

**WORKING PAPER** · NO. 2020-63

# Assessing the Payroll Protection Program: A Framework and Preliminary Results

*John Barrios, Michael Minnis, William Minnis, Joost Sijthoff*

MAY 2020

# **Assessing the Payroll Protection Program: A Framework and Preliminary Results\***

**John Barrios**

University of Chicago Booth School of Business

**Michael Minnis**

University of Chicago Booth School of Business

**William Minnis**

Eastern Illinois University

**Joost Sijthoff**

University of Chicago Booth School of Business

May 13, 2020

## **ABSTRACT**

We develop a simple model to predict requests for the Payroll Protection Program (PPP) and compare these predictions to the actual allocations. The model suggests the amount of requested funds could total \$750 billion, though this is likely a high watermark conditional on several assumptions. The model also generates expectations by industry, state, and firm size, allowing us to assess model performance in the cross-section. The model performs reasonably well. Through the May 1 funding, the state-level cross-sectional model has an  $R^2$  of 99.3% and the average absolute prediction error across states is 6.4%. Interestingly, the prediction errors from the first funding round are significantly negatively correlated to the errors in the second funding round, revealing that the allocations were systematically different in the two rounds. Ultimately, the results suggest that the payroll-based model predicts PPP allocations well and that the funds were allocated as designed. One potential inference from these results is that critique about PPP allocations should be focused on program design rather than program execution. This analysis should be useful for subsequent studies assessing the performance of the PPP.

\*This is still a draft document and we welcome any comments (michael.minnis [at] chicagobooth.edu). We initiated this project by developing a spreadsheet to provide a framework to estimate the magnitude of the PPP, which we posted online on April 14 along with a discussion of factors which could affect the magnitude. We continue to update that tool online as we learn more about the PPP and the actual allocations. This document formalizes and discusses that work. The analysis in this version of the paper uses actual disbursements data from the Small Business Administration as of May 1, 2020. The spreadsheet tool is available online: <https://sites.google.com/view/ppptool/home>. For disclosure, we also provided commentary on the PPP in a Barron's op-ed on April 8, which can be located here: <https://www.barrons.com/articles/small-businesses-without-prior-debt-are-at-risk-of-missing-the-rescue-51586369598>. We thank Pete Lisowsky for initial discussions on this topic.

## 1. INTRODUCTION

As a response to the COVID-19 pandemic, congress passed the Coronavirus Aid, Relief, and Economic Security (CARES) Act on March 27, 2020. The Payroll Protection Program (PPP) is one of the critical policy responses in the act. The PPP is designed to provide funds to help small businesses cover payroll and other expenses (such as rent and utilities) for eight weeks during the pandemic. As part of the first phase of the CARES Act, the PPP received \$349 billion of funding. The depletion of these initial funds led to an additional infusion of \$310 billion to the program on April 24. Despite being allocated approximately \$660 billion in total, there have been two substantive concerns: (1) whether the PPP is sufficiently funded to meet the demands of small businesses and, if not, how much could potentially be requested; and, (2) whether the funds are being allocated according to the design of the program.<sup>1</sup> To provide insights into these two concerns, we use the program rules and publicly available payroll data to develop a simple model to estimate the potential magnitude of PPP requests in total and cross-sectionally by state, industry, and firm size. We then compare our estimates with actual PPP allocation data to examine how funds have been distributed. This exercise suggests that funds from the PPP have broadly been allocated according to the distribution of eligible payroll, essentially as designed. We also identify and examine interesting deviations from the model's predictions.

The design and implementation rules for the PPP are set by the Treasury and the Small Business Administration (SBA).<sup>2</sup> The primary directive of the PPP is to allow any U.S.-based firm with either less than 500 employees or otherwise meeting

---

<sup>1</sup> The range of publicly stated estimates is broad: from \$626 billion (<https://www.sbalenders.com/conclusions-drawn-349-billion-sba-loan-program/>) to over \$1 trillion (Consumer Bankers Association). See the discussion of the ranges in Newmyer, April 22, 2020, here: [https://www.washingtonpost.com/news/powerpost/paloma/the-finance-202/2020/04/22/the-finance-202-economists-and-banking-analysts-debate-whether-new-small-business-money-will-be-enough/5e9f7beb602ff10d49aeae5a/?tid=ss\\_mail](https://www.washingtonpost.com/news/powerpost/paloma/the-finance-202/2020/04/22/the-finance-202-economists-and-banking-analysts-debate-whether-new-small-business-money-will-be-enough/5e9f7beb602ff10d49aeae5a/?tid=ss_mail). For an initial assessment of how funds have been allocated across banks, see Granja et al. (2020): [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3585258](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3585258).

<sup>2</sup> These rules have been regularly updated over time through the use of both "Interim Final Rules" and "Frequently Asked Questions." See <https://home.treasury.gov/policy-issues/cares/assistance-for-small-businesses> for all Treasury documentation about the program.

the usual definition of a small business as established by the SBA to qualify for payroll relief.<sup>3</sup> Specifically, qualified small businesses can request 2.5 times its average annual payroll costs, which include employee benefits. A firm can request a maximum of \$10 million with the constraint that the eligible annual wages for any single employee considered in the payroll costs cannot exceed \$100,000. In addition to payroll costs, profits from both sole proprietors and farmers are eligible based on net income from Schedules C and F, respectively.

We map this set of rules to publicly available payroll and tax data to derive an upper bound PPP request estimate by industry, state, and firm size (which we detail in Section 2). We estimate that total claims could reach \$750 billion, but view this as a likely high-water mark for several reasons.<sup>4</sup> First, we assume a 100% participation rate across all potentially eligible firms. Noting that the Treasury recently emphasized the “economic need” criterion, the participation rate is likely to be lower. Second, we are unable to account for either the \$100K per employee cap on wages or the \$10 million per firm cap on total loan size. Third, the Census payroll data designates a “firm” by aggregating establishments up to the level of each state only, not nationally. Therefore, a firm that has fewer than 500 employees in more than one state will appear to be more than one small business (and thus eligible) when, in fact, it is a larger (ineligible) firm.<sup>5</sup> It should also be noted that our estimate could be pushed higher if, for

---

<sup>3</sup> There are notable exceptions to both the eligibility rules and application amounts which we discuss in the next section. For example, affiliation rules require companies that are controlled by other companies to aggregate all employees across affiliated companies. On the other hand, nonprofit organizations — which are typically not allowed to participate in SBA programs — are eligible to participate in PPP. See the SBA website for details about size standards:

<https://www.sba.gov/document/support--table-size-standards>.

<sup>4</sup> Given the fact that the disbursement rate of round two funding has slowed considerably since May 1, our view that our estimate is a high watermark seems justified ex-post. We also note that our estimate has varied over time as we have learned about the program rules and refinements needed to the model — e.g., as we have learned that finance companies and sole proprietorships are eligible. More recently, for example, our predicted value was \$730 billion but we refined the estimate by increasing the sole proprietor estimate to consider only Schedule Cs with *positive* net income (which the IRS SOI separates), not all aggregate Schedule C net income.

<sup>5</sup> This distinction between firms, establishments, and payroll may be one reason why estimates and analyses vary across research designs. For example, estimates of the PPP based on extrapolating the number of firms may overstate the value because the number of firms is actually smaller once multi-state firms are aggregated. Moreover, examining dollars of allocations per establishment does not alleviate this

example, our adjustments to include benefits or inflation assumptions are too low, or firms substantially game the system by appearing smaller than they are to gain eligibility.

The payroll data allow us to make predictions not only in the aggregate, but also by state, industry, and firm size. We use these cross-sectional predictions and the PPP allocations to-date to test the joint hypothesis that our model correctly incorporates the rules of the program and that the rules were followed in making allocations.<sup>6</sup> The results suggest that based on the distributions reported by the SBA through May 1 (which includes the full \$342bn of first-round allocations and \$176bn of the second round allocations) the PPP loans have been allocated similarly to the model's predictions.<sup>7</sup> For example, using the state-level allocations we find that the average absolute deviation between actual and predicted allocations is only 6.4%. Moreover, half of all states were within 5% of our predicted allocation and 82% of states were within 10%. Nevada has the furthest deviation (24% below our expectation).<sup>8</sup> This deviation may be because the Treasury did not clarify until April 24 that legal gaming operations (which are not typically eligible for SBA loans) would be eligible to participate in the PPP, causing Nevada to have a severe under-allocation in the first round relative to the model's predictions.

Our analysis of cross-sectional deviations of actual allocations from predicted suggests that PPP funds have been allocated in a manner consistent with the model's underlying assumptions based on the program's rules. That is, the PPP funds have

---

concern because, once again, dollars are allocated based on payroll, not establishments, and payroll per establishment varies across states and industries.

<sup>6</sup> Note, at this point, we cannot verify the aggregate predicted level because the program has not been completed.

<sup>7</sup> The Treasury and SBA have now updated the data through May 10, which we have not yet incorporated into this draft. We will do so in future iterations. Because the latest allocation figures have only increased to \$189 billion from the \$176 billion we include in our current analyses, the results will likely only change slightly.

<sup>8</sup> The District of Columbia has strictly the largest prediction error (negative), but the allocations are relatively small. We suspect, but do not test, that at least some of this prediction error is caused by ineligible lobbying organizations that the model does not take into account. Future analyses can consider this explanation.

been allocated based on firms' payroll. In addition to our main finding, we note a few nuances to the results. First, the prediction errors across states of the first round of funding were generally offset by the prediction errors of the second round. For example, New York and California received less than predicted in the first round, but more than predicted in the second. Second, the first round of funding showed that the Manufacturing and Construction industries received more than expected, while the Professional Services and Health Care industries received less. The SBA has not reported the second round allocation by industry; however, preliminary analyses (exploiting state-level differences in industry exposure) suggest that this across-industry disparity potentially reversed in a manner similar to the broad reversal across states. Third, analyzing the actual allocations by firm size is difficult because the SBA has only reported allocations by loan size and not firm size; however, our findings reveal that both tails of the firm size distribution appear to have been under-allocated relative to predictions. The deviation for the largest firms is particularly apparent in the second round of funding which suggests that pressure on larger companies to return funds or to not apply had an effect.<sup>9</sup> Fourth, we find that remaining deviations from modeled expectations are predictable based on the known flaws in the model. For example, because Schedule F (farmer tax schedule) data is not available since 2005, we do not include these amounts in the model. Consistent with this error, we find that states with a higher proportion of GDP deriving from Agriculture receive more PPP than the model predicts. We also find that states with higher average wages receive slightly less than predicted, which we suspect corresponds to those states having a higher proportion of employees who surpass the \$100K wage cap, which, again, the model does not incorporate.

---

<sup>9</sup> The under-allocation for the smallest firms may be more an issue of our prediction amounts than true underfunding. Specifically, sole proprietor income is difficult to predict and allocate by firm size. We discuss this issue in more detail below.

In sum, our findings suggest that the PPP funds were primarily allocated based on 2019's estimated payroll, which was how the PPP was essentially designed.<sup>10</sup> As such, the extent to which other possible objectives were not met (such as targeting certain regions, specific firm sizes, or particular industries more intensively), the issue is likely more the result of how the program was initially designed and less with how the program was ultimately carried out. It is important to emphasize that this does not mean that there was no fraud or that some firms which should not have received funds received them anyway, but rather at a macro level, the data suggests that funds broadly followed historical payroll. It is also the case that differences in disbursements between the first and second funding rounds appear to address allocation concerns relative to predictions. More conclusive inferences related to a specific allocation of funds or, perhaps more importantly, the economic effect of the distributions, will need to wait until the SBA reveals firm-level loan details and employment outcomes manifest.

## **2. ESTIMATING THE POTENTIAL SIZE OF PPP REQUESTS**

To estimate the potential size of the PPP we first identify eligible firms per the SBA's rules and guidelines and then estimate the sum of the maximum loans for these firms. Generally, eligible small businesses are those with less than 500 employees, however, there are several exceptions to this threshold which we discuss below. The maximum loan amount a firm is able to request is 2.5 times monthly payroll for the previous year, using a \$100,000 wage cap per employee, up to a maximum loan of \$10 million per firm.<sup>11</sup> The U.S. Census Bureau provides payroll data by state, industry, and employee number-based firm size in the Statistics of U.S. Businesses (SUSB) tables.<sup>12</sup> The most recent tables are available for the year 2017.

---

<sup>10</sup> Two factors may affect the ultimate distribution of funds — and, in particular, could affect the second round of funds. First, the negative publicity surrounding larger firms, especially restaurants and hotels, could cause firms to reconsider their requests. Second, the Treasury emphasized the “economic need” certification criterion on April 30 in FAQ #31.

<sup>11</sup> Seasonal businesses received an exception to using annual payroll figures.

<sup>12</sup> See: <https://www.census.gov/data/tables/2017/econ/susb/2017-susb-annual.html>.

Table 1 presents our estimated PPP calculations. We start with all payroll for firms with less than 500 employees and then make adjustments for certain industries that have higher limits for employee-based eligibility.<sup>13</sup> Importantly, we allow eligibility for all firms in sector 72, which includes restaurants and hotels, because of an exception granted by the Treasury specifically for this industry.<sup>14</sup> Second, the PPP rules allow for the consideration of benefits in addition to wages in calculating the eligible payroll costs, which the SUSB data does not include. We use the Internal Revenue Service's (IRS) Statistics of Income (SOI) data to estimate a gross-up for eligible benefits.<sup>15</sup> Specifically, we use aggregate tax return data which reports the deductions for total employee costs, Figure which include salaries, wages, retirement, benefits, and costs of labor. To derive the gross-up factor, we divide the total costs by the costs of wages only. Our estimated factor is 110% of the costs of wages. Third, sole proprietor net income is also eligible for PPP, so we add total net income from Schedule C provided by the IRS SOI. Because only individuals with positive Schedule C net income will apply, we use total aggregate net income conditional on those Schedule Cs with positive net income.

The first three steps provide a total annual eligible payroll amount for the year 2017. We then need to adjust this figure to derive our estimate of total potential PPP requests. First, not all firms in an industry are eligible, so we assume an industry participation factor. Specifically, in our default specification we assume 100% participation across all industries except NAICS codes 52 (Finance and Insurance) and

---

<sup>13</sup> The regulations state that all firms with less than 500 employees are eligible or those which are classified as small businesses according to the SBA's size standards. The size standards for approximately 300 six-digit NAICS codes surpass 500 employees. Size standards are also based on revenues and net income, but SUSB data limitations restrict our ability to account for these size standards.

<sup>14</sup> Essentially, the Treasury removed the employee size cap and made an exception to the affiliation rules for firms in industry 72 ostensibly because hotels and restaurants were badly affected by COVID-19 but are often "chains" which have less than 500 employees at any individual location, but collectively more than 500.

<sup>15</sup> See: <https://www.irs.gov/statistics/soi-tax-stats-business-tax-statistics>.

53 (Real Estate) as some of these entities have been deemed ineligible.<sup>16</sup> Next, because the most recent data is from 2017, but our eligible payroll for the PPP calculations is based on 2019 data, we include an inflation factor of 5%.<sup>17</sup> Finally, PPP requests can not exceed 2.5 months of payroll costs, so we divide our total by 12 and multiply by 2.5 to derive our upper bound on the total expected PPP requests. As Table 1 reveals, estimated PPP requests according to the model could total \$750 billion.

It is important to understand the sensitivity around the estimated total. Most issues and assumptions with the model make the estimated value too high, but others could suggest that the value is too low. Factors that suggest that our estimates are too high include: (i) our approach assumes that all firms apply and are eligible based solely on employee size. This does not consider various factors that could drive participation rates lower. For example, the Treasury has emphasized that firms must certify an economic need and that if they are able to get liquidity from other sources, then they are not eligible. Additionally, there are likely many firms that no longer demand funds because they no longer have employees or have otherwise decided not to apply. (ii) The SUSB data aggregates firms to the state level only. This means that some firms which appear to be small businesses (because they employ fewer than 500 employees in the data) are actually too large to participate when considering all establishments across states. (iii) We are also unable to account for affiliation rules. For example, if one person owns several different companies (e.g., a restaurant business and car wash business), these separate firms must be considered collectively for the purposes of

---

<sup>16</sup> To put it bluntly, 50% is a very rough estimate for industries 52 and 53. In our initial model specification we had 0% because typically firms in industries 52 and 53 are completely ineligible to participate in the SBA's 7(a) program. After the first round of PPP, we noted that firms in these industries were getting allocations, but at a lower rate than other industries. The participation factor is also intended to be used to adjust for other program rules, like the \$100,000 per employee and \$10 million in total cap. However, our default model does not make these adjustments because we do not have a good way to estimate them. Therefore, we view these estimates as likely upper bounds. Undoubtedly, one way in which our model is misspecified in the industry participation factor.

<sup>17</sup> The inflationary factor is intended to account for payroll growth from both an increase in wages and an increase in employment. One potential issue is that we assume this factor to be homogenous across states, industries, and firm sizes, which is likely not correct. Also note that the Treasury allows seasonal businesses to use a timeframe other than the full year 2019. Our estimates do not account for this seasonality.

eligibility, but it is likely that these are considered separate entities in the SUSB data. Relatedly, some firms are controlled by venture capital and private equity firms and are not eligible to participate. (iv) The PPP places a cap on the eligible wages per employee (\$100K) and the total requested amount per firm (\$10 million). Because the SUSB data is aggregated, we are unable to account for these caps which inflates our estimates. (v) We include all Schedule C sole proprietor net income, but some portion of this is “side income” incidental to the individual’s main occupation. A big uncertainty is how many Schedule C filers will request funds.

Factors that could push our estimated figure higher include: (i) Our assumption of payroll growth between 2017 and 2019 (which captures both employment and wage growth) could have been higher than our assumed 5%. (ii) Firms may find ways to “game the system” by either making their companies appear smaller and therefore eligible or commit outright fraud and make claims for nonexistent employees. Our estimates attempt to model legitimate claims. (iii) Firms in the finance, insurance, and real estate industries could participate at more than our assumed 50% participation rates. (iv) As noted above, we do not include Schedule F net income in our estimates. Considering all sensitivity factors, we suspect that the \$750 billion represents an upper bound, but there is a wide confidence interval around this figure. We now turn to examine the cross-sectional variation in predictions to assess how the model performs.

### **3. ASSESSING THE MODEL AND ACTUAL PPP ALLOCATIONS**

The SBA provides the actual PPP allocations to-date allowing us to compare the model predictions to actual results. Given the program is not yet complete, we cannot at this point examine total allocations or participation rates, but we can examine cross-sectional predictions. We also note that any test of the allocations is a joint test that the model incorporates the rules of the PPP and that the allocations were made according to the rules. We further note that because the SBA has provided more complete allocation data by state, our cross-sectional analyses focus on the state-level

allocations. However, we do some preliminary analysis of the allocations by industry and firm size.

### 3.1 State-Level Analysis

The SUSB provides payroll data by state allowing us to disaggregate the totals from Table 1 on a state-by-state basis. However, two data limitations prevent us from having the total by state equal the national total in Table 1. First, we cannot adjust for firm sizes which are eligible above 500 employees because employment-based size details are missing in the SUSB files at the state, 6-digit NAICS level.<sup>18</sup> Second, sole proprietor net income data is not available at the state level from the IRS SOI and thus the across state comparisons rely on SUSB payroll data for firms with less than 500 employees. Given that this portion of the sample represents more than 85% of the estimated PPP eligibility (and that the remaining portion is likely highly correlated with the distribution of SUSB payroll figures) and that we are only examining across state variation and not total figures, the across state share approximations should be very close to what they would be if we could use the full data.

Table 2 tabulates the predicted allocation shares for each state. We also provide the actual allocation shares per the SBA for both rounds and the cumulative through May 1. The last column displays the prediction error of our estimate, relative to actual allocations to date, where the prediction error is calculated as  $(\$ \text{ Actual Allocation} - \$ \text{ Predicted Allocation}) / \$ \text{ Predicted Allocation}$ .<sup>19</sup> The first observation to note in Table 2 is that the *cumulative* prediction errors are relatively small for most states, with a few outliers. Nevada has the largest negative deviation in actual allocation versus our prediction — with a substantial negative deviation in the first round and a slightly positive deviation in the second. This is potentially explained by the Treasury not

---

<sup>18</sup> We are in contact with the Census Bureau and we may be able to include this adjustment in future iterations.

<sup>19</sup> The predicted value for each state is a total, so we scale this total by the total amount allocated. For the first round, the scaling factor for each state is  $\$342 \text{ billion} / \$748 \text{ billion} = 45.7\%$ . For the second round, the scaling factor for each state is  $\$176 \text{ billion} / \$748 \text{ billion} = 23.5\%$ .

clarifying until April 24th that legal gaming operations, which are not typically eligible for SBA loans, would be eligible in the PPP. The cumulative absolute average prediction error is 6.4%, and excluding Nevada and the District of Columbia, it is 5.5%. The second observation from Table 2 is that the prediction errors are much larger in each of the individual rounds, compared to the cumulative prediction error. The average absolute prediction errors of rounds 1 and 2 are 20% and 28%, respectively. Moreover — and quite importantly — the prediction errors between rounds are highly negatively correlated, and only eight states have errors with the same signs in both rounds. For example, both California and New York — about which there was much discussion of being under-allocated in the first round — received much more than predicted in the second round. Figure 1 illustrates the prediction errors for both rounds. Note that for most states, the errors from the first round go in the opposite direction from the second.

We more formally analyze the fit of our model in Table 3. We regress the natural log of the actual PPP allocations on the natural log of our model's predicted PPP allocation by state. In Panels A and B, we examine rounds 1 and 2, respectively, and in Panel C we examine the cumulative distribution. First note that the fit of the model is quite high in all panels, but particularly high ( $R^2 = 99.3\%$ ) when examining the cumulative distribution to date. We further illustrate the fit across states in Figures 2 (A, B, and C), noting that most states fall very close to the prediction, particularly in Figure 2C, which is the cumulative amount.<sup>20</sup>

In each of the panels in Table 3, we also explore why states deviate from expectations. These variables are included to examine potential shortcomings of our model and how allocations changed across rounds. First, our model is unable to account for farmers who file Schedule F to claim PPP funds. Note in Table 2 that states typically associated with farming (e.g., Indiana, Iowa, Kansas, Minnesota, Nebraska,

---

<sup>20</sup> We also note that the Pearson correlation of the residuals between Rounds 1 and 2 is -90%, reflecting the reversal of the allocation disparity across rounds.

North Dakota, and South Dakota) have positive cumulative prediction errors.<sup>21</sup> In Table 3, we include a cross-sectional variable which is the Log(Gross Farm Income).<sup>22</sup> For Round 1 (Panel A) and cumulatively (Panel C), the coefficient on this variable is significantly positive, supporting the idea that at least some of the “abnormal” PPP allocated to these states is the result of farm income allocations. Another known problem with the prediction model is that we cannot account for the \$100K per employee wage cap. We predict that states with higher wages will appear to be under-allocated as a result of these caps. We include Log(Average Pay per Employee) to proxy for wage levels across states. Table 3, Panel C shows that the coefficient on this variable is significantly negative consistent with the higher wages creating a binding constraint in some states. Our model is also not able to account for the \$10 million cap in funds allocated per firm. States with a relatively higher proportion of large firms will, therefore, appear to be under-allocated. However, there is also an alternative hypothesis for this firm size explanation. It could also be that if large firms were able to more easily gain access to funds, it could be that states with relatively more large companies would appear to be overallocated. To examine this, we include a variable which is the percentage of total state payrolls which are in firms of 400 to 500 employees. Consistent with the assertions that larger companies were more successful in the first round but that the second round went predominantly to small companies, we see that the coefficient on the firm size variable is significantly positive (negative) in Panel A (B), but insignificant in Panel C, suggesting that the early advantage for larger firms was at least partially offset in the second round.

We conduct one more state-level cross-sectional test in Table 3. States differ in the relative industry concentration; moreover, we noted in the first round that certain industries appeared to be overallocated relative to our expectations. Therefore, some of

---

<sup>21</sup> The allocation based on farming is also likely one reason for initial reports of overallocation per establishment to states such as Nebraska and North Dakota versus New York and California in the first round of funding. As long as Schedule F farmers are not included in the number of establishments (e.g., they are not included in SUSB data), then allocations per establishment will appear high in states overrepresented with agriculture.

<sup>22</sup> See: <https://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics/>

the across-state differences in PPP allocations could be driven by heterogeneity in industry shares. To test this, we include the ratio of the payroll of NAICS sectors 23 (Construction) plus 31-33 (Manufacturing) divided by 54 (Professional Services) and 62 (Health Care) to capture the relative dominance of these industries across states. In Table 3, Panel C we see that this variable is significantly negative suggesting that some of the across state variation is related to industry heterogeneity. However, we also note that this appears to be reversing in Round 2 allocations.

Overall, we find that actual allocations to date correspond well to payroll-based predictions while we observe variation in the prediction errors in rounds one and two separately. We further note that the reversal between rounds could be the result of several factors. One possible explanation is simply that this is a natural mean reversion. Some states had more firms that were “ready to go” in the first round and in the second round those states just did not have as many firms needing to apply. Another possibility is that the changes the Treasury and SBA made between rounds could have affected the distribution. Yet another possibility is that larger banks, which were shown to have made an under allocation of funds in the first round (Granja et al. 2020), were better prepared in the second round. This, in turn, affected the allocations. Regardless of the explanation, the cumulative distribution to-date tracks very closely to our expectations because of reversals in the second round compared to the first. Collectively, our state-level analysis suggests that the cumulative actual allocations closely correspond to payroll-based predicted allocations, suggesting that PPP funds have been allocated following to the design of the program (i.e., according to the distribution of payroll records).

### **3.2 Industry-Level Analysis**

We next briefly examine the predicted and actual allocations across industries. Unfortunately, the SBA has only provided the actual industry-level allocations for the first round, thus limiting our analysis. Table 4 shows the amounts and shares of predicted and actual PPP allocations by industry. Excluding sectors 11 and 99, the

average absolute prediction error is 25%, which is similar to the errors we observed at the state level in the first round. A few sectors have substantial deviations. As mentioned previously, our model lacks data for farmers with Schedule F and indeed the Agricultural sector reports a large positive prediction error. In addition, the real estate sector has a large positive prediction error, which is likely because we made a particularly conjectural assumption on the participation rate of this sector. Finally, the model performs poorly in Sector 99, but this is a particularly small sector. Figure 3 graphically illustrates the PPP allocations and predictions by industry.

### 3.3 Loan-Size Analysis

We conclude our analysis by examining the allocations by firm size. Unfortunately, the SBA only provides data by loan size and not by company size. Therefore, we have to estimate a mapping from the firm size categories in the SUSB (which is based on the number of employees) to the SBA's actual PPP data (based on loan amounts). To create this mapping, we estimate the likely loan size category by using the average payroll per employee and the range of the number of employees in each industry by state by firm size category in the SUSB data set.<sup>23</sup> In addition to the SUSB payroll figures, we must allocate sole proprietor net income data across loan sizes. Unfortunately, the IRS SOI does not provide this data by employee size, so we simply assign half of the sole proprietor PPP eligible amount to the two smallest loan size categories.

---

<sup>23</sup> An example of our estimation procedure is as follows: for the size category 30-34 employees with an average annual payroll of \$50k, annual eligible payroll is between \$1.5m and \$1.7m, and maximum loan size between \$312,500 and \$354,167. In the current case, 90% of observation's payroll is attributed to \$150k - \$350k, and 10% to \$350k - \$1m. In addition to the allocation of payroll to the loan size categories, we are required to make several other assumptions in our estimation procedure. Specifically, we attribute the full amount of PPP eligible payroll of small business firms with more than 500 employees to loans in the larger than \$5m category. Moreover, depending on a company's salaries and wages, it is possible for such a firm to have a maximum loan amount lower than \$5m. At exactly 500 employees with no employees making over \$100k, the maximum loan amount is under \$5m if the average payroll is less than \$48k. In the SUSB data, the average payroll of firms with more than 500 employees is under \$48k for only 3 out of 358 state x 2-digit NAICS observations which make it unlikely this possibility substantially affects our estimates. Additional details of this distribution are available upon request.

Table 5 reports the allocations by loan size for both rounds and the model’s predictions based on the estimated mapping (Figure 4 also illustrates the distribution). Consistent with widespread reporting, the actual allocation across firm sizes changed substantially from the first round to the second. For example, while smaller companies (i.e., the first two loan size categories) received approximately 32% of the funds in the first round, they received 54% in the second. Comparing the allocation relative to expectations, it appears that both tails of the firm size distribution are under-allocated — i.e., both the largest and smallest firms appear to be getting less than expected. For larger firms, this may be the result of pressure to return funds or not participate; whereas smaller firms may have had more difficulty participating initially. Moreover, some may be reluctant to participate if they no longer have their employees. It is important to note regarding the small companies, though, that sole proprietor participation may also be responsible for some of this apparent under-allocation. For example, we assume 100% participation of sole proprietors and then evenly distribute these amounts to the smallest two buckets. If we remove sole proprietors, then the allocations to the small firms are consistent with expectations.

#### **4. CONCLUSION AND DISCUSSION**

The Payroll Protection Program is a massive governmental response to the COVID-19 epidemic. Two initial key questions are how big could the requests for funds be and how will those funds be allocated based on the way the program has been designed? We develop a simple model using publicly available payroll and tax data to establish a framework to answer these questions. Our preliminary findings lead to three points for future analysis of the PPP to consider. First, total requests could reach \$750 billion, surpassing the currently allocated \$660 billion, though our estimate has a high degree of uncertainty because of several assumptions we have to make. Our current view is that this figure is a high water market given that several assumptions bias this figure upward, but it could also be that some features — such as growth, benefit expenses, gaming the system by firms, etc. — could push the figure higher.

Unfortunately, testing the model’s predictions on the total allotment will be difficult if the \$660 billion is fully allocated and the program is not replenished. Moreover, we emphasize that the model’s \$750 billion estimate is not intended to be an expectation, but rather a framework for considering the amounts that get ultimately distributed.

Second, based on the publicly available data to date (for activity through May 1), the model works well in predicting the cross-sectional variation in PPP approvals. In short, the first order finding is that the program’s allocations — across states, industries, and firm sizes — generally follows the program’s design. The PPP has generally been allocated to firms based on payroll. From a broad perspective, this appears to be how the program was designed. We think, therefore, that such a model should serve as a reasonable launching point for future evaluation of the PPP. One potential inference from the preliminary results is that to the extent that policymakers are dissatisfied with how the funds were allocated, the initial place to look will be in the original program design and to a lesser extent in the program’s execution.<sup>24</sup>

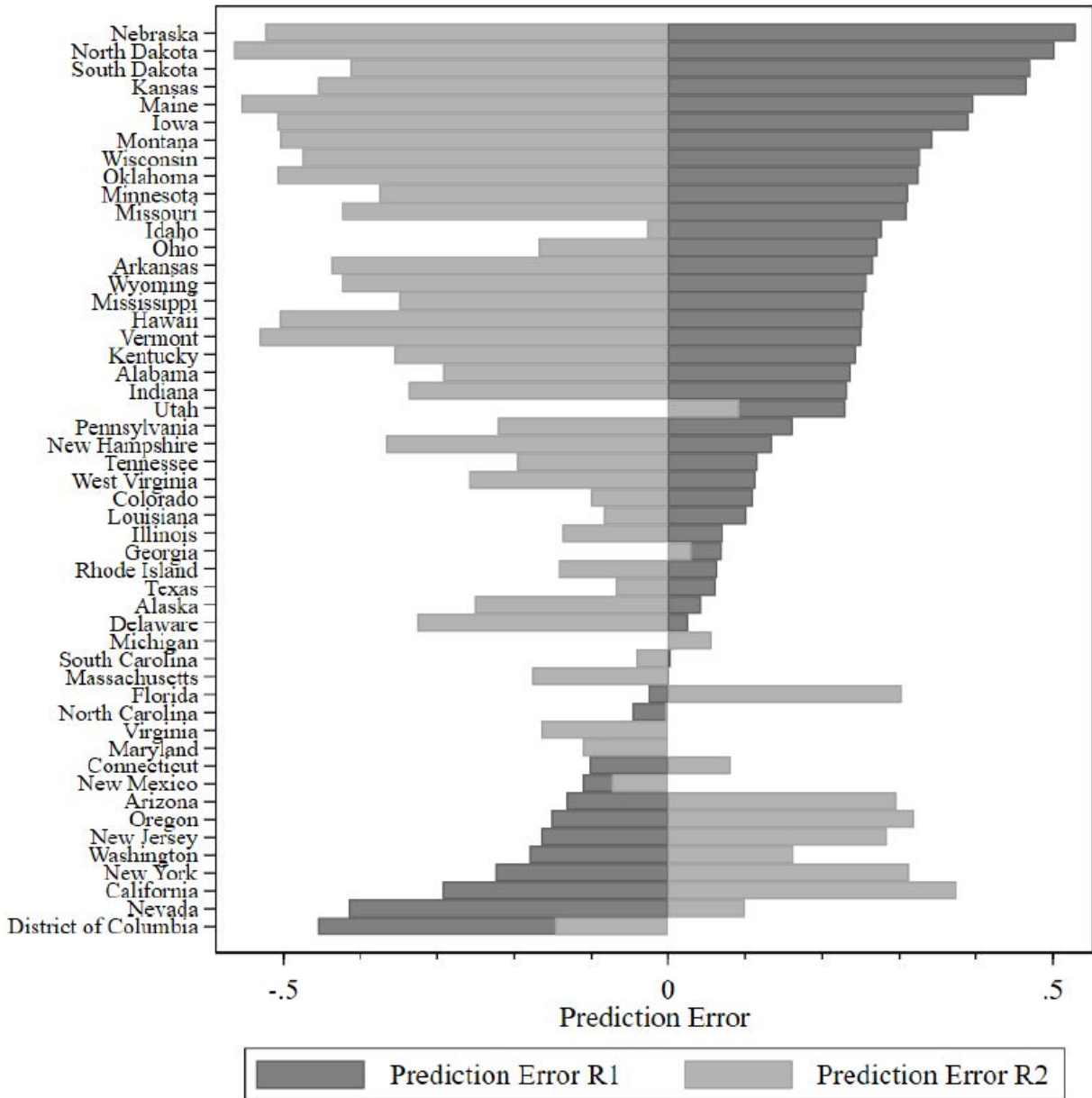
Third, our analysis has very little to say at this point about aspects such as fraud or even normative questions such as “how should the funds be allocated?” or “were they allocated to the right places?” Nor does this analysis answer very likely the more important questions such as “was the program effective?” or “did it prevent job losses?” Our objective to this point has been much more modest in trying to assess what the size of the program could be and the extent to which the funds were allocated given the set of rules that the Treasury and SBA provided. This should serve as a starting point to begin answering the other questions as data becomes available.

---

<sup>24</sup> Of course, the execution of the program is still open to critique and analysis, for example, with respect to how quickly the funds were disbursed, among other issues. Our point is simply that the ultimate allocations appear to be broadly in line with how the program was designed based on high-level estimates.

**Figure 1: Prediction errors by state and funding round**

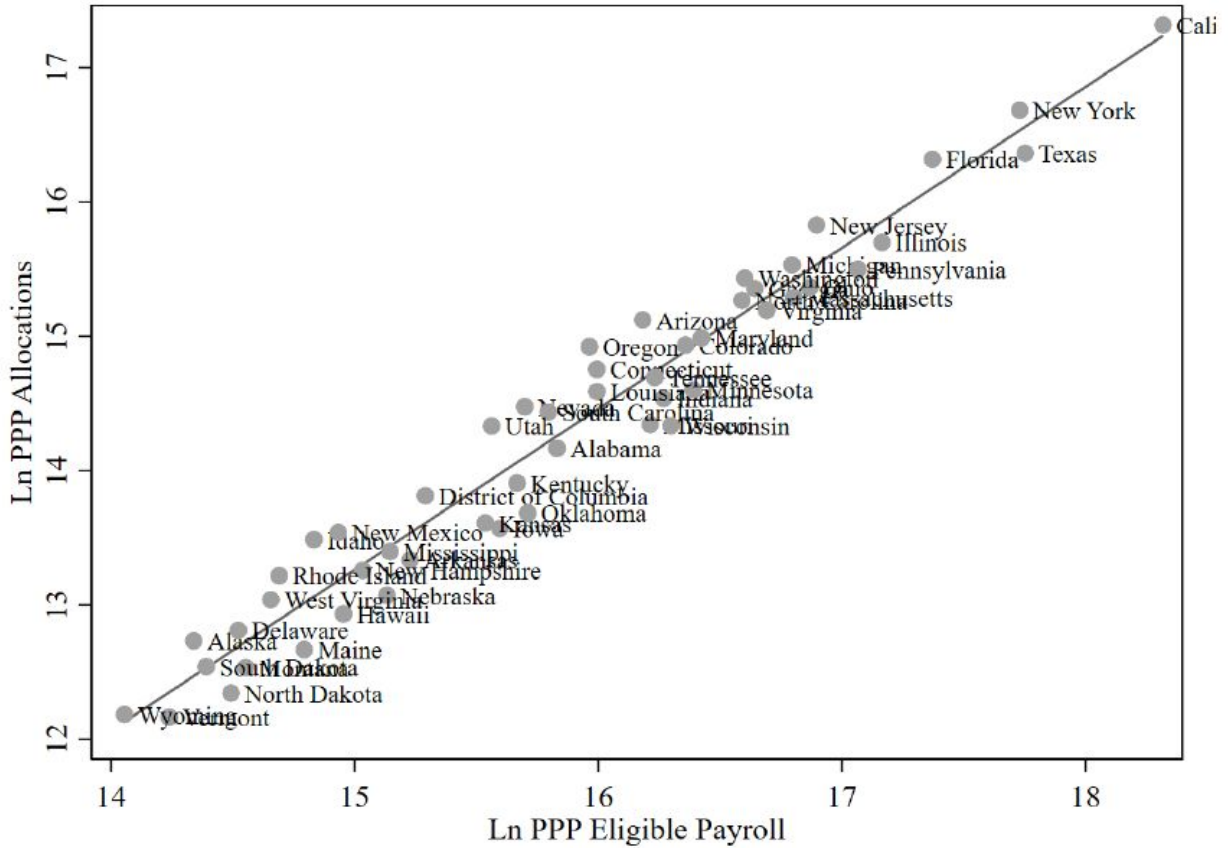
This figure shows the prediction errors (calculated as  $(\$ \text{ Actual} - \$ \text{ Predicted}) / \$ \text{ Predicted}$ ) for the first round (dark) and second-round (light). Bars to the right indicate states which appear to be “overfunded” relative to expectations, while bars the left indicate states which appear to be “underfunded.”





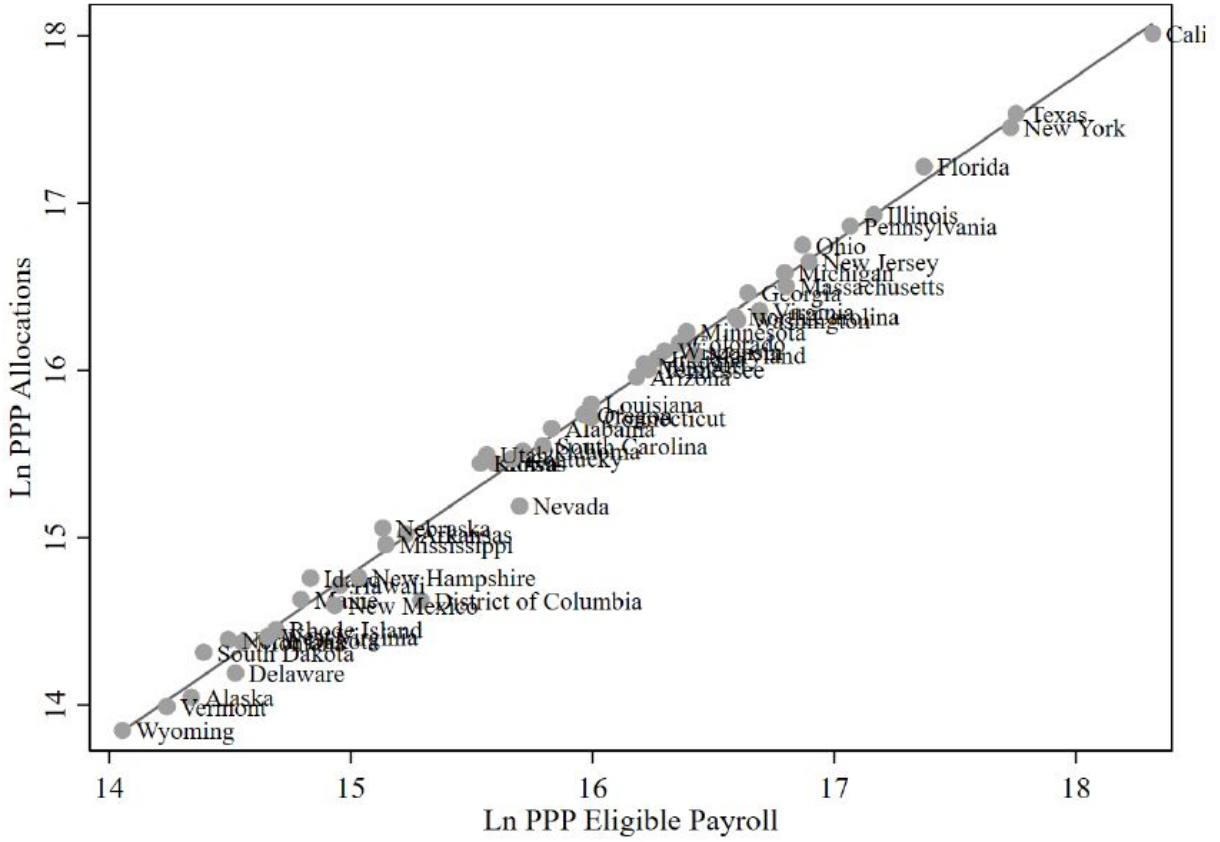
**Figure 2B: Predicted vs. actual by state in round 2**

This figure is a scatterplot of Log(\$ Allocated) on the y-axis and Log(\$ Predicted) on the x-axis by state in the second round of funding. States above (below) the linear fit received relatively more (less) funding than expected in the second round.



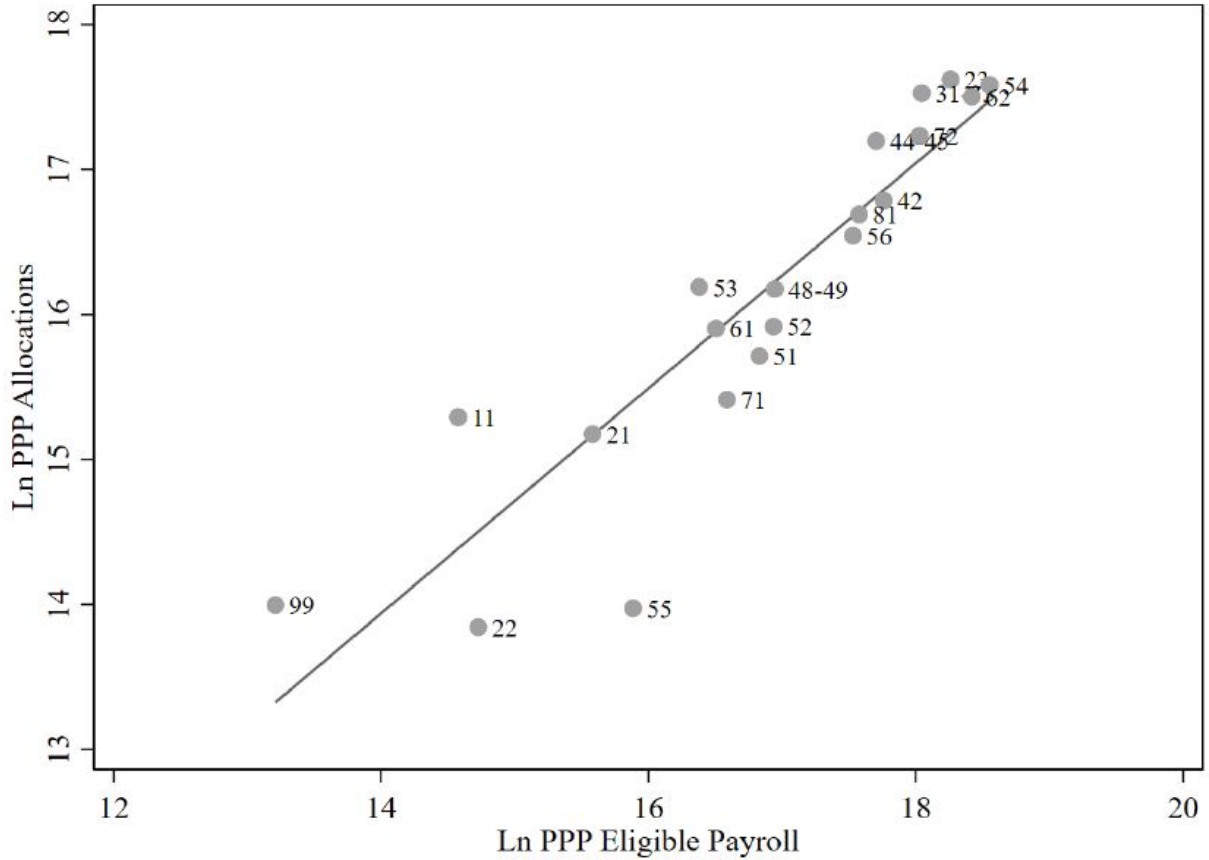
**Figure 2C: Predicted vs. actual by state in both rounds**

This figure is a scatterplot of Log(\$ Allocated) on the y-axis and Log(\$ Predicted) on the x-axis by the state for both cumulative rounds of funding. States above (below) the linear fit received relatively more (less) funding than expected in total.



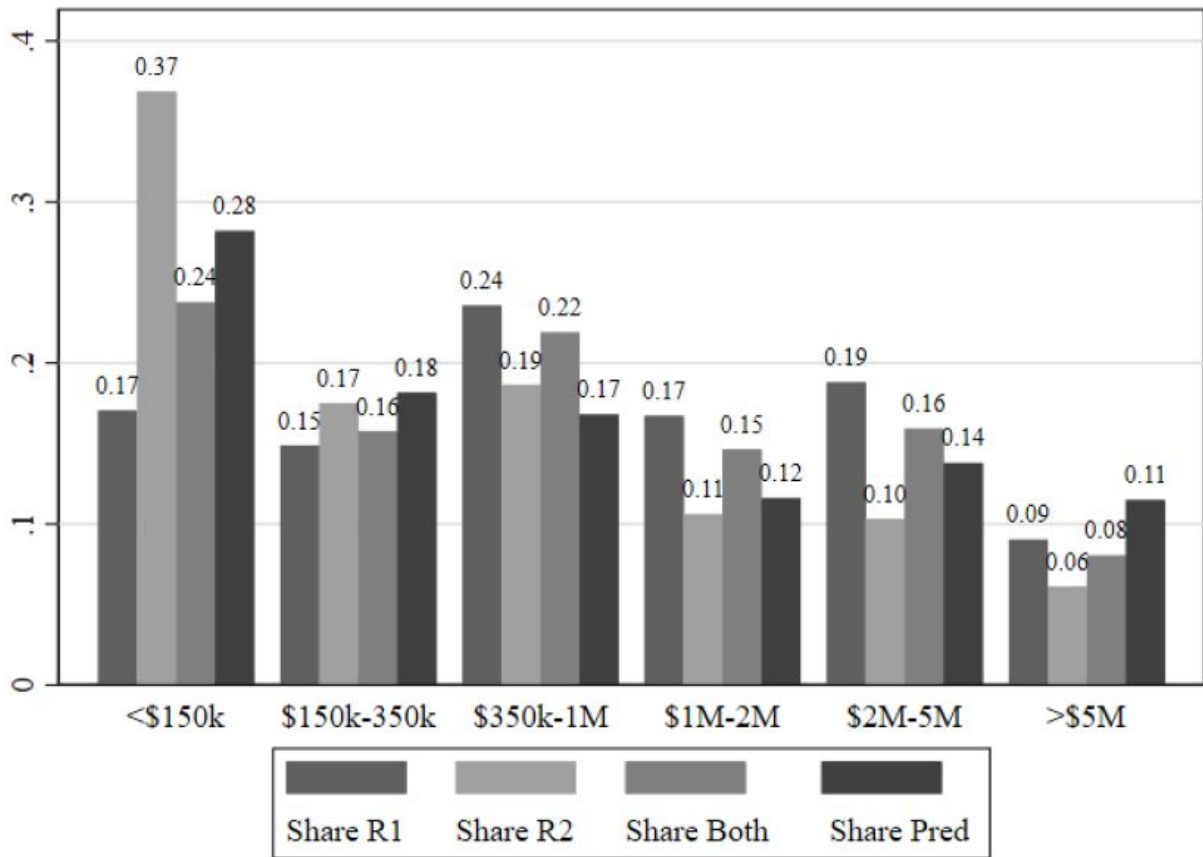
**Figure 3: Predicted vs. actual by industry in round 1**

This figure is a scatterplot of Log(\$ Allocated) on the y-axis and Log(\$ Predicted) on the x-axis by the industry for the first round of funding. Industries above (below) the linear fit received relatively more (less) funding than expected in the first round.



**Figure 4: PPP allocation by loan and firm size**

This figure presents the share of PPP funds allocated by loan size. The actual data for rounds 1 and 2 are provided by the SBA. The predictions are derived from our model. Because the model predicts amounts by firm size and not by loan size, we have to map firm sizes into loan sizes based on data about employee sizes and average wages per employee by industry and state. See the text for details.



### **Table 1: PPP Prediction Model Summary**

This table presents the estimated calculations of the potential total PPP eligibility requests. We begin with payroll data from the SUBS for firms with less than 500 employees and add to this payroll for firms in certain industries with an exception for greater than 500 employees. We then assume a gross up factor of 10% of payroll for benefits (e.g., healthcare and retirement) which we estimate using IRS SOI data. We then add sole proprietor net income per the IRS SOI data for Schedule Cs with positive net income. We further assume an eligible participation rate of 100% for all industries except finance, insurance, and real estate, for which we assume 50% eligibility. Because all figures are based on 2017 data, we assume 5% inflation for higher wages and employment in 2019. Finally, we divide by 12 and multiply by 2.5 to calculate the estimated PPP eligibility total.

NAICS	Industry	Payroll SUSB (in \$mm)	Payroll Adj>500 (in \$mm)	Benefits	Sole Prop Net Income (in \$mm)	Parti- cipation	Inflation	Annual Payroll (in \$mm)	Amount Predicted (in \$mm)
11	agriculture, forestry, fishing and hunting	5,849		10%	3,356	100%	5%	10,279	2,142
21	mining, quarrying, and oil and gas extraction	19,280	3,037	10%	2,222	100%	5%	28,109	5,856
22	utilities	8,673	1,476	10%	220	100%	5%	11,953	2,490
23	construction	309,515		10%	48,711	100%	5%	408,636	85,133
31-33	manufacturing	248,745	31,367	10%	5,923	100%	5%	329,750	68,698
42	wholesale trade	208,798		10%	6,596	100%	5%	248,088	51,685
44-45	retail trade	184,989		10%	19,947	100%	5%	234,607	48,876
48-49	transportation and warehousing	73,523	1,183	10%	22,297	100%	5%	109,697	22,854
51	information	75,056	5,551	10%	4,497	100%	5%	97,823	20,380
52	finance and insurance	160,853	4,184	10%	25,577	50%	5%	108,736	22,653
53	real estate and rental and leasing	73,808		10%	37,692	50%	5%	62,412	13,003
54	professional, scientific, and technical services	385,580	5,261	10%	91,412	100%	5%	547,405	114,043
55	management of companies and enterprises	32,928		10%	0	100%	5%	38,031	7,923
56	admin, support and waste management and remediation services	144,299		10%	28,948	100%	5%	197,061	41,054
61	educational services	55,567		10%	6,225	100%	5%	70,716	14,733
62	health care and social assistance	368,437		10%	51,167	100%	5%	479,270	99,848
71	arts, entertainment, and recreation	51,764		10%	16,184	100%	5%	76,780	15,996
72	accommodation and food services	274,536		10%	6,846	100%	5%	324,278	67,558
81	other services (except public administration)	145,903		10%	35,817	100%	5%	206,126	42,943
99	industries not classified	315		10%	2,151	100%	5%	2,622	546
Total:		2,828,417	52,059		415,789			3,592,380	748,413

**Table 2: Predicted and allocated PPP shares by state**

This table presents the predicted share of PPP eligibility by state and the actual allocations for both round 1 and round 2 through May 1. The prediction errors have been calculated as  $(\$ \text{ actual} - \$ \text{ predicted}) / \$ \text{ predicted}$ , where  $\$ \text{ predicted}$  has been scaled by a factor to account for the fact that only \$342 billion was allocated in the first round and \$176 billion in the second thus far.

State	PPP Share Predicted	PPP Share R1	Prediction Error R1	PPP Share R2	Prediction Error R2	PPP Share Both	Prediction Error Both
Alabama	0.0115	0.0142	0.2375	0.0081	-0.2926	0.0122	0.0582
Alaska	0.0026	0.0027	0.0421	0.0019	-0.2511	0.0024	-0.0571
Arizona	0.0164	0.0142	-0.1322	0.0212	0.2962	0.0166	0.0127
Arkansas	0.0063	0.0080	0.2663	0.0035	-0.4374	0.0065	0.0282
California	0.1384	0.0979	-0.2931	0.1903	0.3749	0.1291	-0.0671
Colorado	0.0195	0.0217	0.1098	0.0175	-0.1009	0.0203	0.0385
Connecticut	0.0136	0.0122	-0.1027	0.0146	0.0809	0.0130	-0.0406
Delaware	0.0031	0.0032	0.0265	0.0021	-0.3255	0.0028	-0.0926
D. of Columbia	0.0067	0.0037	-0.4553	0.0057	-0.1474	0.0044	-0.3511
Florida	0.0537	0.0523	-0.0263	0.0700	0.3030	0.0583	0.0851
Georgia	0.0259	0.0277	0.0691	0.0267	0.0303	0.0274	0.0559
Hawaii	0.0048	0.0060	0.2514	0.0024	-0.5050	0.0048	-0.0045
Idaho	0.0042	0.0054	0.2768	0.0041	-0.0282	0.0050	0.1736
Illinois	0.0437	0.0468	0.0713	0.0377	-0.1369	0.0437	0.0008
Indiana	0.0178	0.0219	0.2315	0.0118	-0.3376	0.0185	0.0390
Iowa	0.0091	0.0126	0.3898	0.0045	-0.5076	0.0099	0.0863
Kansas	0.0086	0.0126	0.4652	0.0047	-0.4553	0.0099	0.1538
Kentucky	0.0098	0.0122	0.2440	0.0063	-0.3567	0.0102	0.0408
Louisiana	0.0136	0.0149	0.1017	0.0124	-0.0836	0.0141	0.0390
Maine	0.0041	0.0057	0.3959	0.0018	-0.5545	0.0044	0.0744
Maryland	0.0208	0.0191	-0.0805	0.0185	-0.1107	0.0189	-0.0907
Massachusetts	0.0304	0.0303	-0.0008	0.0250	-0.1769	0.0285	-0.0604
Michigan	0.0302	0.0304	0.0071	0.0319	0.0564	0.0309	0.0238
Minnesota	0.0201	0.0264	0.3118	0.0126	-0.3747	0.0217	0.0796
Mississippi	0.0058	0.0073	0.2532	0.0038	-0.3490	0.0061	0.0495
Missouri	0.0169	0.0221	0.3102	0.0097	-0.4243	0.0179	0.0617
Montana	0.0032	0.0043	0.3427	0.0016	-0.5047	0.0034	0.0560
Nebraska	0.0057	0.0088	0.5292	0.0027	-0.5239	0.0067	0.1730
Nevada	0.0101	0.0059	-0.4150	0.0111	0.1000	0.0077	-0.2408
New Hampshire	0.0052	0.0059	0.1342	0.0033	-0.3666	0.0050	-0.0352
New Jersey	0.0334	0.0279	-0.1644	0.0429	0.2844	0.0330	-0.0126
New Mexico	0.0047	0.0042	-0.1109	0.0043	-0.0735	0.0042	-0.0983
New York	0.0768	0.0596	-0.2244	0.1009	0.3130	0.0736	-0.0426
North Carolina	0.0246	0.0234	-0.0457	0.0245	-0.0042	0.0238	-0.0316
North Dakota	0.0030	0.0045	0.5019	0.0013	-0.5642	0.0034	0.1413
Ohio	0.0325	0.0413	0.2715	0.0270	-0.1691	0.0365	0.1225
Oklahoma	0.0102	0.0135	0.3245	0.0050	-0.5076	0.0106	0.0430
Oregon	0.0131	0.0111	-0.1520	0.0173	0.3191	0.0132	0.0074
Pennsylvania	0.0396	0.0460	0.1611	0.0308	-0.2216	0.0408	0.0316
Rhode Island	0.0037	0.0039	0.0632	0.0032	-0.1430	0.0037	-0.0065
South Carolina	0.0111	0.0112	0.0037	0.0107	-0.0411	0.0110	-0.0114
South Dakota	0.0027	0.0040	0.4709	0.0016	-0.4127	0.0032	0.1720
Tennessee	0.0172	0.0192	0.1154	0.0138	-0.1968	0.0173	0.0098
Texas	0.0786	0.0834	0.0615	0.0733	-0.0676	0.0800	0.0178
Utah	0.0088	0.0108	0.2296	0.0096	0.0915	0.0104	0.1829
Vermont	0.0023	0.0029	0.2504	0.0011	-0.5308	0.0023	-0.0138
Virginia	0.0272	0.0255	-0.0614	0.0227	-0.1641	0.0246	-0.0961
Washington	0.0249	0.0204	-0.1800	0.0289	0.1620	0.0233	-0.0643
West Virginia	0.0036	0.0040	0.1129	0.0026	-0.2586	0.0035	-0.0128
Wisconsin	0.0184	0.0244	0.3271	0.0096	-0.4757	0.0194	0.0555
Wyoming	0.0019	0.0025	0.2576	0.0011	-0.4244	0.0020	0.0269

**Table 3: State level regressions examining the actual distribution of PPP**

This table shows regressions of the Log(\$ Actual) on Log(\$ Predicted) and other covariates by state. The additional covariates are Log(Gross farm income) per the USDA, Log(Average wages) per SUSB, percentage of total state payrolls which are in firms of 400 to 500 employees, and ratio of the payroll of NAICS sectors 23 (Construction) plus 31-33 (Manufacturing) divided by 54 (Professional Services) and 62 (Health Care). Panel A (B) includes data for the first (second) round, while Panel C includes data for the cumulative funding through May 1.

Panel A: Round 1

Outcome Variable	Ln PPP Amount 1	Ln PPP Amount 2	Ln PPP Amount 3	Ln PPP Amount 4	Ln PPP Amount 5
Ln Amount Pred	0.903*** -0.022	0.850*** -0.021	0.932*** -0.018	0.853*** -0.028	0.914*** -0.017
Ln Gross Farm Income		0.080*** -0.016			
Ln Avg. Pay			-0.602*** -0.157		
Large Firm Payroll Ratio				12.154* -4.793	
Ind Heterog Ratio					0.436*** -0.071
Constant	0.983** -0.343	0.612 -0.348	2.818*** -0.614	1.393*** -0.367	0.458 -0.29
Observations	50	50	50	50	50
Adj R2	0.963	0.974	0.967	0.967	0.97

Std Errors in Parentheses: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Panel B: Round 2

Outcome Variable	Ln PPP Amount 1	Ln PPP Amount 2	Ln PPP Amount 3	Ln PPP Amount 4	Ln PPP Amount 5
Ln Amount Pred	1.201*** -0.029	1.261*** -0.035	1.181*** -0.031	1.289*** -0.036	1.189*** -0.026
Ln Gross Farm Income		-0.089*** -0.025			
Ln Avg. Pay			0.411 -0.28		
Large Firm Payroll Ratio				-21.141** -6.415	
Ind Heterog Ratio					-0.475** -0.149
Constant	-4.758*** -0.472	-4.343*** -0.454	-6.012*** -0.998	-5.470*** -0.508	-4.184*** -0.433
Observations	50	50	50	50	50
Adj R2	0.955	0.962	0.955	0.963	0.96

Std Errors in Parentheses: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Panel C: Cumulative

Outcome Variable	Ln PPP Amount 1	Ln PPP Amount 2	Ln PPP Amount 3	Ln PPP Amount 4	Ln PPP Amount 5
Ln Amount Pred	0.987*** -0.01	0.962*** -0.01	1.002*** -0.008	0.975*** -0.013	0.992*** -0.008
Ln Gross Farm Income		0.037*** -0.007			
Ln Avg. Pay			-0.308*** -0.069		
Large Firm Payroll Ratio				2.838 -2.463	
Ind Heterog Ratio					0.197*** -0.051
Constant	-0.011 -0.162	-0.183 -0.148	0.928** -0.265	0.085 -0.167	-0.249 -0.138
Observations	50	50	50	50	50
Adj R2	0.993	0.995	0.994	0.994	0.995

Std Errors in Parentheses: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

**Table 4: Predicted and allocated PPP shares by industry**

This table presents the predicted share of PPP eligibility by industry and the actual allocations for round 1. The prediction errors have been calculated as  $(\$ \text{ actual} - \$ \text{ predicted}) / \$ \text{ predicted}$ , where  $\$ \text{ predicted}$  has been scaled by a factor to account for the fact that only \$342 billion was allocated in the first round.

Industry	NAICS	Amount Predicted (in \$mn)	PPP Share Predicted	PPP Share R1	Prediction Error R1
agriculture, forestry, fishing and hunting	11	2,142	0.0029	0.0128	3.4664
mining, quarrying, and oil and gas extraction	21	5,856	0.0078	0.0114	0.4543
utilities	22	2,490	0.0033	0.0030	-0.0977
construction	23	85,133	0.1138	0.1312	0.1534
manufacturing	31-33	68,698	0.0918	0.1196	0.3025
wholesale trade	42	51,685	0.0691	0.0569	-0.1755
retail trade	44-45	48,876	0.0653	0.0859	0.3161
transportation and warehousing	48-49	22,854	0.0305	0.0310	0.0140
information	51	20,380	0.0272	0.0195	-0.2838
finance and insurance	52	22,653	0.0303	0.0239	-0.2107
real estate and rental and leasing	53	13,003	0.0174	0.0314	0.8067
professional, scientific, and technical services	54	114,043	0.1524	0.1265	-0.1699
management of companies and enterprises	55	7,923	0.0106	0.0034	-0.6769
admin, support, waste management and remediation services	56	41,054	0.0549	0.0447	-0.1859
educational services	61	14,733	0.0197	0.0236	0.1966
health care and social assistance	62	99,848	0.1334	0.1165	-0.1264
arts, entertainment, and recreation	71	15,996	0.0214	0.0144	-0.3248
accommodation and food services	72	67,558	0.0903	0.0891	-0.0128
other services (except public administration)	81	42,943	0.0574	0.0517	-0.0984
industries not classified	99	546	0.0007	0.0035	3.7929
Total		748,413	1	1	

**Table 5: Predicted and allocated PPP shares by loan and firm size**

This table presents the share of PPP funds allocated by loan size. The actual data for rounds 1 and 2 are provided by the SBA. The predictions are derived from our model. Because the model predicts amounts by firm size and not by loan size, we have to map firm sizes into loan sizes based on data about employee sizes and average wages per employee by industry and state. See the text for details.

Loan Size	Amount R1 (in \$mn)	Amount R2 (in \$mn)	Amount Both (in \$mn)	Amount Predicted (in \$mn)	Share R1	Share R2	Share Both	Share Predicted
<\$150k	58,322	64,748	123,069	210,440	0.1704	0.3684	0.2376	0.2819
\$150k-350k	50,926	30,739	81,665	135,530	0.1488	0.1749	0.1576	0.1815
\$350k-1M	80,628	32,766	113,395	125,365	0.2356	0.1864	0.2189	0.1679
\$1M-2M	57,188	18,620	75,808	86,574	0.1671	0.1059	0.1463	0.1160
\$2M-5M	64,315	18,122	82,437	102,914	0.1879	0.1031	0.1591	0.1379
>\$5M	30,898	10,749	41,647	85,730	0.0903	0.0612	0.0804	0.1148