I. Introduction

The COVID-19 pandemic led to an unprecedented decrease in economic activity. As a response, various countries implemented loan guarantee programs. Benmelech and Tzur-Ilan (2020) report that government loan guarantees amounted to an average of 2.73% of GDP in the year 2020 across 85 countries, while total fiscal spending (excluding these guarantees) averages 4.97% of GDP. In particular, the United States Congress created the Paycheck Protection Program (PPP) as part of the larger Coronavirus Aid, Relief, and Economic Security (CARES) Act. The main goal of the program was to preserve jobs of SMEs that were substantially affected by COVID-19. In 2020 and 2021, the total volume of loans and grants made through the program was $800 billion. To speed up the delivery of loans, the government used financial institutions to make decisions on applications, but the loans were ultimately guaranteed by the government.[1] This arrangement gave banks the ability to disburse loans according to their incentives, particularly at the beginning of the program, when the demand for funds largely exceeded the supply.

The empirical literature provides robust evidence of targeting of PPP loans. Granja et al. (2022), Bartik et al. (2021), Doniger and Kay (2023), and Joaquim and Netto (2022) show that the earliest PPP loans were made to larger firms, firms with more preexisting debt, firms less affected by the pandemic, and firms that would have had a higher probability of survival without PPP loans. In this paper, we develop a model that captures the key mechanisms of the program and that is consistent with the empirical evidence. Using our model, we discuss what should have been the optimal target of PPP loans, examine the distortions caused by allocating these loans through the banking system, and discuss alternative policies that could have been implemented to minimize the misallocation of PPP loans. We also show how the model can be extended to accommodate informational differences between the government and banks, and the determinants of the optimal size of the program. Finally, while the paper is mainly focused on the PPP, we discuss how our framework can be mapped to other loan guarantee programs.

Our theoretical framework has three agents: firms, banks, and the government. Firms

[1] Throughout the paper, we refer to these intermediaries as banks for simplicity.

are heterogeneous in their number of employees, cash-on-hand, debt payments, and in how affected they are by the pandemic. For the latter, each firm in our model faces a random cost shock that must be paid for with its current cash-on-hand and potential borrowing from the PPP. We assume that, at the moment of application and loan disbursement, firms, banks, and the government know the distribution, but not the realization, of the shock.[2] In our benchmark model there are no application costs, and no asymmetric information or operational capacity differences between banks and the government. We also assume that banks are not lending outside of the program for firms that are eligible for the PPP.[3] We relax all of these assumptions in the extensions of our model.

We analyze the optimal allocation for a government whose objective is to maximize the number of preserved jobs, and characterize the optimal allocation of PPP funds in two steps.[4] First, we focus on the case where the government can choose how much to lend to each individual firm given the size of the program, which we denote as constrained first-best allocation. Second, we explore the case where the government must allocate loans to firms following the rules of the program, that is, only choosing which applications (that conform to the rules of program) to approve. Then, we assess potential misallocation from deploying these loans through banks.

In the constrained first-best case, the optimal allocation targets firms for which each additional dollar's marginal effect on the firm's survival probability is higher. For firms that face the same distribution of the cost shock, this implies the government finds that it is optimal to allocate money to firms with low cash-on-hand or high debt levels. On the other hand, when firms have similar financial conditions but face different cost shock distributions (due to different sectoral or regional exposure), the optimal allocation across

[2]This is a reasonable assumption, since the PPP was first introduced on March 27, 2020, long before the full magnitude of the pandemic was known. For example, a survey of small businesses in Bartik et al. (2020) shows that there was substantial disagreement and uncertainty regarding the duration of the crisis.

[3]This assumption is broadly consistent with the evidence in Li and Strahan (2022) that throughout 2020 virtually all of the change in banks' outstanding C&I loans can be accounted for by the PPP. In other countries, however, there is substantial lending outside of loan guarantee programs (e.g. Altavilla et al. (2022), Jimenez et al. (2022)). We provide an extension of our model in Appendix E which incorporates bank lending outside of the loan guarantee program, and show that it is consistent with the empirical evidence in Jimenez et al. (2022).

[4]The stated objective of the program was to "provide a direct incentive for small businesses to keep their workers on the payroll". See https://www.sba.gov/funding-programs/loans/covid-19-relief-options/paycheck-protection-program/first-draw-ppp-loan.

firms depends on the size of the program. When the program is relatively small, the government finds it optimal to allocate funds in an inverted U-shaped pattern with respect to shock exposure; that is, firms moderately affected by the pandemic will be targeted first. Intuitively, saving the most affected firms is not cost effective (as those funds could be used alternatively to save a larger number of firms which are only moderately affected). As the program grows larger, the government allocates more funds to the firms that are most affected. This analysis highlights that the target of the PPP is not ubiquitous: It depends on the nature of the shock faced by firms, the size of the program, and the dimensions in which firms are heterogeneous. The optimal allocation is a joint product of firms' financial positions and the nature of the shock distribution, and thus does not have a distribution or model-free ranking.

We then consider the case where the government can only choose which applications to approve subject to a constraint on the amount of loans allocated to the program. In this case, it is optimal to allocate loans to firms with the highest difference between their probabilities of survival with and without a PPP loan, that is, those with the highest treatment effects. In our model, these firms are those that are moderately affected by the pandemic. Firms that are expected to be severely affected by the pandemic will likely shut down, while firms not meaningfully affected by the pandemic shock will likely survive regardless of the allocation of PPP loans. Additionally, we also find that the government wants to allocate funds to firms in better or worse pre-COVID financial condition depending on the underlying distribution of the pandemic shock and the maximum loan size per firm. For instance, if the distributions has a lot of weight on the right tail, that is, large pandemic shocks are relatively more likely, the government prefers to allocate funds to those firms in better financial conditions ex-ante. The intuition is that these are the firms with the highest gain from receiving a loan, since firms in bad financial conditions would likely not survive the pandemic regardless of receiving or not a loan.

Next, we focus on the allocation of PPP loans through banks. The use of banks to distribute loans makes the allocation sensitive to banks' incentives — even in a program like the PPP where loans are fully guaranteed by the government. We assume that banks are choosing which loans to approve to maximize its profits. The banks' profit function is different from the government objective for two reasons. First, banks have outstanding loans with firms that will default if those firms do not survive the pandemic. As a result, banks distort the allocation toward firms with whom they already have outstanding loans. Second, banks value future relationships with current and potential clients. As a result, banks distort the allocation toward less affected firms, which are more likely to survive and thus be profitable clients in the future. Both of these distortions are consistent with the empirical literature. For instance, Bartik et al. (2021) show that approval rates are higher for less distressed firms and borrowers with larger outstanding debt.

To illustrate how our framework can be combined with micro data to generate insights about the misallocation of funds, we provide an empirical application of our model. For that, we combine loan level data from the SBA with treatment effect estimates from Autor et al. (2022*a*) and Dalton (2021). We find that the program was initially 50% less effective than it could have been under the optimal allocation. This reflects an excessive demand for loans early in the program, when banks played an important role in the allocation of PPP loans. This difference in effectiveness shrunk as more loans were made. By the end of the program, the bank based allocation represents a decrease in 17% in effectiveness relative to the optimal allocation. Additionally, we compute what the misallocation would have been if the government allocated loans to the smallest firms earlier. We find that if smallest firms were prioritized the program would have been much more effective, very close to optimal even in the first round of the program. This follows from the fact that smallest firms are those with the highest treatment effects (Bartlett and Morse (2021), Dalton (2021)), but they received loans later in program.

Finally, we consider various extensions of our model. First, we consider a situation in which banks have more information and operational capacity to timely disburse loans relative to the government, and assume firms without access to banks have larger treatment effects. Within this framework, we characterize situations in which the government would still prefer to allocate funds itself, rather than to rely on banks. In particular, if the cost of delayed disbursement is relatively low compared to the asymmetric information problem, then the government would rather disburse funds by itself. Second, we explore the determinants of the the optimal size of the program. We find that the optimal size depends on the tradeoff between the governments' cost of funds and benefit of lower misallocation from smaller effects of targeting over time. Intuitively, the more firms receive funds, the smaller the misallocation problem that arises when the program is too small and only a specific subset of firms is able to obtain funds. Lastly, we discuss how characteristics of loan guarantee programs implemented in other countries could be incorporated in our framework. We discuss how these differences affect both the subset of firms who apply for loans and the incentives of banks. We provide a detailed discussion of one specific difference that is relevant for various loan guarantee programs: the amount of loans made outside of the program and its substitutability with guaranteed loans. We show that a simple extension of our model that features lending outside of the program captures the key empirical facts of Altavilla et al. (2022) and Jimenez et al. (2022) for loan guarantee programs implemented in Europe.

Our paper contributes to the literature on the optimal design of COVID-19 corporate support policies. The literature has focused on how government support can address various frictions and externalities, such as excessive business liquidation, congestion in the bankruptcy system, aggregate demand, corporate and financial sector doom loops, COVID-19 transmission, and others.[5] Closer to our paper, Gourinchas et al. (2022) develop a macro framework with heterogeneous firms and show that the effectiveness of lending programs from a macro perspective depends fundamentally on the micro allocation of loans across different types of firms. While this literature has provided various insights based on broad abstractions of these corporate support programs, we take a different approach and build a framework that zooms in on the details and implementation of a particular program. By focusing on a specific program, we can understand the observed allocation of funds, how it depends on the incentives of banks (which in most cases intermediated the disbursement of funds), and precisely point out how changes in the policy design could have optimized the program's resource allocation.

[5]See, for instance, Elenev, Landvoigt and Nieuwerburgh (2020), Kahn and Wagner (2021), Philippon (2021), Goodhart, Tsomocos and Wang (2023), Wang and Wang (2021). Additionally, Crouzet and Tourre (2021) and Brunnermeier and Krishnamurthy (2020) link pandemic driven interventions to corporate debt overhang. Finally, Greenwood, Iverson and Thesmar (2020), Blanchard, Philippon and Pisani-Ferry (2020) and Hanson et al. (2020) discuss various broad based policies, such as the Primary and Secondary Market Credit Facilities, wage subsidies, unemployment benefits and corporate restructuring.

This paper also joins the large literature exploring the economic impact of policy responses to the COVID-19 pandemic, in particular the impact of the PPP.[6] Among those empirical studies, Doniger and Kay (2023), Bartik et al. (2021), Granja et al. (2022), Li and Strahan (2022), Joaquim and Netto (2022) and Glancy (2023) show evidence of systematic heterogeneous allocation of PPP loans. For instance, Bartik et al. (2021) show that banks prioritize loans to their pre-existing customers. Moreover, Granja et al. (2022) show that PPP loans did not flow to the areas most affected by the pandemic. Our contribution to this literature is twofold. First, we provide a theoretical model of the program that is consistent with the empirical evidence discussed above, and that allows us think about counterfactual analysis and alternative policy designs. Second, the empirical literature on the PPP has so far focused on the observed allocation of PPP loans and its effects. In contrast, we focus on what is the optimal allocation of loans and grants made through the program and, thus, what *could have been* the effect of the program had it been designed and implemented differently.

Finally, our paper is also related to the literature on loan guarantee programs, a common form of intervention in credit markets (Beck, Klapper and Mendoza, 2010). Benmelech and Tzur-Ilan (2020) and Hong and Lucas (2023) show that these programs were widely implemented in response to the pandemic. As a result, there is a burgeoning literature that studies these programs in various settings, such as Altavilla et al. (2022), Huneeus et al. (2022), Cirera et al. (2021), Core and Marco (2022) and Jimenez et al. (2022). We bring to the forefront the role of bank incentives in the disbursement of guaranteed loans and how this can lead to misallocation. Similar to our paper, Martín, Mayordomo and Vanasco (2024) also study banks' incentives to allocate loan guarantees across heterogeneous firms, and highlight how these incentives can lead to misallocation. In their project misallocation arises as guarantees distort bank incentives to promote entrepreneurial effort through lower interest rates. In contrast, our model has misallocation along the extensive margin arising from banks targeting a suboptimal set of firms due to

[6] A non-exhaustive list of the growing literature exploring the PPP includes Barrios et al. (2020), Chetty et al. (2020), Cororaton and Rosen (2020), Hassan et al. (2020), Hubbard and Strain (2020), Neilson, Humphries and Ulyssea (2020), Balyuk, Prabhala and Puri (2021), Bartlett and Morse (2021), Chodorow-Reich et al. (2021), Dalton (2021), Duchin and Hackney (2021), Faria-e-Castro (2021), Faulkender, Jackman and Miran (2021), Autor et al. (2022*a*), Erel and Liebersohn (2022) and Rabetti (2022).

their incentives to use guaranteed loans to reduce their credit risk. Moreover, by explicitly modelling a scenario where banks and the government have different information and operational capacity, we shed light on the circumstances under which guarantees should be allocated through banks.

II. The Paycheck Protection Program

Created on March 27, 2020, as part of the Coronavirus Aid, Relief, and Economic Security (CARES) Act, the Paycheck Protection Program (PPP) was designed to address liquidity shortages that could lead to employment losses from small businesses. The Small Business Administration (SBA) oversaw the program. To guarantee a timely disbursement of funds, firms applied for a loan through qualified financial intermediaries.

Through 2020 and 2021, the PPP disbursed loans in two separate draws. The first draw ran from April 3 through August 8, 2020, and it is the one we consider in this paper.[7] Given the PPP's small-business focus, only firms with fewer than 500 employees were eligible to apply, and each firm could apply for no more than one loan in the first draw of the program.[8] The maximum loan amount was 2.5 times the firm's average monthly payroll costs in the preceding year, up to $10 million. PPP loans have an interest rate of 1 percent, deferred payments for six months, and maturity of two years for loans issued before June 5 and five years for loans issued after June 5, 2020. Finally, PPP loans do not require collateral or personal guarantees.

A PPP loan was fully forgiven if funds are used for the purpose of payroll maintenance. To obtain full loan forgiveness, businesses were required to use most of the loan amount on payroll expenses and to maintain pre-crisis employment headcount and wage levels. The amount forgiven is reduced if wages or full-time headcount decreases. Initially, funds had be used to pay for these costs over the eight-week period following the disbursement of the loan. This period was eventually extended to 24 weeks in June 2020.

[7] In December 2020, Congress authorized an additional $284 billion in funding for the program as part of the $900 billion Coronavirus stimulus package. The PPP resumed disbursement in 2021, including second-draw loans for some of the firms that had received a PPP loan in the first draw.

[8] Exceptions to the size limit were firms in the restaurant and hospitality sectors (NAICS code 72), which were allowed to apply as long as they had no more than 500 employees in each location.

Recent SBA data indicates that over 90% of the funds disbursed in the program were forgiven.

Applications were processed by financial intermediaries (e.g. federally insured depository institutions and credit unions), which were responsible for checking application documents (for simplicity, we refer to these intermediaries as *banks*). Banks were paid a fee by the government to cover these processing costs. PPP loans are fully guaranteed by the government and carry zero risk weight for the calculation of risk-weighted assets, with the purpose of minimizing the impact on banks' capital requirements.[9] The program was designed to allow many institutions to process loan requests while minimizing impacts on their balance sheet structure.

Figure 1 shows the approval date of PPP loans through August 8, 2020, the application deadline of the first draw of the PPP. This first draw of the PPP, was composed of two separate rounds. The first round of the program ran from April 3 to April 16, 2020. During the first round, PPP loan demand vastly exceeded supply. We see in Figure 2 that 72 percent of firms reported applying for the program, but only 36 percent reported receiving a PPP loan at the end of the first round. This excess loan demand gave banks a significant role in deciding the allocation of PPP funds.[10]

As a result of the enormous demand, the PPP ran out of money on April 16, 2020. In response, on April 24, Congress enacted the Paycheck Protection Program and Health Care Enhancement Act, which appropriated an additional $321 billion (for a total of $670 billion) for the PPP. The second round of the program ran from April 27 to August 8, 2020. In the following two weeks after banks resumed accepting applications on April 27, an additional $176 billion in PPP loans was approved, and by May 15, 95 percent of the funds allocated in the program had already been dispersed. After that, demand for PPP loans subdued and led to excess supply of PPP funds, reducing the role of banks in the allocation of PPP loans in the second round. The program stopped accepting first

[9] Federal Reserve Banks were authorized to provide liquidity to banks through the Paycheck Protection Program Lending Facility (PPPLF). This allowed Federal Reserve Banks to extend loans to institutions that were eligible to make PPP loans using such loans as collateral.

[10] Press accounts in the early days of the PPP show that banks gave preference to their existing clients and larger firms. See, for instance, Emily Flitter and Stacy Cowley, "Banks Gave Richest Clients Concierge Treatment for Pandemic Aid"*New York Times*, April 22, 2020, updated Oct. 11, 2021. https://www.nytimes.com/2020/04/22/business/sba-loans-ppp-coronavirus.html.

draw applications on August 8, 2020, with \$144 billion remaining from the Paycheck Protection Program and Health Care Enhancement Act appropriation.

Table 1 reports aggregate statistics of the first draw of the PPP. The program funded 5.147 million loans at a total value of \$526 billion. At the end of the first draw, the (cumulative) average loan size was around \$100,000 for firms that self-reported, on average, 11.8 jobs supported by the program. Overall, firms reported more than 61.1 million jobs as being supported (in a universe of 70 million jobs at firms eligible for the program; see Autor et al. (2022*a*)). Most of these loans were not made by the top four banks (in terms of 2019 assets). Together, these banks were responsible for around 36 percent of small-business loans (less than \$ 1 million) before COVID-19, but accounted for only 13% of the amount disbursed in the PPP.

III. The Optimal Allocation of PPP Funds

In this section, we describe the setting of our model, discuss the optimal allocation of a program's funds from the perspective of the government, and the potential misallocation from using the banking sector to disburse these funds.

III.1. Model Setup

We consider a continuum of firms indexed by j. Each firm has N_j workers. We define our model in terms of *per worker* variables. Firm j's available cash *per worker* before the onset of the pandemic and the launch of the lending program is given by (1)

$$c_j \equiv \rho_j - b_j, \tag{1}$$

where b_j represents the firm's debt payments per worker, and ρ_j is the remainder of the cash-on-hand. Without loss of generality, we normalize N_j such that $\int_j N_j dj = 1$. We model the pandemic following Guerrieri et al. (2022). Each firm faces a reduction ν_j in cash flow (revenue shortfalls, extra costs to remain open). That is, the per-worker magnitude of the shock is ν_j, with cumulative distribution denoted by Φ and probability distribution given by ϕ, both parameterized by η_j (we define the specific functional form for the distribution below).

A firm that borrows ω_j from the lending program *can* survive the pandemic if

$$\nu_j < c_j + \omega_j \equiv \Gamma_j(\omega_j),$$

where $\Gamma_j(\omega)$ is defined as the available funds per worker to guarantee firm survival. We assume that $\Gamma_j(0) > 0,\ \forall j$; that is, all firms across all sectors and regions are profitable enough before the pandemic to remain open.[11] Throughout the paper, we suppose all firms apply for the program, and do so for the maximum amount of funds per worker available, which we denote by φ. Given the small application costs and forgiveness rules, this is a reasonable assumption in our setting. In Appendix D, we endogeneize firms' application choice and the amount that they apply for in the program.[12]

For tractability, we follow Guerrieri et al. (2022) and assume that the cumulative distribution of the fixed-cost shock distribution is given by (2)

$$\Phi(\nu\ ;\ \eta) = \begin{cases} 0, \text{ if } \nu < 0 \\ \left(\frac{\nu}{c_0}\right)^{\eta}, \text{ if } \nu \leq c_0 \\ 1, \text{ if } \nu > c_0 \end{cases}, \text{ with } \eta > 0. \tag{2}$$

The distribution in (2) has two characteristics that greatly simplify our analysis while still allowing us to focus on the difference between bank and government incentives. First, the shape parameter η controls the concavity of the cumulative distribution function (CDF), and thus we have a monotonic probability density function (PDF), which is increasing if $\eta > 1$ and decreasing if $\eta < 1$. Second, a distribution with higher η first-order stochastically dominates a distribution with lower η, making it easier to compare more affected (higher η) with least affected (lower η) firms. We assume in our benchmark model that banks and the government also observe η_j (the parameter of the distribution)

[11] As our focus is on the allocation of funds across firms, it is natural to assume that firms that are not profitable before the pandemic will shut down and will not receive any funds from the PPP.

[12] For another approach to modeling firm demand for PPP loans, which includes costs of switching across banks, see Glancy (2023).

but not ν_j (the realization of the shock). We relax this assumption and solve a version of the model where the government has imperfect information and limited operational capacity to disburse loans in Section V.1.

We simplify the notation by defining $\Phi_j(\varphi) \equiv \Phi\left(\Gamma_j(\varphi)\,;\,\eta_j\right)$, which is the probability a firm survives the pandemic if it receives an amount φ from the program. Furthermore, to analyze the effects of the program on participating firms, we introduce two key variables in our model. Let θ_j be the probability of survival of a firm without receiving no loans from the PPP, and T_j be the treatment effect for firms of type j of receiving φ rather than 0 of loans from the PPP as in (4):

$$\theta_j \equiv \Phi_j(0) \tag{3}$$

$$T_j \equiv \Phi_j(\varphi) - \Phi_j(0). \tag{4}$$

The variable T_j plays a special role as it measures the expected effectiveness of the PPP program for each type of firm j.

Dynamics and Timing of PPP Allocation. Our model of the PPP is a static model in which we ask which firms should receive PPP funds. However, given the massive size of the program, a lot of the variation in PPP disbursements came from not only *which* firms received PPP loans, but also *when* those loans were received (see Doniger and Kay (2023), for instance). This additional time dimension also raises the question of *which* operational constraint prevents all firms that would be chosen to receive funds to obtain funds right at the onset of the program. Nevertheless, most of our results can be easily extended for the timing of allocations.

First, if we assume that firms' ability to survive or treatment effects do not change over time, the problems are equivalent. More broadly, we can assume that firms' ability to survive and treatment effects change over time but do so proportionally for all firms, and our results of who should get PPP loans easily translate to who should get loans earlier in the program. We return to this discussion in our quantitative exercise in Section IV and show how small changes to the framework can accommodate a dynamic setup.

III.2. Constrained First-Best Allocation

Our first theoretical result is based on the problem of the government when it can choose the amount ω_j^G per worker to lend to each firm. The objective of the government is to maximize the number of preserved jobs.[13] Since there is no intensive margin adjustment at the firm level (such as downsizing), we model this objective as maximizing the number of surviving jobs. The government observes the types of firms j but not their actual realization of the pandemic shock, so we denote the solution of this problem as the *Constrained First-Best*, since the government is constrained by its information set.

We show that the government wants to allocate funds to where the marginal effect is the highest, which does not necessarily correspond to the places most/least affected by the shock. The marginal effect depends on the distribution of the expected cash-flow shock, on the size of the program, and on the initial financial condition of the affected firms. The optimal target of the lending program thus is neither obvious nor invariant to the nature of the economic shock (as more information on the depth of the pandemic becomes available). For instance, if the shock to a firm (or region/sector) is large enough, the government does not always find it optimal to save this firm, as the opportunity cost of not allocating these funds to other firms is too high. Our analytical results below formalize this intuition and characterize what is then the government's optimal allocation.

Let the total amount of loans in the program be denoted by M. The problem of a planner who aims to provide a direct incentive for small businesses to keep their workers employed is given by:

$$\max_{\{\omega_j^G\}} \int N_j \cdot \Phi\left[\Gamma_j(\omega_j^G) \mid \eta_j\right] \; dj \;\; \text{s.t.} \quad \int N_j \cdot \omega_j^G \, dj = M. \tag{5}$$

The solution to the maximization problem in (5) is what we denote as the *Constrained First-Best*. In what follows we separately compare firms with the same shock exposure (η_j)

[13]This is a reasonable assumption, since employment losses would likely increase the effects of liquidity shocks, worsening economic outcomes.

but heterogeneous financial position first, and then firms with the same financial position (c_j) but different shock exposures. The motivation for this approach is twofold. First, it highlights the key channels of the PPP allocation in our model for different sources of heterogeneity across firms. Second, it speaks directly to the empirical literature that generally tries to control for either of these factors (with firm controls, fixed effects, etc.) and focus solely on one of them at a time (for instance, in Bartik et al. (2021)). Our main analytical result of this section is Lemma 1, which considers the case where $\eta_j < 1, \ \forall j$; that is, all firms face a concave distribution of the pandemic shock:

Lemma 1. ***Constrained first-best allocation with*** $\eta_j < 1$. *Let* $\eta_j < 1$, $\bar{c} \equiv \int_j N_j c_j dj > 0$ *and* c_0 *sufficiently large.*[14] *The solution to (5) entails an equal marginal probability of survival across firms, that is, for firms* i, j

$$\Phi\left[\Gamma\left(\omega_j^G\right)\right] = \Phi\left[\Gamma\left(\omega_i^G\right)\right], \ \forall i, j. \tag{6}$$

Using the distribution in (2), we have that

$$N_j \omega_j^{G,*} = N_j \tau(\eta_j, \overline{M}) + M - \left[N_j c_j - \bar{c}\right], \tag{7}$$

where $\overline{M} \equiv M + \bar{c}$ [15] *and* $\tau(\eta, \overline{M})$ *is an exposure-based per-worker transfer that sums to zero; that is,* $\int_j N_j \tau(\eta_j, \overline{M}) = 0$. *Furthermore, we have that:*

- $\overline{M}$ *small:* $\tau(\eta, \overline{M})$ *is inverted U-shaped in* $\eta \Rightarrow$ *funds should flow to intermediately affected firms.*
- $\overline{M}$ *large:* $\tau(\eta, \overline{M})$ *is strictly increasing in* $\eta \Rightarrow$ *funds should flow to most affected firms.*

Lemma 1 implies that for two firms i, j with the same shock exposure $\eta_j = \eta_i$, we have that $N_j \omega_j^{G,*} - N_i \omega_i^{G,*} = N_i c_i - N_j c_j$, while for two firms with the same financial position $c_j = c_i$, we have that $N_j \omega_j^{G,*} - N_i \omega_i^{G,*} = N_j \tau(\eta_j, \overline{M}) - N_i \tau(\eta_i, \overline{M})$. Intuitively, Lemma 1

[14] We assume that c_0 is sufficiently large to focus on the interior solution of the problem of the government. A sufficient condition for this interior solution is $c_0 > M + \bar{\pi}$.

[15] This is the relative size of the program. M represents the total funds to be allocated, and $\bar{c}$ represents the financial conditions of the firms at the onset of the shock.

shows that (i) the optimal policy equalizes the marginal probability of survival across firms who receive funds (Eq. (6)), and (ii) the firm specific allocation can be decomposed in the cash flow needs of firm j relative to the average cash flow needs in the economy and a transfer based on the size of the program and exposure to the shock (Eq. (7)). Therefore, firms in a more fragile financial situation (as are small firms) would receive more funds from the PPP. Moreover, if the relative size of the program is large, the government can allocate enough funds to the most affected firms to significantly increase their probability of survival. On the other hand, if the program is relatively small, the government must focus on firms that are intermediately affected by the pandemic. In this second case, the relative cost of saving the most affected firms is too high, as they are not the firms whose treatment effects are the largest.

Sectoral/regional allocation. Our result in Lemma 1 is also useful for analyzing the optimal allocation of funds across different sectors and regions in the country, as there is evidence that the funds did not flow to the most affected regions (Granja et al. (2022)). The result that optimal policy will equate the marginal probability of survival across firms in Eq. (6) is still true for the marginal probability of survival across sectors and regions. For instance, if different sectors have different initial levels of debt per worker, sectors with relatively *more* debt per worker should receive more of the funds, since the probability that a firm in this sector survives the pandemic absent the government program is small, hence the marginal effect of funds on survival probability is large. However, if sectors or regions have shocks with different distributions (that is, different exposures to the pandemic), the optimal transfers across sectors are given by $\tau(\eta_j, \overline{M})$, and should not necessarily go to the most affected sectors or regions.

Other cases. In Lemma 1, we focus on the case where $\eta_j < 1, \ \forall j$. In Appendix B.1, we show in a simple example that when $\eta_j > 1$, the problem of the government is convex and the solution is to allocate funds to either the firms with the lowest or highest π or η; that is, the government is indifferent to the choice between allocating funds to the least or most affected firms as long as all of the funds flow to either. More generally, take any distribution $\Upsilon\left(\pi_j + \omega_j \mid \theta_j\right)$ parameterized by θ_j. We show in Appendix B.2 that the government should allocate money to the highest θ_j (most affected) if Υ is supermodular in ω, θ (and to the least affected if it is submodular). Overall, this shows that the optimal target of PPP funds depends on the relative benefit of the marginal dollar rather than the funding needs of each individual firm.

III.3. Optimal Allocation of Loans under PPP Rules

In Section III.2, we considered the government allocation when the government can choose how much to lend to each firm. To make a direct comparison with the bank allocation, we now focus on the optimal government allocation under the same rules as the bank allocation in the PPP, conditional on the set of firms that in fact applied for the PPP. In this case, the government can choose to accept or reject applications from firms, but it cannot change the loan allocation at the intensive margin. This problem ensures that we are comparing the bank allocation with a government allocation equally constrained by the program (instead of the constrained first-best), and thus that any difference comes from banks' incentives.

The problem of the government is to choose the probability $l_j^G \in [0,1]$ to accept the application from firm j, as in Eq. (8):

$$\max_{\{l_j^G \in [0,1]\}} \int N_j \left[l_j^G \Phi_j^\Gamma(\varphi) + (1 - l_j^G) \Phi_j^\Gamma(0) \right] dj \;\; \text{s.t.} \;\; \int N_j l_j^G \, dj = \frac{M}{\varphi} \tag{8}$$

Distribution of firms/workers in the population. Note that so far we have used the shorthand notation of dj to represent the integral over the distribution of firms, but we haven't defined how types of firms are present in the population. To provide clarity, we discuss what is implicitly behind this notation. Let $G(\rho, b, \eta, N)$ be the joint distribution of ρ, b, η, N in the population of firms. For any variable at the firm level that is not a function of the number of employees, $x(\rho, b, \eta)$, we can write:

$$\int N x(\rho, b, \eta) dG(\rho, b, \eta, N) = \int x(\rho, b, \eta) \overline{N}(\rho, b, \eta) dG(\rho, b, \eta), \tag{9}$$

where $\overline{N}(\rho, b, \eta) \equiv \int_N N dG(N \mid \rho, b, \eta)$. The term $\overline{N}$ is the average number of employees of firms of a given type $\{\rho, b, \eta\}$, and it acts in our model as a shifter in the distribution of firms of type $\{\rho, b, \eta\}$. What matters in our model is not the marginal distribution of firms, but rather the marginal distribution of the firm variables at the *job* level. Thus, our model can encompass various other channels highlighted in the literature that focus on jobs and not firms, such as how firm survival varies by firm size (Bartlett and Morse (2021)).

As in our model, the treatment effect T_j is not a direct function of N_j, and the resource constraint is linear in it. The optimal allocation l_j^G is also not a function of N_j, where dj in this case represents the integration of all firms of types $\{b_j, \rho_j, \eta_j\}_j$, with the cumulative distribution of type j given by $G(\rho, b, \eta) \times \overline{N}(\rho, b, \eta)$. Following the argument in Eq. (9), we can write the problem of the government as Eq. (10)

$$\max_{\{l_j^G \in [0,1]\}} \int l_j^G T_j \, dj \;\; \text{s.t.} \;\; \int l_j^G \, dj = \frac{M}{\varphi}, \tag{10}$$

where dj in this case represents the integration of all firms of types $\{b_j, \rho_j, \eta_j\}_j$ with the cumulative distribution of type j given by $G(\rho, b, \eta) \times \overline{N}(\rho, b, \eta)$, that is, the job-weighted distribution, as in Eq. (9). The exact same argument can be made for everything that follows in this paper, thus, in what follows, we leave the dependence on N_j implicit.

Importantly, our empirical application does incorporate treatment effect heterogeneity for firms of different sizes. This can be reconciled with our theoretical setup by assuming other firm characteristics, such as firms' debt per worker b, financial position ρ, or their exposure to the shock η, which are relevant determinants of T_j in our setup, correlate with firm size. In other words, while T_j is not a *direct* function of N_j, it will nevertheless be related to size as firms of different sizes have different underlying characteristics (e.g. smaller firms are more levered, and thus are more financially constrained).

From Eq. (10), it is clear that the government wants to approve the applications of firms with the *highest treatment effect*, that is the solution to the government problem is to lend to firms with $T_j \geq \underline{T}$, where $\underline{T}$ is pinned down by the resource constraint. The key question is, which firms are those in the case of the PPP? In the first result of Lemma 2, we show that for firms with the same η, the government wants to allocate loans to firms with high levels of debt b_j (or low ρ_i) if $\eta < 1$, and to firms with low levels of debt b_i if $\eta > 1$. If the shock is likely to be relatively small ($\eta < 1$), the government can try to save the firms that have the lowest probability of survival, which are those with high levels of debt per worker (*ceteris paribus*). On the other hand, for shocks that are most likely large, the government prefers simply to allocate the loans to firms with relatively low levels of debt, as those are the ones the government can still save in the face of the pandemic. The insight here is that the treatment effects are a joint product of firms' financial positions and the nature of the shock distribution, and thus they do not have a distribution or model-free ranking.

Lemma 2. ***Government PPP allocation.***
Debt heterogeneity. *Consider that all firms in the economy are equal except for their level of debt* b_j*; that is,* $c_j = c > 0$*, and* $\eta_j = \eta$*. The solution to Eq. (10) implies that* $\exists!\, b^*$ *such that: (i) for* $\eta < 1$*,* $l_j^G = 1$ *if* $b_j > b^{*,G}$ *and* $l_j^G = 0$ *otherwise, and (ii) the opposite for* $\eta > 1$*.*[16]
Shock exposure heterogeneity. *Consider that all firms are equal except for their shock exposure* η_j*; that is,* $c_j = c > 0$*. The solution to Eq. (10) implies that* $\exists!\, \underline{\eta}_G, \overline{\eta}_G$ *such that the government chooses* $l_j^G = 1$ *if* $\eta_j \in [\underline{\eta}_G, \overline{\eta}_G]$ *and* $l_j^G = 0$ *otherwise.*

In the second result of Lemma 2, we show that for firms with the same financial position c_j, the government wants to allocate loans to firms with intermediate exposure to the pandemic shock, that is, the treatment effects are inverse U-shaped in η_j. Intuitively, the most affected firms won't survive with the extra φ, while the least affected firms will likely survive regardless, such that φ is too much to allocate to them (Gourinchas et al. (2022)). Here, contrary to the constrained first-best, this is not a function of the total size of the program, M, since the amount at the intensive margin that the government can allocate to each firm is fixed. This second result of Lemma 2 is shown in Figure 3, where we compare the optimal allocation of the government with that of the private banking sector.

[16] Note that $b^{*,G}$ is different if η is < or > than 1. Moreover, for $l_j^G = b^{*,G}$, the government is indifferent regarding the allocation.

III.4. Banks' Optimal Allocation

We now focus on the private banking sector allocation in the PPP. As in the government optimization problem Eq. (8), banks can choose to accept or reject applications from firms to maximize their profits. We focus on the problem of a single representative bank.[17]

Banks receive positive profits from making more PPP loans and thus will make as many loans as possible in the program. If the banks accepts a PPP application, there are two possible scenarios. If the firm survives, the bank recovers b_j of the current loan payments and a present value of $\psi_F b_j + \beta_j$ from potential future loans to this firm. This captures the notion that lenders benefit from higher probability of issuing future loans to its relationship borrowers (Bharath et al. (2007)), and that firms with larger debt are more bank dependent and will likely demand more future loans. The firm specific component β_j captures the notion that firms without any bank debt also present future lending opportunities to banks. If the firm does not survive, the bank receives a share $\delta \in (0,1)$ of the current payments and no value from potential future loans to this firm. We allow for the value of potential future loans to the firm to depend on current debt b_j. The same two scenarios of survival and bankruptcy are possible when the bank rejects the PPP application. However, we additionally assume that if the firm survives after having its application denied, there is a probability $\psi_C < 1$ that the firm switches bank providers.[18] Finally, while applications for the PPP do not require firms to disclose the entirety of their debt, each bank would know how much debt b_j a firm owes to itself, therefore influencing their incentives to lend to firm j, as incorporated in our model.

Let $l_{j,t}^B \in \{0,1\}$ be the choice of a bank to approve the application of a firm. The profit Π_j^B per firm j a bank receives is

$$\begin{aligned}\Pi_j^B \equiv & \left[\Phi_j^\Gamma(\varphi)\left((1+\psi_F)b_j+\beta_j\right)+\left(1-\Phi_j^\Gamma(\varphi)\right)\delta b_j\right]l_j^B \\ & +\left[\Phi_j^\Gamma(0)\left((1+(1-\psi_C)\psi_F)b_j+(1-\psi_C)\beta_j\right)+\left(1-\Phi_j^\Gamma(0)\right)\delta b_j\right](1-l_j^B) \qquad (11)\end{aligned}$$

[17] For a model with multiple banks and bank heterogeneity, see Joaquim and Netto (2022).

[18] For instance, see Peter Rudegeair, "When Their PPP Loans Didn't Come Through These Businesses Broke Up with Their Banks," *The Wall Street Journal*, July 31, 2020. https://www.wsj.com/articles/when-their-ppp-loans-didnt-come-through-these-businesses-broke-up-with-their-banks-11596205736.

Simplifying the profit function and removing the constant terms, we can write the problem of the bank as

$$\max_{\{l_j^B \in \{0,1\}\}} \int \Omega_j l_j^B \, dj \quad \text{s.t.} \quad \int l_j^B \, dj = \frac{M}{\varphi}. \tag{12}$$

where

$$\Omega_j \equiv T_j \Big[(1 - \delta + \psi_F) b_j + \beta_j \Big] + \theta_j \psi_C \left(\psi_F b_j + \beta_j \right). \tag{13}$$

The misallocation comes from $\Omega_j \neq T_j$, that is, the difference between the treatment effect and profits from allocating PPP loans to a given firm.[19] In our setting, there are two channels through which profits of the banking sector deviate from the objective function of the government, which we explore in Lemma 3.

Lemma 3. ***Banks' PPP allocation.***

Debt heterogeneity. *Consider that all firms in the economy are the same except for their level of debt b_j; that is, $c_j = c$, and $\eta_j = \eta$. The solution to Eq. (12) is such that banks give preference to firms with higher b_j.*

Shock exposure heterogeneity. *Consider that all firms are the same except for their exposure η_j. The solution to Eq. (12) implies that $\exists!\ \underline{\eta}_B, \overline{\eta}_B$ such that the bank chooses $l_j^{*,B} = 1$ if $\eta_j \in [\underline{\eta}_B, \overline{\eta}_B]$ and $l_j^{*,B} = 0$ otherwise. Additionally, $\underline{\eta}_B < \underline{\eta}_G$ and $\overline{\eta}_B < \overline{\eta}_G$; that is, banks distort the allocation toward firms with a higher probability of survival without a PPP loan.*

First, the banking sector already has a heterogeneous exposure from firms that have outstanding loans and potential future loans to be made to this firm, which is captured by $(1 - \delta + \psi_F) b_j + \beta_j$. Everything else being equal, this implies that compared with the government, banks allocate loans to firms with more pre-shock debt per worker.[20] This

[19] One obtains $\Omega_j \neq T_j$, as firms without debt would have $\Omega_j = 0$ *even if* under the assumption that PPP loans do not provide any benefits from lenders beyond their ability to recover b_j (that is, if we set $\psi_F = \beta_j = 0$).

[20] Note that under the conditions of the first part of Lemma 3, Ω_j is strictly increasing in b_j. However, if we take into account that b_j can also enter into the probability of survival without PPP loans, $\Phi_j(0)$, we can show that Ω_j is hump shaped in b_j. This means that compared with the government allocation, banks want to allocate loans to firms with more debt but not necessarily to firms with the highest levels of pre-pandemic debt, as the probability of survival for some of those firms is too small. We opt here for the simpler statement of Lemma 3, as it captures the channels we want to highlight and is consistent with the empirical evidence in Bartik et al. (2021).

result is consistent with the empirical findings in the literature. In particular, Bartik et al. (2021) show that conditional on the set of firms with a relationship with a bank (at the extensive margin), banks approved more loans to firms with higher preexisting debt with those banks at what the authors call a “striking magnitude."

Second, banks are also concerned about the probability of survival of the firms θ_j, as those are clients that might switch banks if they do not receive PPP loans. This implies that, everything else being equal, banks allocate loans with a higher probability of survival without a PPP loan. This incentive can be particularly perverse for the effectiveness of the program, since the firms that do receive loans are exactly those that could have survived without the loans. In the second result of Lemma 3, we show that the banking sector distorts the optimal government allocation toward firms with lower η's and thus a higher probability of survival without a PPP loan. We illustrate this second result of Lemma 3 in Figure 3. Intuitively, this result comes from the second term of Ω_j in Eq. (13), that is, the fact that banks also derive larger profits from firms that have a higher probability of survival ex ante. This result is also consistent with the evidence in the empirical literature, which finds that banks accept more applications from less distressed firms. For instance, Joaquim and Netto (2022) and Bartlett and Morse (2021) find that firms that experience a revenue decrease in the COVID-19 crisis were more likely to apply for, but less likely to receive, a PPP loan in the first round.

IV. Empirical Application: Firm Size Heterogeneity

In this section, we provide an empirical application of our model. Although our model highlights several dimensions of firm heterogeneity, we focus in this section on firm size. We do so because firm size is directly observed in the PPP data (while the other characteristics we highlight are not) and there are various estimates of the treatment effects of the program on firms based on their size. We extend our model to account explicitly for the timing of loan disbursement in the program and introduce a welfare metric to evaluate different allocations of the program's funds. We compute the effects of the program under four allocation mechanisms: (i) optimal allocation from the government; (ii) allocation by banks (that is, the observed one); (iii) random allocation; and (iv) allocation based on firm size that allocates funds first to small firms. In our benchmark analysis, we consider that each of these counterfactual allocations is constrained by the program being able to disburse the same amount of dollars the PPP program disbursed each given date in the data. Alternatively, we also consider a case where the constraint is based on the number of loans made by the PPP at each given date in the data.

Our findings suggest that the the first round of the program is substantially less effective than it could otherwise have been due to the allocation of funds through banks. As banks allocated funds to the largest firms early, and those are the ones with the smallest treatment effects, the allocation of funds through banks performs worse than a random allocation or one that favours smaller firms. The second round of the program, however, was much more effective per dollar invested given the profile of the recipients, and we find a small degree of misallocation by the end of the program in August of 2020. This is a consequence of the fact that the supply of funds was larger than the demand for funds by the end of the program, so every firm that was approved for a loan could obtain one.

PPP timing and welfare. So far, our model has focused explicitly on which firms receives PPP loans at a given moment in time. Now, we take into account PPP timing in our welfare computation. For that, we include a discount factor in the welfare function of the government. This discounting is a reduced-form way to capture that the program is more effective if firms have access to funds earlier. We calibrate this discount rate based on Barrot and Nanda (2020), who show that a policy that accelerated government payments to small businesses increased employment at the firm level. Mathematically, let $l_{j,t} = 1$ if firm j receives a loan at time $t \leq T$. The welfare of the government from allocation $\boldsymbol{l} \equiv \{l_{j,t}\}_{j,t}$ is given by

$$\mathcal{W}(\boldsymbol{l}, T) = \sum_{t}^{T} (1 - \xi)^t \int_j l_{j,t} T_j dj, \tag{14}$$

where $\xi \in (0, 1)$ is a discount rate. We calibrate $\xi = .0037$ to match the evidence in Barrot and Nanda (2020) that finds a 5.7 percent increase in employment for a payment of 100 percent of payroll 15 days earlier. This value of ξ implies that a loan allocated to a firm

30 days later is approximately 90 percent as effective as a loan allocated today. Our misallocation measure of an allocation $\boldsymbol{l}$ relative to the government optimal allocation $\boldsymbol{l}^{*,G}$ is given by

$$\mathcal{M}(\boldsymbol{l},\boldsymbol{l}^{*,G}) = \frac{\mathcal{W}(\boldsymbol{l}) - \mathcal{W}(\boldsymbol{l}^{*,G})}{\mathcal{W}(\boldsymbol{l}^{*,G})}. \tag{15}$$

We interpret $\mathcal{M}(\boldsymbol{l},\boldsymbol{l}^{*,G})$ as a reduction in program effectiveness. For instance, if $\mathcal{M}(\boldsymbol{l},\boldsymbol{l}^{*,G})$, the time-discounted treatment effect relative to the optimal policy, is -0.5, we interpret this as the PPP being 50 percent less efficient than it could have otherwise been by that time.

Empirical Estimates of the Heterogeneous Treatment Effect. We obtain estimates for the treatment effect from Dalton (2021) and compare our approach to the estimates from Autor et al. (2022*a*). Dalton (2021) uses the microdata from the Quarterly Census of Employment and Wages (QCEW) and timing of receipt to recover the employment effects of the program for firms of different sizes (1-10 employees, 10-50 employees, 50-100 employees, and more 100 employees). We use these estimates from Dalton (2021) as they leverage a large, representative dataset on small business employment, thus allowing for an analysis of the heterogeneous treatment effects of the program. We then combine these estimates of treatment effects for firms of different sizes with data from the SBA/Treasury release of loans made in the PPP (January 2023 version). More specifically, we take the estimates from Dalton (2021), together with the underlying distribution of the number of employees in the program, and do a piecewise linear interpolation to find the treatment effects for firms based on their number of employees. The results are shown in Figure 4.

To understand how this approach compares to other estimates of size sensitive treatment effects, we compare our estimates with the estimates from Autor et al. (2022*b*), which complements Autor et al. (2022*a*) and uses an approach similar to Dalton (2021) to compute the treatment effects for firms between 1-50 employees. Since all that matters in the misallocation measure of Eq. (15) are *relative* treatment effects, we compare the ratios of the treatment effects for firms with 1-50 employees to the overall recipients of the

PPP in both papers.[21] In Dalton (2021), firms with 1-50 employees have treatment effects that are 30% larger than the average firm, while in Autor et al. (2022*b*) this number is approximately 60%. Since our results highlight the misallocation coming from firm-size, by following Dalton (2021) we are choosing the conservative estimate of the differences in treatment effects.

Counterfactual policies. We start the analysis of counterfactual policies by considering different allocations of PPP funds. First, we consider the observed allocation, which we denote as the bank allocation $\boldsymbol{l}^{*,B}$. Second, we consider the optimal government allocation under the PPP rules, $\boldsymbol{l}^{*,G}$. Third, we consider the allocation of PPP loans when the government allocates funds to firms following an increasing firm-size order (i.e., it starts with the smallest firms), $\boldsymbol{l}^{*,I}$. Finally, we consider a random allocation of PPP loans across firms, $\boldsymbol{l}^{*,R}$.

For a given objective function, we compute $\boldsymbol{l}_t^*$, the allocation that maximizes the objective function. For instance, for the problem of the government at time t, we compute $l_t^{*,G}$ as:

$$l_t^{*,G} = \arg\max_{l_{t,j}\in\{0,1\}} \int l_{t,j} T_j^{emp} dj \quad \text{s.t.} \quad \int l_{t,j} N_j \varphi_j dt = M_t, \tag{16}$$

where $T_j^{emp} = T_j \times N_j$ is the effective treatment effect of firm j, φ_j is the per worker amount of loans given to firm j, and M_t is the total amount distributed in the program until time t observed in the data.

Results. We begin by looking at the average firm size on the optimal and the observed allocations. Figure 5 panel (a) illustrates how the optimal allocation targets smaller firms relative to the observed bank allocation. Since treatment effects decreases with firm size, the government has incentives to prioritize small firms. Next, we compare the optimal allocation with the observed bank based allocation and with the size-targeted allocation. Panel (b) shows that at the end of the first round (vertical line), the misallocation from allocating funds through banks is large—the program would have been 50 percent less effective. A size-based policy would have reduced this misallocation by 45 percentage points. At the end of the first draw, a size-based program would have been only 5 per-

[21] We opt for this ratio since it is the one we can compute for the two papers.

cent less effective compared with the optimal allocation. During the first round, banks targeted firms that were larger and less affected by COVID-19 and thus are far from maximizing the effect of the program. As more funds are appropriated for the PPP in late April, and eventually the supply of PPP loans exceeds demand, the role of banks in the allocation of funds is reduced and the high-treatment-effect firms receive PPP loans. Given the excess supply of funds by the end of the program, the misallocation by August of 2020 comes from our discount factor, ξ, which captures misallocation due differential timing of PPP loans, as analyzed by Doniger and Kay (2023).

Constraint on the *Number* of loans. So far we have focused on a volume constraint, that is, banks and government must choose their allocation subject to a given dollar amount of dollars that can be disbursed at a given moment in time. However, a volume constraint might not be the relevant constraint faced by banks when approving loan applications. For instance, if loans of different size are equally costly to process, allocations might be constrained by the number of loans an intermediary can approve in a certain time period. In Appendix C we discuss in detail the case where the allocation is constrained by the number of loans, regardless of their volume.

When the allocation is constrained by the number of loans, the total treatment effect of allocating a loan to a firm is their own treatment effect times the number of employees. Therefore, even though treatment effects are decreasing in firm size, it pays off in our calibration for the government to allocate funds to the largest firms early on (even more so than what is observed in the data based on banks' allocation). A random or size-based allocation are inefficient relative to the banks' allocation, as neither is able to reach a large number of employees during most of the duration of the program. This exercise illustrates that the optimal allocation is not only a function of the government's objective function, but also of the relevant constraint banks and the government face in the actual implementation of the program.

V. Extensions

In this section we consider two extensions to our baseline model. First, we discuss the case where banks have an information and operational advantage relative to the government. This introduces a tradeoff between allocating funds through private banks, which are more operationally efficient and know firm fundamentals, and the government, who minimizes misallocation. Second, we study the determinants of the optimal size of the program. The optimal program size if banks are allocating the loans can be larger or smaller relative to the optimal size of the program under the optimal allocation. The key determinant is how large is the marginal cost of the additional dollar in the program. If the marginal cost is large, then it is better to do a smaller program if banks are allocating funds to avoid misallocation. If the marginal cost is low, then it is better to do a larger program that offers loans to the majority of firms and thus reduces the misallocation (which was the case of the PPP).

V.1. Asymmetric Information and Capacity Constraints

In the benchmark version of our model, banks and the government have the same information and same ability to disburse loans. Therefore, there is no reason for the government to intermediate the allocation through the private banking system. We therefore augment our model so that banks have more information and can disperse loans more quickly than the government, which is the core reasoning behind loan guarantee programs being intermediated by the banking system (e.g., Jimenez et al. (2022), Bartik et al. (2021)). The key question we focus in this extension is under which conditions does it make sense for the government to delegate the program to banks. We start with firms' treatment effects and banks' profits when lending to that firm as reduced form objects. We do so for exposition purposes, and return to the relationship between these reduced form objects and our microfounded model later on.

We augment our model in two ways. First, we introduce differential information between banks and the government. We assume that the government observes the true treatment effect for a share of firms and a random treatment effect for the remaining firms. We suppose that the government observes for every firm j a signal T_j^G of the true treatment effect T_j. The signal is given by $T_j^G = T_j$ for a share μ of the population and $\tilde{T}_j$ for a share $1-\mu$, where $\tilde{T}_j$ is independent of T_j but has the same distribution. The parameter $\mu \in [0,1]$ indicates the degree of asymmetric information: $\mu = 0$ means that the government has no information, while $\mu = 1$ means that the government has the same information as banks (as in our benchmark model). Second, we introduce the operational capacity distortion. We assume that because of delays in the disbursement of funds by the government, the treatment effects are only $\iota \in [0,1]$ as effective as they would otherwise be.[22, 23]

We link the government and banks allocation as follows. Suppose that banks' profits by lending to firm j, Ω_j, and firm j's treatment effect, T_j, are identically distributed and related with through the following equation:

$$T_j - \overline{T} = \rho(\Omega_j - \overline{\Omega}) + \epsilon_j, \tag{17}$$

where ϵ_j is a shock that satisfies $\mathbb{E}[\epsilon_j|\Omega_j] = 0$. The parameter ρ is the correlation between Ω_j and T_j. This correlation ρ thus captures the degree of misallocation in this reduced form setting. For $\rho = 1$, there is no misallocation: governments and banks have identical incentives. For $\rho = -1$, banks have exactly the opposite incentives as the government.

Our main result of this section is Lemma 4. We show that there is a cutoff ι^* such that if the government has sufficient information (high μ) or the misallocation problem is low (high ρ) the government prefers to allocate loans through banks. Information and operational capacity are substitutes in this setting. If there are no costs of delay ($\iota = 1$), then the government delegates the program to banks if, and only if, $\rho > \mu$. In words, this condition means that the government delegates the program to banks if the degree of misallocation is *lower* than then degree of information asymmetry.

Lemma 4. ***Delegation versus misallocation: efficiency-information frontier.*** *Let $\iota^*(\mu;\rho)$*

[22] Here we assume for simplicity that all of the loss of effectiveness and discounting happens through the ι term, and not through a discount rate as in our empirical application.

[23] The evidence that this delay actually matters is Barrot and Nanda (2020), and Doniger and Kay (2023) for the PPP.

be given by:

$$\iota^*(\mu,\rho) = \frac{(1-\rho)m_T + \rho\mathbb{E}[T|T \geq \underline{T}]}{(1-\mu)m_T + \mu\mathbb{E}[T|T \geq \underline{T}]} \tag{18}$$

where m_T is the average treatment effect in the population. Then, if $\iota < \iota^*(\mu;\rho)$, the government prefers to allocate PPP loans through banks.

One additional result that can be easily derived from Lemma 4 is that if T_j and Ω_j are normally distributed, a higher variance in firms' treatment effects implies that it is more likely for the government to delegate the program to the banks only if $\rho > \mu$, that is, if there is less misallocation than asymmetric information. The average effect of the program across all firms would not matter in the delegation decision. Intuitively, if there is more variation in treatment effects, the cost of delaying the program is relatively lower than the cost of the allocation deviating from the optimal. In this case, more variation in treatment effects implies that the best choice between allocating funds through the banks or directly by the less informed government will move towards whichever is closer to the optimal.

So far in this extension we have focused in the case where treatment effects and banks' profits are reduced form objects related by Eq. (17). In our microfounded model, however, we have from Eq. (13) that we can write (up to a normalization):

$$\Omega_j = \alpha_M b_j \cdot \left[T_j + \alpha_\theta \theta_j\right] + [T_j + \psi_c \theta_j] \cdot \beta_j \tag{19}$$

for some constants α_M, $\alpha_\theta > 0$. Although there is no direct mapping between Eqs. (17) and (19), our model allow us to understand what firms characteristics would generate the disparity between T_j and Ω_j. For instance, if most of the heterogeneity across firms comes from pre-COVID debt, the relationship between Ω_j and T_j will be weaker (and, depending on the relationship between b_j and T_j it can be negative). Similarly, if there is no heterogeneity in pre-COVID debt, we have that a higher α_θ increases the misallocation. In our model, a larger probability of a firm switching banks in case the PPP application is denied, ψ_C, is one of the factors that increases α_θ. A higher chance of losing clients causes banks to prioritize firms with larger ex-ante probability of survival, and thus moves the banks' and government objectives away from each other.

The Efficiency-Information Frontier in the Empirical Application. Within the context of our empirical application, we can compute the function i^* at any point in time. We do so for Figure 6 shows the efficiency-information frontier on April 16 and May 15, 2020. These two dates correspond, respectively, to the end of the first round and to the point where 95% of the funds in the program had already been disbursed. On April 16, even if the government had *no* information about firms' treatment effects, that is, $\mu = 0$, and thus did a completely random allocation of funds, the government could still be 85 percent as efficient in disbursing the funds relative to banks and still it would be optimal for the government not to delegate the disbursement of funds. On May 15, however, this figure increases to 96 percent. At this point, most firms have already received loans in the program, and thus the effects of misallocation and differences in information are muted. This example highlights that the delegation decision is not independent of other program characteristics such as program size, which we study next.

V.2. Optimal Program Size

So far we have taken the total amount available under the program, M, as given. In this section, we discuss the implications of bank incentives for the optimal size of the program. To do so, we suppose that there is an exogenous cost function that fully captures the welfare cost of allocating extra dollars to the program given by $\frac{c_M}{2}M^2$. The problem of the government is to choose a program size M that maximizes welfare, subject to either the government or banks allocating the loans:

$$M_P^* \equiv \arg\max_M \int l_j^P(M) T_j \, dj - \frac{c_M}{2\varphi} M^2, \text{ where } P \in \{G, B\}, \tag{20}$$

where $P \in \{G, B\}$ denotes if the government or bank is allocating the loans, respectively, and allocation $\boldsymbol{l}^P(M) \equiv \{l_j\}_j$ is the optimal allocation of whichever party does the allocation. We analyze the difference between M_G^* and M_B^*, that is, the optimal program size when loans are allocated through the government and when they are allocated through banks.

The marginal benefit of increasing the program size is equal to the treatment effect of the marginal firm receiving a loan. Let $\underline{T}_G(M)$ and $\underline{T}_B(M)$ denote these firms under the government and bank allocations, respectively. The M_P^* that is a solution to Eq. (20) satisfies Eq. (21), that is, equates the marginal cost with the marginal benefit of the additional dollar allocated to the program.[24]

$$\underline{T}_P(M_P^*) = c_M M_P^*, \;\; P \in \{G, B\} \tag{21}$$

We focus the analysis of optimal program size until May 2020 since after this point the program is large enough and there is no excess demand for PPP loans. We plot the difference $\underline{T}_B$ and $\underline{T}_G$ in our empirical example in Figure 7. In the first round of the program (small M), $\underline{T}_G(M) > \underline{T}_B(M)$: banks are allocating loans, at the margin, to firms with lower treatment effects relative to the government allocation. This difference flips by the end of the second round: banks, relative to the government, are allocating loans to firms with higher treatment effects (since the government prioritized firms with high treatment effects).

Combining Eq. (21) with our evidence in Figure 7, we can arrive at the following result: If the marginal cost of raising funds is high (high c_M), such that the optimal program size when banks allocate loans is small (e.g., smaller than the first round of the PPP), we have that $M_G^* > M_B^*$. At M_B^*, the marginal firm entering the program through the government allocation has a higher treatment effect than the marginal firm entering the program due to bank allocation, and thus $\underline{T}_G(M_B^*) > \underline{T}_B(M_B^*) = c_M M_B^*$. The opposite is true if the cost of the marginal dollar in the program is low (low c_M). Intuitively, we have two competing forces in establishing the optimal size of the program when banks allocate loans. On the one hand, we have that funds are costly. On the other hand, as the program grows, the effect of banks' propensity to misallocate loans toward firms with low treatment effects is reduced. The analysis shows that if the marginal cost is high, the former dominates, and

[24] Here we assume implicitly that the first-order condition of Eq. (20) in Eq. (21) is sufficient to characterize the global optimum. This assumption is true for the problem of the government, given that $\underline{T}_G(0) > 0$, $\underline{T}_G(M)$ is strictly decreasing in M and $c_M M$ is strictly increasing in M. For the problem of the bank, we are implicitly assuming here that $\underline{T}_B(0) > 0$, that is, that the bank does not choose a firm with a zero treatment effect (in case a firm with zero treatment effect exists), and that it crosses the marginal cost curve $c_M M$ once and from above. This assumption is satisfied, for instance, if $\underline{T}_B(M)$ is increasing or concave.

if the marginal cost is low, the latter dominates.

VI. Other Pandemic Lending Schemes

We have so far focused on the PPP throughout the paper. However, debt relief policies were implemented in several countries, and while each program's rules are unique, there are common characteristics across programs, and in how they differ from the PPP. For example, Altavilla et al. (2022) describes how the EU Comission Regulation provided overall guidance in terms of firm eligibility, aiming to exclude insolvent firms. Moreover, guaranteed amounts often were partial and decreasing in firm size.[25] In the context of our model the main differences are: (i) the guaranteed amount can be partial and sensitive to firm size; (ii) relief loans were not forgivable and had positive interest rates ($r_G > 0$); (iii) tighter solvency requirements were included, to prevent firms that were illiquid before the pandemic from accessing relief funds (firms with $c_j < 0$ could not apply). Throughout this section, we illustrate how some of these changes could be included in our framework.

Size-dependent partial loan guarantees can be incorporated in the bank's optimization problem by assuming the bank recovers a fraction $q(N_j) \leq 1$ of each loan in case of firm default. This adds an additional term to Eq. (12) reflecting expected losses of providing a loan conditional on firm default. This additional term implies banks can reject loan applications if the added benefit in terms of repayment of existing debt is smaller than the expected cost of losses in case of default. If $q(N_j)$ is decreasing in N_j, banks will be more likely to supply loans to smaller firms, conditional on then same amount of debt per worker. If treatment effects correlate negative with firm size, as observed empirically, then partial guarantees would *reduce* misallocation by directing funds for more affected firms.

The other two differences, non-grant status and tighter eligibility criteria, would affect firm choices, and are explored in more detail in Appendix D. Importantly, both requirements imply a smaller subset of firms would apply for relief loans. In this case, the program would be larger relative to firm demand, further reducing misallocation, as il-

[25] See also Cirera et al. (2021), Core and Marco (2022), Jimenez et al. (2022) and Huneeus et al. (2022) for programs with similar characteristics.

lustrated by Figure 5. It would also imply a larger share of funds would flow to the most affected firms, which is consistent with findings from empirical studies exploring other COVID lending schemes (e.g. Altavilla et al. (2022)).

VI.1. Bank Lending Outside of the Program

One important caveat of our model is that we do not consider lending outside of the Paycheck Protection Program. The reason we do so is that, given the rules and size of the PPP, there was little incentive for banks to lend outside of the program. Li and Strahan (2022) show that almost all of the growth in C&I lending during the period can be attributed to PPP loans. On the other hand, different designs in other loan guarantee programs are such that loans outside of the program are still made, and understanding the substitution between loans made through the program and those outside of the program is key. Empirically, several papers which explore other loan guarantee programs identify patterns in how banks choose between guaranteed and non-guaranteed credit (e.g. Altavilla et al. (2022), Jimenez et al. (2022)). They find that banks direct guaranteed loans to riskier, more affected firms, firms were more likely to obtain guaranteed loans from banks with whom they had larger exposures, and banks that participate more in guarantee schemes gain market share.

Appendix E includes an extension of our baseline model in which we allow for guaranteed and outside lending to co-exist. For that, we modify our model in two ways: we suppose that there is a limited amount of guaranteed loans that can be made, and that banks can charge higher rates (and thus have higher profits) for loans made outside of the program. In this setup, banks' choice of providing guaranteed or non-guaranteed loans hinges crucially on the probability of survival of firms to the pandemic shock. The benefits of supplying guaranteed credit are greater for more affected firms, which is exactly the opposite of what we find in the PPP. The intuition is that more affected firms are likelier to default, thus making a guaranteed loan to them insulates the bank from the credit risk. In the PPP, however, since all loans are guaranteed, this channel is not present, and the value of the future relationship dominates in the allocation. Similarly to the PPP, we find that banks are more likely to allocate funds to ex-ante indebted firms.

More broadly, we show in Appendix E how including outside lending in our model can match the key empirical findings from Altavilla et al. (2022) and Jimenez et al. (2022) for loan guarantee programs implemented in Europe.

VII. Conclusion

In response to the COVID-19 crisis, the US government created the Paycheck Protection Program, a large debt relief policy designed to preserve jobs in small and medium-sized firms. In this paper, we develop a theoretical framework that incorporates the relevant characteristics of the PPP to address three questions. First, we study what should have been the optimal allocation of PPP loans from the perspective of the government. Second, we highlight what distortions were caused by the allocation of these loans through the banking system. Third, we discuss potential changes in the PPP that could have been made to minimize the misallocation of funds. We find that the optimal allocation is not trivial and depends on the nature of the shock and the design of the program. Banks have incentives to distort the allocation to firms that have more debt outstanding from that bank and are less exposed to COVID and/or more likely to survive any adverse shock. Finally, we illustrate empirically how a small change—targeting smaller firms—would have substantially reduced the distortions introduced by bank incentives. Misallocation can be minimized by identifying how banks target firms with relatively smaller treatment effects, and explicitly imposing requirements to ensure that funds flow to these firms earlier. These are valuable lessons for policy makers when designing emergency lending programs, which are now part of the crisis response toolkit.

References

Altavilla, Carlo, Andrew Ellul, Marco Pagano, Andrea Polo, and Thomas Vlassopoulos. 2022. "Loan Guarantees, Bank Lending and Credit Risk Reallocation."

Autor, David, David Cho, Leland D Crane, Mita Godar, Byron Lutz, Joshua Montes, William B Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. 2022*a*. "An Evaluation of the Paycheck Protection Program Using Administrative Payroll Microdata." *Journal of Public Economics*, 211.

Autor, David, David Cho, Leland D. Crane, Mita Goldar, Byron Lutz, Joshua Montes, William B. Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. 2022*b*. "The $800 Billion Paycheck Protection Program: Where Did the Money Go and Why Did It Go There?" *Journal of Economic Perspectives*, 36: 55–80.

Balyuk, Tetyana, Nagpurnanand Prabhala, and Manju Puri. 2021. "Small Bank Financing and Funding Hesitancy in a Crisis: Evidence from the Paycheck Protection Program."

Barrios, John Manuel, Michael Minnis, William C. Minnis, and Joost Sijthoff. 2020. "Assessing the Payroll Protection Program: A Framework and Preliminary Results."

Barrot, Jean-noël, and Ramana Nanda. 2020. "The employment effects of faster payment: evidence from the federal quickpay reform." *Journal of Finance*, 75: 3139–3173.

Bartik, Alexander W, Marianne Bertrand, Zoe Cullen, Edward L Glaeser, Michael Luca, and Christopher Stanton. 2020. "The impact of COVID-19 on small business outcomes and expectations."

Bartik, Alexander W., Zoe E. Cullen, Edward L. Glaeser, Michael Luca, Christopher T. Stanton, and Adi Sunderam. 2021. "The Targeting And Impact of Paycheck Protection Program Loans to Small Businesses."

Bartlett, Robert P, and Adair Morse. 2021. "Small Business Survival Capabilities and Policy Effectiveness: Evidence from Oakland." *Journal of Financial and Quantitative Analysis*, 56.

Beck, Thorsten, Leora F Klapper, and Juan Carlos Mendoza. 2010. "The Typology of Partial Credit Guarantee funds around the World." *Journal of Financial Stability*, 6: 10–25.

Benmelech, Efraim, and Nitzan Tzur-Ilan. 2020. "The Determinants of Fiscal and Monetary Policies During the Covid-19 Crisis." *NBER working Paper No. 27461*.

Bharath, Sreedhar, Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan. 2007. "So what do I get? The bank's view of lending relationships." *Journal of Financial Economics*, 85: 368–419.

Blanchard, Olivier, Thomas Philippon, and Jean Pisani-Ferry. 2020. "Policy Brief 20-8: A New Policy Toolkit Is Needed as Countries Exit COVID-19 Lockdowns." *PIIE Policy Briefs*.

Brunnermeier, Markus, and Arvind Krishnamurthy. 2020. "Corporate debt overhang and credit policy." *Brookings Papers on Economic Activity*.

Buffington, Catherine, Carrie Dennis, Emin Dinlersoz, Lucia Foster, Shawn Klimek, et al. 2020. "Measuring the Effect of COVID-19 on US Small Businesses: The Small Business Pulse Survey."

Chetty, Raj, John N Friedman, Nathaniel Hendren, Michael Stepner, and Opportunity Insights Team. 2020. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built from Private Sector Data."

Chodorow-Reich, Gabriel, Olivier Darmouni, Stephan Luck, and Matthew Plosser. 2021. "Bank liquidity provision across the firm size distribution." *Journal of Financial Economics*.

Cirera, Xavier, Marcio Cruz, Elwyn Davies, Arti Grover, Leonardo Iacovone, Jose Ernesto L Cordova, Denis Medvedev, Franklin Okechukwu Maduko, Gaurav Nayyar, Santiago Reyes Ortega, and Jesica Torres. 2021. "Policies to Support Businesses through the COVID-19 Shock: A Firm Level Perspective." *The World Bank Research Observer*, 36: 41–66.

Core, Fabrizio, and Filippo De Marco. 2022. "Information Technology and Credit: Evidence from Public Guarantees."

Cororaton, Anna, and Samuel Rosen. 2020. "Public Firm Borrowers of the US Paycheck Protection Program."

Crouzet, Nicolas, and Fabrice Tourre. 2021. "Can the cure kill the patient? Corporate credit interventions and debt overhang."

Dalton, Michael. 2021. "Putting the Paycheck Protection Program into Perspective: An Analysis Using Administrative and Survey Data."

Doniger, Cynthia, and Benjamin Kay. 2023. "Long-lived employment effects of delays in emergency financing for small businesses." *Journal of Monetary Economics*.

Duchin, Ran, and John Hackney. 2021. "Buying the Vote? The Economics of Electoral Politics and Small-Business Loans." *Journal of Financial and Quantitative Analysis*, 56: 2439–2473.

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh. 2020. "Can the Covid Bailouts Save the Economy?"

Erel, Isil, and Jack Liebersohn. 2022. "Can FinTech reduce disparities in access to finance? Evidence from the Paycheck Protection Program." *Journal of Financial Economics*, 146: 90–118.

Faria-e-Castro, Miguel. 2021. "Fiscal Policy during a Pandemic." *Journal of Economic Dynamics and Control*.

Faulkender, Michael W., Robert Jackman, and Stephen Miran. 2021. "The Job Preservation Effects of Paycheck Protection Program Loans."

Glancy, David. 2023. "Bank Relationships and the Geography of PPP Lending."

Goodhart, Charles A.E., Dimitrios P. Tsomocos, and Xuan Wang. 2023. "Support for small businesses amid COVID-19." *Economica*, 90: 612–652.

Gourinchas, Pierre-Olivier, Sebnem Kalemli-Özcan, Veronika Penciakova, and Nick Sander. 2022. "Estimating SME Failures in Real Time: An Application to the COVID-19 Crisis."

Granja, João, Christos Makridis, Constantine Yannelis, and Eric Zwick. 2022. "Did the Paycheck Protection Program Hit the Target?" *Journal of Financial Economics*, 144: 725–761.

Greenwood, Robin, Ben Iverson, and David Thesmar. 2020. "Sizing Up Corporate Restructuring in the COVID-19 Crisis." *Brookings Papers on Economic Activity*.

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Iván Werning. 2022. "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?" *American Economic Review*, 112.

Hanson, Samuel G., Adi Sunderam, Jeremy C. Stein, and Eric Zwick. 2020. "Business Credit Programs in the Pandemic Era." *Brookings Papers on Economic Activity*.

Hassan, Tarek Alexander, Stephan Hollander, Laurence van Lent, Markus Schwedeler, and Ahmed Tahoun. 2020. "Firm-level Exposure to Epidemic Diseases: Covid-19, SARS, and H1N1."

Hong, Gee Hee, and Deborah Lucas. 2023. "COVID Credit Policies Around the World: Size, Scope, Costs, and Consequences." *BPEA Conference Draft*.

Hubbard, R. Glenn, and Michael R. Strain. 2020. "Has the Paycheck Protection Program Succeeded?"

Huneeus, Federico, Joseph Kaboski, Mauricio Larrain, Sergio L Schmukler, and Mario Vera. 2022. "The Distribution of Crisis Credit: Effects on Firm Indebtedness and Aggregate Risk." *NBER Working Paper No. 29774*.

Jimenez, Gabriel, Luc A. Laeven, David Martinez-Miera, and Jose-Luis Peydro. 2022. "Public Guarantees, Relationship Lending and Bank Credit: Evidence from the COVID-19 Crisis."

Joaquim, Gustavo, and Felipe Netto. 2022. "Bank Incentives and the Effect of the Paycheck Protection Program."

Kahn, Charles M., and Wolf Wagner. 2021. "Liquidity provision during a pandemic." *Journal of Banking Finance*, 133: 106152.

Li, Lei, and Philllip Strahan. 2022. "Who Supplies PPP Loans (And Does It Matter)? Banks, Relationships And The COVID Crisis." *Journal of Financial and Quantitative Analysis*, 56: 2411–2438.

Martín, Alberto, Sergio Mayordomo, and Victoria Vanasco. 2024. "Banks vs. Firms: Who Benefits from Credit Guarantees?"

Neilson, Christopher, John Eric Humphries, and Gabriel Ulyssea. 2020. "Information Frictions and Access to the Paycheck Protection Program." *Journal of Public Economics*, 190.

Philippon, Thomas. 2021. "Efficient Programs to Support Businesses During and After Lockdowns." *The Review of Corporate Finance Studies*, 10: 188–203.

Rabetti, Daniel. 2022. "Non-Information Asymmetry Benefits of Relationship Lending."

Wang, Tianxi, and Xuan Wang. 2021. "Designing Financial Support for SMEs during Crises: the Role of Bank Lending."

TABLES

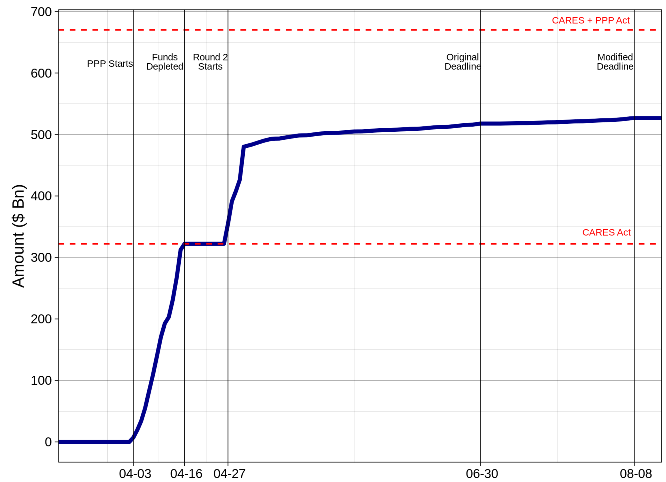
Table 1: Summary Statistics of the Paycheck Protection Program

	Apr-16	May-15	Aug-08
Loan Amount ($, Billions)	322.2	498.4	526.6
# Loans (,000)	1,619.7	4,209.7	5,147.6
Jobs Supported (Million)	33.2	57.3	61.1
Average Loan Size ($,000)	198.96	118.40	102.30
Average Jobs Supported	20.5	13.6	11.8
Top-4 Share - # Loans	0.03	0.18	0.17
Top-4 Share - Volume	0.05	0.13	0.13

Note: Data from the SBA/Treasury January 2023 release. Loan amounts (in billions of dollars) and number of loans (in thousands) accumulated after the start of the program (April 3, 2020). See Figure 1 for details. Average loan size is the ratio of the cumulative loan amount over the cumulative number of loans. Jobs supported were reported by the firms during the PPP application. The top four banks (by assets in December 2019) are (i) J.P. Morgan Chase, (ii) Bank of America, (iii) Wells Fargo, and (iv) Citibank, N.A.

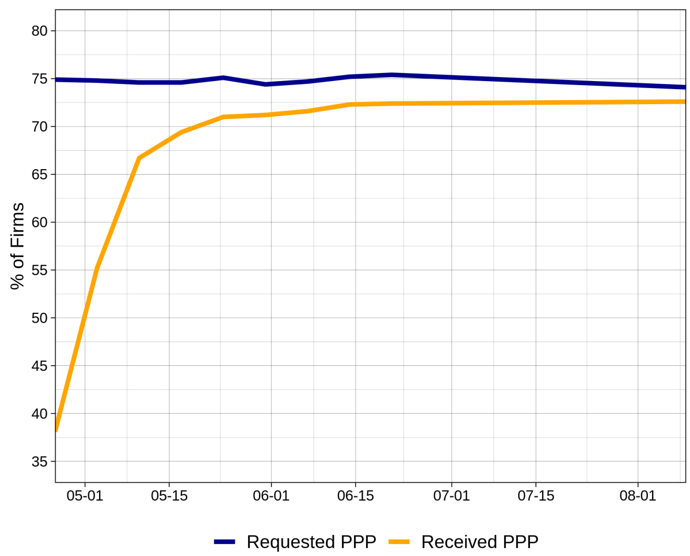
FIGURES

Figure 1: Cumulative PPP Disbursement over Time ($, Billions)



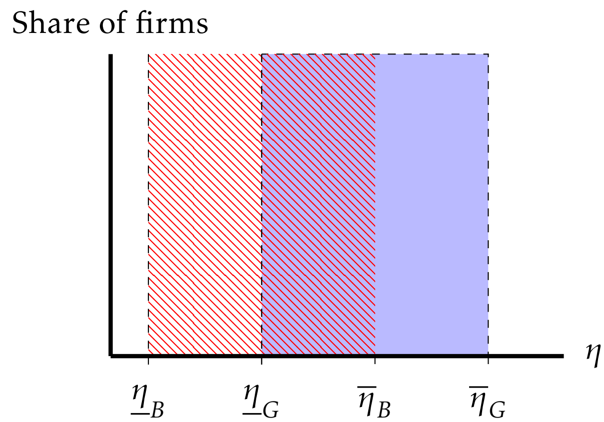
Note: Our primary source for data on the PPP comes from the SBA/Treasury). The data set includes information self-reported by the borrower (name, address, Zip code, NAICS code and jobs supported) as well as loan amount, approval date, and lender name. We analyze the loans made in the first draw or the program (April 3 through August 8, 2020). No loans were made in the program from August 8, 2020 through January 11, 2021 (when the second draw of PPP loans began). The date of a loan is the date of approval (according to the rules of the program, loans had to be disbursed within 10 calendar days of approval). Billions of dollars of PPP loans approved by day, from April 3, 2020, (CARES Act) through August 8, 2020 (modified deadline for second-round applications). Dashed horizontal lines represent the cumulative capacity of the program.

Figure 2: Small Business Pulse Survey: PPP Application vs. PPP Receipt (% of Firms)



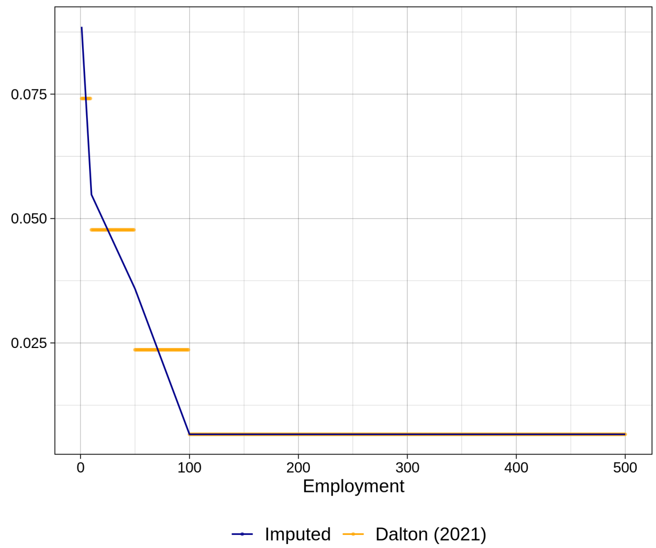
Note: US-level data from the Small Business Pulse Survey (SBPS) collected weekly from April 26 through June 21, 2020). In this figure, we show the percentage of firms that report applying and receiving PPP loans. For details, see Buffington et al. (2020). The SBPS was designed to collect real-time information from small businesses during the pandemic. According to the Census, the target populations is "all nonfarm, single-location employer businesses with 1 to 499 employees and receipts of $1,000 or more". The blue line denotes the percentage of firms that reported applying for a PPP loan. The yellow line denotes the percentage of firms that reported receiving a PPP loan.

Figure 3: Credit Allocation under the PPP: Firm Heterogeneous Shock Exposure (η_j)



Note: This figure is the visual representation of Lemmas 2 and 3. The solid blue rectangle is the government allocation. The dotted red allocation is the banking sector allocation. The η_G's are the lower and upper bounds of the regions/sectors for the government, and the η_B's are those for the banking sector.

Figure 4: Firm Size and Imputed Treatment Effect

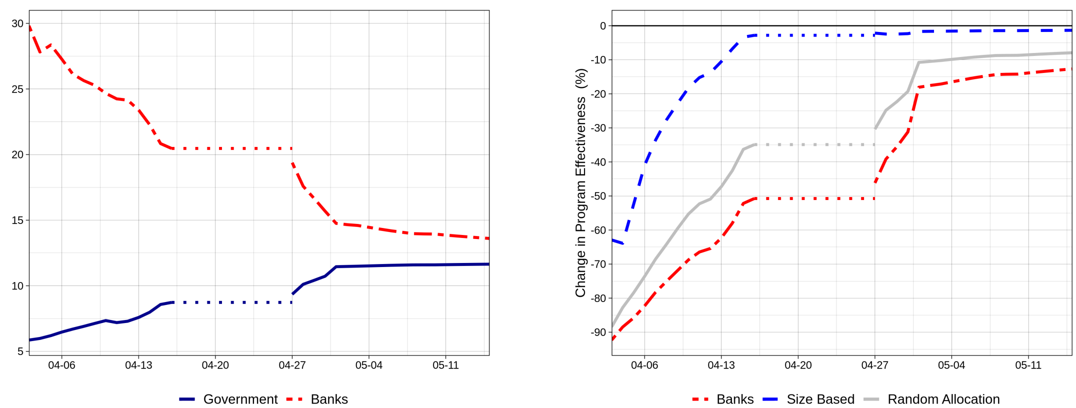


Note: This figure presents the estimates of the program's treatment effect by firm size from Dalton (2021).

Figure 5: Optimal, Observed and Alternative Policies: Firm-Size and Relative Effectiveness

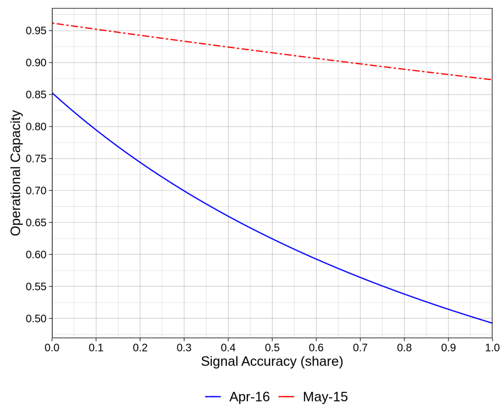
(a) Firm Size

(b) Relative Effectiveness



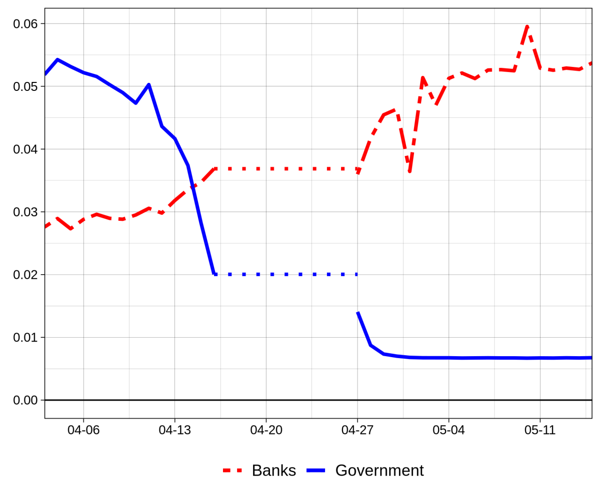
Note: The optimal allocation is computed from the problem in Eq. (16). **Panel A.** Average firm size that received (banks) or that would have received (government) a loan by a given moment in time. **Panel B.** The welfare relative to the government's optimal allocation, as in Eq. (15).

Figure 6: Efficiency-Information Frontier on April 16 and May 15, 2020



Note: Efficiency-information frontier from Lemma 4, Eq. (18) on April 16 and May 15, 2020. Welfare measures for Eq. (18) come from the use of Eq. (14) with the allocation given by Eq. (16) .

Figure 7: Treatment Effect at the Marginal Firm



Note: This figure shows the treatment effect of the marginal firm when funds are allocated by banks or the government. The bank problem solution is the observed allocation in the data. The counterfactual government allocation is that given by Eq. (16).

Appendix

A. Proofs and Derivations

A.1. Lemma 1

Proof. Let λ be the Lagrange multiplier on the constraint that $\int N_j \omega_j^G dj = M$. Taking the FOC of Eq. (5) w.r.t. ω_j^G

$$N_j \phi\left[\Gamma\left(\omega_j^G\right)\right] \cdot \frac{\partial \Gamma_j}{\partial \omega_j^G} - N_i \lambda = 0 \Rightarrow \phi\left[\Gamma\left(\omega_i^G\right)\right] = \phi\left[\Gamma\left(\omega_j^G\right)\right], \ \forall i, j,$$

where we use that $\frac{\partial \Gamma_j}{\partial \omega_j^G} = 1$ of the last equation. Let $\tilde{\lambda} \equiv \lambda \cdot c_0^{\eta}$. Using the equation for the distribution $\phi(\nu)$ in Eq. (2):

$$\eta_j \left[c_j + \omega_j^{G,*}\right]^{\eta - 1} - \tilde{\lambda} = 0 \Rightarrow \omega_j^{G,*} = \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} - c_j \Rightarrow M + \overline{c} = \int N_j \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} dj, \tag{22}$$

where the last equality comes from integrating $N_j \omega_j^{G,*}$ across firms to solve for λ. This is the unique global maximum of the problem, as the constraint is linear and the objective function is strictly concave.

Note that the RHS is (i) strictly *decreasing* in $\tilde{\lambda}$ since $\eta_j < 1$, (ii) goes to infinity with $\tilde{\lambda} \to 0$, and (iii) goes to zero with $\tilde{\lambda} \to \infty$, so there is always a unique solution for $\tilde{\lambda}$ from Eq. (22). We can use (22) in the individual firm j equation to obtain:

$$N_j \omega_j^{G,*} = M - \left[N_j c_j - \overline{c}\right] + N_j \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} - \int N_j \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} dj. \tag{23}$$

Thus, we have that $N_j \tau(\eta_j, \overline{M}) \equiv N_j \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} - \int N_j \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} dj$. Therefore:

$$\frac{\partial \omega_j^{G,*}}{\partial \eta_j} = \frac{\partial \tau(\eta_j, \overline{M})}{\partial \eta_j} = -\left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j - 1}} \frac{1}{(\eta_j - 1)^2} \left[\ln(\tilde{\lambda}) + 1 - \frac{1}{\eta_j} - \ln(\eta_j)\right]. \tag{24}$$

Let $f(\eta) = 1 - \eta^{-1} - \ln(\eta)$. We know that $f(1) = 0$ and $f^{'}(\eta) = \eta^{-2} - \eta^{-1}$. Thus, $f(\eta) < 0$ for $\eta < 1$.

Case 1. If $\tilde{\lambda} < 1$, we have that $\frac{\partial \omega_{i,j}^{G,*}}{\partial \eta_j} > 0$. To see this, note that $f(\eta_j) < 0$ and $\ln(\tilde{\lambda}) < 0$, thus in Eq. (24) the RHS is positive. For $\tilde{\lambda} < 1$, we need $M + \bar{c} > \int N_j \left(\eta_j\right)^{\frac{1}{1-\eta_j}} \, dj$.

Case 2. if $\tilde{\lambda} > 1$, we have that $\frac{\partial \tau(\eta_j, \overline{M})}{\partial \eta_j}$ is positive for $\eta_j < \bar{\eta} < 1$ and negative otherwise (that is, the transfer function τ is concave in η_j). To see this, note that $\lim_{\eta \to 0^+} f(\eta) + \ln(\tilde{\lambda}) = -\infty$, $f(1) + \ln(\tilde{\lambda}) > 0$ and $f(\eta) + \ln(\tilde{\lambda})$ always increasing. Therefore, by the intermediate value theorem, we have that $\exists! \; \bar{\eta} < 1$ s.t. $\frac{\partial \tau(\eta_j, \overline{M})}{\partial \eta_j} > 0 \Leftrightarrow \eta_j < \bar{\eta}$. For $\tilde{\lambda} > 1$, we need $M + \bar{c} < \int N_j \left(\eta_j\right)^{\frac{1}{1-\eta_j}} \, dj$. ∎

A.2. Lemma 2

Proof. Let $G(\{l_j^G\}_j)$ be the Lagrangian of the problem of the government in Eq. (8). The derivative of the Lagrangean $G(.)$ with respect to l_j^G, that is, the marginal allocation

$$G_l \equiv \frac{\partial G}{\partial l_j^G} = T_j - \varphi \lambda,$$

where λ is the Lagrange multiplier in the resource constraint.

Case 1. Debt heterogeneity. Consider that all firms in the economy are the same except for their level of debt b_j. Then:

$$T_j = c_0^{-\eta} \left[\rho - b_j + \varphi\right]^{\eta} - c_0^{-\eta} \left[\rho - b_j\right]^{\eta} \Rightarrow \frac{\partial T_j}{\partial b_j} = -c_0^{-\eta} \eta \left(\left[\rho - b_j + \varphi\right]^{\eta - 1} - \left[\rho - b_j\right]^{\eta - 1} \right).$$

For $\eta < 1$, T_j is thus increasing in b_j. For $\eta > 1$, T_j is decreasing in b_j.

Case 2. Shock exposure Heterogeneity. Consider that all firms are the same except for their shock exposure η_j. Define $\tilde{c} \equiv \frac{c}{c_0}$ and $\tilde{\varphi} \equiv \frac{\varphi}{c_0}$. Then:

$$\frac{\partial T_j}{\partial \eta_j} = (\tilde{c} + \tilde{\varphi})^{\eta_j} \cdot \ln(\tilde{c} + \tilde{\varphi}) - \tilde{c}^{\eta_j} \ln(\tilde{c}) > 0 \Leftrightarrow (\tilde{c} + \tilde{\varphi})^{\eta_j} \cdot \ln(\tilde{c} + \tilde{\varphi}) > \tilde{c}^{\eta_j} \ln(\tilde{c}).$$

Which implies:

$$\eta_j \ln\left(1+\frac{\varphi}{c}\right) + \ln(-\ln(\tilde{c}+\tilde{\varphi})) < \ln(-\ln(\tilde{c})) \Leftrightarrow \eta_j < \eta_G^* \equiv \frac{\ln\left(\frac{\ln(\tilde{c})}{\ln(\tilde{c}+\tilde{\varphi})}\right)}{\ln(1+\frac{\varphi}{c})} > 0.$$

Therefore, T_j is strictly increasing up to $\eta_G^* > 0$ and strictly decreasing afterward. The optimal allocation is thus $l_j^G = 1$ if $\eta_j \in [\underline{\eta}_G, \overline{\eta}^G]$, where $T_{\underline{\eta}} = T_{\overline{\eta}}$ and $\int_{\underline{\eta}_G}^{\overline{\eta}^G} \varphi dj = M$, which (i) exists, since the resource constraint is binding and (ii) is unique, since T_j is quasi-concave in η_j. ∎

A.3. Lemma 3

Proof. We will proceed as in the proof of Lemma 2. Let $B(\{l_j^B\}_j)$ be the Lagrangian of the problem of the government in Eq. (8). The derivative of the Lagrangean of $B(.)$ with respect to l_j^B, that is, the marginal allocation

$$B_l \equiv \frac{\partial B}{\partial l_j^B} = \Omega_j - \varphi\lambda,$$

where λ is the Lagrange multiplier in the resource constraint.

Case 1. Debt heterogeneity. When firms are heterogeneous only in b_j, we have $\frac{\partial \Omega_j}{\partial b_j} = T_j(1-\delta+\psi_F) + \theta_j \psi_C \psi_F > 0$; that is, banks want to allocate loans to the firms with the highest levels of pre-pandemic debt per worker.

Case 2. Shock exposure heterogeneity. Consider that all firms are the same except η_j. Then:

$$\frac{\partial \Omega_j}{\partial \eta_j} = \left[\kappa(\tilde{c}+\tilde{\varphi})^{\eta_j} \cdot \ln(\tilde{c}+\tilde{\varphi}) - (\kappa - \tilde{\psi})\tilde{c}^{\eta_j} \ln(\tilde{c})\right],$$

where $\kappa \equiv (1-\delta+\psi_F)b + \beta$ and $\tilde{\psi} \equiv \psi_C(\psi_F b + \beta)$. Therefore

$$\frac{\partial \Omega_j}{\partial \eta_j} > 0 \Leftrightarrow (c+\tilde{\varphi})^{\eta_j} \cdot \ln(\tilde{c}+\tilde{\varphi}) > \left[1 - \frac{\tilde{\psi}}{\kappa}\right] \tilde{c}^{\eta_j} \ln(\tilde{c})$$

Which implies:

$$\eta_j \ln\left(1+\frac{\varphi}{c}\right)+\ln(-\ln(\tilde{c}+\tilde{\varphi}))<\ln\left(-\left[1-\frac{\tilde{\psi}}{\kappa}\right]\ln(c)\right) \Leftrightarrow \eta_j<\eta_B^* \equiv \frac{\ln\left(\left[1-\frac{\tilde{\psi}}{\kappa}\right]\frac{\ln(\tilde{c})}{\ln(\tilde{c}+\tilde{\varphi})}\right)}{\ln(1+\frac{\varphi}{c})}$$

since $\tilde{\psi}<\kappa$.

Therefore, B_j is strictly increasing up to $\eta_B^*>0$ and strictly decreasing afterward. The optimal allocation is thus $l_j^B=1$ if $\eta_j \in [\underline{\eta}^B, \overline{\eta}^B]$, where $B_{\underline{\eta}_j^B}=B_{\overline{\eta}_j^B}$ and $\int_{\underline{\eta}_B}^{\overline{\eta}_B} \varphi dj = M$, which (i) exists, since the resource constraint is binding and (ii) is unique, since B_j is quasi-concave.

Finally, we will show that: $\overline{\eta}_B<\overline{\eta}_G$ and $\underline{\eta}_B<\underline{\eta}_G$. By contradiction, assume that $\overline{\eta}_B \geq \overline{\eta}_G$. In this case, $\overline{\eta}_B \geq \overline{\eta}_G$ (from the resource constraint). The strategy of the proof is to take an alternative η smaller but sufficiently close to $\underline{\eta}_G$ and show that the profit at this point is higher than at $\overline{\eta}_G$. The T_j at this point will be closer to a point at $\overline{\eta}^G$, but the probability of survival will be much higher, and thus this point will offer a much higher profit for the bank. Mathematically, given that T_η is a continuous function at $\eta>0$, we have that $\forall\, \varepsilon>0, \exists\, \zeta>0$

$$|\eta-\underline{\eta}_G|<\zeta \Rightarrow |T_\eta - T_{\underline{\eta}_G}|<\varepsilon$$

Take $\varepsilon<\kappa^{-1}\tilde{\psi}\left[\Phi_{\underline{\eta}_G}(0)-\Phi_{\overline{\eta}_G}(0)\right]$. Then, there $\exists\, \eta=\underline{\eta}_G-\zeta$, with $\zeta>0$ such that:

$$\kappa T_{\overline{\eta}_G}+\tilde{\psi}\Phi_{\overline{\eta}_G}(0)=\kappa T_{\underline{\eta}_G}+\tilde{\psi}\Phi_{\overline{\eta}_G}(0)<\kappa T_\eta+\tilde{\psi}\Phi_{\underline{\eta}_G}(0)<\kappa T_\eta+\tilde{\psi}\Phi_\eta(0).$$

Therefore, $\overline{\eta}_B \geq \overline{\eta}_G$ cannot be optimal for the bank. ■

A.4. Lemma 4

Proof. The government prefers to allocate loans instead of delegating to banks if

$$\iota[(1-\mu)m_T+\mu\mathbb{E}[T|T\geq \underline{T}]]>[(1-\rho)m_T+\rho\mathbb{E}[T|T\geq \underline{T}]] \tag{25}$$

The left hand side comes from the welfare under the government allocation. Since T_j and T_j^G have the same distribution, the problem of the government is the same and it allocates funds based on the signal as it would allocate based on types in the benchmark problem. The right hand side of the equation comes from the analogous version of the bank problem and the assumption that T_j and Ω_j are identically distributed, and thus $\mathbb{E}[T|T \leq \underline{T}] = \mathbb{E}[\Omega_j|\Omega \leq \underline{\Omega}]$. Therefore, we can define $\iota^*(.)$ as:

$$\iota^*(\mu,\rho) = \frac{(1-\rho)m_T + \rho\mathbb{E}[T|T \geq \underline{T}]}{(1-\mu)m_T + \mu\mathbb{E}[T|T \geq \underline{T}]} \tag{26}$$

■

B. Constrained First-Best: Extensions

B.1. Example of Constrained First-Best with $\eta_j > 1$.

Suppose that there are two equally present types of firms in the economy, H and L. All firms have zero pre-pandemic profits $\pi_j = 0$. However, each firms has its own η_F, $F \in \{L,H\}$, with $\eta_H > \eta_L$. Let the total amount of the program be $M = 1$. The CFB problem in this case can be written as:

$$\max_{d\in[0,1]} d^{\eta_H} + (1-d)^{\eta_L}$$

For $\eta_H > \eta_L > 1$, this function is maximized with $d = 0$ or $d = 1$. The government in this case is indifferent to choice between allocating loans to the most or least affected firms.

■

B.2. Super and Submodular Distributions

Let $\nu_j \sim \Upsilon(\pi_j + \omega_j, \theta_j)$, where $\theta_j \in \Theta$, a complete lattice, parameterizes the distribution Υ and can be different across firms and that, as in the text, a higher θ implies that a firm is more affected, that is, the distribution parametrized by a higher θ first order stochastically dominates that characterized by a lower θ. Take $j, \hat{j}$ such that $\theta > \hat{\theta}$. Let $\omega^*, \hat{\omega}^*$ be candidates for an optimum for these two types.

Suppose by contradiction that for $\omega^* \leq \hat{\omega}^*$. For the strict inequality, since Υ is strictly supermodular

$$\Upsilon(\hat{\omega}^*,\theta)-\Upsilon(\omega^*,\theta)>\Upsilon(\hat{\omega}^*,\hat{\theta})-\Upsilon(\omega^*,\hat{\theta})\Leftrightarrow\Upsilon(\omega^*,\theta)+\Upsilon(\hat{\omega}^*,\hat{\theta})<\Upsilon(\hat{\omega}^*,\theta)+\Upsilon(\omega^*,\hat{\theta})$$

For the case where $\omega^*=\hat{\omega}^*$, $\forall\varepsilon>0$: $\Upsilon(\omega^*-\varepsilon,\theta)+\Upsilon(\omega^*+\varepsilon,\hat{\theta})<\Upsilon(\omega^*+\varepsilon,\theta)+\Upsilon(\omega^*-\varepsilon,\hat{\theta})$. The same argument applies for the submodular case. ■

C. Constraint on the Number of Loans

Throughout the paper and in our empirical exercises, we considered that the allocation of PPP loans is constrained by the total amount that can be disbursed, that is, we used the following constraint:

$$\int l_j dj = \frac{M}{\varphi}$$

This might not be the relevant constraint faced by banks or the government when approving loan applications. For instance, it can be the case that the constraint is that the system that receives applications can only accept a number of applications per day (regardless of the volume). In this case, if loans of different sizes are equally costly to process for the bank, the key constraint in the allocation problem is the number of loans. Mathematically, then, we would have that the problem of the government

$$\max_{l_j^{G,\#} \in [0,1]} \int N_j T_j l_j^{G,\#} dj \text{ s.t.} \int l_j^{G,\#} = L \tag{27}$$

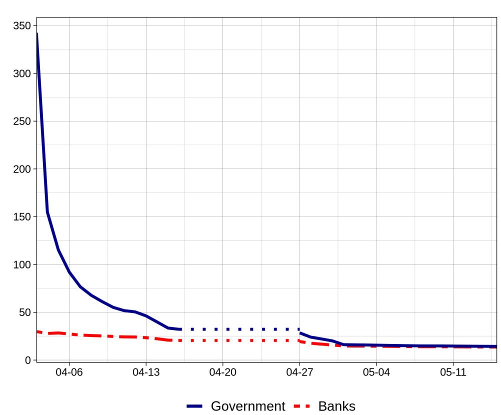
where we use the superscript # to denote the problem based on the number of loans and L denotes the maximum number of loans that can be made. The integration here is done over the distribution of firms, that is, the joint distribution of ρ, b, η, N, and not the distribution of workers (as we do in the benchmark version of our model in the main text). We can define an analogous problem for banks by replacing T_j with Ω_j.

This change in the constraint does not affect the results of Lemmas 2 or 3 if we compare two firms of the same size. An additional result that would come in this formulation, however, is that everything else constant (that is, same b, c, η), both the banks and the government would prefer to allocate funds to larger firms first. This is a consequence of the fact that with the same b, c, η, firms would have the same treatment effect per employee and thus the larger firm would have a larger effect on total employment. In our benchmark model, this does not happen as lending to the larger firm would require a larger amount, while here the constraint is based on the number of loans.

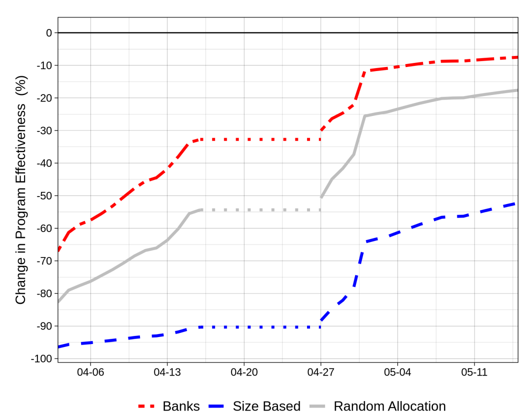
We perform the same empirical exercise under this alternative formulation. The results are shown in Figure C.1. Panel (a) shows that the optimal allocation prioritizes large firms. Even though treatment effects are decreasing in firm size, it pays off in calibration to prioritize large firms if the government is restricted by *how many* applications it can accept, since by doing so the program would reach the largest number of employees. Panel (b) shows that the banks' allocation performs relatively better relative to the random and size-based allocations now. The use of an alternative constraint allows us to illustrate an important point: the magnitude of the misallocation depends on what are the constraints the government and banks face when screening and approving PPP applications. Allocations constrained by total amount internalize the fact that applications from larger firms are "more expensive" as they apply for larger amounts, while allocations from small firms are "less expensive", as they apply for smaller amounts. Conversely, an allocation based on the number of loans effectively each PPP application has the same processing cost, regardless of firm size.

Figure C.1: Optimal, Observed and Alternative Policies: Constraint given by the Number of Loans

(a) Firm Size



(b) Relative Effectiveness



Note: This Figure is equivalent to Figure 5 in the main text. The difference is how we compute the government allocation. Instead of solving Eq. (16), we solve the dynamic version of Eq. (27) for the government. **Panel A.** Average firm size that received (banks) or that would have received (government) a loan by a given moment in time. **Panel B.** The welfare relative to the government's optimal allocation, as in Eq. (15).

D. Firms' Choice

In this appendix we discuss firms' choice in applying for the PPP and the amount they apply for in the program. Suppose that applying for the PPP has a fixed cost of F, and firms either choose to apply ($a_j = 1$) or not ($a_j = 0$) for the program. Each firm also chooses ω_j, the amount it applies for from the program per worker, subject to a program limit based on the firm's current employment level of φN_j.

A firm that borrows ω_j from the PPP *wants* to survive the pandemic if

$$\nu_j < c_j - r_G \omega_j + \pi_j^{LR} \equiv \Pi_j(\omega_j),$$

where π_j^{LR} is the perpetuity value of long-run profits of the firm, and Π_j is the total profit of the firm (both per worker). We assume that all firms that *can* survive *want* to survive—that is, $\Gamma_j \leq \Pi_j$, $\forall j$.[26] Note that this assumption does not prevent the existence of zombie firms in our model. It is possible to have firms that are not profitable without PPP funds surviving due to the program. For example, consider that $\pi_j^{LR} = 0$ and $r_G = -1$, such that $\Gamma_j = \Pi_j = c_j + \omega_j$. We can have a firm where $c_j < 0$ and $c_j + \varphi > \nu_j$ as long as the program is sufficiently generous; that is, $\varphi > \nu_j - c_j$.

The problem of the firm is given by (28), where each firm chooses to apply or not apply for the program ($a_j \in \{0, 1\}$), and the amount to request from the program ($\omega \in [0, \varphi]$) is:

$$\max_{a \in \{0,1\}, \omega \in [0,\varphi]} \int_0^{\Gamma_j(a\omega)} N_j \left[\Pi_j(a\omega) - \nu \right] d\Phi(\nu \,;\, \eta_j). \tag{28}$$

In (28), we assume that the firm chooses ω_j before observing the realization of ν_j, which is consistent with the fact that the firm does not know the extent of the pandemic or of its own exposure to it ex ante, but it knows the distribution of shocks it can face.[27]

The objective function of the firm can be rewritten as (29). The expected profit is given

[26] Lending programs are designed as short-term sources of finance for these firms, such that it is expected that $\pi_j^{LR} \geq (1 + r_G)\varphi$.

[27] This is a reasonable assumption given the uncertainty regarding the depth and duration of a pandemic. For instance, in a survey of more than 5,800 small businesses, Bartik et al. (2020) show that there is substantial disagreement on the expected duration of the COVID-19 crisis across small businesses, and the reported levels of confidence in their expected duration is low.

by the probability of survival multiplied by the expected profit conditional on survival, minus the application cost (if the firm chooses to apply):

$$\max_{\omega\in[0,\varphi],a\in\{0,1\}} \underbrace{\Phi_j(a\omega)}_{\text{Prob. Survival}} \cdot \underbrace{\left[\Pi_j(a\omega) - \mathbb{E}\left(\nu_j \mid \nu_j \le \Gamma_j(a\omega)\right)\right]}_{\text{Expected Profit}} - aF. \tag{29}$$

In (29), the problem of the firm is to balance borrowing to increase the probability of survival with reduced profitability in the future and the application cost.

All else being equal, firm j is more likely to apply for the program if $(c_j + \pi_j^{LR})T_j$ is higher, that is, if the *increase* in expected profits is higher. Additionally, applying for the PPP increases the expected cost to be paid in terms of survival, that is, $\mathbb{E}\left(\nu_j \mid \nu_j \le \Gamma_j(\varphi)\right) > \mathbb{E}\left(\nu_j \mid \nu_j \le \Gamma_j(0)\right)$, and the loan will have to be repaid at r_G. We characterize the subset of firms that apply in Lemma 5:

Lemma 5. ***Firm's Choice in the PPP.*** *If $r_G \le 0$, then all firms apply for the maximum amount of PPP funds; that is, $\omega_j^* = \varphi$. Firms apply for the PPP ($a_j^* = 1$) if:*

$$T_j\Pi_j(0) - T_j\mathbb{E}\left[\nu_j \mid \nu_j \in [\Gamma_j(0), \Gamma_j(\varphi)]\right] - \Phi_j(\varphi) r_G \varphi > \frac{F}{N_j}. \tag{30}$$

For the distribution in (31)

$$\left[\frac{1}{\eta_j+1}c_j + \pi_j^{LR}\right]T_j - \Phi_j(\varphi)\left(\frac{\eta_j}{\eta_j+1} + r_G\right)\varphi > \frac{F}{N_j}. \tag{31}$$

Proof. See the end of this section. ∎

We solve the problem in steps. First, consider a firm that has chosen to apply for the program. If the interest rate r_G is too high, then the firm does not want to borrow from the program, as Π_i^{LR} is decreasing and linear in r_G, and thus $\omega_j^* = 0$. On the other hand, if $r_G \le 0$ - as it is the case in the PPP given the implicit grants in the program - then borrowing increases the probability of survival *and* increases profits in the future. Therefore, conditional on applying, $\omega_i^* = \varphi$. Second, a firm applies for the program if the benefits of applying are larger than the fixed cost F as in (30). Using our specific distribution, this can be written as (31). Intuitively, firms with more workers (N_j) apply more often given a smaller per-worker cost of applying. This is consistent with survey evidence from Neilson, Humphries and Ulyssea (2020), who show that small businesses were less likely to be aware of, and apply for, in the first days the PPP program. We do not focus on this dimension in our paper since this differential application evidence quickly disappears by the end of the first week of the program.

Other relief programs and firm choice: As discussed in section VI, other COVID-19 lending schemes had $R_G > 0$ and prevent firms with $c_j < 0$ from applying. Firm choices in this alternative setup would be affected in two ways: First, $r_G > 0$ can change the choice of optimal amount in case a firm decides to apply. In particular, if r_G is sufficiently large, then the maximization problem described by Equation 28 when $a = 1$ has a interior solution $\omega_j \in (0, \varphi_j)$. Second, and more importantly, imposing $c_j > 0$ implies a smaller subset of firms would apply for guaranteed loans. Relative to the PPP the expected treatment effect T_j of firms who apply for guaranteed loans would be *smaller* than the expected treatment effect of firms who apply to PPP loans when $\eta < 1$. Additionally, firm eligibility constraints put a boundary on how indebted is the subset of firms that apply for guaranteed loans, reducing distortions introduced by bank incentives.

Proof of Lemma 5. First, we show the following auxiliary result. For the distribution in Eq. (2), we have that $\mathbb{E}[\nu \mid \nu \leq X] = \frac{\eta}{\eta+1} X$

$$\mathbb{E}[\nu \mid \nu \leq X] = \left(\frac{X}{c_0}\right)^{-\eta} \int_0^X \eta t \frac{1}{c_0} \left(\frac{t}{c_0}\right)^{\eta-1} dt = (X)^{-\eta} \eta \int_0^X t^\eta dt = X^{-\eta} \eta \frac{X^{\eta+1}}{\eta+1} = \frac{\eta}{\eta+1} X$$

Our proof proceeds in two steps. First, we consider the case where $a = 1$ (firm applies), and then compute the amount of funds in the application, ω. Then, we focus on which firms choose to apply.

Step 1: Choice of ω given $a = 1$. From the problem of the firm in Eq. (28), we can take the FOC w.r.t. ω when $a = 1$ to obtain

$$\phi_j(\omega) \cdot \left[\Pi_j(\omega) - \frac{\eta_j}{\eta_j + 1} \Gamma_j(\omega)\right] - \Phi_j(\omega) \cdot r_G > 0,$$

from $\Phi_j(\omega) \geq 0$ and $\Pi_j(\omega) > \Gamma_j(\omega) > \frac{\eta_j}{\eta_j+1}\Gamma_j(\omega)$.

Step 2: Choice of a**.** From the firm objective function in Eq. (29), a firm chooses to apply if

$$\Phi_j(\varphi)\Big(\Pi_j(\varphi) - \mathbb{E}\Big[\nu_j \mid \nu_j \leq \Gamma_j(\varphi)\Big]\Big) - \Phi_j(0)\Big(\Pi_j(0) - \mathbb{E}\Big[\nu_j \mid \nu_j \leq \Gamma_j(0)\Big]\Big) > \frac{F}{N_j}$$

Therefore, $a_j^* = 1$ if :

$$T_j\Pi_j(0) - \Phi_j(\varphi)r_G\varphi - \int_{\Gamma_j(0)}^{\Gamma_j(\varphi)} \nu d\Phi(\nu \mid \eta_j) > \frac{F}{N_j},$$

which delivers Eq. (30). Using the distribution in Eq. (2):

$$T_j\Pi_j(0) - T_j\mathbb{E}\Big[\nu_j \mid \nu_j \in [\Gamma_j(0), \Gamma_j(\varphi)]\Big] = (c_j + \pi_j^{LR})T_j - \Phi_j(\varphi)\frac{\eta_j}{\eta_j+1}(c_j+\varphi) + \Phi_j(0)\frac{\eta_j}{\eta_j+1}c_j, \tag{32}$$

which delivers Eq. (31). ■

E. Bank Lending Outside of the Program

We extend our model to allow banks to lend to firms outside of the program. The main idea of developing this extension to illustrate the flexibility of our framework and how it can account for the empirical findings in other loan guarantee programs, and, therefore, can be used to study the optimal allocation of relief funds in various settings. In particular, we focus on three key findings in Jimenez et al. (2022) for the loan guarantee program implemented in Spain after the pandemic:

1. Firms that were riskier and more affected by the pandemic were more likely to receive a loan through the loan guarantee programs and part of this increase was due to substitution of pre-pandemic non-guaranteed loans.

2. Firms are more likely to obtain a guaranteed loan from banks to which they have larger pre-COVID exposures.

3. Banks that participate more in the public credit guarantee scheme gain market share.

Note that the first two findings are the opposite of what the empirical literature documents in the PPP. PPP loans were allocated first to larger firms in the less affected sectors and areas in the economy. This extension shows that loans made outside of the loan-guarantee program can account for this difference.

For simplicity and exposition purposes, we make various simplifying assumptions that could be relaxed in an in-depth analysis of any specific program. First, we assume that a bank can make a loan to firms in the program or outside of the program, but not both, and that the loan size in either case is the same and exogenously determined. Second, we assume that the bank receives applications for both guaranteed and non-guaranteed loans and it is again only accepting or rejecting them (that is, the loan characteristics such as rates, maturities, etc.. are not endogenously determined). Let G denote a guaranteed loan and NG non-guaranteed loan, and let Ω_j^G and Ω_j^{NG} be the bank profit at lending to firm j under each alternative, respectively. We return to the link between Ω_j^G and Ω_j^{NG} with our benchmark model after discussing the solution to the bank problem.

The extended problem of the bank is given by

$$\max_{\{l_j^G, l_j^{NG}\}_j} \int \Omega_j^G l_j^G dj + \int \Omega_j^{NG} l_j^{NG} dj \tag{33}$$

subject to

$$\int l_j^G dj \leq M^G \text{ and } \int l_j^{NG} dj \leq M^{NG}, \tag{34}$$

and

$$l_j^G + l_j^{NG} \leq 1, \tag{35}$$

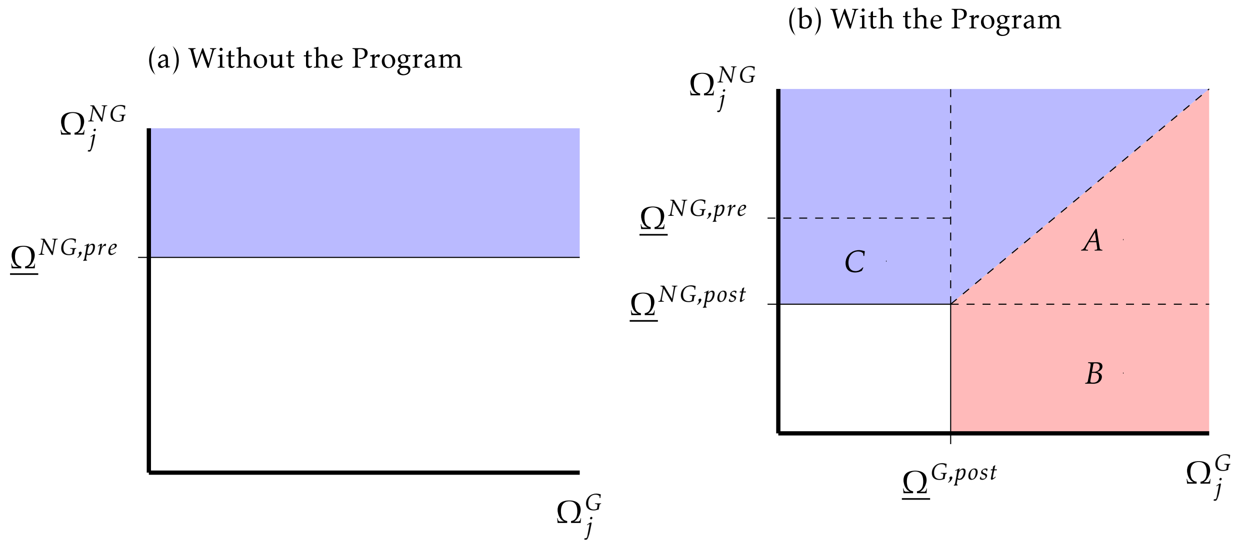
where $l_j^G, l_j^{NG} \in \{0,1\}$ indicate whether a bank allocates a guaranteed or non-guaranteed loan (or neither) to firm j, and M_G and M_{NG} are the exogenous amount of loans the bank wants to make in each category.

Our comparative statics exercise will be the introduction of a loan guarantee program. Following this introduction, both M^G and M^{NG} will adjust. Before the program, we have that $M^{G,pre} = 0$ (no program) and denote the amount of lending as $M^{NG,pre}$. Post-COVID, we have that $M^{G,post} > 0$ and $M^{NG,post} < M^{NG,pre}$, but the total amount of lending after the program is such that $M^{NG,post} + M^{G,post} > M^{NG,pre}$, that is, there is partial, but not complete, substitution between guaranteed and non-guaranteed lending, as in the empirical evidence of Jimenez et al. (2022) and Altavilla et al. (2022).

The solution of the problem of the bank can be easily summarized in Figure E.1. For each firm, let $\Delta_j = \Omega_j^G - \Omega_j^{NG}$. Without the government guarantee program, the bank simply chooses a cutoff $\underline{\Omega}^{NG,pre}$ and lends to those firms above this cutoff. These are the firms in the blue are in panel (a). With the program, the banks chooses a point $(\underline{\Omega}^{NG,post}, \underline{\Omega}^{G,post})$ such that if $\Omega_j^G \geq \underline{\Omega}^{G,post}$ or $\Omega_j^{NG} \geq \underline{\Omega}^{NG,post}$. The loan to firm j is guaranteed if $\Delta_j \geq \underline{\Delta}$, where $\underline{\Delta} \equiv \underline{\Omega}^{G,post} - \underline{\Omega}^{NG,post}$ is a function of the underlying firm distribution and M_G and M_{NG}. Firms in the blue are in panel (b) get a non-guaranteed loan, while firms in the red are get a guaranteed loan.[28]

[28]To see that this is the solution of the bank problem there are two steps. First, note that if j gets a non-guaranteed loan, then all firms k such that $\Omega_k^{NG} > \Omega_j^{NG}$ should also have gotten a non-guaranteed loan (and similarly for guaranteed loans). Second, within the set of firms that receive a loan, the bank is indifferent between allocating a guaranteed and non-guaranteed loans to firms in a line defined by $\Omega_{NG} = \Omega_G + \text{constant}$. Putting both together, we get the solution in Figure E.1.

Figure E.1: Credit Allocation With and Without the Loan Guarantee Program



Note: This figure is the visual representation of the banks' optimal allocation with and without a loan guarantee program. The firms in the blue area are those that receive non-guaranteed loans. The firms in the red area are those that receive guaranteed loans.

To understand which firms are those with high or low Ω_j^G and Ω_j^{NG}, we now return to the relationship between these objects and our benchmark model. Suppose that the bank changes r^{NG} on non-guaranteed loans and r^G on guaranteed loans, with $r^{NG} > r^G$, exogenously determined.[29] We can write Ω_j^G and Ω_j^{NG} as

$$\Omega_j^{NG} = \Omega_j + \varphi r^{NG} - (1 - \Phi_j^{\Gamma}(\varphi))(1-\delta)\varphi r^{NG} \quad \text{and } \Omega_j^G = \Omega_j + \varphi r^G \tag{36}$$

where Ω_j is as given in Eq. (13), φ is the loan size and δ is the recovery share in the case of default, as in the main text. In the PPP, banks do not fund or keep the risk of any given loan, so their objective function depends only on firm characteristics (such as debt, risk, how affected these firms are by COVID etc.). In this extension, the profit of the bank depends on the same channels as it did in the PPP, but it also depends on the repayment of the guaranteed or non-guaranteed loans. For a non-guaranteed loan the bank can charge a higher rate r^{NG}, but faces the risk of default, given by the probability of default times the unrecovered share of the loan. For a guaranteed loan the bank can

[29]We discuss how the choice of interest rates affect firms' odds of applying for different programs in Appendix D.

charge a lower rate r^G, but does not face any risk of default.

This framework accounts for the three empirical findings we highlighted at the beginning of this section. As in our benchmark model, we keep all but one firm characteristic constant and evaluate differences in allocation based on the heterogeneity over this characteristic. For two firms j and k we say that j is more likely to receive a guaranteed loan if $\Delta_j > \Delta_k$. This condition is trivially sufficient if we condition on firms that received a loan after the implementation of the program (that is, highlighted in Panel (b) of Figure E.1). This is the case in our setting, since only firms that receive a loan in Panel (b) would be in a hypothetical empirical sample. Firms that do not received a loan after the implementation of the program also did not receive a loan before its implementation, and thus wouldn't show up in the credit registry in either period.

First, we study firms that are heterogeneously affected by the pandemic or that have different baseline financial conditions. suppose that firms j and k have the same total debt (same b), the same financial conditions (same c), but firm j is more affected by the pandemic, that is, $\eta_j > \eta_k$. In this case, we have that

$$\Delta_j - \Delta_k = (1-\delta)\varphi r^{NG}\left[\Phi_k^{\Gamma}(\varphi) - \Phi_j^{\Gamma}(\varphi)\right] > 0 \tag{37}$$

and, therefore, firm j is *more* likely to receive a guaranteed loan relative to firm k. The argument is the same if we instead assume that j and k are equally affected by the pandemic (same η), but j has worse financial conditions at the moment of the shock, that is, $c_j < c_k$. As in the empirical findings of Jimenez et al. (2022) and Altavilla et al. (2022), part of this channel is through substitution of pre-pandemic non-guaranteed credit (area A in Figure E.1b) and part of it is the inclusion of new firms (area B in Figure E.1b).

Second, we study the role of within-firm bank heterogeneity. Suppose that firms j and k have the same financial conditions (same c), are equally affected by the pandemic (same η), and have the same level of *total* debt (b). Although we have so far focused on the case of a representative bank, we need to introduce some extra notation to account for the second empirical fact. Let b_j^A and b_j^B be the debt of firm j with bank A and B, respectively, such that $b_j = b_j^A + b_j^B$. Suppose, additionally, that banks A and B are identical: they make the same amount of loans in and out of the program to the same pool of clients, their only difference is the share of loans they make to specific firms. In the bank problem, the total amount of debt enters in the probability of survival function, $\Phi_j(.)$, but only b_j^B enters as the pre-COVID bank exposure to this firm. In this case, we have that if $b_j^A > b_k^A$, and , thus, $b_j^B < b_k^B$

$$\Delta_j^A - \Delta_k^A = \text{constant} \cdot \left[b_j^A - b_k^A \right] > 0 \quad \text{and} \quad \Delta_j^B < \Delta_k^B \tag{38}$$

and, therefore, firm j is more likely than firm k to receive a guaranteed loan from bank A. Thus, firms are more likely to obtain a guaranteed loan from banks to which they have larger pre-COVID exposures in our model.

Finally, firms in areas B and C represent the market share gains of banks that participate in the program in terms of their portfolio of, respectively, non-guaranteed and guaranteed loans to new borrowers, that is, in the extensive margin.[30] Therefore, banks that participate more in the public credit guarantee scheme gain market share.

[30]Since this model has no intensive margin, we do not have a channel through which banks gain market share by extending more non-guaranteed loans to their existing clients.