# Internet Appendix for Racial Disparities in the Paycheck Protection Program

This internet appendix discusses the linking of Florida restaurant licenses, corporate records, voter registration, and PPP loans.

The appendix also reports the following additional results:

1. Table IA1 shows the robustness of the results in the paper to a number of alternative specifications.
2. Table IA2 reports the results of linear probability model regressions of receiving different types of emergency loans on the interaction between female owner and county-level gender-career bias.
3. Table IA3 shows the robustness of the results in Table 6 to controlling for the interaction between *Black* and county characteristics.
4. Table IA4 reports the full output of the regressions in Table 5.

## Linking approved PPP loans to Florida corporate records

1. We start with the sample of borrower who according to PPP data are located in Florida. Some of these will turn out to be firms registered in other states and will be dropped after linking to Florida corporate records.
2. Because our focus in on for profit incorporated businesses, we drop non-profits, cooperatives, employee stock ownership plans (ESOPs), joint ventures, partnerships, professional associations, and a few other rare business types. Business types remaining in the sample are
   - Corporation
   - Independent Contractors
   - Limited Liability Company (LLC)
   - Limited Liability Partnership

- Self-Employed Individuals
- Sole Proprietorship
- Subchapter S Corporation

We keep independent contractors, self-employed individuals, and sole proprietorships at this point because while most of these borrowers report individual names, some are in fact set up as firms. We drop a small number of borrowers whose names are missing or identify borrowers simply as an Uber or Lyft driver or an independent contractor without providing any other information.

3. Match exactly on firm name after removing punctuation and standardizing legal form (for example replacing "Corporation" with "Corp").

4. Match exactly on firm name after removing white space.

5. Match after removing legal form.

6. Set aside borrowers that have not been matched up to this point and that can be matched to Florida voter registration or white pages data. We parse borrower name into first and last names, and try to match these to voter registration data. If there is at least one match in voter registration data, we tag borrower as an individual.

7. Fuzzy match based on borrower name, ZIP, and street address. We require an exact match on ZIP and first five letters of firm name.

## Linking restaurants to corporate records

1. Match restaurant license holder name to firm name in corporate records after removing punctuation and standardizing legal form. To break ties between multiple potential matches, we use the following screens.

   (a) If there is a match that is currently active according to corporate records, drop inactive matches.

   (b) If there is a match that has the same ZIP code, drop firms that do not match on ZIP code. We compare the principal address in corporate records with the restaurant location and the mailing address in corporate records with the license holder mailing address.

   (c) If there is a match with the same city name, drop firms that do not match on city name.

(d) Keep only those licenses that match to a single corporate record. Three-quarters of all restaurants are matched in this step.

2. Remove white space from license holder and firm names and repeat step 1 on unmatched restaurants.

3. Remove legal form from license holder and firm names and repeat step 1 on unmatched restaurants.

4. Try matching restaurant name to firm name in corporate records after removing punctuation and legal form. Apply the same procedure for breaking ties as in step 1.

5. Fuzzy match between license holder name and firm name.

    (a) Match on ZIP code and first three letters of license holder and firm names.

    (b) Calculate similarity score between license holder and firm names. We use bigram string matching method and calculate the Jaccard score. Specifically we split each name into bigrams, i.e., all possible sequences of two letters. The similarity score between two strings is then $\frac{m}{\sqrt{s_1 \times s_2}}$, where $m$ is the number of bigrams the two strings have in common, and $s_1$ and $s_2$ are the numbers of birgram in the two strings.

    (c) Consider the firm with the highest similarity score to be a match if the similarity score is greater than 0.80 and the difference compared to the next highest similarity score is at least 0.10.

6. Repeat step 5 but matching on city instead of ZIP code.

7. Fuzzy match between license holder name and first officer name in corporate records. Because a first officer can be affiliated with multiple firms in corporate records we keep only those cases where the first officer is affiliated with only one firm. We follow the same general approach as in step 5 except that we

    (a) Match on the first five letters of license holder and first officer names.

    (b) Require similarity score of the best match to be at least 0.95 and at least 0.05 larger than the similarity score of the next best match.

8. Match restaurant name to Florida fictitious records data that report owners of fictitious names including restaurants. Many owners either do not register their fictitious names or fail to update their registration. We implement a fuzzy match, first matching on

ZIP code and first five letters of the restaurant/fictitious name. We require similarity score of the best match to be at least 0.95 and at least 0.05 larger than the similarity score of the next best match. Fictitious names owned by individuals are excluded.

## Linking restaurant owners to voter registration

We combine two snapshots of Florida voter registration data: January 31, 2017 and January 31, 2021.[24] The most recent snapshot includes 89% of all voter records. The older snapshot includes voters who may have been deregistered for various reasons.

While we have street address for both restaurant owners and voters, we do not match on street address for two reasons. First, officer's address in corporate records can be a business instead of residential address. Second, corporate records report the most recent address, while voter registration data report the address at the time of registration. If a voter moved within a county and did not update voter registration, the address in voter registration data will be an old address that will not match the address in corporate records, which is updated at least annually. For these reasons we use street address only to break ties between multiple potential matches.

1. Drop restaurants whose first officer is not a person or whose first officer has an out-of-state mailing address.
2. Parse officer name into components: last, first, middle, suffixes and credentials.
3. Match based on county, first and last names.[25]
   (a) If both owner and voter have middle initials, check that the initials agree. If they do not, drop potential matches. Note that we do not rule out matches where only one of the names includes middle initial or name.
   (b) Similarly, if both owner and voter names include a suffix like Jr or II, check for agreement and drop potential that disagree.
   (c) If there is a unique match, keep it.

[24] https://dos.myflorida.com/elections/data-statistics/voter-registration-statistics/voter-extract-disk-request/

[25] We map ZIP codes into counties using HUD USPS ZIP Code Crosswalk Files. In rare cases where a ZIP belongs to multiple counties, we use the counties that accounts for the largest share of the ZIP code.

4. If there are multiple matches, try to break ties using city and street address.
   (a) Drop potential matches whose city does not match.
   (b) Calculate similarity score between street addresses. Keep the best match if its similarity score is at least 0.80 and is at least 0.05 larger than the next highest similarity score.
5. If after the previous step there are still multiple potential matches, but they all report the same race, use that race.

## Table IA1
## Robustness

This table shows the robustness of the results to a number of alternative specifications. Columns 1 and 4 restrict the sample to single-restaurant firms. Columns 2 and 5 classify a firm as having received PPP loans if any of its affiliates receive PPP loans. Columns 3 and 6 use 4,183 Census block group instead of 820 ZIP code fixed effects. Robust standard errors are reported. *, **, and *** indicate statistical significance at 10%, 5%, and 1%.

| | PPP | | | Bank PPP | | |
|---|---|---|---|---|---|---|
| | Single restaurant firms | Loans to affiliates | Census block group | Single restaurant firms | Loans to affiliates | Census block group |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Black | −0.088*** | −0.125*** | −0.080* | −0.155*** | −0.181*** | −0.155*** |
| | (0.029) | (0.028) | (0.042) | (0.028) | (0.027) | (0.040) |
| Hispanic | −0.037** | −0.047*** | −0.001 | −0.035** | −0.043*** | −0.016 |
| | (0.015) | (0.014) | (0.020) | (0.016) | (0.015) | (0.021) |
| Asian | 0.021 | −0.001 | 0.023 | −0.051*** | −0.070*** | −0.047* |
| | (0.016) | (0.015) | (0.023) | (0.018) | (0.017) | (0.025) |
| Other | −0.016 | −0.023 | −0.017 | −0.039 | −0.053** | −0.043 |
| | (0.025) | (0.024) | (0.036) | (0.026) | (0.025) | (0.036) |
| Female | −0.006 | −0.019* | 0.001 | −0.014 | −0.028*** | −0.008 |
| | (0.011) | (0.010) | (0.015) | (0.011) | (0.011) | (0.015) |
| $N$ | 9,271 | 9,980 | 9,980 | 9,271 | 9,980 | 9,980 |
| $R^2$ | 0.192 | 0.192 | 0.501 | 0.200 | 0.207 | 0.507 |
| Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| ZIP FEs | ✓ | ✓ | | ✓ | ✓ | |
| Census block group FEs | | | ✓ | | | ✓ |

## Table IA2
## Gender-Career Bias

This table reports the results of linear probability model regressions of receiving different types of emergency loans on the interaction between female owner and county-level gender-career bias:

$$Emergency\ loan_{c,f,r,z} = \alpha_z + \beta \cdot Minority_f + \delta \cdot Female_f + \theta \cdot Female_f \times Gender\ bias_c + \gamma' X_{f,r} + \varepsilon_{f,c,r,z},$$

where $c$ indexes counties, $f$ indexes firms, $r$ indexes restaurants, and $z$ indexes ZIP codes. Gender-career bias is based on the responses to the Gender-Carreer Implicit Association Test, which measures the implicit association between family and females and between career and males. County-level averages of gender-career bias are standardized so that the interaction terms represent the effect of a one standard deviation increase in bias. Regressions are weighted by the number of respondents to the test during the 2005–2019 period. Controls include all restaurant characteristics included in the regression in column 2 of Table 5. Standard errors are adjusted for clustering by county. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at 10%, 5%, and 1%. Continuous variables are standardized so that their coefficients represent the effect of a one standard deviation change. $N = 9,980$

| | PPP | | Bank PPP | | Nonbank PPP | | EIDL | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Black | −0.205*** | −0.121** | −0.263*** | −0.170*** | 0.058 | 0.049 | 0.017 | 0.026 |
| | (0.044) | (0.054) | (0.024) | (0.022) | (0.048) | (0.050) | (0.038) | (0.040) |
| Hispanic | −0.064** | −0.027 | −0.058* | −0.018 | −0.005 | −0.009 | 0.068*** | 0.071*** |
| | (0.027) | (0.024) | (0.032) | (0.029) | (0.006) | (0.006) | (0.012) | (0.012) |
| Asian | −0.019 | 0.003 | −0.077*** | −0.052** | 0.058*** | 0.055*** | 0.003 | 0.012 |
| | (0.012) | (0.012) | (0.018) | (0.023) | (0.017) | (0.018) | (0.037) | (0.038) |
| Other | −0.020 | 0.001 | −0.060* | −0.037 | 0.041 | 0.038 | 0.062*** | 0.068*** |
| | (0.040) | (0.041) | (0.034) | (0.037) | (0.025) | (0.025) | (0.017) | (0.018) |
| Female | −0.028*** | −0.012 | −0.054*** | −0.036*** | 0.026** | 0.024** | −0.006 | −0.004 |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.010) | (0.010) | (0.011) | (0.012) |
| Female × Bias | −0.040 | −0.020 | −0.131** | −0.111*** | 0.091** | 0.090** | −0.045 | −0.032 |
| | (0.031) | (0.027) | (0.051) | (0.037) | (0.045) | (0.044) | (0.075) | (0.077) |
| $R^2$ | 0.112 | 0.171 | 0.118 | 0.179 | 0.090 | 0.095 | 0.086 | 0.098 |
| ZIP FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | ✓ | | ✓ | | ✓ | | ✓ |

## Table IA3
## Robustness to Controlling for County Characteristics

This table shows the robustness of the results in Table 6 to controlling for the interaction between *Black* and county characteristics. Regressions are weighted by the number of white respondents to the Race Implicit Association Test during 2008–2019 period. Education gap is the difference between the share of white population with at least a high school and the share of Black population with at least a high school degree. Standard errors are adjusted for clustering by county. *, **, and *** indicate statistical significance at 10%, 5%, and 1%. Continuous variables are standardized so that their coefficients represent the effect of a one standard deviation change. $N = 9,980$

| | PPP | | Bank PPP | | Nonbank PPP | | EIDL | |
|---|---|---|---|---|---|---|---|---|
| | Explicit | Implicit | Explicit | Implicit | Explicit | Implicit | Explicit | Implicit |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Black | −0.183*** | −0.202*** | −0.255*** | −0.217*** | 0.071* | 0.014 | 0.034 | 0.023 |
| | (0.054) | (0.047) | (0.033) | (0.033) | (0.038) | (0.042) | (0.044) | (0.044) |
| Hispanic | −0.031 | −0.031 | −0.021 | −0.021 | −0.011 | −0.011 | 0.060*** | 0.060*** |
| | (0.022) | (0.022) | (0.028) | (0.028) | (0.006) | (0.006) | (0.017) | (0.017) |
| Asian | 0.002 | 0.002 | −0.058*** | −0.058*** | 0.060*** | 0.060*** | 0.004 | 0.004 |
| | (0.014) | (0.014) | (0.021) | (0.021) | (0.016) | (0.016) | (0.036) | (0.036) |
| Other | −0.007 | −0.007 | −0.041 | −0.041 | 0.034 | 0.034 | 0.053** | 0.053** |
| | (0.035) | (0.035) | (0.031) | (0.030) | (0.022) | (0.022) | (0.022) | (0.022) |
| Female | −0.010 | −0.010 | −0.027*** | −0.027*** | 0.016* | 0.016* | −0.003 | −0.003 |
| | (0.007) | (0.007) | (0.007) | (0.007) | (0.009) | (0.009) | (0.012) | (0.012) |
| Black × Bias | 0.027 | 0.091 | −0.126** | −0.125* | 0.154* | 0.216** | −0.039 | 0.098 |
| | (0.097) | (0.104) | (0.053) | (0.074) | (0.078) | (0.095) | (0.052) | (0.097) |
| Black × White population share | 0.105 | 0.081 | 0.051 | 0.057 | 0.054 | 0.023 | 0.052 | 0.006 |
| | (0.071) | (0.078) | (0.047) | (0.056) | (0.040) | (0.044) | (0.054) | (0.055) |
| Black × Ln(Population) | 0.236** | 0.219** | 0.107* | 0.132* | 0.129* | 0.087 | 0.024 | 0.006 |
| | (0.096) | (0.107) | (0.054) | (0.067) | (0.075) | (0.074) | (0.052) | (0.059) |
| Black × Income gap | 0.002 | −0.004 | −0.004 | −0.011 | 0.005 | 0.007 | 0.122** | 0.106** |
| | (0.049) | (0.050) | (0.036) | (0.037) | (0.036) | (0.034) | (0.046) | (0.046) |
| Black × Education gap | −0.096*** | −0.112*** | −0.058* | −0.040 | −0.038* | −0.072** | 0.006 | −0.015 |
| | (0.028) | (0.035) | (0.034) | (0.036) | (0.021) | (0.028) | (0.040) | (0.038) |
| Black × Unemployment gap | −0.014 | −0.007 | 0.010 | −0.001 | −0.023 | −0.006 | −0.022 | −0.016 |

*(Continued)*

**Table IA3**—*continued*

| | PPP | | Bank PPP | | Nonbank PPP | | EIDL | |
|---|---|---|---|---|---|---|---|---|
| | Explicit | Implicit | Explicit | Implicit | Explicit | Implicit | Explicit | Implicit |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | (0.046) | (0.043) | (0.050) | (0.049) | (0.033) | (0.039) | (0.052) | (0.049) |
| Ln(Seats) | 0.040*** | 0.040*** | 0.044*** | 0.044*** | −0.005 | −0.005 | 0.003 | 0.003 |
| | (0.006) | (0.006) | (0.006) | (0.006) | (0.004) | (0.004) | (0.005) | (0.005) |
| Ln(Firm age) | 0.034*** | 0.034*** | 0.041*** | 0.041*** | −0.007* | −0.007* | −0.015* | −0.015* |
| | (0.005) | (0.005) | (0.004) | (0.004) | (0.003) | (0.003) | (0.008) | (0.008) |
| Accepts credit cards | 0.166*** | 0.166*** | 0.136*** | 0.136*** | 0.030*** | 0.030*** | 0.089* | 0.088* |
| | (0.033) | (0.033) | (0.035) | (0.035) | (0.010) | (0.010) | (0.048) | (0.048) |
| Missing credit cards | −0.020 | −0.020 | −0.005 | −0.005 | −0.015 | −0.015 | −0.006 | −0.005 |
| | (0.033) | (0.033) | (0.025) | (0.025) | (0.014) | (0.014) | (0.012) | (0.012) |
| Ln(Reviews) | 0.036*** | 0.036*** | 0.035*** | 0.035*** | 0.000 | 0.001 | 0.018 | 0.018 |
| | (0.007) | (0.007) | (0.010) | (0.010) | (0.004) | (0.004) | (0.013) | (0.013) |
| Average rating | 0.000 | 0.000 | 0.002 | 0.002 | −0.001 | −0.001 | 0.043** | 0.043** |
| | (0.017) | (0.017) | (0.015) | (0.014) | (0.006) | (0.006) | (0.019) | (0.019) |
| No reviews | −0.070 | −0.070 | −0.057 | −0.056 | −0.014 | −0.014 | 0.093* | 0.095* |
| | (0.078) | (0.078) | (0.064) | (0.064) | (0.032) | (0.032) | (0.048) | (0.048) |
| Ln(Page views) | 0.005 | 0.005 | 0.005 | 0.005 | 0.001 | 0.001 | −0.006 | −0.006 |
| | (0.012) | (0.012) | (0.009) | (0.009) | (0.010) | (0.010) | (0.014) | (0.014) |
| Ln(Photos) | 0.039*** | 0.039*** | 0.049*** | 0.049*** | −0.010 | −0.010 | −0.009 | −0.009 |
| | (0.009) | (0.009) | (0.008) | (0.008) | (0.008) | (0.008) | (0.012) | (0.012) |
| Bank UCC loan | 0.040*** | 0.040*** | 0.063*** | 0.063*** | −0.023*** | −0.023*** | 0.054*** | 0.054*** |
| | (0.015) | (0.015) | (0.019) | (0.019) | (0.006) | (0.006) | (0.011) | (0.011) |
| Nonbank UCC loan | 0.048*** | 0.048*** | 0.013 | 0.013 | 0.035*** | 0.035*** | 0.106*** | 0.106*** |
| | (0.016) | (0.016) | (0.014) | (0.014) | (0.012) | (0.012) | (0.012) | (0.012) |
| $R^2$ | 0.171 | 0.171 | 0.180 | 0.180 | 0.103 | 0.103 | 0.098 | 0.098 |
| ZIP FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |