Bank Incentives and the Impact of the Paycheck Protection Program *

Gustavo Joaquim †
FRB of Boston

Felipe Netto
Columbia University

October 2, 2020

Abstract

The Paycheck Protection Program (PPP) administered hundreds of billions of dollars of loans and grants to small business through private banks. In this paper, we explore the optimal allocation of funds across firms and the distortions caused by allocating these funds through the banking system. We show that it can be optimal to allocate funds to the least or most affected firms depending on the nature of the shock, the firms' financial position and program design. Bank incentives distort the allocation towards firms with more pre-pandemic debt per-employee and a higher probability of survival *ex-ante*. We show that even in an idealized experiment, the distortion from bank incentives implies that firm-level regressions can *overestimate*, while regional regressions potentially *underestimate* the effect of the PPP even when controlling for PPP demand. Moreover, we show that if bank incentives are heterogeneous across banks, a bank based instrumental variable approach will likely yield biased results. Our model thus provides a unifying framework that reconciles some of the conflicting results found in the empirical literature and guides future empirical work.

*We thank Falk Bräuning, Joe Peek, Joanna Stavins, Christina Wang, Paul Willen for valuable comments and discussions. We also thank Morgan Klaeser for outstanding research assistance. The views expressed herein are those of the author and do not indicate concurrence by the Federal Reserve Bank of Boston, the principals of the Board of Governors, or the Federal Reserve System.

†Contact: gustavo.joaquim@bos.frb.org

1 Introduction

The COVID-19 pandemic has led to an unprecedented decrease in economic activity, affecting in particular small businesses (Bartik et al. (2020*a*)). In April of 2020, small business revenues decreased by more than 40% compared to January of the same year, and are still 20% down in August (Chetty et al. (2020)). As a response, Congress created the novel Paycheck Protection Program (PPP) as part of the larger CARES Act. The main goal of the program was to preserve jobs of small and medium business that were substantially affected by COVID-19. Over $650 billion dollars were allocated in the first two tranches of the program, almost two thirds of *all* interventions in the financial crisis (Bartik et al. (2020*b*)). To speed up the delivery of funds to businesses, the government used the private banking system to make decisions on applications. This gave banks the ability to target funds to preferred borrowers, particularly in the first tranche, where demand for funds overwhelmingly exceeded supply.[1]

In this paper, we develop a theoretical framework to understand the role of banks in the allocation of PPP funds. We focus on three main questions. First, we characterize the optimal allocation from the perspective of the government of PPP funds, that is, which firms/sectors/regions should be the target of the PPP. Second, we explore how bank incentives distort the allocation of these funds, that is, if a program designed like the PPP hit this target. Finally, and most importantly, we explore how the distortion in the allocation of PPP through private banks impacts the estimation of the effect of the program at the firm and regional levels. We show that the same channel can lead to overestimation of the effect of PPP at the firm level and underestimation at the regional level. Although we focus mostly on the PPP, we highlight the generalizable lessons from our paper in the design and evaluation of other lending programs.

The starting point of our theoretical framework are firms potentially in need of external finance to survive the pandemic shock. Firms have a fixed cash flow and debt level, and face a fixed cost shock to stay open as a result of the pandemic (as in Guerrieri et al. (2020)). As our focus is on the incentives of the banking system, we assume throughout the paper that firms, banks and the government have the same information. All agents know the parameters that characterize the distribution of the pandemic shock, but not the actual realization of this shock for each individual firm. This is a reasonable assumption as the first tranche of PPP

[1] According to the Census Small Business Pulse Survey, by the last week of April, 2020, 75% of eligible firms requested financial assistance, but less than 40% received it.

loans was introduced in March 27th, 2020, long before the full dimension of the pandemic was known. As shown in a survey of small business in Bartik et al. (2020*a*), for instance, there is substantial disagreement and uncertainty among small business of the duration of crisis. Firms must choose the loan amount in the application for the program, limited to a multiple of their total payroll, to balance the trade-off between borrowing more — and thus increasing their debt burden in the future — with an increase in the probability of survival. We show that if interest rates are below zero (for instance, as in the PPP where a large part of the loans are grants), all firms want to max their lending program amount subject to the rule.

Our theoretical results are divided in four parts. In the first part, we focus on the optimal allocation of the government that maximizes the number of preserved jobs if it could choose how much to lend to each individual firm given a total amount of money in the program (decided exogenously), which we denote the constrained first best allocation. Overall, our results suggest that the constrained first best PPP target is where each additional dollar is most effective, i.e. where each marginal dollar is most effective, which is not necessarily the firms that are most affected by the pandemic. When the shock has a concave distribution, the optimal allocation of funds equalizes the marginal probability of survival across different firms. For firms equally likely to be affected by the pandemic, this implies that the government finds it optimal to allocate money toward firms with a lower cash-flow (for instance, with low cash-on-hand or high debt levels), as those are the ones most in need of funds. If firms are affected differently by the pandemic (as they are in different sectors, regions etc.), the government finds it optimal to allocate funds to sectors/regions most affected by the pandemic when the program is large, but follows a inverted U-shaped pattern when the program is smaller. Intuitively, for a large program, there is enough space for the government to try to save the most affected firms, while it must choose the most cost-effective firms for a smaller program (which are exactly those intermediately affected by the pandemic).

Given our results for the constrained first best, in the second part of our model we focus on the government's allocation under the rules of the PPP, that is, when the government is subject to the same rules as the banking system (can only accept or reject applications for firms and not choose the specific amount of how much to lend to each individual firm). Contrary to the constrained first best, the government finds it optimal to allocate money toward more firms with a worse financial position if, and only if, the shock is likely to be small.

Intuitively, for relatively small shocks, the government can try to save the firms with a low probability of survival (worse financial position), but for large shocks it is more effective to focus on firms with a higher probability of survival (better financial position). Moreover, for any shock distribution, the government finds optimal to allocate funds to firms likely to be intermediately affected by the pandemic. The intuition here is the same as in the constrained first best. As the government is constrained by the rules of the program, it must choose in the extensive margin the firms with the highest *gain* in the probability of survival. Firms that are likely to be extremely affected by the pandemic will probably shut down, while firms that are not likely to be affected will probably survive — regardless of the allocation PPP funds.

In the third part of our paper, we focus on the allocation of the PPP funds when deployed through the banking system. Banks have incentives that are different from the government for two reasons. First, banks have outstanding loans with various firms that will default in case the firm does not survive the pandemic. Second, banks potentially lose clients if they don't provide them with PPP funds, even if they are not the ones where the marginal gain is the highest.[2] We show that this has two key implications. First, the banking sector distorts the government allocation toward firms with a higher ex-ante probability of survival, that is, funds allocated through the banking sector flow to firms that are less affected by the pandemic than the government finds optimal. Second, the banking sector distorts the allocation toward more indebted firms. Both of these effects are consistent with the empirical findings in Bartik et al. (2020*b*), who shows that approval rates are higher for less distressed firms and borrowers with larger existing debt. We show that this misallocation problem is worse with larger per firm lending amounts, which allows the bank to increase the probability of survival of their indebted clients by more.

Finally, we focus on what the allocation of the funds through the banking system implies for empirical work estimating firm and regional level regressions. For that, we conduct two idealized experiment within our model. We assume that banks have more information than the econometrician on the probability of survival of each individual firm.[3] We start by comparing firms within the same region-sector-bank and level of indebtedness, controlling for firms that actually apply for the PPP. Since banks select firms with an unobserved higher probability of survival *ex-ante*, this leads to a selection bias that overestimates the effect of

[2] See, for instance, this WSJ article.

[3] We assume that the econometrician perfectly observes the bank-firm relationship and thus can control for it, that is, only the first distortion in the previous paragraph is relevant empirically.

the PPP compared to a situation in which funds are either randomly allocated or optimally allocated by the government. Afterwards, we consider various different regions as the one described in the first experiment, but with a random PPP allocation across regions. In this case, comparing the different regions leads to an underestimation of the average treatment effect of the PPP, but correctly estimates the treatment effect on the treated. As the banks select firms with a higher probability of surviving, the share of firms surviving in the region with and without the PPP will be smaller than the average gain in the probability of survival with the PPP. Our framework shows that the approach in Granja et al. (2020) can recover a causal parameter if PPP exposure is not correlated with firm survival and employment. We discuss situations where this assumption can be violated. Our last result analyses the bank instrumental variable approach of Bartik et al. (2020*b*) in the framework of our model. If banks approve PPP applications at different levels at random, instrumenting PPP allocation through bank dummies corrects for the selection bias we just described. However, if both the PPP effect is heterogeneous across firms and the channels we highlight in this paper are heterogeneous across banks, the bank IV strategy can yield biased results.

Literature Review. This paper joins the growing literature exploring the economic impact and policy response following the COVID-19 pandemic, in particular the impact of the PPP. Autor et al. (2020) and Chetty et al. (2020) use the 500 employee eligibility cut-off to run a difference-in-differences analysis at the firm level. Autor et al. (2020) finds that PPP increased employment at the firm level by 2 to 4.5%, which corresponds to an effect on aggregate employment between 1.4-3.2 millon jobs through the first week of June, 2020. Neilson, Humphries and Ulyssea (2020) focus on the informational differences among small and large firms in terms of terms of application and approval rates. Erel and Liebersohn (2020) shows that there is a significant level of substitutability between traditional banks and fintechs in the PPP. Elenev, Landvoigt and Van Nieuwerburgh (2020) takes a more structural approach to understand the effects of the PPP and other policies enacted in March and April of 2020.

Closest to our paper, Bartik et al. (2020*b*) empirically investigates the targeting of funds in the PPP through the banking system. Using an IV strategy that leverages previous bank-firm relationships and heterogeneity in PPP processing at the bank level, Bartik et al. (2020*b*) shows that applications from firms more distressed business (for instance, those with less cash on hand, were less likely to be approved, despite large effects of the PPP for this set of firms. Moreover, banks favored firms with closer relationships to the bank, rather than those in greater distress.

While the rest of the literature focuses on firm level evidence from surveys or private data providers, Granja et al. (2020) takes a regional approach to understand the effect of the PPP. Granja et al. (2020) shows that the PPP funds did not flow to the areas most economically affected by the pandemic. Moreover, the authors show that there is a substantial heterogeneity at the bank level in terms of approval and disbursement, which gives rise to their identification strategy of using regional exposure to the PPP by weighting the relative disbursement of PPP banks at the national level with each bank's pre-pandemic market share in deposits by region. Using this strategy, Granja et al. (2020) find precisely estimate insignificant results in terms of employment and business shutdowns from the program.

Our contribution to this literature is twofold. First, we propose a model of the PPP that allows us to characterize the optimal allocation of PPP funds from the government and understand the sources of misallocation of using the private banking system to disburse these funds. Our model is consistent with the empirical evidence of Bartik et al. (2020*b*), where banks distort the allocation towards less distressed firms and to those with previous relationships with banks. Importantly, our model shows that this distortion is not monotonic. For instance, in our framework banks prefer to allocate funds to firms more indebted when compared to the government allocation, but not monotonically to the firms with the highest levels of pre-pandemic debt. This suggests that the strategy of constructing dummies at the median of these variables adopted by Bartik et al. (2020*b*) can be underscoring the true magnitude of these channels.

Second, as Autor et al. (2020) stresses, understanding why the firm and regional level approaches yield different conclusions is an important line of research. Our model provides a unifying framework where bank incentives distort the allocation toward firms with a higher probability of survival *ex-ante*. If this probability is unobserved, we show analytically that firm-level analysis will overestimate the Average Treatment Effect and the Average Treatment effect on the Treated (ATT), while the regional regression will overestimate the ATE and, in our idealized setting, recover the ATT. We also show that the bank IV strategy of Bartik et al. (2020*b*) can solve for the bank induced selection bias under some homogeneity conditions. If the channels we highlight in our theoretical model are heterogeneous across banks and the effect of the PPP is heterogeneous across firms (as shown empirically by Bartik et al. (2020*b*)), we show that it is possible for the bank IV strategy to yield biased results. Our framework serves as a map for empirical work on the area, as one can estimate which channels and sources of heterogeneity are important (at the bank, firm or both?) and understand what types of bias are potentially present.

The paper proceeds as follows. The next section describes the details and design of the PPP program and, more generally, the CARES act. In the next section we then focus on the model of the PPP, where we characterize the solution of the government and banking sector optimal allocations. In the fourth section we discuss the implications of our model to the estimation of the PPP effect in firm and regional level data, as well as the bank IV identification strategy. Finally, the fifth section concludes the paper and points for directions to quantify the channels we discuss and the design of lending programs in a post-crisis world.

2 The Paycheck Protection Program

Created in early April 2020 as part of the CARES act, the Paycheck Protection Program (PPP) was designed to address liquidity shortages that could lead to employment losses from small businesses. The PPP is a loan which incorporates mechanisms to incentivize firms to keep their labor force, as it should be used to pay for employee wages and other fixed costs such as rent. Funds must be used to pay for these costs over the eight-week period following the provision of the loan. The allocation of funds for the program worked on a first come, first served basis. The initial provision of funds for the PPP was of $ 349 billion, and due to high demand, got depleted by April 16. Additional $300 billion were approved by late April to increase the support of the program.

The PPP is fully forgiven if funds are used for the specified purposes of employment maintenance. Originally, to obtain full loan forgiveness, businesses were required to use at least 75% of the amount on payroll expenses and to maintain employment headcount and wage levels. This percentage was reduced to 60% after the Flexibility Act was passed in June. Importantly, loan forgiveness would be reduced if wages decrease or full time headcount decreases.

Given its small business focus, only firms with less than 500 employees were eligible to apply,[4] and each firm could apply to at most one loan from the PPP program, with a maximum of $10 million for each loan or 2.5 times the firm's average monthly payroll costs. PPP

[4]The exception were firms in restaurant and hospitality sectors (NAICS code 72), which were allowed to apply as long as they had at most 500 employees in each location.

loans have an interest rate of 1%, deferred payments for six months, and maturity of two years for loans issued in the first phase of the program and five years for loans issued after June 5. Moreover, PPP loans do not require collateral or personal guarantees.

Loans processing is performed by authorized banks, for example federally insured depository institutions and credit unions, which are responsible for checking documentation submitted by applicants, and are paid fees to cover these processing costs. Importantly, loans from the PPP carry zero risk weight for the calculus of risk weighted assets, with the purpose of minimizing the impact on banks' capital requirements. Additionally, Federal Reserve Banks were authorized to provide liquidity to banks through the Paycheck Protection Program Lending Facility (PPPL Facility). This allows Federal Reserve Banks to extend loans for institutions which are eligible to make PPP loans using such loans as collateral. Overall, the program was designed to allow a large number of institutions to process loan requests while minimizing impacts on their balance sheet structure.

As a result of the facilitated access structure, the PPP attracted a substantial number of applications. As of the beginning of August, more than 5 million loans were granted, with a total amount of about $525 billion when the program closed and stopped receiving applications. In short, the PPP was an important liquidity tool for small firms with mechanisms that incentivize employment maintenance and broad access, facing substantial demand from borrowers since its creation. For a more detailed description and history of the PPP program, see the excellent summaries in Bartik et al. (2020*b*) and Autor et al. (2020).

3 A Theoretical Model of the PPP

In this section we describe the setting of our model, discuss the optimal allocation of PPP funds for the government and the potential misallocation of using the banking sector to allocate these funds.

3.1 Firms

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

$$\pi_j \equiv \rho_j - b_j \tag{1}$$

where b_j are their debt payments per-worker to be made and ρ_j is the remainder of the cash flow per-worker (includes productivity, wages, cash-on-hand hand etc.). Without loss of generality, we normalize N_j such that $\int_j N_j dj = 1$. We assume that applying for the PPP is costless and, as consequence, all firms apply to the program.[5] Each firms chooses ω_j, the amount they apply for in the program per-worker, subject to a program limit based on their current employment level of φN_j.

We model the pandemic following Guerrieri et al. (2020). Each firm receives a fixed cost shock that must be paid in order for the firm to remain open. The per-worker magnitude of the shock is ν_j with c.d.f. denoted by Φ and p.d.f. ϕ parametrized 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

$$\nu_j < \rho_j - b_j + \omega_j \equiv \Gamma_j(\omega_j)$$

where $\Gamma_j(\omega)$ corresponds to the available funds per employee 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 pre-pandemic to remain open.[6] A firm that borrows ω_j from the lending program *wants* to survive the pandemic if

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

where π_j^{LR} is the perpetuity value of long-run profits of the firm (per-employee) and Π_j is the total profit of the firm (per-employee). We assume that all firms that *can* survive *want* to survive - that is, $\Gamma_j < \Pi_j$, $\forall j$. The motivation for this assumption is twofold. First, the lending programs are designed as a short-term source of finance for these firms, such that it is expected that $\pi_j^{LR} > (1 + r_G)\varphi$ (which guarantees that all firms that can survive want to survive). Second, we are analyzing how to allocate funds among firms. Firms that do not

[5] In our setting, we will discuss bank incentives given the set of firms that applied for the program. As most empirical analysis is conditional on application or the number of applicants, we opt for the simplicity of keeping the application decision given. It is possible to extend the model with a fixed application cost, for instance, at a heavy simplicity loss and little insight gain.

[6] As our focus is between the allocation of funds across firms, it is natural to assume that firms that are not profitable before the pandemic will shut-down and won't receive any funds from the program.

want to survive won't be part of the program as is and thus won't participate in the optimal allocation regardless.

The problem of the firm is given by (2), where each firm chooses the application amount given the rules of the program ($\omega \in [0, \varphi]$):

$$\max_{\omega \in [0,\varphi]} \int_0^{\Gamma_j(\omega)} N_j \left[\Pi_j(\omega) - \nu_i \right] d\Phi(\nu_i) \tag{2}$$

In (2), we assume that the firm chooses ω_j before observing the realization of ν_j, which stands for the fact that the firm does not know the extent of the pandemic or of its own exposure to it ex-ante, but do know the distribution of shocks they can face. This is a reasonable assumption given the uncertainty in terms of depth and duration of pandemic. For instance, in a survey of over 5,800 small business, Bartik et al. (2020*a*) shows that there is substantial disagreement on the expected duration of the crisis across small business and the reported levels of confidence in their expected duration is low. Moreover, as we want to highlight here solely the role of bank incentives, we also assume that banks and government also observe η_j (the parameter of the distribution), but not ν_j (the realization of the shock).

For tractability, we follow Guerrieri et al. (2020) and assume that the c.d.f. of the fixed cost shock distribution is given by (3)

$$\Phi(\nu \mid \eta) = \begin{cases} 0, \text{ if } \nu < 0 \\ \left(\frac{\nu}{c_0}\right)^{\eta}, \text{ if } \nu \leq c_0 \\ 1, \text{ if } \nu > c_0 \end{cases}, \text{ with } \eta > 0 \tag{3}$$

The distribution in (3) has two key characteristics that great simplify our analysis while still allowing us to focus on the different between bank and government incentives. First, The shape parameter η controls the concavity of the c.d.f., and thus we have a monotonic p.d.f., 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 the comparison between more affected (higher η) and least affected (lower η) firms. It is worth noting that for $\eta < 1$, the distribution in (3) is a truncated Pareto distribution (and converges to Pareto as $c_0 \to \infty$).

The objective function of the firm can be rewritten as (4). The expected profit is given by the probability of survival (firm wants to and can survive) multiplied by the expected profit conditional of survival,

$$\max_{\omega \in [0,\varphi]} \underbrace{\Phi\Big(\Gamma_j(\omega)\Big)}_{\text{Prob. Survival}} \cdot \underbrace{\Big[\Pi_j(\omega) - \mathbb{E}\Big(\nu_j \mid \nu_j \leq \Gamma_j(\omega)\Big)\Big]}_{\text{Expected Profit}} \tag{4}$$

For notation purposes, define ω_j^* as the solution to the problem in (4). In (4), the problem of the firm is to balance borrowing to increase the probability of survival with reduced profitability in the future. 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 $r_G < 0$, 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, thus $\omega_i^* = \varphi$.

Firm's Choice in the PPP.[7] *If $r_G < 0$, then all firms apply for the maximum amount of PPP funds, that is, $\omega_j^* = \varphi$.*

3.2 Constrained First Best

In this section we focus on the problem of the government when it can choose the amount ω_j^G per-worker to lend to each firm to maximize the number of surviving jobs outside of the rules of the PPP program. We show that the government want to allocate funds to where their marginal effect is the highest, which does not necessarily corresponds to the places most/least affected by the shock. In particular, we show that the marginal effect depends on the shock distribution, the size of the program, and the initial financial condition of the affected firms. The target of the PPP thus is neither obvious nor fixed (as more information on the depth of the pandemic becomes available). For instance, it is intuitive that if the shock for a firm (or region/sector) is large enough, the government does not always find optimal to save this firm as the opportunity cost of not allocating these funds for other firms is too high. Our analytical results below formalize this intuition and characterizes what is then the government optimal allocation.

We separately compare firms with the same shock exposure (η_j) but heterogeneous financial position first and then firms with the same financial position (π_j) and different shock exposures separately in this section and throughout the paper. The motivation for this is twofold. First, it highlights the key channels of the PPP allocation in our model for different

[7] See Appendix A.1 for details.

sources of heterogeneity across firms. Second, it speaks more directly with the empirical literature which generally tries to control for either of these factors (with firms controls, fixed effects etc.) and focus solely on of them at a time (for instance, in Bartik et al. (2020*b*)).

We assume that the government observes for each firm j their cash-flow per-employee π_j and the number of workers N_j. The government does not observe the cost shock to each firm, ν_j, and thus can only choose the amount ω_j^G per-worker to lend to each firm based on the distribution of the shock. For now we assume the amount of lending in the program, denoted by M, is fixed. [8] The problem of the government is given by (5)

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

We denote the solution to the maximization problem in (5) as the *Constrained First Best* (CFB). In the constrained first best, we have that the solution is to allocate funds to where their marginal effect is higher, that is, where the *marginal* dollar will increase the probability of survival of firm j the most.

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. Let $\overline{\pi} \equiv \int_j N_j \pi_j dj$ denote the average pre-shock financial position of all firms. Define $\overline{M} \equiv M + \overline{\pi}$, which is the relative size of the lending program. The higher M is, more can be done within the program and the higher $\overline{\pi}$ is, the least needs to be done. All proofs and derivations are in Appendix A.

Lemma 1. ***Constrained First Best with*** $\eta_j < 1$***.*** *The solution to (5) entails an equal gain in the probability of survival across firms, that is, for firms i, j*[9]

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

Using the distribution in (3), we have that

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

where $\tau(\eta, \overline{M})$ is a exposure-based transfer that sums to zero, $\sum_j \tau(\eta_j, \overline{M}) = 0$. Furthermore, we

[8] We return to the choice of M in Section 3.4

[9] This is the interior solution to the problem of the government when c_0 is sufficiently large (e.g. $c_0 > M + \overline{\pi}$ is sufficient).

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 $\omega_j^{G,*} - \omega_i^{G,*} = -\pi_j$, while for two firms with the same financial position $\pi_j = \pi_i$, we have that $\omega_j^{G,*} - \omega_i^{G,*} = \tau(\eta_j, \overline{M}) - \tau(\eta_i, \overline{M})$. Intuitively, Lemma 1 shows that (i) the optimal policy maximizes the marginal probability of survival across firms (eq. (6)) and (ii) this can be decomposed in the cash flow needs from 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 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. Firms that are strongly affected would cost too much to save, while the least affected firms can likely survive without PPP funds.

Sectoral/Regional Allocation. Our result in Lemma 1 is also useful to analyze 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. (2020)). The result that optimal policy will equate the marginal probability of survival across firms in (6) is still true across sectors and regions. For instance, if different sectors have different initial levels of debt per-employee, sectors with relatively *more* debt per-employee should receive more of the funds since the probability that a firm of 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/regions have shocks with a different distributions (that is, different exposures to the pandemic), the optimal transfers across sectors are those given by $\tau(\eta_j, \overline{M})$, and shouldn't necessarily go to the most affect regions.

Other cases. In Lemma 1, we focused in the case where $\eta_j < 1, \ \forall j$. In Appendix A.3, we show in a simple example that when $\eta_j > 1$, the problem of the government is convex and the solution will be to allocate funds to either the firms with the lowest or highest π or η, that is, the government is indifferent 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 $\Upsilon\left(\pi_j + \omega_j \mid \theta_j\right)$ parametrized by θ_j. We show in Appendix A.4 that the government wants to 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 benefit of the marginal dollar rather than the funding needs of each individual firm.

3.3 Government's Optimal Allocation under the PPP Rules

In Section 3.2 we considered the government allocation when the government could choose how much to lend to each firm. To make the comparison with the bank allocation direct, we now focus on the government allocation under the same rules as the banks in the PPP. In this case, governments can choose to accept or reject applications from firms, but can't change the loan allocation in the intensive margin. This problem guarantees that we are comparing the bank allocation with an equally constrained by the program government allocation (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 (8):

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

where we define $\Phi_j^\Gamma(\varphi) \equiv \Phi\left(\Gamma_j(\varphi) \mid \eta_j\right)$. Alternatively, we can write the problem of the government as (9):

$$\max_{\{l_j^G \in [0,1]\}} \int N_j l_j^G T_j \, dj \;\; \text{s.t.} \;\; \int N_j l_j^G \, dj = \frac{M}{\varphi} \tag{9}$$

where T_j is the treatment effect for firms of type j between receiving φ or 0 of loans from the PPP as in (10):

$$T_j \equiv \Phi\left(\Gamma_j(\varphi)\right) - \Phi\left(\Gamma_j(0)\right) \tag{10}$$

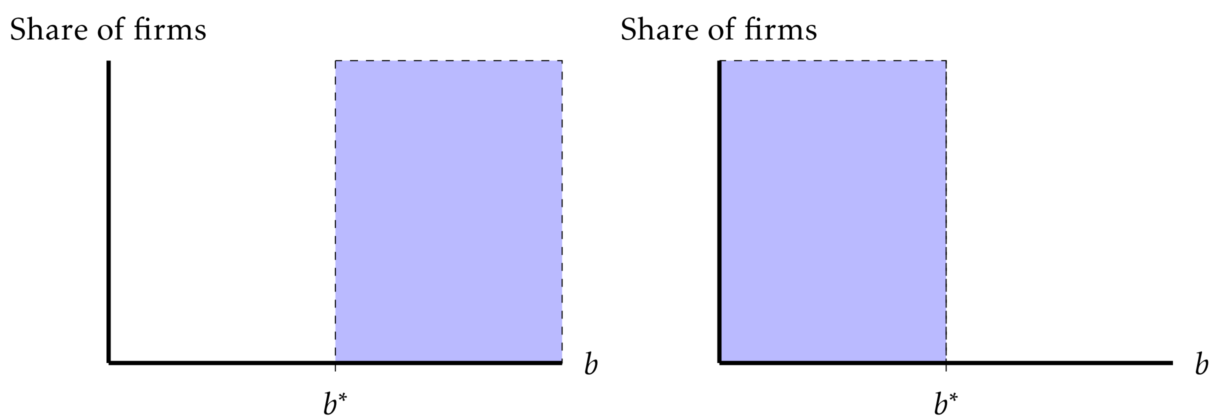
From (9), it is clear that the government wants to approve the application of firms with the *highest treatment effect*. The key question is which firms are those in the case of the PPP. In the first result of Lemma 2, we show that for firms with the same η, the government wants to allocate funds to 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 simply prefers to allocate the funds to firms with relatively low levels of debt, as those are the ones the government can still save in face of the pandemic. The insight here is that the treatment effects are a joint product of the firm financial position and the nature of the shock distribution, and thus do not have a distribution or model free ranking. This first result of Lemma 2 is illustrated in Figure 1.

Lemma 2. ***Government PPP allocation.***

Debt heterogeneity. *Consider that all firms in the economy are equal except for their level of debt* b_j*. The solution to* (9) *implies that* $\exists! b^*$ *such that: (i) for* $\eta < 1$, $l_j^G = 1$ *if* $b_j > b^*$ *and* $l_j^G = 0$ *otherwise (and the opposite for* $\eta > 1$*, and (ii) the opposite for* $\eta > 1$.*[10]

Shock Exposure Heterogeneity. *Consider that all firms are equal except for their shock exposure* η_j*. The solution to* (9) *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.*

Figure 1: Optimal Government Allocation under the PPP for Firms with Different Debt Per-worker b: $\eta < 1$ (left) vs $\eta > 1$ (right)



Note: This figure is the pictorial representation of Lemma 2. We denote b as the pre-pandemic debt per-worker of each firm.

In the second result of Lemma 2, we show that for firms with the same financial position π, the government wants to allocate funds to firms with intermediate exposure to the pandemic

[10]Note that b^* is different if η is < or > than 1. Moreover, for $l_j^G = b^*$, the government is indifferent in terms of the allocation. As there is a continuum of firms, the allocation to this specific set of firms is irrelevant at the aggregate level.

shock. This result has the same intuition as the case where M is small in the constrained first best. The most affected firms won't survive with the extra φN_j, while the least affected firms will likely survive regardless, such that φN_j is too much to allocate them. Here, contrary to the constrained first best, this is not a function of the total size of the program, M, since the amount in the intensive margin the government can allocate to each firm is fixed. Pictorially, this second result of Lemma 2 is in Figure 3, where we compare the optimal allocation of the government vs the private banking sector (Section 3.4). This result reinforces the idea that the optimal target of the PPP is not the most affected firms (or sectors and regions), since the return can be lower given the likely magnitude of the shock.

3.4 Banks' Allocation in the PPP program

We now focus on the private banking sector allocation in the PPP program. As in the government optimization problem (8), banks can choose to accept or reject applications from firms, but can't change the loan allocation in the intensive margin. The difference is that the banking sector is interested in maximizing their profits — and not firm and job survival. As small business lending is regional and specialized (Granja, Leuz and Rajan (2018)), we assume that there is a representative bank.[11] Banks receive positive profits from making more loans (and thus the constraint on the total amount available for the program will be binding).

If the banks accepts the PPP application of a firm, there are two possible scenarios. If the firm survives, the bank recovers $N_j b_j$ of the current loan and a present value of $\psi_F N_j$ of potential future loans to this firm. If the firm does not survive, the bank receives a share $\delta \in (0,1)$ of their outstanding loans, that is, $\delta N_j b_j$. If the bank rejects the PPP application, there are same two possible scenarios. However, we additionally assume that if the firm survives, the bank loses a share of ψ_C of future clients that switch bank providers. Mathematically, the problem of the bank is given by (11)

$$\begin{aligned}\max_{\{l_j^B \in [0,1]\}} \int N_j \left\{ \Phi_j^\Gamma(\varphi) \left[1 + \frac{\psi_F}{b_j} \right] + \left[1 - \Phi_j^\Gamma(\varphi) \right] \delta \right\} b_j l_j^B \, dj \\ + \int N_j \left\{ \Phi_j^\Gamma(0) \left[1 + (1-\psi_C) \frac{\psi_F}{b_j} \right] + \left[1 - \Phi_j^\Gamma(0) \right] \delta \right\} b_j (1 - l_j^B) \, dj\end{aligned} \tag{11}$$

[11] One can relax this assumption to have a representative bank for a subset of firms, such as firms in a sector-region pair, but we prefer to model as a single representative bank for simplicity.

or, alternatively,

$$\max_{\{l_j^B \in [0,1]\}} \int N_j \left\{ T_j \left[(1-\delta) b_j + \psi_F \right] + \psi_C \psi_F \Phi_j^\Gamma(0) \right\} l_j^B \, dj \tag{12}$$

in either case subject to the total lending constraint $\int N_j l_j^B \, dj = \frac{M}{\varphi}$.

In our setting, there are two channels through which profits of the banking sector deviate from the objective function of the government. First, the banking sector already has an heterogeneous exposure to firms that they have outstanding loans and potential future loans to be made to this firm, which is captured by a combination of the default recovery δ and future profits ψ_F. Second, banks are also concerned about the probability of survival of the firm $\Phi_j^\Gamma(0)$, as those are potential clients to switch banks if they don't receive PPP loans.

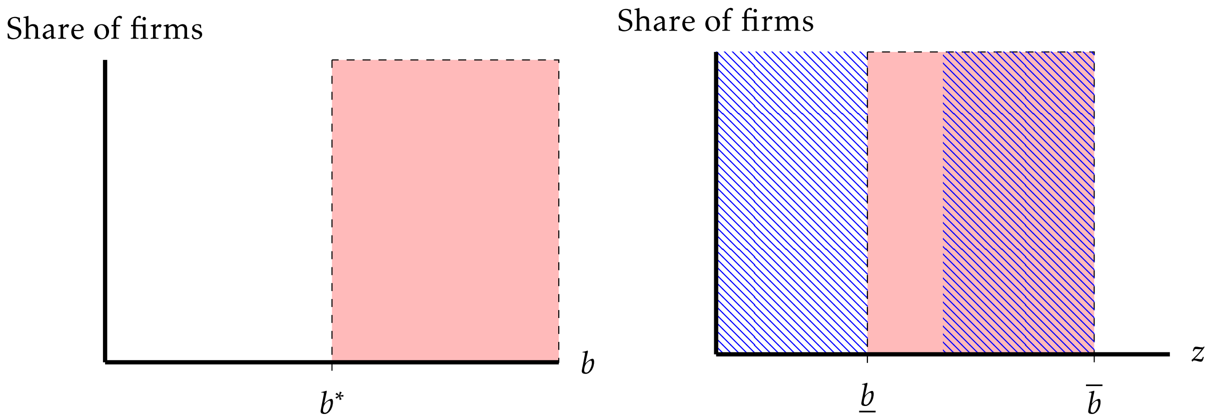
Together, these two differences imply that the optimal bank allocation is different from the optimal allocation from the government under the rules of the program. We show in the first part Lemma 3 that if firms have heterogeneous levels of debt, banks distort the allocation towards firms with higher pre-pandemic debt per employee in the case $\eta > 1$. Note that the banking sector does not want to necessarily save those firms with the highest levels of debt, given that their chance of survival may be too small to compensate the higher level of outstanding loans. For $\eta < 1$, both banks and the government want to save firms with the highest levels of pre-pandemic debt per employee. Importantly, this shows that it is not always the case that the allocation using the banking sector is different than the optimal government allocation, such that there is no tradeoff in the speed of delivery of funds and incentives of the banking sector. The pictorial representation of this result is in Figure 2. In the left panel ($\eta < 1$), there is no misallocation. In the right panel, the misallocation is given by the dotted area, where banks distort the allocation towards firms with more debt. The latter case is consistent with the empirical findings of Bartik et al. (2020*b*). In particular, Bartik et al. (2020*b*) shows that conditional on the set of firms with the relationship with a bank (in the extensive margin), banks approved more loans of firms with higher pre-existing debt at what the authors call a "striking magnitude".

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. The solution to (12) implies that: (i) for $\eta < 1$, banks' allocation coincides with government allocation, $l_j^B = l_j^G$, and (ii) $l_j^B = 1$ if $b_j \in [\underline{b}, \overline{b}]$ and $l_j^B = 0$ otherwise, that is, banks give preference*

to firms with higher b_i than the government.

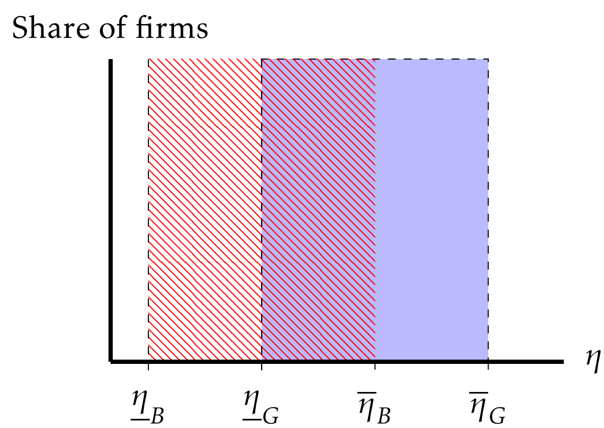
Shock Exposure Heterogeneity. *Consider that all firms are the same except for their exposure η_j. The solution to (12) implies that $\exists!\ \underline{\eta}_B, \overline{\eta}^B$ such that the bank chooses $l_j^G = 1$ if $\eta_{rs} \in [\underline{\eta}, \overline{\eta}]$ and $l_j^G = 0$ otherwise. Additionally, $\underline{\eta}_B < \underline{\eta}_G$ and $\overline{\eta}_B < \overline{\eta}_G$, that is, bank's distort the allocation to firms with a higher probability of ex-ante survival.*

Figure 2: Bank Allocation under the PPP: $\eta < 1$ (left) vs $\eta > 1$ (right)



Note: The shaded area in the RHS figure denotes the misallocation, i.e., the difference from the government to the bank solution. There is no misallocation in the case where $\eta < 1$(Lemma 3).

Figure 3: Credit Allocation under the PPP for firms with Heterogeneous Shock Exposure (η_j): Government (blue, solid) vs Banks (red, dotted)



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

In the second result of Lemma 3, we show that for firms with the same financial position π, the banking sector distorts the optimal government allocation towards firms less affected by the pandemic, that is, those with lower η's and a higher probability of survival *ex-ante*.

Pictorially, this second result of Lemma 2 is in Figure 3. This result is also consistent with the evidence in Bartik et al. (2020*b*), which finds that banks accept more applications from less distressed firms.

In also show that the channels highlighted in Bartik et al. (2020*b*) may not be monotone. Lemma 3 compares the allocation of the private banking sector with the government's allocation, but now with a random allocation of funds across firms. Take the second result in Lemma 3, for instance. We show that if banks want to allocate funds to less distressed firms compared to the government's allocation, but that the relationship is non-monotone with the level of distress in the population. Therefore, the strategy of Bartik et al. (2020*b*) of separating the sample at the median level of distress can lead to an underestimation of the bank targeting channel.

Misallocation and program design. A natural question from from Lemma 3 is: given the role of banks in allocating funds, what is the optimal program design in terms of the size of the program M and maximum amount per firm φ? To answer this question, let the social welfare evaluated at a feasible allocation be given by (13)

$$\mathcal{W}(\{l_j\}) \equiv \int N_j l_j T_j \, dj \tag{13}$$

and let $\Delta\mathcal{W} = \mathcal{W}(\{l_j^G\}) - \mathcal{W}(\{l_j^B\})$, that is, $\Delta\mathcal{W}$ represents the misallocation of PPP funds through the allocation using the private banking system. For instance, if $\Delta\mathcal{W}$ is increasing in φ, the optimal φ will be smaller due to the allocation of funds through the banking system. The proofs for the results in this section are in Appendix A.7, and we focus here on the intuition of the results.

First, consider the case where there is misallocation due to debt heterogeneity. For M large enough,[12] as in the case of the PPP, the misallocation is decreasing in M and increasing in φ, that is:

$$\frac{\partial \Delta\mathcal{W}}{\partial M} < 0 \text{ and } \frac{\partial \Delta\mathcal{W}}{\partial \varphi} > 0$$

Intuitively, when the program is larger, the government will expand the optimal allocation towards more indebted firms, and banks will expand their allocation in both directions, increasing the set of firms both the banking sector and government want to receive PPP funds.

[12] Technically, we need $b^* > \underline{b}$ in Figure 2, that is, that there is a positive measure of firms which both the government and banking sector want to fund through the program.

On the other hand, when φ increases, banks can direct even more funds toward clients with pre-existing debt and increase their likelihood of survival while the government would rather have these funds go the the least indebted firms.

Second, consider the case where there is misallocation due to heterogeneous exposure. In this case, our model shows that the relationship between M, φ and misallocation is ambiguous. For instance, consider the case of φ. Both the government and the private banking sector will move their optimal allocation towards the middle of their respective intervals in Figure 3, which can increase or decrease the difference among the two and have an ambiguous effect in terms of the treatment effect of individual firms.

4 Firm and Regional Level Regressions

In this section we evaluate how the bank incentives we highlighted in Section 3.4 affect the estimation of the PPP effect at the firm and regional levels. We start our analysis considering an estimation that controls for demand for loans (for instance, that conditions on applications for the program, sector-region fixed effects etc.). We focus on the case where the econometrician observes the debt level of the firm (and thus control for it), but does not observe the true probability of survival. If this probability is unobserved, the effect at the firm level will underestimate the ATE and ATT, while under some assumptions, the analysis at the regional level (as in Granja et al. (2020)) will underestimate the ATE but correctly estimate the ATT.

Finally, we consider the identification strategy of Bartik et al. (2020*b*) in our framework. Given that banks have approved PPP loans at heterogeneous rates, Granja et al. (2020)) use bank indicators as an instrument for PPP approval on a sample of firms that applied for the program. We show that their strategy corrects the bias we highlight if either the channels of Lemma 3 are homogeneous across banks or the effect of the PPP is homogeneous across firms. If both are heterogeneous, we show that it the bank IV strategy can also lead to biased results. The bias comes from the correlation of the heterogeneity of bank allocation and higher/lower firm effects.

Notation. Before proceeding with our results, we need to introduce some notation. We assume now that there are i subsets of firms, each with a share α_i. All firms of the the same subset i have a common treatment effect T_i. In this case, we define

$$ATE \equiv \int_i \alpha_i T_i \, di, \;\; ATT \equiv \int_i \alpha_i^B T_i \, di, \;\; \text{and} \;\; ATG \equiv \int_i \alpha_i^G T_i \, di \tag{14}$$

The ATE is the average treatment effect using the population weights, the ATT is the treatment effect of firms that actually receive funding through the program, that is, those that $l_i^B = 1$, while ATG is the treatment effect if the government was allocating the funds (instead of the private banking sector). Moreover, define the conditional probability of a firm being the type i in the set of firms that receive the PPP in either the government(G) or bank (B) allocations as (15)

$$\alpha_i^P = \mathbb{P}\left[i \mid l_i^P = 1\right], \;\; P \in \{B, G\} \tag{15}$$

Idealized Experiments. We consider two idealized experiments in our model. First, we focus in a single region. In this region, all firms are the same (same sector, debt, productivity etc.), except for an unobserved part in η, that is, if $\eta_i = \tilde{\eta}_i + \zeta_i$, the researcher only observes $\tilde{\eta}_i$. We consider the following firm level regression:

$$y_j = \alpha + \beta^F PPP_j + \varepsilon_j \tag{16}$$

where y_j is a dummy that is equal to one if the firm survives and PPP_j is a dummy if firm j receives PPP funds. In our exercise, we don't need to include region-industry-week fixed effects, firm fixed effects and controls as we assume here we are comparing firms within the same sector-region-time and that are equal in everything other than an unobserved part in η. For simplicity, we assume that the size of the PPP is such that half of the firms receive funds. We show in Appendix A.8 that

$$\begin{aligned}\mathbb{E}\left[\widehat{\beta^F}\right] &= ATE + \text{Cov}\left\{\alpha_i^B - \alpha_i, \, \Phi_i^\Gamma(\varphi) + \Phi_j^\Gamma(0)\right\} \\ &= ATG + \text{Cov}\left\{\alpha_i^B - \alpha_i^G, \, \Phi_i^\Gamma(\varphi) + \Phi_i^\Gamma(0)\right\} \\ &= ATT + 2\,\text{Cov}\left\{\alpha_i^B - \alpha_i, \, \Phi_i^\Gamma(0)\right\}\end{aligned} \tag{17}$$

From (17), we can show that the OLS estimator of β in (16) will overestimate all of the treatment effects when the bank distorts the allocation toward firms with a higher probability of survival. Intuitively, the researcher does not observe ζ_i, and a higher ζ_i has two implications: a lower probability of survival regardless of PPP allocation (lower ε_j) and a lower probability of receiving a PPP loan (from the problem of the bank in the second part of Lemma 3), we have a positive correlation between the shock ε_j and PPP_j.

In the second idealized experiment, we consider case where PPP funds are randomly allocated across regions and a researcher runs the following regression at the regional level

$$y_r = \alpha + \beta^F PPP_r + \varepsilon_r, \tag{18}$$

where y_r is there share of firms in region r that survive and PPP_r is the share of firms in r that receives PPP funds . Similarly to the firm level regression, we don't include here other controls or normalize PPP_r by the deposit exposure as we assume that PPP_r is exogenous. In particular, we assume that there is a large number of regions with the exact same composition of firms and that half of them receive PPP funds and half of them do not. For those that do receive PPP funds, we assume for simplicity, as in our first idealized experiment, that the size of the PPP is such that half of the firms receive funds. We show in Appendix A.8 that

$$\begin{aligned}\mathbb{E}\left[\widehat{\beta}^R\right] &= ATE + \operatorname{Cov}\left\{\alpha_i^B - \alpha_i,\ T_i\right\} \\ &= ATG + \operatorname{Cov}\left\{\alpha_i^B - \alpha_i^G,\ T_i\right\} \\ &= ATT\end{aligned} \tag{19}$$

From (19), we can show that the OLS estimator of β in (18) will underestimate the ATE and ATG, but will correctly estimate the ATT for the PPP. Intuitively, the difference between the two regions of firm survival will be smaller than the ATE since banks will divert funds to firms that would have survived regardless of the PPP allocation. Lemma 4 formalizes these results.

Lemma 4. ***Bank Incentives and Bias with Independent Regions.*** *Suppose that* $\min_j \eta_j \in [\underline{\eta}_B, \underline{\eta}_G]$ *and* $\max_j \eta_j \in [\overline{\eta}_B, \overline{\eta}_G]$ *(for instance, M and ψ_C relatively large) such that either a bank or the government wants to allocate money to every firm that applies for the program. Then:*

$$\mathbb{E}\left[\widehat{\beta}^F\right] > \max\{ATE, ATT, ATG\} \quad \textit{and} \quad \mathbb{E}\left[\widehat{\beta}^R\right] < \max\{ATE, ATG\}$$

Example. To strengthen the intuition of our results in Lemma 4, we provide an example. Consider that in each region there are two types of firms: H and L (unobserved for the econometrician). A share α of firms has a type L and $1/2$ of the firms receive loans through the program. We assume that L firms will survive with probability one if they receive PPP funds and will not survive otherwise, that is $\Phi_i(\varphi; L) = 1$. We assume that H firms will survive with probability one regardless of the PPP allocation $\Phi_i(\varphi; H) = \Phi_i(0; H) = 1$. In this case, the $ATE = \alpha$. Consider the extreme case where banks distort the allocation in way that only H firms receive loans. This yields $\mathbb{E}\left[\beta^F\right] = 1$, that is, all firms that receive PPP funds survive and those who don't receive don't survive. Simultaneously, this yields $\mathbb{E}\left[\beta^R\right] = ATT = 0$, since the firms that received PPP loans would have survived regardless.

Interconnected Regions. Our result in Lemma 4 suggests if the measure of regional exposure in the PPP used in Granja et al. (2020) is in fact exogenous, than the ATT effect of the PPP is indeed small. One strong assumption we made to reach this result is that regions are not connected, that is, there is no across region relocation of funds of PPP loans. This is an assumption that can be empirically tested using the bank-time fixed effect strategy of Drechsler, Savov and Schnabl (2017), which we intend to do in future versions of this paper. If this assumption of independent regions is violated, we are back to our firm-level regressions where banks will distort the allocation toward places with less affected firms, which affects the estimation of the PPP impact.

4.1 Bank Instrumental Variables

We now explore the implications of our model of using bank approval rates as an instrument for PPP approval, as in Bartik et al. (2020*b*). In the version of the model we discussed so far, the bank IV strategy corrects for the biases discussed in the previous section. Too see that, consider a situation where there are two banks b, $\hat{b}$, each in completely separate but otherwise identical regions. The only difference is that bank b accepts a share $s_b > s_{\hat{b}}$ of the applications made. Therefore, as in Bartik et al. (2020*b*), one can instrument for PPP_i using the exposure to bank b, $\mathbb{1}_b$ and run (16) in the set of firms that apply for the PPP to recover β. In this case, as firms are similar except for their banks handling of the PPP, we have that $\mathbb{E}\left[\mathbb{1}_b, \varepsilon_j\right] = 0$

However, given that the focus of our paper and Bartik et al. (2020*b*) is on the heterogeneous treatment effects of PPP loans and there is evidence that larger banks (which potentially have a lower ψ_C) approved less loans in the PPP, we consider the case where the treatment effect is heterogeneous across firms and ψ_C is heterogeneous across banks.

Suppose that $\psi_C^b > \psi_C^{\hat{b}}$ and that each region has a continuum of firms. For simplicity, we assume that the exposure of firms to the shock is parametrized as $\xi \in [0,1]$ and that the probability of survival of firm i is given by (20)

$$y_i = \alpha \xi_i + [\beta + \gamma \xi_i] PPP_i + \varepsilon_i \tag{20}$$

where ε_i is true exogenous shock i.i.d. across firms. Using the same proof as the second part of Lemma 3, it is easy to show that a lower ψ_C will dislocate the allocation interval of bank b, $[\underline{\xi}_b, \overline{\xi}_b]$, to the right. Mathematically, this implies that

$$\mathbb{E}[\xi | PPP = 1, b] < \mathbb{E}[\xi \mid PPP = 1]$$

Given that $\text{Cov}(PPP_i, \xi_i) \neq 0$, the OLS estimation of the treatment effect in (21) will yield biased results

$$y_i = \alpha_0 + \beta_0 PPP_i + \epsilon_i \tag{21}$$

To understand the role of the instrument in this case, first consider the case where $\gamma = 0$. Then the expected value of the IV estimator in (21) is

$$\mathbb{E}\left[\hat{\beta}^{IV}\right] = \frac{\text{Cov}(y_i, \mathbb{1}_b)}{\text{Cov}(PPP_i, \mathbb{1}_b)} = \beta + \alpha \frac{\text{Cov}(\xi_i, \mathbb{1}_b)}{\text{Cov}(PPP_i, \mathbb{1}_b)} = \beta + \frac{\alpha}{2} \frac{\mathbb{E}[\xi_i \mid b] - \mathbb{E}[\xi_i]}{\text{Cov}(PPP_i, \mathbb{1}_b)}$$

Therefore, we have that $\mathbb{E}\left[\hat{\beta}^{IV}\right] = \beta$ if there is no composition effect through which more or less affected firms have a relationship with a specific bank conditional on applying for the PPP. This assumption can be violated if, due to different ψ_C's, banks encourage or make easier for differently affected firms to apply.[13] Then even if all banks face the same potential set of customers, conditional on application, we will have a bias. If bank b, that accepts more applications, is such that $\psi_C^b > \psi_C^{\hat{b}}$, we will have that lower risk people will be encouraged to apply more often, and thus $\mathbb{E}[\xi_i \mid b] < \mathbb{E}[\xi_i]$.

Second, consider the case where $\alpha = 0$. In this case:

$$\begin{aligned}
\mathbb{E}\left[\hat{\beta}^{IV}\right] &= \frac{\text{Cov}(y_i, \mathbb{1}_b)}{\text{Cov}(PPP_i, \mathbb{1}_b)} = \beta + \gamma \frac{\text{Cov}(\xi_i PPP_i, \mathbb{1}_b)}{\text{Cov}(PPP_i, \mathbb{1}_b)} \\
&= \beta + \gamma \mathbb{E}[\xi_i | PPP_i = 1] + 2\gamma \frac{s_b}{s_b - s_{\hat{b}}} \{\mathbb{E}[\xi_i \mid b, PPP_i = 1] - \mathbb{E}[\xi_i \mid PPP_i = 1]\}
\end{aligned}$$

Intuitively, if each bank has a different ψ_C, it will distort the allocation heterogeneously,

[13] See andecdotal evidence of this, for instance, here.

which introduces a correlation between ϵ_i and s_b. Note that we need both $\gamma \neq 0$ and ψ_C different across banks to generate an under/overestimation of the ATT. Together, both sources of heterogeneity generate a correlation between bank exposure and the unobserved shock that affects probability of survival.

5 Concluding Remarks and Next Steps

This paper provides a framework to understand the role of the private banking sector in the allocation of the Paycheck Protection Program (PPP), a large novel crisis response adopted by the U.S. as a response to the Covid-19 crisis. We consider three main dimensions. First, what is the optimal target for the PPP? Second, how far are we from this optimal by using the private banking sector to allocate these funds? Third, what is the effect of bank incentives in the estimation of the effect of the PPP at the firm and regional levels?

We find that the target of the PPP is firms for which the treatment effect is the highest, which are not necessarily those most or least affected by the pandemic. We show that the optimal allocation can go from helping the least to the most affected firms depending on the nature of the shock, the size of the lending program and the financial position of these firms. Moreover, in our setting, as banks already have outstanding loans to firms and potentially lose clients that they don't provide PPP loans, banks distort the allocation towards more indebted firms and firms with a higher probability of survival *ex-ante*, as in the empirical evidence of Bartik et al. (2020*b*). The latter distortion can cause results using firm level regressions to overestimate the effect of the PPP if the probability of survival is not perfectly observed by the econometrician. At the regional level, the unobserved probability of survival causes the effect to capture the ATT, but underestimates the ATE. The bank IV strategy of Bartik et al. (2020*b*) corrects for the selection bias we highlight, but can yield problematic results if bank incentives are heterogeneous across bank and the PPP effect is heterogeneous across firms.

Our ultimate goal is to use the same data as Granja et al. (2020) to estimate the relevance of our channels empirically. We plan to estimate the parameters of our model by comparing PPP allocation across branches of the same bank in different locations (that is, using the Drechsler, Savov and Schnabl (2017) identification strategy).

References

Autor, David, David Cho, Leland D Crane, Byron Lutz, Joshua Montes, William B Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. 2020. "An Evaluation of the Paycheck Protection Program Using Administrative Payroll Microdata *." 5, 6, 8

Bartik, Alexander W., Marianne Bertrand, Zoe Cullen, Edward L. Glaeser, Michael Luca, and Christopher Stanton. 2020*a*. "The impact of COVID-19 on small business outcomes and expectations." *Proceedings of the National Academy of Sciences*, 117(30): 17656–17666. 2, 3, 10

Bartik, Alexander W., Zoe E. Cullen, Edward L. Glaeser, Michael Luca, Christopher T. Stanton, and Adi Sunderam. 2020*b*. "The Targeting And Impact of Paycheck Protection Program Loans to Small Businesses." NBER Working Paper No. 27623. 2, 4, 5, 6, 8, 12, 17, 19, 20, 23, 25

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." 2, 5

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2017. "The deposits channel of monetary policy." *Quarterly Journal of Economics*, 132(4): 1819–1876. 23, 25

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh. 2020. "Can the Covid Bailouts Save the Economy?" National Bureau of Economic Research Working Paper 27207. 5

Erel, Isil, and Jack Liebersohn. 2020. "Does FinTech Substitute for Banks? Evidence from the Paycheck Protection Program." National Bureau of Economic Research Working Paper 27659. 5

Granja, João, Christian Leuz, and Raghuram Rajan. 2018. "Going The Extra Mile: Distant Lending And Credit Cycles." NBER Working Paper No.25196. 16

Granja, João, Christos Makridis, Constantine Yannelis, and Eric Zwick. 2020. "Did the Paycheck Protection Program Hit the Target?" 5, 6, 13, 20, 23, 25

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Iván Werning. 2020. "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?" National Bureau of Economic Research, Cambridge, MA. 2, 9, 10

Neilson, Christopher, John Eric Humphries, and Gabriel Ulyssea. 2020. "Information Frictions and Access to the Paycheck Protection Program." National Bureau of Economic Research Working Paper 27624. 5

Appendix

A Proofs and Derivations

Auxiliary Result. For the distribution in (3), we have that $\mathbb{E}[\nu \mid \nu \leq X] = \frac{\eta}{\eta+1}X$

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

A.1 Firm's Choice in the PPP

From the problem of the firm in (2), we can take the FOC w.r.t. ω when $a = 1$ to obtain

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

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

A.2 Lemma 1

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

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

where we use that $\frac{\partial \Gamma_i}{\partial \omega_i^G} = 1$ of the last equation. Let $\tilde{\lambda} \equiv \lambda \cdot c_0^\eta$. Using the equation for the distribution $\phi(\nu)$ in (3):

$$\eta_j \left[\pi_j + \omega_j^{G,*}\right]^{\eta-1} - \tilde{\lambda} = 0 \Rightarrow \omega_j^{G,*} = \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_{rs}-1}} - \pi_j = M - \left[\pi_j - \overline{\pi}\right] \Rightarrow M + \overline{\pi} = \int_j \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j-1}} drds \tag{22}$$

where the last equality comes from integrating $N_i \omega_j^{G,*}$ 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. $\blacksquare$

Note that the RHS is strictly *decreasing* in $\tilde{\lambda}$, goes to zero as $\lambda \to \infty$ and goes to $+\infty$ when $\lambda \to 0$, so there is always a unique solution for $\tilde{\lambda}$ from (22). We can manipulate the individual firm j equation to obtain:

$$\omega_j^{G,*} = \overline{M} - \left[\pi_j - \overline{\pi}\right] + \left(\frac{\tilde{\lambda}}{\eta_j}\right)^{\frac{1}{\eta_j-1}} - \overline{M} \tag{23}$$

Thus, we have that $\tau(\eta_{rs}, \overline{M}) \equiv \left(\frac{\tilde{\lambda}}{\eta_{rs}}\right)^{\frac{1}{\eta_{rs}-1}} - \overline{M}$. 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 $\eta < 1$ and $f(\eta) < 0$ for $\eta > 1$.

Case 1. If $\tilde{\lambda}$ small enough, we have that $\frac{\partial \omega_{i,j}^{G,*}}{\partial \eta_j} > 0$. Too see this, note that $f(\eta_j) < 0$ and $\ln(\tilde{\lambda}) < 0$, thus in (24) the RHS is positive. For $\tilde{\lambda}$ small, we must have a large $M + \overline{\pi}$ from (22).

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 $\lim_{\eta \to 0^+} f(\eta) = -\infty$. Therefore, any positive $\ln(\tilde{\lambda})$ is simply a shifter downward of this function. We still have that $\lim_{\eta \to 0^+} f(\eta) + \ln(\tilde{\lambda}) = -\infty$, $f(\eta) + \ln(\tilde{\lambda}) > 0$ and $f(\eta) + \ln(\tilde{\lambda})$ always decreasing. 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 $\tilde{\lambda} > 1$, it must be the case where $\overline{M}$ large, in particular: $M + \overline{\pi} > \int_j \left(\eta_j\right)^{\frac{1}{1-\eta_j}} \, dj$. ■

A.3 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 in this case:

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

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

A.4 Super and Submodular Dsitributions.

Let $\nu_j \sim \Upsilon(\pi_j + \omega_j, \theta_j)$, where $\theta_j \in \Theta$, a complete lattice, parametrizes 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 in a FOSD. 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. ■

A.5 Lemma 2

Proof. Let $G(\{l_j^G\}_j)$ be the objective function of the problem of the government in (8). To make the problem consistent across firms of different sizes N_j, we will re-write the choice variable as $\tilde{l}_j^G \equiv l_j^G N_j$, that is, the marginal dollar allocated to firms of type j or, alternatively, the number of employees of firm j that are saved. We thus can re-write the problem in (9) as

$$\max_{\{l_j^G \in [0,N_j]\}} \int \tilde{l}_j^G T_j \, dj \text{ s.t. } \varphi \int \tilde{l}_j^G \, dj = M \tag{25}$$

The problem in (25) is linear in the choice variables, and thus we will compute the optimal allocation by focusing at which points the derivative of G (which is not a function of $\tilde{l}_i^G$ is the highest. Ignoring the constraint that $\tilde{l}_j^G \in [0, N_j]$, the derivative of the objective function of $G(.)$ with respect to $N_j l_j^G$, that is, the marginal allocation[14]

$$G_j \equiv \frac{\partial G}{\partial \tilde{l}_j^G} = \left[\Phi_j^\Gamma(\varphi) - \Phi_j^\Gamma(0)\right] - \varphi\lambda$$

where λ is the Lagrange multiplier in the resource constraint. Therefore, the government allocates the money to firms j with the highest T_j, up to the point of $\tilde{l}_j^G = N_j \Leftrightarrow l_j = 1$.

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

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

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

Case 2. Regional and Sectoral heterogeneity. Consider that all firms are the same except for a region-sector specific shock exposure η_{rs}. Then:

$$\frac{\partial T_{rs}}{\partial \eta_{rs}} = (\pi + \varphi)^{\eta_{rs}} \cdot \ln(\pi + \varphi) - \pi^{\eta_{rs}} \ln(\pi) > 0 \Leftrightarrow (\pi + \varphi)^{\eta^{rs}} \cdot \ln(\pi + \varphi) > \pi^{\eta_{rs}} \ln(\pi)$$

Which implies:

$$\eta_{rs} \ln\left(1 + \frac{\varphi}{\pi}\right) + \ln(-\ln(\pi + \varphi)) < \ln(-\ln(\pi)) \Leftrightarrow \eta_{rs} < \eta_G^* \equiv \frac{\ln\left(\frac{\ln(\pi)}{\ln(\pi+\varphi)}\right)}{\ln(1 + \frac{\varphi}{\pi})} > 0$$

Therefore, T_{rs} is strictly increasing up to $\eta_G^* > 0$ and strictly decreasing afterwards. The optimal allocation is thus $l_{rs}^G = 1$ if $\eta_{rs} \in [\underline{\eta}_G, \overline{\eta}^G]$, 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_{rs} is quasi-concave. ■

A.6 Lemma 3

Proof. Let $B(\{\tilde{l}_j^B\}_j)$ be the objective function of the problem of the bank in (11), but, as in the case of the government in Lemma 2, in terms of $\tilde{l}_j^B \equiv l_j^B N_j$. As in the problem of the government, the problem of the bank is linear in the choice variables, and thus we will proceed as in the case of the government. Taking the derivative w.r.t. $N_j l_j^B$:

$$B_j \equiv \frac{\partial B}{\partial \tilde{l}_j^B} = T_j\left[(1-\delta)b_j + \psi_F\right] + \psi_C \psi_F \Phi_j^\Gamma(0) - \varphi\lambda$$

[14] For simplicity, we normalize ρ_j, b_j, φ by the scaling variable c_0.

where λ is the Lagrange multiplier in the resource constraint. Therefore, the bank allocates the money to firms j with a combination of the highest T_j weighted by their pre-existing levels of debt, b_j, and an extra term that captures a preference for firms with a high *ex-ante* probability of survival.

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

$$\frac{\partial B_j}{\partial b_j} = -(1-\delta)\eta\left(\left[\rho - b_j + \varphi\right]^{\eta-1} - \left[\rho - b_j\right]^{\eta-1}\right)b_j + (1-\delta)\left(\left[\rho - b_j + \varphi\right]^{\eta} - \left[\rho - b_j\right]^{\eta}\right) - \psi_C\psi_F[\rho - b_j]^{\eta-1}$$

For $\eta < 1$: $\frac{\partial B_j}{\partial b_j} > 0$, thus it is optimal to allocate funds to firms with the highest levels of b_j. For $\eta > 1$, we have that:

$$\frac{\partial B_j}{\partial b_j} < 0 \Leftrightarrow b_j > \frac{g(\eta)}{\eta g(\eta-1)\kappa} \tag{26}$$

where $g(\eta) \equiv \left[\rho - b_j + \varphi\right]^{\eta} - \left[\rho - b_j\right]^{\eta}$. We have that the RHS of (26) starts at zero and is increasing with a rate of one. The LHS starts at a positive number and increases at a rate smaller than one, which implies that $\frac{\partial B_j}{\partial b_j} > 0$ at zero and decreasing after some $\tilde{b}$, since it must be the case that when the LHS and RHS meet it is with the RHS crossing from above. Therefore, the optimal allocation is an interval around $\tilde{b}$ such that $\int_{\underline{b}}^{\overline{b}} dj = \frac{M}{\psi}$ and $B_{\underline{b}} = B_{\overline{b}}$ with $\underline{b} > 0$ if the resource constraint is binding. Note that at $b_j = \rho$, the LHS is ρ, and the RHS $=\eta^{-1}\varphi$. Therefore, if $\eta < \frac{\varphi}{\rho}$, $\overline{b} = \max_j b_j$.

Case 2. Shock Exposure Heterogeneity. Consider that all firms are the same except η_j. Then:

$$\frac{\partial B_j}{\partial \eta_j} = \left[\kappa(\pi+\varphi)^{\eta_j} \cdot \ln(\pi+\varphi) - (\kappa - \psi_C\psi_F)\pi^{\eta_j}\ln(\pi)\right]$$

Therefore

$$\frac{\partial B_j}{\partial \eta_j} > 0 \Leftrightarrow (\pi+\varphi)^{\eta_j} \cdot \ln(\pi+\varphi) > \left[1 - \frac{\psi_C\psi_F}{\kappa}\right]\pi^{\eta_j}\ln(\pi)$$

Which implies:

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

Therefore, B_j is strictly increasing up to $\eta_G^* > 0$ and strictly decreasing afterwards. The optimal allocation is thus $l_j^G = 1$ 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: $\overline{\eta}_B \geq \overline{\eta}_G$ and $\underline{\eta}_B \geq \underline{\eta}_G$. 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 η smaller, but sufficiently close to $\underline{\eta}_G$. 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 $\varepsilon < \psi_F \psi_C \left[\Phi_{\underline{\eta}_G}(0) - \Phi_{\overline{\eta}_G}(0)\right]$. Then, there $\exists\ \eta = \underline{\eta}_G - \zeta$, with $\zeta > 0$ such that:

$$\kappa T_{\underline{\eta}_G} + \psi_F \psi_C \Phi_{\overline{\eta}_G}(0) = \kappa T_{\overline{\eta}_G} + \psi_C \psi_F \Phi_{\overline{\eta}_G}(0) < \kappa T_{\eta} + \psi_F \psi_C \Phi_{\underline{\eta}_G}(0) < \kappa T_{\eta} + \psi_F \psi_C \Phi_{\eta}(0)$$

where $\kappa \equiv (1-\delta)b$. Therefore, $\overline{\eta}_B \geq \overline{\eta}_G$ cannot be optimal for the bank. ■

A.7 Misallocation and Program Design

We can write the misallocation as the difference between the social welfare given the government and bank allocations in the PPP.

Step 1: Debt Heterogeneity, changes in φ For the case where firms are only heterogeneous in terms of their debt-levels (as in Figure 2), we have that:

$$\Delta W = \int_0^{\min\{\underline{b}, b^*\}} T_j \, dj - \int_{\max\{\underline{b}, b^*\}}^{\overline{b}} T_j \, dj \tag{27}$$

Taking the derivative:

$$\frac{\partial \Delta W}{\partial \varphi} = T_{b^*} \frac{\partial b^*}{\partial \varphi} + T_{\underline{b}} \frac{\partial \underline{b}}{\partial \varphi} - T_{\overline{b}} \frac{\partial \underline{b}}{\partial \varphi} + \int_0^{\min\{\underline{b}, b^*\}} \partial_\varphi T_j \, dj - \int_{\max\{\underline{b}, b^*\}}^{\overline{b}} \partial_\varphi T_j \, dj$$

For the case where $\eta > 1$ (when there is the misallocation, as shown in Lemma 3), and noting that from the resource constraint $\frac{\partial b^*}{\partial \varphi} = \frac{\partial \overline{b}}{\partial \varphi} - \frac{\partial \underline{b}}{\partial \varphi}$, we have that:

$$b^* > \underline{b} \Rightarrow \frac{\partial \Delta W}{\partial \varphi} > \int_0^{\min\{\underline{b}, b^*\}} \partial_\varphi T_j \, dj - \int_{\max\{\underline{b}, b^*\}}^{\overline{b}} \partial_\varphi T_j \, dj > 0$$

since

$$[T_{b^*} - T_{\overline{b}}] \frac{\partial b^*}{\partial \varphi} + [T_{\underline{b}} - T_{\overline{b}}] \frac{\partial \underline{b}}{\partial \varphi} > 0$$

and $\int_0^{\min\{\underline{b}, b^*\}} dj = \int_{\max\{\underline{b}, b^*\}}^{\overline{b}} dj$, and the fact that $\partial_\varphi T_j = \eta\,[\rho - b_i + \varphi]^{\eta-1} - \eta\,[\rho - b_i]^{\eta-1}$ is *decreasing* in b_j for $\eta > 1$.

Step 2: Debt Heterogeneity, changes in M. Now taking the derivative w.r.t. M in (27)

$$\frac{\partial \Delta W}{\partial \varphi} = T_{b^*} \frac{\partial b^*}{\partial M} + T_{\underline{b}} \frac{\partial \underline{b}}{\partial M} - T_{\overline{b}} \frac{\partial \overline{b}}{\partial M}$$

For the case where $\eta > 1$ (when there is the misallocation, as shown in Lemma 3), and noting that from the resource constraint $\frac{\partial b^*}{\partial M} = \frac{\partial \overline{b}}{\partial M} - \frac{\partial \underline{b}}{\partial M}$, we have that:

$$b^* > \underline{b} \Rightarrow \frac{\partial \Delta W}{\partial M} = [T_{b^*} - T_{\overline{b}}] \frac{\partial b^*}{\partial M} + [T_{\underline{b}} - T_{\overline{b}}] \frac{\partial \underline{b}}{\partial M} < 0$$

Similarly

$$b^* < \underline{b} \Rightarrow \frac{\partial \Delta W}{\partial M} > T_{b^*}\frac{\partial b^*}{\partial M} + T_{\underline{b}}\frac{\partial \underline{b}}{\partial M} - T_{\underline{b}}\frac{\partial \underline{b}}{\partial M} = [T_{b^*} - T_{\underline{b}}]\frac{\partial b^*}{\partial M} > 0$$

Step 3: Shock Exposure Heterogeneity. For the case where firms are only heterogeneous in terms of their debt-levels (as in Figure 2), we have that:

$$\Delta W = \int_{\max\{\overline{\eta}_B, \underline{\eta}_G\}}^{\overline{\eta}_G} T_j \, dj - \int_{\underline{\eta}_B}^{\min\{\overline{\eta}_B, \underline{\eta}_G\}} T_j \, dj \tag{28}$$

Taking the derivative w.r.t φ:

$$\begin{aligned}\frac{\partial \Delta W}{\partial \varphi} = T_{\underline{\eta}_G}\frac{\partial \underline{\eta}_G}{\partial \varphi} - T_{\overline{\eta}_G}\frac{\partial \overline{\eta}_G}{\partial \varphi} - \left[T_{\underline{\eta}_B}\frac{\partial \underline{\eta}_B}{\partial \varphi} - T_{\overline{\eta}_B}\frac{\partial \overline{\eta}_B}{\partial \varphi}\right] \\ + \int_{\underline{\eta}_B}^{\underline{\eta}_G} \partial_\varphi T_j \, dj - \int_{\overline{\eta}_B}^{\overline{\eta}_G} \partial_\varphi T_j \, dj\end{aligned}$$

From the resource constraint $\frac{\partial \overline{\eta}_G}{\partial \varphi} - \frac{\partial \underline{\eta}_G}{\partial \varphi} = \frac{\partial \overline{\eta}_B}{\partial \varphi} - \frac{\partial \underline{\eta}_B}{\partial \varphi}$.

$$\frac{\partial \Delta W}{\partial \varphi} > \underbrace{\left[\frac{\partial \overline{\eta}_G}{\partial \varphi} - \frac{\partial \underline{\eta}_G}{\partial \varphi}\right] \cdot \left[T_{\underline{\eta}_B} - T_{\underline{\eta}_G}\right]}_{\equiv I} + \underbrace{\int_{\underline{\eta}_B}^{\underline{\eta}_G} \partial_\varphi T_j \, dj - \int_{\overline{\eta}_B}^{\overline{\eta}_G} \partial_\varphi T_j \, dj}_{\equiv II}$$

We can show that both I, II can be positive or negative for $\overline{\eta}_B > \underline{\eta}_G$. For I, see the proof in Step 4. For II, note that:

$$\partial_\varphi T_j = \eta \left[(\rho - b_i + \varphi)\right]^{\eta - 1}$$

is a non-monotone function of η.

Step 4: Shock Exposure Heterogeneity with changes in M. Taking the derivative w.r.t to M in (28)

$$\frac{\partial \Delta W}{\partial M} = T_{\overline{\eta}_G}\frac{\partial \overline{\eta}_G}{\partial M} - T_{\underline{\eta}_G}\frac{\partial \underline{\eta}_G}{\partial M} - T_{\overline{\eta}_B}\frac{\partial \overline{\eta}_B}{\partial M} + T_{\underline{\eta}_B}\frac{\partial \underline{\eta}_B}{\partial M}$$

From the resource constraint $\frac{\partial \overline{\eta}_G}{\partial \varphi} - \frac{\partial \underline{\eta}_G}{\partial \varphi} = \frac{\partial \overline{\eta}_B}{\partial \varphi} - \frac{\partial \underline{\eta}_B}{\partial \varphi}$. Moreover, we have that $T_{\underline{\eta}_G} = T_{\overline{\eta}_G}$ and $T_{\underline{\eta}_B} < T_{\overline{\eta}_B}$, since $\overline{\eta}_B \in [\underline{\eta}_G, \overline{\eta}_G]$ and $\underline{\eta}_B \notin [\underline{\eta}_G, \overline{\eta}_G]$. Therefore:

$$\overline{\eta}_B < \underline{\eta}_G \Rightarrow \frac{\partial \Delta W}{\partial M} = \left[\frac{\partial \overline{\eta}_G}{\partial M} - \frac{\partial \underline{\eta}_G}{\partial M}\right] \cdot \left[T_{\underline{\eta}_G} - T_{\overline{\eta}_B}\right] + \frac{\partial \underline{\eta}_B}{\partial M} \cdot \left[T_{\underline{\eta}_B} - T_{\overline{\eta}_B}\right] > 0$$

which may be positive or negative for $\overline{\eta}_B > \underline{\eta}_G$. ■

A.8 OLS estimation and Treatment Effects

First, note that the type distribution among the firms that do not receive the PPP is given by

$$\alpha_i^P = \mathbb{P}\left[i \mid l_i^P = 1\right] \Rightarrow \mathbb{P}\left[i \mid l_i^P = 0\right] = 2\alpha_i - \alpha_i^P, \;\; P \in \{B, G\}$$

Step 1: ATE. In the regression in (16), the OLS estimator for $\widehat{\beta^F}$ is such that

$$\mathbb{E}\left[\widehat{\beta^F}\right] = \int_i \alpha_i^B \Phi_i^\varphi \, di - \int_i (\alpha_i + \alpha_i - \alpha_i^B)\Phi_i^0 \, di = ATE + \int_i \left(\alpha_i^B - \alpha_i\right)\left[\Phi_i^\varphi + \Phi_i^0\right] di$$

In the regression in (18), the adjusted OLS estimator for $\widehat{\beta^R}$ is such that

$$\mathbb{E}\left[\widehat{\beta^R}\right] = 2\left[\frac{1}{2}\int_i \alpha_i^B \Phi_i^\varphi \, di + \frac{1}{2}\int_i (\alpha_i + \alpha_i - \alpha_i^B)\Phi_i^0 \, di - \int_i \alpha_i \Phi_i^\varphi \, di\right] = ATE + \int_i \left(\alpha_i^B - \alpha_i\right) T_i \, di$$

Step 2: ATT. The OLS estimator $\widehat{\beta^F}$ is such that

$$\mathbb{E}\left[\widehat{\beta^F}\right] = \int_i \alpha_i^B \Phi_i^\varphi \, di - \int_i (\alpha_i + \alpha_i - \alpha_i^B)\Phi_i^0 \, di = ATT + 2\int_i \left(\alpha_i^B - \alpha_i\right)\Phi_i^0 \, di$$

Similarly, the adjusted OLS estimator for $\widehat{\beta^R}$ is such that

$$\mathbb{E}\left[\widehat{\beta^R}\right] = 2\left[\frac{1}{2}\int_i \alpha_i^B \Phi_i^\varphi \, di + \frac{1}{2}\int_i (\alpha_i + \alpha_i - \alpha_i^B)\Phi_i^0 \, di - \int_i \alpha_i \Phi_i^\varphi \, di\right] = ATT$$

Step 3: ATG. In comparison to the average treatment effect of the government allocation:

$$\mathbb{E}\left[\widehat{\beta^F}\right] = \beta_G + \int_i \left(\alpha_i^B - \alpha_i^G\right)\left[\Phi_i^\varphi + \Phi_i^0\right] di$$
$$\mathbb{E}\left[\widehat{\beta^R}\right] = \beta_G + \int_i \left(\alpha_i^B - \alpha_i^G\right) T_i \, di$$

Step 4: Covariance. Note in all cases that for any A_i function of i:

$$\int_i \left(\alpha_i^B - \alpha_i\right) A_i \, di = \mathrm{Cov}\left\{\left(\alpha_i^B - \alpha_i\right), A_i\right\} + \int_i \left(\alpha_i^B - \alpha_i\right) di \cdot \int_i A_i \, di = \mathrm{Cov}\left\{\left(\alpha_i^B - \alpha_i\right), A_i\right\}$$

■

A.9 Lemma A.9

Proof. Whenever banks distort the optimal allocation toward firms with a lower η (Lemma 3) and $\min_j \eta_j \in [\underline{\eta}_B, \underline{\eta}_G]$ and $\max_j \eta_j \in [\overline{\eta}_B, \overline{\eta}_G]$:

$$\int_i \left(\alpha_i^B - \alpha_i^G\right)\left[\Phi_i^\varphi + \Phi_i^0\right] = \int_i \left(\alpha_i^B - \alpha_i^G\right)^+ \left[\Phi_i^\varphi + \Phi_i^0\right] - \int_i \left(\alpha_i^G - \alpha_i^B\right)^- \left[\Phi_i^\varphi + \Phi_i^0\right]$$
$$> \left[\Phi_{\underline{\eta}_G}^\varphi + \Phi_{\underline{\eta}_G}^0\right] \int_i \left(\alpha_i^B - \alpha_i^G\right)^+ - \left[\Phi_{\overline{\eta}_B}^\varphi + \Phi_{\overline{\eta}_B}^0\right] \int_i \left(\alpha_i^G - \alpha_i^B\right)^- > 0$$

and

$$\int_i \left(\alpha_i^B - \alpha_i^G\right) T_i = \int_i \left(\alpha_i^B - \alpha_i^G\right)^+ T_i - \int_i \left(\alpha_i^G - \alpha_i^B\right)^- T_i < T_{\underline{\eta}_G} \int_i \left(\alpha_i^B - \alpha_i^G\right)^+ - T_{\overline{\eta}_G} \int_i \left(\alpha_i^G - \alpha_i^B\right)^- = 0$$

The same argument applies for α_i's instead of α_G's and for the ATTs and ATGs. ■