than using self-identified race, the model appears to work reasonably well. Among the sample of borrowers who self-identify as Black and for which we observe the name, we successfully predict 78% to be Black. The same figures for individuals who self-identify as Hispanic, Asian, and White are 84%, 95%, and 99%, respectively. It is worth noting that the signal may contain differentially useful information than self-identified race. For example, two of the prediction algorithm's "errors" are individuals whose last names are Huang and Rodriguez, who identify as Black but are not predicted to be Black. A name may send a different signal to a loan officer than actual race or ethnicity.

Comfortingly, the rate of Black ownership in the self-identified data, at 3.2%, is very similar to the rate in the predicted sample, which is 3.3%. The predicted rates for Hispanic, Asian, and White business owners are 7.6%, 10.2%, and 79%, respectively. These are also very close to the self-identified proportions, which are 7.8%, 10.9%, and 78%, respectively.

IV Results

Figure 1 shows preliminary estimates using our preferred predictor of Black ownership within the subset of firms for which information on owner name allows us to predict race. Panel A shows the share of loans for each lender that went to a business that is predicted to be Black-owned. Panel B includes controls for business location, industry, and loan amount, and reflects the shares of loans for each lender type to Black-owned businesses not explained by these factors. Both panels show that fintechs tend to make a substantially larger share of their PPP loans to Black-owned businesses than other lender types, at just over 6% including the full suite of controls (Panel B). The next highest rates are among CFDIs, nonprofits, credit unions, and MDIs, at 3%. Small- and medium-sized banks are least likely to lend to Black-owned businesses, with large banks and the top four banks close behind.

A similar pattern emerges at the community level in Figure 2, where we analyze the share of the population that is Black in the zip codes of PPP borrowers, using all PPP loans to all firms (the SBA data includes firm location for all firms). Fintechs, CFDIs, MDIs, and nonprofits made a larger share of their loans in ZIP codes with a higher share of black residents.

We explore differences in lending to Hispanic-owned businesses in Figure 3. Without controls, MDIs have by far the highest rate of lending to Hispanic-owned businesses. When we add controls, however, the ranking changes; fintechs and Wells Fargo are most likely to lend to

Hispanic-owned businesses, while MDIs along with small and medium-sized banks are least likely to do so (Panel B). The change for MDI reflects the inclusion of controls for city; that is, models with controls for loan amount or industry are similar to Panel A, but a model with only location controls is similar to Panel B. This highlights the importance of controlling for location in understanding disparities across lender types.

Lending to Asian-owned businesses, shown in Figure 4, shows markedly different patterns across both specifications. MDIs are almost twice as likely as any other lender type to lend to Asian-owned businesses; in raw averages, over 40% of their loans are to this group (Panel A). MDIs are followed by top four banks, with around 20% of loans to Asian-owned businesses. Finally, consistent with the implications of the previous figures, Figure 5 shows that small- and medium-sized banks were most likely lend to White-owned businesses, with roughly 90% of their loans to this group (Panel B).

We find similar patterns when using alternative measures of identifying loans to minority-owned businesses. Most importantly, we find similar results using only those PPP loans with self-identified data on race (Figures A.1 and A.2). Also, the results are similar with alternative controls, including narrower definitions of location. For example, Figure A.3 shows that fintech's higher propensity to lend to Black-owned businesses is more striking in models that include fixed effects for the zip code (Panel A) or Census tract (Panel B).

Overall, we find that, relative to other lenders, fintech lenders make a substantially larger share of their loans to Black- and Hispanic-owned businesses. They also made far more loans than MDIs, nonprofits, and CDFIs traditionally used to reach underserved communities. Fintech lenders, therefore, appear to have played an important role in extending PPP loans to Black- and Hispanic-owned businesses.

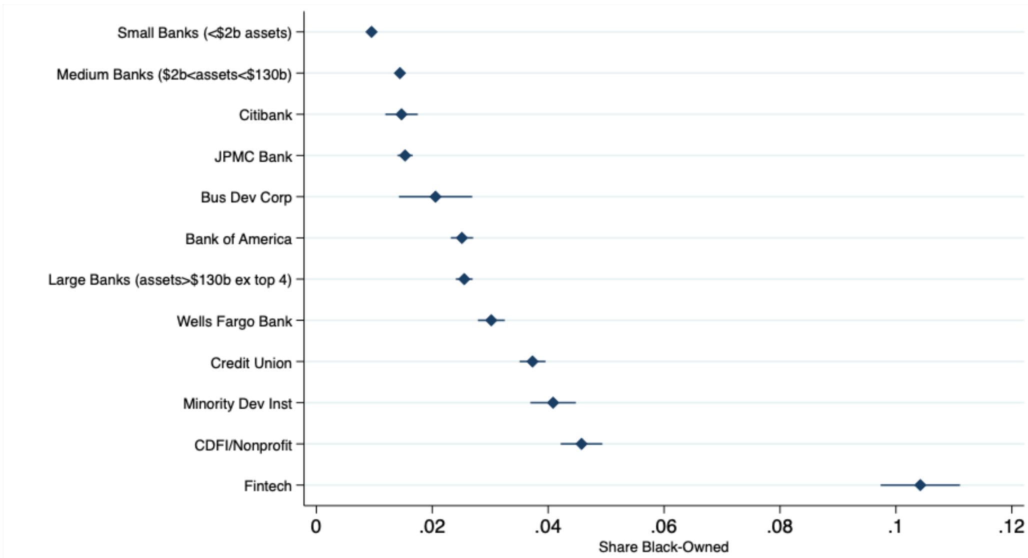
Importantly, these results thus do not necessarily indicate discriminatory lending practices by some lenders. In particular, the observed patterns could, among other things, reflect higher demand from Black and Hispanic business owners for fintech loans. Alternatively, traditional lenders may have lent at lower rates to Black or Hispanic businesses because they prioritized customers who were less likely to be Black or Hispanic (for example, because they had no pre-existing business relationship with these borrowers).[5]

In ongoing work with data from Ocrolus, Lendio, Enigma, and Middesk, we hope to shed more light on mechanisms. For example, we are linking business checking account data to the PPP lending data to explore the role of pre-existing customer relationships with different types of financial institutions. We are also gathering data on credit and debit card monthly revenue for all PPP borrowers, which will help us to explore how any disparities are related to firms' real outcomes.

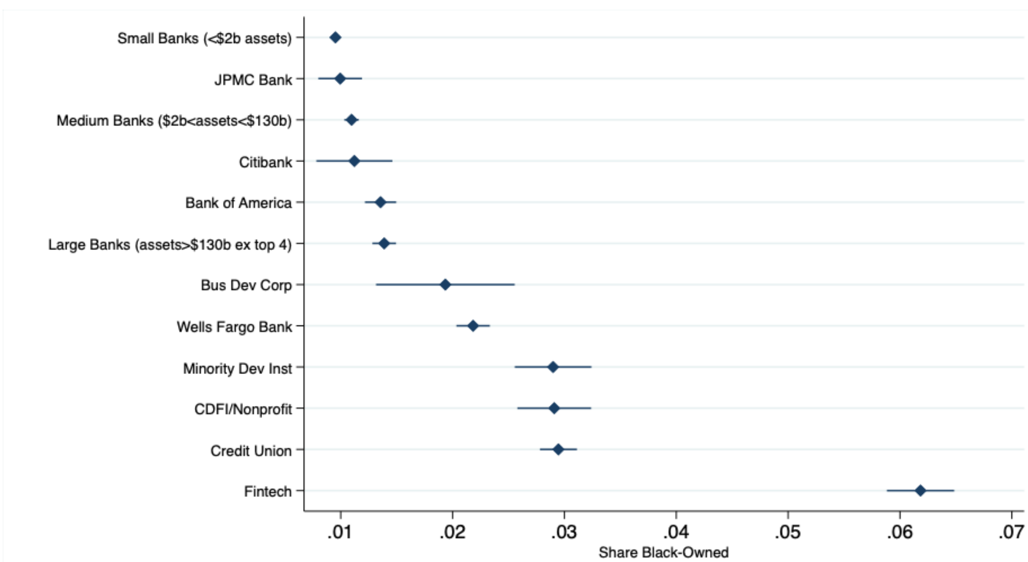
[5] For loan applications below $350,000, which includes the vast majority of applications, lenders were incentivized to prioritize larger loans since they receive a fee of 5% of the loan value. Since banks also have an incentive to keep their existing borrowers in good financial health, they also have an incentive to prioritize their clients with existing credit relationships (Granja et al., 2020).

Figure 1: **Black-Owned Business PPP Lending by Institution Type**

A: No Controls



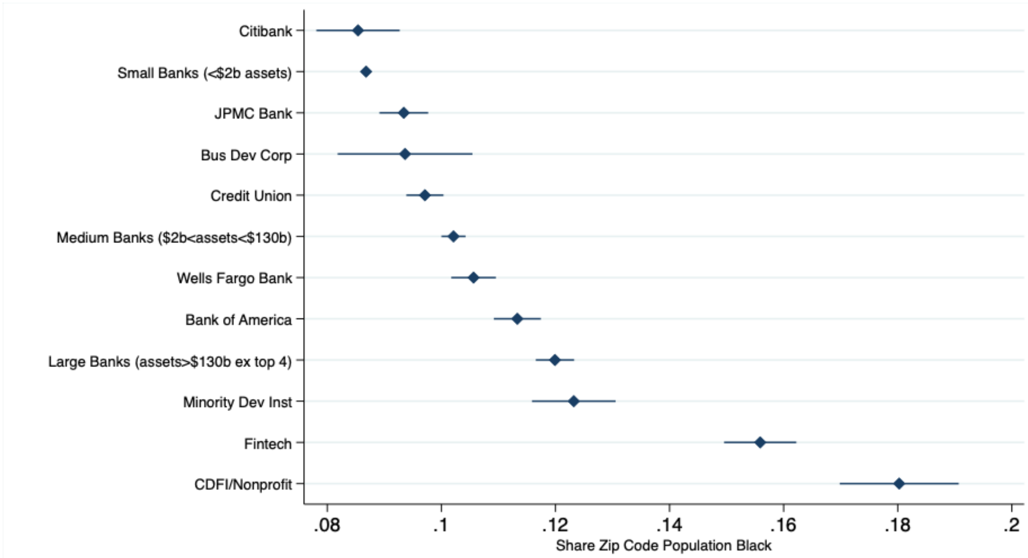
B: Full Controls



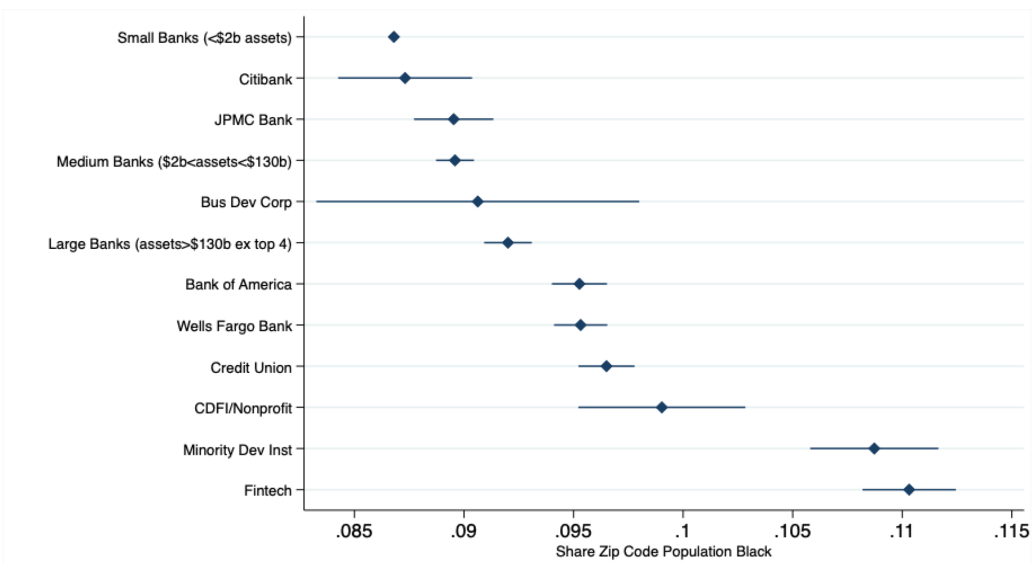
Note: This figure shows shares of PPP loans made to businesses predicted to be Black-owned by lender type. We classify business owners as Black if our algorithm predicts that the individual has a higher probability of being this race than any being other race. Included in the sample are loans to self-employed individuals and sole proprietorships for which we observe business owner name in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is Black-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. The bottom graph includes fixed effects for borrower city, state, and NAICS3 industry, and loan amount percentiles. Standard errors are clustered by zip code. N=3,343,396.

Figure 2: **Share of Black Population in Businesses' Communities by Institution Type**

A: No Controls



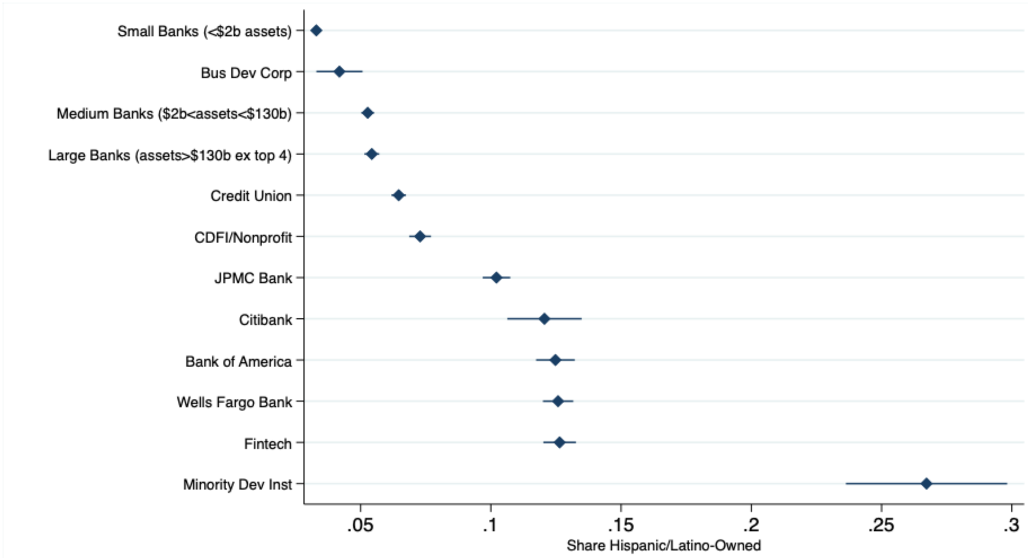
B: Full Controls



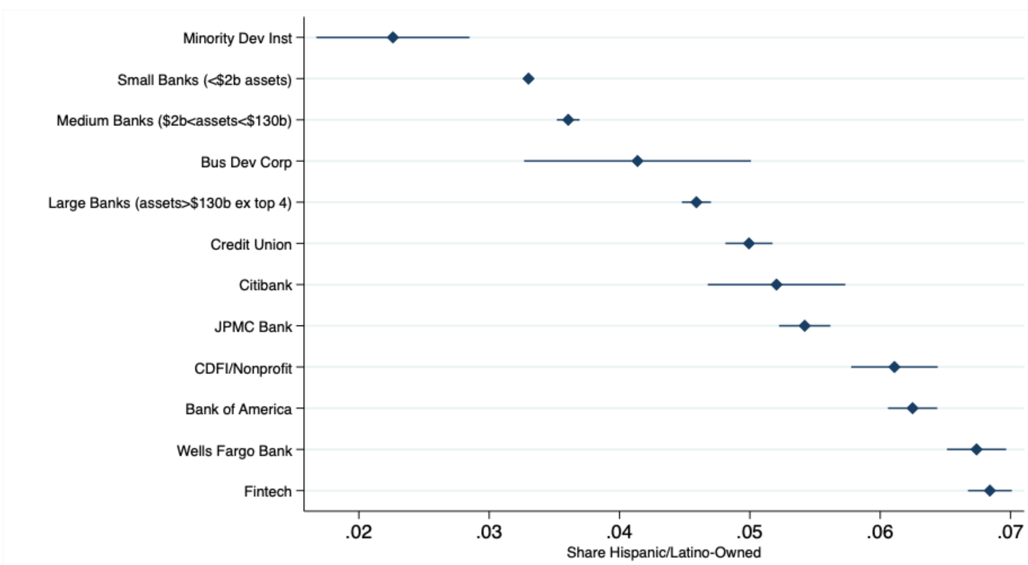
Note: This figure shows the share of Black residents in the ZIP codes of PPP loan recipients by lender type. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is the share of the zip code's population that is Black, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. The bottom graph includes fixed effects for borrower city, state, and NAICS3 industry, and loan amount percentiles. Standard errors are clustered by zip code.

Figure 3: **Hispanic-Owned Business PPP Lending by Institution Type**

A: No Controls



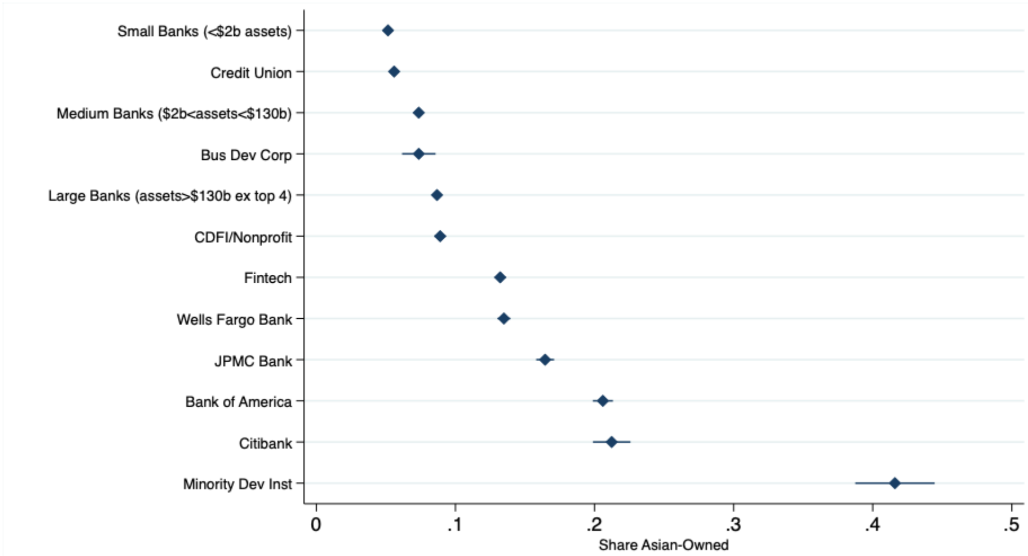
B: Full Controls



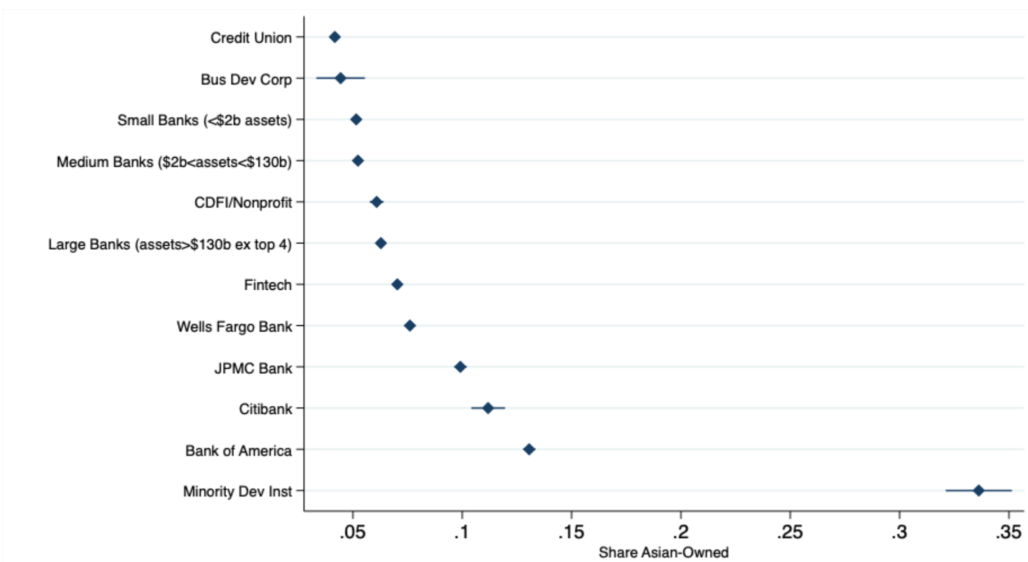
Note: This figure shows shares of PPP loans made to businesses predicted to be Hispanic-owned by lender type. We classify business owners as Black if our algorithm predicts that the individual has a higher probability of being this race than any being other race. Included in the sample are loans to self-employed individuals and sole proprietorships for which we observe business owner name in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is Hispanic-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. The bottom graph includes fixed effects for borrower city, state, and NAICS3 industry, and loan amount percentiles. Standard errors are clustered by zip code. N=3,343,396.

Figure 4: **Asian-Owned Business PPP Lending by Institution Type**

A: No Controls



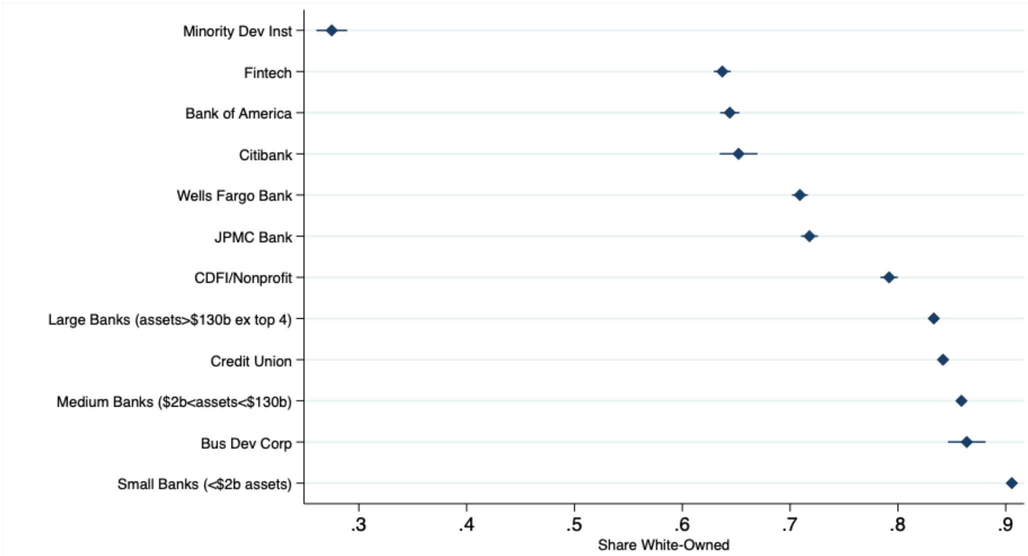
B: Full Controls



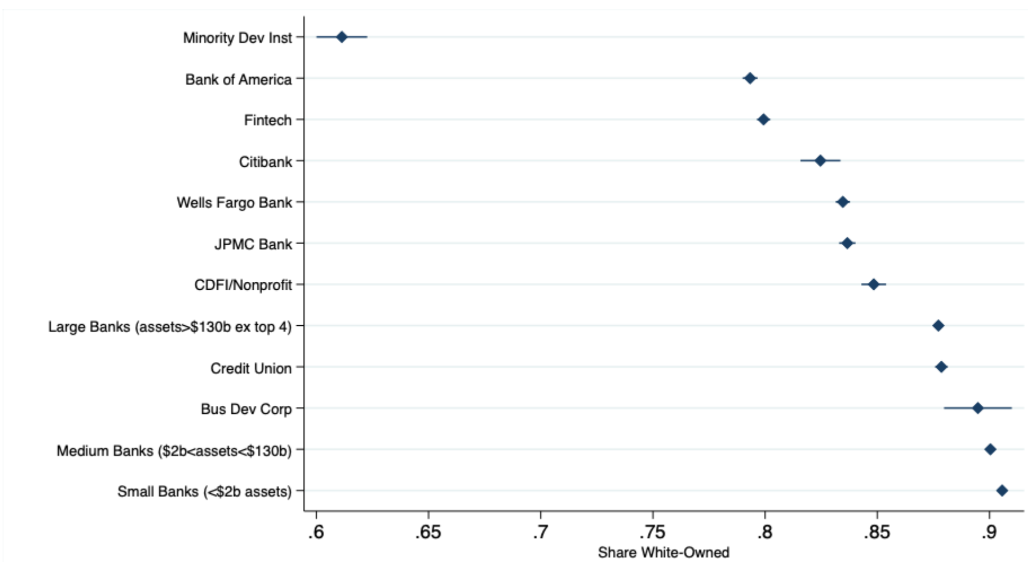
Note: This figure shows shares of PPP loans made to businesses predicted to be Asian-owned by lender type. We classify business owners as Black if our algorithm predicts that the individual has a higher probability of being this race than any being other race. Included in the sample are loans to self-employed individuals and sole proprietorships for which we observe business owner name in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is Asian-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. The bottom graph includes fixed effects for borrower city, state, and NAICS3 industry, and loan amount percentiles. Standard errors are clustered by zip code. N=3,343,396.

Figure 5: **White-Owned Business PPP Lending by Institution Type**

A: No Controls



B: Full Controls



Note: This figure shows shares of PPP loans made to businesses predicted to be White-owned by lender type. We classify business owners as White if our algorithm predicts that the individual has a higher probability of being this race than any being other race. Included in the sample are loans to self-employed individuals and sole proprietorships for which we observe business owner name in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is White-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. The bottom graph includes fixed effects for borrower city, state, and NAICS3 industry, and loan amount percentiles. Standard errors are clustered by zip code. N=3,343,396.

References

Alekseev, G., S. Amer, M. Gopal, T. Kuchler, J. W. Schneider, J. Stroebel, and N. Wernerfelt, "The Effects of COVID-19 on US Small Businesses: Evidence from Owners, Managers, and Employees," Technical Report, NBER Working Paper No. w27833 2020.

Bartik, Alexander W, Marianne Bertrand, Zoe Cullen, Edward L Glaeser, Michael Luca, and Christopher Stanton, "The impact of COVID-19 on small business outcomes and expectations," *Proceedings of the National Academy of Sciences*, 2020, *117* (30), 17656–17666.

Beer, Tommy, "Minority-Owned Small Businesses Struggle To Gain Equal Access To PPP Loan Money," https://www.forbes.com/sites/tommybeer/2020/05/18/minority-owned-small-businesses-struggle-to-gain-equal-access-to-ppp-loan-money/?sh=66237cd05de3 2020. [Online; accessed 09-Dec-2020].

Bipartisan Emergency COVID Relief Act

***Bipartisan Emergency COVID Relief Act**, https://assets.documentcloud.org/documents/20421761/12920-emergency-covid-relief-act-of-2020-framework-summary.pdf 2020. [Online; accessed 09-Dec-2020].*

Fairlie, R., *"The impact of COVID-19 on small business owners: Evidence from the first three months after widespread social-distancing restrictions," Journal of Economics & Management Strategy, 2020, 29, 727–740.*

Granja, J., C. Makridis, C. Yannelis, and E. Zwick, *"Did the Paycheck Protection Program Hit the Target?," Technical Report, NBER Working Paper No. w27095. 2020.*

Greenwood, Robin M, Benjamin Charles Iverson, and David Thesmar, *"Sizing up corporate restructuring in the covid crisis," NBER Working Paper, 2020, (w28104).*

Humphries, J. E., N. Mader, D. Tannenbaum, and W van Dijk, *"Does eviction cause poverty? quasi-experimental evidence from Cook County, IL.," Technical Report, NBER Working Paper No. w26139 2019.*

Imai, K. and K. Khanna, *"Improving ecological inference by predicting individual ethnicity from voter registration records," Political Analysis, 2016, pp. 263–272.*

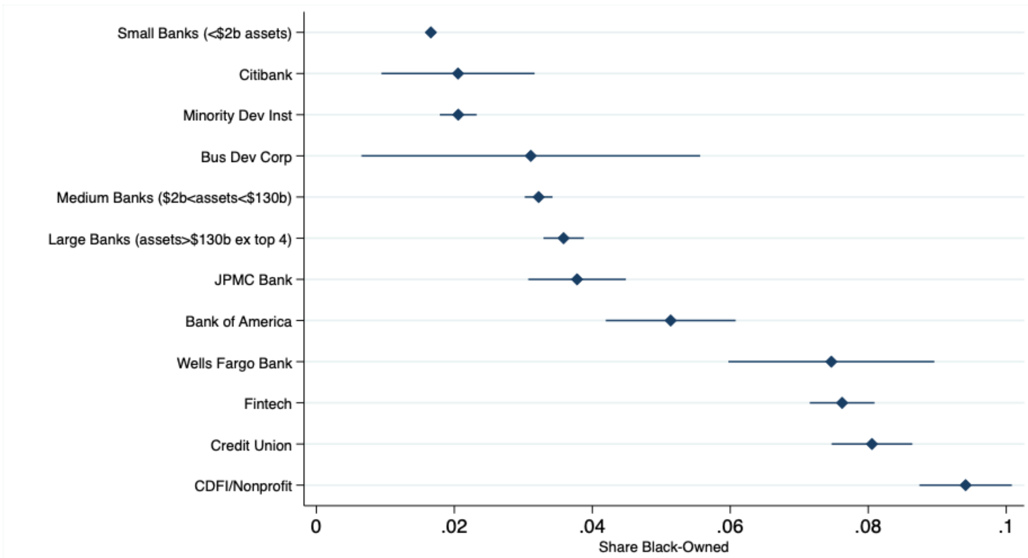
Tzioumis, Konstantinos, *"Demographic aspects of first names," Scientific data, 2018, 5, 180025.*

Zhou, Li, *"The Paycheck Protection Program failed many Black-owned businesses," https://www.vox.com/2020/10/5/21427881/paycheck-protection-program-black-owned-businesses 2020. [Online; accessed 09-Dec-2020].*

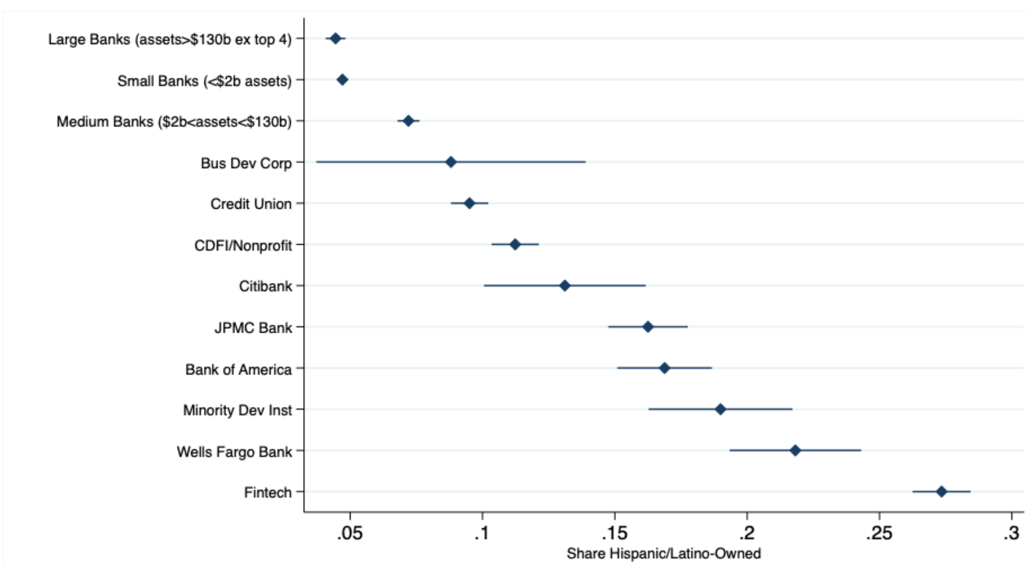
Appendix

Figure A.1: **Black- and Hispanic-Owned Business PPP Lending by Institution Type (Self-Identified)**

A: Self-identify as Black



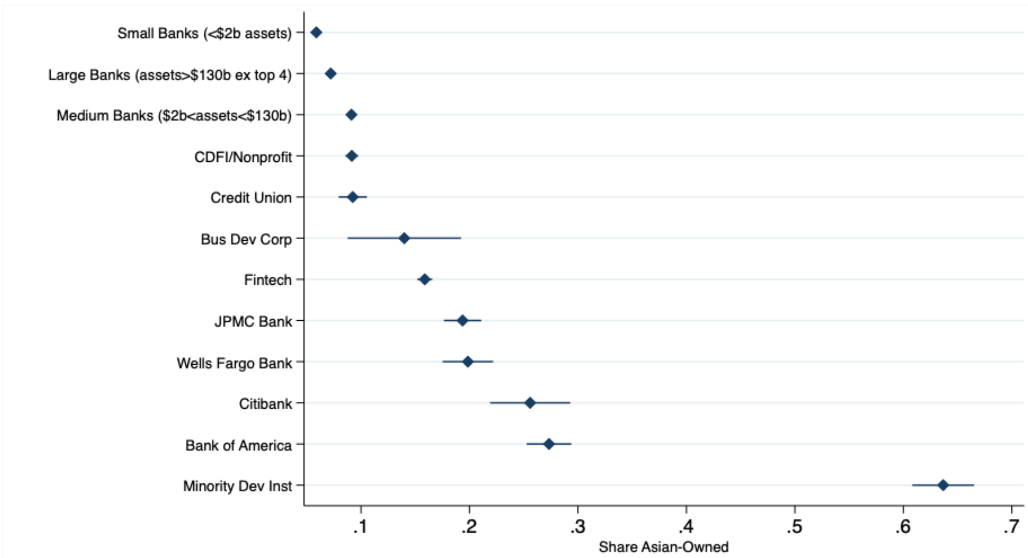
B: Self-identify as Hispanic



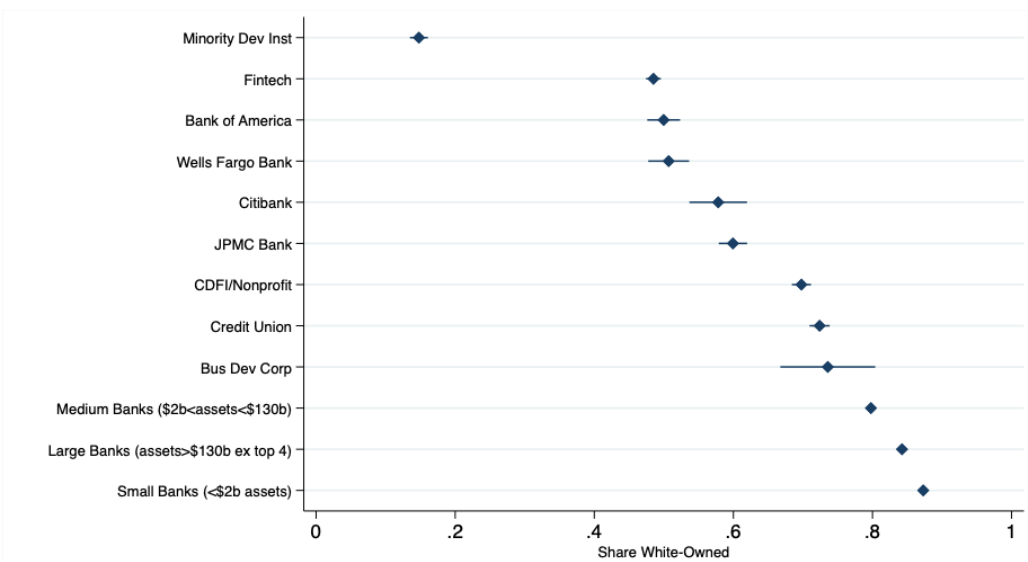
Note: This figure shows shares of PPP loans made to businesses that self-identify as Black- or Hispanic-owned by lender type. Included in the sample are loans to business for which self-identified race is included in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is Black- or Hispanic-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. Standard errors are clustered by zip code. N=535,839.

Figure A.2: **Asian- and White-Owned Business PPP Lending by Institution Type (Self-Identified)**

A: Self-identify as Asian



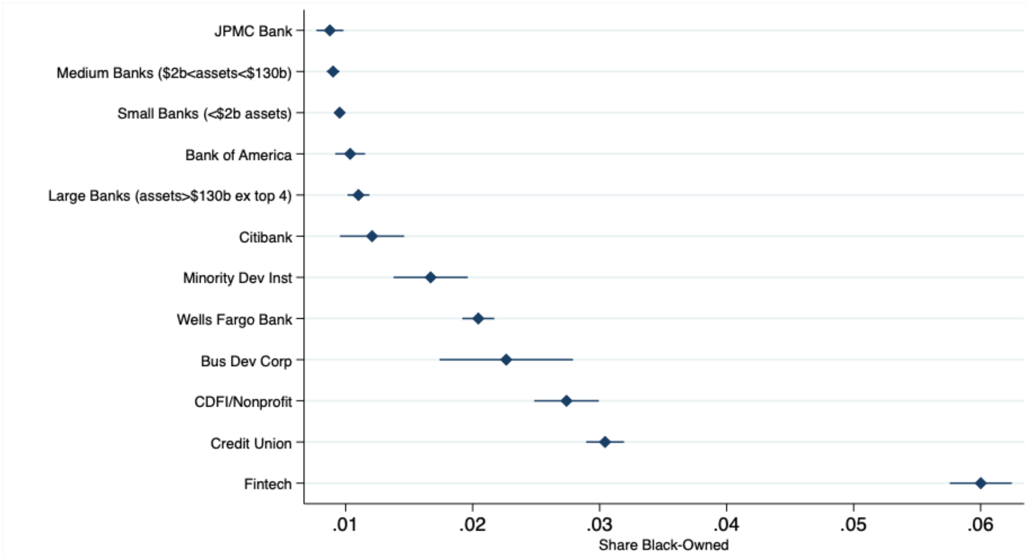
B: Self-identify as White



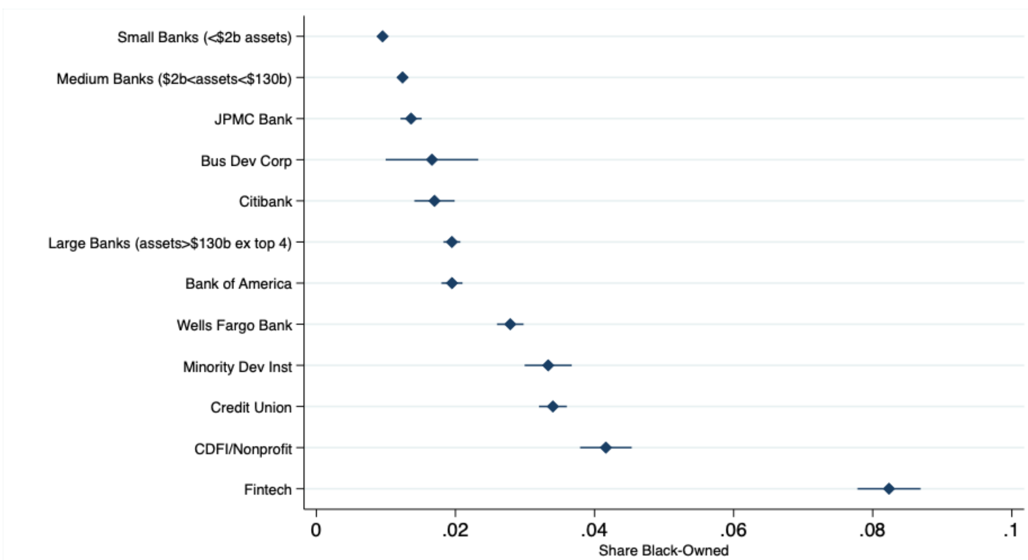
Note: This figure shows shares of PPP loans made to businesses that self-identify as Asian- or White-owned by lender type. Included in the sample are loans to business for which self-identified race is included in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is Asian- or White-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. Standard errors are clustered by zip code. N=535,839.

Figure A.3: **Black-Owned Business PPP Lending by Institution Type with Zip Code and Census Tract Controls**

A: Zip Code Fixed Effects



B: Census Tract Fixed Effects



Note: This figure shows shares of PPP loans made to businesses predicted to be Black-owned by lender type. We classify business owners as Black if our algorithm predicts that the individual has a higher probability of being this race than any being other race. Included in the sample are loans to self-employed individuals and sole proprietorships for which we observe business owner name in the data. Each graph presents coefficients from a regression of equation 1, where we add back the mean of the dependent variable among the excluded lender category. The dependent variable is an indicator for whether the business is Black-owned, and the independent variables are indicators for mutually exclusive categories for PPP lender type with small banks as the base category. The top panel includes fixed effects for the borrower zip code, while the bottom panel includes fixed effects for the borrower Census tract. Standard errors are clustered by zip code. N=3,343,396 in the top panel and 2,972,560 in the bottom panel.

Table A.1: List of Fintechs

Lender	Num Loans	Median Loan Amt (Thou)
CRF Small Business Loan Company	2395	20800.000
Celtic Bank Corporation	147223	9941.330
Centerstone SBA Lending	898	70600.000
Cross River Bank	194422	19062.750
Customers Bank	69295	13544.000
Evolve Bank and Trust	705	27400.000
FC Marketplace	6134	25994.000
Fountainhead SBF	2766	75850.000
Fund-Ex Solutions Group	1404	75000.000
Fundbox	14236	12506.560
Grow America Fund	704	31391.000
Harvest Small Business Finance	5346	77844.500
Intuit Financing	18509	17922.000
Kabbage	161174	11002.000
Lendio	100352	22500.000
Loan Source	193	26600.000
NBKC Bank	484	18700.000
Newtek Small Business Finance	11554	20800.000
Radius Bank	5480	24723.000
Readycap Lending	34261	23726.000
Sunrise Banks National Association	1840	26750.000
The Bancorp Bank	1288	66750.000
VelocitySBA	108	78750.000
WebBank	76424	14000.000
immito	294	30550.000

Note: This table lists the lenders we classify as fintechs, along with the number of PPP loans they issued as well as the median loan amount. It is important to note that the Lendio loans are only residual Lendio loans after removing loans that Lendio intermediated as a matching platform but were issued by another fintech. That is, the loans assigned to Lendio are those that Lendio sourced but were originated by a non-fintech lender. A second note is that we classify Customers Bank, Celtic and Cross River as fintech because the vast majority of their PPP loans were sourced and processed by a fintech, which partners with them to provide banking services. For example, Cross River and Customers are each the originator for about 20% of Kabbage's PPP loans. That is, the loans ascribed to Kabbage in this table represent only about 60% of Kabbage's total fintech loans; the remainder are among the loans ascribed to Cross River and Customers.