The Cost of Capital and Misallocation in the United States*

Miguel Faria-e-Castro — FRB of St. Louis
Julian Kozlowski — FRB of St. Louis
Jeremy Majerovitz — University of Notre Dame

June 11, 2025

Abstract

We develop a methodology to estimate the cost of capital using credit registry microdata, and apply it to study capital allocation efficiency in the United States. Our measure incorporates the contractual interest rate, expected default probability, recovery rate, and expectations of future rates. We estimate three distinct rates: (i) the lender's discount rate, (ii) the firm's cost of capital, and (iii) the social cost of capital. We derive a sufficient statistic for misallocation based on the first and second moments of the social cost of capital. Dispersion in this rate captures both heterogeneity in lender discounting and the presence of financial frictions. Normal times feature modest amounts of misallocation, corresponding to an output loss of 0.5%. However, during the 2020–2021 period, misallocation increased to 1.1%, primarily due to the underpricing of riskier loans.

JEL Classification: D24, E22, E44, O47, O51.

Keywords: cost of capital, misallocation, macrofinance.

*We thank Vladimir Yankov for an excellent discussion and Juan Sanchez, François Gourio, seminar participants at Notre Dame, the St. Louis Fed, the Midwest Macro, and the Federal Reserve Macrofinance Workshop for thoughtful questions and feedback. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis, the Board of Governors of the Federal Reserve, or the Federal Reserve System.

1 Introduction

How much does it cost a firm to obtain capital? Economic models often simplify by assuming that all firms can borrow in a competitive market at a common rate. In reality, however, the cost of capital varies significantly across firms. This variation stems not only from differences in contractual interest rates but also from firm-specific factors such as default probabilities, loan terms, and lender cost of funds. Such heterogeneity in the cost of capital has profound implications: it can distort the allocation of capital across firms, leading to inefficiencies in economic output (Gilchrist et al., 2013; Hsieh and Klenow, 2009). Understanding these inefficiencies is critical for policymakers and researchers seeking to design effective financial and economic policies.

This paper makes two key contributions to the literature. First, we develop a novel methodology that leverages a corporate finance model and uses credit registry microdata to measure the dispersion of the cost of capital. This methodology allows us to quantify how these variations contribute to capital misallocation. The theory implies a sufficient statistic for misallocation that can be directly measured with credit registry data. Second, we apply this methodology to U.S. data and uncover two primary insights. While the cost of capital is heterogeneous across firms, the implied misallocation is surprisingly small, suggesting that U.S. capital markets are close to allocative efficiency. However, this efficiency deteriorated in the aftermath of the COVID-19 pandemic, driven by the underpricing of risky loans.

The methodology we develop offers several advantages. Unlike traditional approaches that require solving structural models computationally, our approach uses sufficient statistics derived directly from moments of the data. This not only simplifies implementation but also provides more robust identification of the sources of misallocation without heavy reliance on calibration assumptions.

Section 2 develops the dynamic corporate finance model that provides the foundation for the derivation of the sufficient statistic for misallocation and our empirical measurement of the cost of capital. The model captures firm-level borrowing, investment, and default decisions in the presence of idiosyncratic shocks, such as productivity fluctuations and fixed operating costs. Each firm borrows from a single lender who discounts future cash flows at a match-specific rate ρ. While we refer to ρ as the lender's discount rate, it more broadly reflects match-specific efficiency—capturing variation in loan pricing that is not fully explained by observable loan terms. Using the firm's optimality conditions, we show how ρ influences the firm's internal cost of capital, which in turn determines its expected marginal revenue product of capital. This link forms the basis for our measure of misallocation.

In Section 3, we quantify the efficiency costs arising from heterogeneity in the cost of capital. We compare the decentralized equilibrium to the allocation chosen by a planner who reallocates capital across firms to maximize aggregate output, subject to the same aggregate capital stock and taking firms' default decisions as given. We define the social cost of capital, which is the marginal value of allocating an extra unit of capital for each firm from the point of view of the planner. We show that this measure is approximately equal to the sum of two terms: the lender's discount rate plus a term that reflects financial frictions related to limited liability and recovery in case of default in the spirit of Cooley and Quadrini (2001). At the optimum, the planner would like to equate the social cost of capital across firms. This insight allows us to derive a sufficient statistic for the output loss from misallocation that depends only on the mean and variance of the social cost of capital. This statistic is robust to firm-level heterogeneity in production technologies and does not rely on structural estimation.

Section 4 describes how we map the model to U.S. credit registry data. We define the lender's discount rate as the internal rate of return that satisfies the lender's break-even condition, accounting for both repayment probabilities and expected losses in default. To compute this rate at the firm level, we require loan-level data on contractual interest rates, loan maturities, borrower-specific probabilities of default, and loss given default (LGD). In the case of floating-rate loans, we also need forward-looking benchmark interest rate expectations. Using these variables as well as the equations of the model, we estimate three distinct rates: the lender's discount rate, the firm's cost of capital, and the social cost of capital. We then apply the sufficient statistic to the social cost of capital to estimate the output loss from misallocation.

Finally, in Section 5, we present our empirical findings. Using data from over sixty thousand loans originated between 2014 and 2024, we show that the average measures of cost of capital closely track the five-year U.S. Treasury rate. We estimate three distinct rates. First, the *lender's discount rate* is specific to each borrower-lender pair and captures the efficiency of the credit match. It has a mean of 3.8% and a variance of 1.8%. Second, the *firm's cost of capital*—defined as the expected payment by the firm conditional on no default—has a mean of 2.9% and a variance of 6.5%. This rate is lower than the lender's discount rate because it does not account for repayment in case of default. Finally, the *social cost of capital* reflects the total return on capital from a social perspective, incorporating expected recovery in the event of default. It has a mean of 3.6% and a variance of 2.5%. For comparison, the average five-year U.S. Treasury rate over this period was 2.2%.

At the optimum, the planner seeks to equalize the social cost of capital across firms. Our sufficient statistic provides a mapping from the variance of the social cost of capital—2.5% on average—to output losses due to misallocation. We estimate that, under normal conditions, the implied output loss from capital misallocation is modest, around 0.5%. However, this loss increased significantly during the COVID-19 pandemic (2020–2021), rising to 1.9% at its peak.

The increase in misallocation during 2020-2021 is predominantly explained by rising heterogeneity in lender discount rates, rather than by a worsening of financial frictions related to default. Specifically, we find a marked rise in the coefficient of variation of the lender discount rates, reflecting greater dispersion in borrowing conditions across firms. This increase in dispersion is not driven by higher expected credit losses—as default probabilities and losses given default remain stable—but rather by changes in the distribution of contractual interest rates. We trace these changes to the underpricing of very risky loans, which expanded access to cheap credit for low-quality borrowers.

Our hypothesis is that the broad fiscal and monetary interventions enacted during the COVID-19 crisis—such as the Paycheck Protection Program (PPP), the Main Street Lending Program (MSLP), and the corporate credit facilities (PMCCF and SMCCF)—played a key role in distorting credit markets. These programs, while aimed at stabilizing the economy, effectively supported distressed firms. This may have led lenders to perceive implicit guarantees for riskier loans, generating a moral hazard problem that incentivized them to extend credit to lower quality borrowers under the expectation of bailouts in the event of default. This may have also contributed to zombie lending. The central implication is that these guarantees may have misaligned credit pricing, resulting in capital misallocation. In the absence of such implicit guarantees, lenders would have priced risk more accurately, leading to a more efficient allocation of credit across firms.

Literature Review. Our paper contributes to the broader literature on measuring misallocation. Following seminal work by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), there has been significant progress in quantifying misallocation across various settings (see Hopenhayn (2014) and Restuccia and Rogerson (2017) for comprehensive reviews). A key challenge in this literature is measuring misallocation without imposing strong assumptions on firms' production technologies. Haltiwanger et al. (2018) emphasize that standard approaches are only valid under restrictive assumptions, such as a common Cobb-Douglas production function with firm-specific productivity shifters.

One strand of the literature focuses on specific sources of distortions. For instance, Kaymak and Schott (2024) study corporate tax asymmetries and find that heterogeneity in effective marginal tax rates can distort capital and labor allocation, reducing aggregate productivity. Alternatively, recent work has sought to directly estimate marginal products using (quasi-)experimental variation, allowing for richer production heterogeneity (e.g., (Carrillo et al., 2023; Hughes and Majerovitz, 2025). However, such approaches have only been applied in narrow contexts where experimental variation is available.

Our paper measures heterogeneity in the marginal product of capital by exploiting firm-level variation in the cost of capital, allowing us to assess misallocation across a much broader set of firms while remaining agnostic to functional form assumptions. Closest to our approach, Gilchrist et al. (2013) develop a tractable framework to quantify misallocation arising from dispersion in borrowing costs. We emphasize two key differences. First, a portion of the credit spread dispersion in Gilchrist et al. (2013) reflects variation in default probabilities and recovery rates, whereas our framework explicitly models corporate default. Second, their analysis relies on corporate bond data, which restricts attention to large firms. In contrast, we use bank loan data that encompasses a significantly wider range of firms, including small, medium, and large-sized enterprises.

We also contribute to a literature that estimates heterogeneity across firms in interest rates and/or the cost of capital. Banerjee and Duflo (2005) summarize early evidence for substantial heterogeneity in interest rates across borrowers in developing countries, arguing that this heterogeneity implies significant misallocation. Recent work by Gormsen and Huber (2023, 2024) analyzes transcripts of firm earnings calls to extract information on the discount rates and cost of capital that firms use. Cavalcanti et al. (2021) use credit registry data to study heterogeneity in interest rates for borrowing firms in Brazil. They find substantial heterogeneity across firms and use a dynamic structural model with financial frictions to infer the cost of capital. This paper also builds on the findings of Faria-e-Castro et al. (2024), who analyze the dispersion in borrowing rates for U.S. firms using a comprehensive database of loans and bonds. Their study highlights significant heterogeneity in borrowing costs, even within firms, and demonstrates the persistent impact of borrowing costs on firm-level investment and borrowing behaviors.

Relative to this previous literature, our paper makes two key methodological contributions. First, we provide a methodology to estimate a firm's cost of capital from credit registry data. This is not as simple as measuring the interest rate because the cost of capital depends on the ex-ante repayment probability and expected losses given default. Second, we show how to use moments of the distribution of the cost of capital to develop sufficient statistics that allow us to measure the cost of misallocation non-parametrically in a dynamic, stochastic model.

2 Corporate Finance Model

This section outlines the core components of the model, which captures the interactions between borrowers and lenders. We demonstrate how the model's optimality conditions can be derived and integrated with microdata on loan characteristics to estimate the lender's discount rate and the firm's cost of capital. These rates provide insights into the firm's expected marginal product of capital, a critical metric for assessing misallocation.

Time is discrete and indexed by $t = 0, 1, \ldots$. The economy is populated by firms that borrow and invest, and by lenders who finance those firms. There is a unit mass of firms, indexed by i, who exit over time. We assume that every firm that exits is replaced by a firm with identical characteristics that does not produce in the current period, such that the mass of firms is constant and equal to 1. We now describe the decision problem of the firm, and its interaction with the lenders.

Borrowers. The borrowers in the model are firms operating in the nonfinancial sector. These firms operate under limited liability and make decisions regarding production, investment, and borrowing. Output (net of non-capital costs) is generated using a production function $f(k_i, z_i)$, where k_i represents capital and z_i denotes a vector of shocks that affect firm net output. Note that since z_i can be a vector, this accommodates productivity shocks, stochastic fixed costs as well as rich heterogeneity in the production function. To sustain or expand their operations, firms invest in capital and issue long-term defaultable debt b_i. In the event of default, lenders recover a fraction $\phi_i(k_i)$ of the firm's existing assets k_i.

Lenders. Lenders finance firms, with each firm matched to a single lender. Upon matching, the borrower-lender pair draws a realization of ρ_i, which represents the efficiency of the match.[1] We refer to ρ_i as the lender's discount rate. Loans are priced so that lenders break even using ρ_i as their discount rate, taking into account firm-specific characteristics and risk assessments.

Firm's Problem. Firms determine their investment and borrowing strategies to maximize their value, taking into account the possibility of future default. The value of repayment for a firm is expressed as:

$$V_i(k_i, b_i, z_i) = \max_{k_i', b_i'} \pi(k_i, b_i, z_i, k_i', b_i') + \beta \mathbb{E} \left[\max \left\{ V_i(k_i', b_i', z_i'), 0 \right\} | z_i \right],$$

where $\pi(k_i, b_i, z_i, k_i', b_i')$ denotes the firm's profit function, and β represents the discount factor.

The profit function captures the firm's net return from production and financing decisions:

$$\pi(k_i, b_i, z_i, k_i', b_i') = f(k_i, z_i) + (1-\delta)k_i - k_i' - \theta_i b_i + Q_i(k_i', b_i', z_i)(b_i' - (1-\theta_i)b_i).$$

Here, $f(k_i, z_i)$ represents the firm's output as a function of capital k_i and productivity z_i, $(1-\delta)k_i$ accounts for the depreciated value of current capital, and k_i' denotes new capital investment. The term $\theta_i b_i$ reflects repayment on existing debt, while $Q_i(k_i', b_i', z_i)$ captures the price of new debt, with $b_i' - (1-\theta_i)b_i$ representing the net new borrowing.

Debt Pricing. Lenders are risk-neutral and price debt based on their cost of capital, ρ_i. The price of debt $Q_i(k_i', b_i', z_i)$ is determined as:

$$Q_i(k_i', b_i', z_i) = \frac{\mathbb{E} \left[\mathcal{P}_i(k_i', b_i', z_i') \left(\theta_i + (1-\theta_i) Q_i(k_i'', b_i'', z_i') \right) + \left(1 - \mathcal{P}_i(k_i', b_i', z_i')\right) \frac{\phi_i(k_i')}{b_i'} \middle| k_i', b_i', z_i \right]}{1 + \rho_i}, \tag{1}$$

where $\mathcal{P}_i(k_i', b_i', z_i')$ is an indicator function that is equal to 1 if the firm repays, and 0 otherwise, and $\phi_i(k_i')/b_i'$ is the recovery rate in the event of default, per dollar lent.

[1]This variable captures both lender- and borrower-specific factors that lie outside the scope of the model, such as lender financing costs, risk appetite, or the dynamics of relationship lending. While we do not provide a specific microfoundation for the heterogeneity in ρ_i, we focus on analyzing its implications.

The Firm's Cost of Capital. We define the firm's cost of capital, r_i^{firm}, as the ratio of the expected value of future repayments adjusted for the probability of repayment, $\mathbb{E}\left[\mathcal{P}_i'(\theta_i + (1-\theta_i)Q_i')\right]$, relative to the current price of borrowing, Q_i.[2] The firm's cost of capital is the implicit interest rate that it pays on its debt. Formally, it is expressed as:

$$1 + r_i^{firm} = \frac{\mathbb{E}\left[\mathcal{P}_i'(\theta_i + (1-\theta_i)Q_i') \middle| \, k_i', b_i', z_i\right]}{Q_i}. \tag{2}$$

This equation captures how the firm's borrowing cost depends on repayment probabilities and debt maturity. The firm's cost of capital is one of the key components of the firm's first order condition with respect to capital. Intuitively, we show how to measure r_i^{firm} in the data, and this will give us information about the marginal revenue product of capital.

Proposition 1 characterizes the firm's cost of capital. All proofs are in Appendix A.

Proposition 1 (Firm's Cost of Capital)**.** *The firm's cost of capital can be written as:*

$$1 + r_i^{firm} = \frac{1+\rho_i}{1+\Lambda_i}, \qquad \Lambda_i := \frac{\mathbb{E}\left[(1-\mathcal{P}_i')\,\phi_i(k_i')/b_i' \middle| \, k_i', b_i', z_i\right]}{\mathbb{E}\left[\mathcal{P}_i'\,(\theta + (1-\theta_i)Q_i') \middle| \, k_i', b_i', z_i\right]}.$$

The term Λ_i represents the wedge between the borrower's cost of capital, r_i^{firm}, and the lender's discount rate, ρ_i. This wedge arises due to lender recovery in the event of default. When there is no recovery ($\phi_i = 0$), the wedge disappears ($\Lambda_i = 0$), and the firm's cost of capital equals the lender's discount rate ($r_i^{firm} = \rho_i$). On the other hand, when the lender can recover some value after default ($\phi_i > 0$), the wedge becomes positive ($\Lambda_i > 0$), and the firm's cost of capital r_i^{firm} is lower than ρ_i. This reduction in perceived borrowing cost occurs because the borrower only accounts for states where repayment occurs.

Marginal Revenue Product of Capital. The firm's investment decision follows a standard first-order condition, which equates the firm's cost of capital with its expected marginal revenue product of capital. Formally, this condition is expressed as:[3]

$$(1 + r_i^{firm})\mathcal{M}_i = \mathbb{E}[\mathcal{P}_i'(f_k(k_i', z_i') + 1 - \delta) | \, k_i', b_i', z_i]. \tag{3}$$

[2] We use $\mathcal{P}_i'$ as a shorthand for $\mathcal{P}_i(k_i', b_i', z_i')$ and similar for Q_i'.
[3] We provide a derivation of this equation in Appendix A.

The left-hand-side of equation (3) represents the cost of raising more capital. This includes the firm's cost of capital, r_i^{firm}, adjusted by the price feedback multiplier, $\mathcal{M}_i$, which captures the effect of the firm's borrowing and investment on the price of debt. The price feedback multiplier $\mathcal{M}_i$ is given by:

$$\mathcal{M}_i := \frac{1 - \gamma_i \times \frac{Q_i \cdot b_i'}{k_i'} \times \frac{\partial \log Q_i}{\partial \log k_i'}}{1 + \gamma_i \times \frac{\partial \log Q_i}{\partial \log b_i'}}, \qquad \gamma_i := \frac{b_i' - (1-\theta_i) b_i}{b_i'},$$

where γ_i measures the share of debt tomorrow that will be newly purchased. The numerator of $\mathcal{M}_i$ reflects the feedback from changes in capital on the price of debt, while the denominator incorporates the feedback from changes in borrowing. Together, these terms provide a comprehensive characterization of how price dynamics influence the firm's cost of capital.

The right-hand-side of equation (3) represents the expected marginal revenue product of capital. This term includes the marginal productivity of capital, $f_k(k_i', z_i')$, and the depreciation factor, $1 - \delta$, weighted by the states of the world in which the firm repays, $\mathcal{P}_i'$.

3 Measuring Misallocation

When financial markets are efficient, all firms face the same cost of capital. However, in the data we find that the cost of capital varies across firms. How does this inefficiency in financial markets translate into an inefficiency in the real economy? We now consider the aggregation of output and investment across firms in order to study the steady-state costs of misallocation arising from dispersion in the cost of capital.

3.1 The Aggregate Economy and Welfare

We begin by setting up the aggregate environment in order to study both the decentralized equilibrium and the planner's problem. The firm's problem will be the same as before. There is no aggregate risk, so aggregates are not stochastic. Firms make undifferentiated products and take the price of their output as given. There is some initial stock of capital K_0, and future capital depends on investment and depreciation through the standard law of motion.

We introduce the notation $\omega_{i,t}$, which is equal to one if firm i is still operating at time t, and zero if it has exited. Note that $\mathbb{E}_{t-1}[\omega_{i,t}] = \mathcal{P}_{i,t}$. Aggregate output is given by:

$$Y_t = \int_0^1 \underbrace{\omega_{i,t} \cdot f(k_{i,t}, z_{i,t})}_{\text{Output if Operates}} - \underbrace{(1-\omega_{i,t}) \cdot ((1-\delta) k_{i,t} - \phi_i(k_{i,t}))}_{\text{Losses if Defaults}} di \tag{4}$$

Note that we have defined output, Y_t, so that it includes both the firm's output in the event of production, $f(k_{i,t}, z_{i,t})$, and the losses from liquidation, $(1-\delta) k_{i,t} - \phi_i(k_{i,t})$, in the event of default. This allows us to define aggregate investment simply:

$$I_t = K_{t+1} - (1-\delta) K_t \tag{5}$$

Finally, aggregate capital is given by:

$$K_t = \int_0^1 k_{i,t} di \tag{6}$$

The planner wishes to maximize welfare, U, controlling each firm's capital and exit decision. However, the planner is subject to the same information constraints as the firm: $k_{i,t}$ must be decided in period $t-1$, without yet knowing the productivity or operating costs that will prevail in that period. Exit decisions are made after $z_{i,t}$ is revealed, but with the values for future periods still unknown.

There is a representative household that obtains utility from consumption: we abstract from inequality to focus on productive efficiency. The household's utility is additively separable over time. Consumption is equal to aggregate output minus investment. Thus, welfare in this economy is given by:

$$U = \sum_{t=0}^{\infty} \beta^t \cdot u(Y_t - I_t)$$

where β is the household's discount rate and u is the utility it gets from consumption.

3.2 The Planner's Problem

Let $S_i^t := \{z_{is}\}_{s=0}^t$ denote the entire history of states, through period t.[4] Define $S^t := \{S_i^t\}_{i\in[0,1]}$ as the collection of all firms' histories. We can use this notation to set up the appropriate constraints to the planner's problem: the planner must set $k_{i,t}$ as a function of S^{t-1}, and $\omega_{i,t}$ as a function of S^t.[5] The planner's problem is:

$$U^* = \max_{\left\{\{k_{i,t}(S^{t-1}),\omega_{i,t}(S^t)\}_{i\in[0,1]}\right\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \cdot u\left(Y_t - I_t\right)$$

$$s.t.$$

$$\omega_{i,t}\left(S^t\right) \in \{0,1\}\,\forall i$$

$$\omega_{i,t+1}\left(S^{t+1}\right) \geq \omega_{i,t}\left(S^t\right)\,\forall S^t \subset S^{t+1}, \forall i$$

and Equations 4, 5, and 6 hold

where the inequality $\omega_{i,t+1}\left(S^{t+1}\right) \geq \omega_{i,t}\left(S^t\right)$ notes that if the firm exits, it cannot subsequently re-enter. In period $t = 0$, all firms operate and capital is set exogenously.

We can rewrite the planner's problem as a nested maximization problem, to isolate the intensive-margin choice of capital, holding aggregate capital and the extensive margin fixed. Note that $I_t = K_{t+1} - (1-\delta)K_t$, and so it depends only on aggregate capital (not the allocation across firms). We can thus rewrite the planner's problem in the following nested form:

$$U^* = \max_{\left\{K_t,\{\omega_{i,t}(S^t)\}_{i\in[0,1]}\right\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \cdot u\left(\left(\max_{\left\{\{k_{i,t}(S^{t-1})\}_{i\in[0,1]}\right\}_{t=1}^\infty} Y_t\right) - I_t\right)$$

with the same constraints as before.

[4] Note that the only shock in our model is $z_{i,t}$, so this is the full history of states.

[5] In practice, since there is no aggregate risk, the planner will only need to use the individual firm's state histories to make decisions.

3.3 The Cost of Misallocation

We can now turn our attention to the inner problem. Note that the inner problem is separable across time periods, allowing us to separate it into a sequence of static problems. We focus on the cost of misallocation in terms of output. Simplifying our notation, we can rewrite the problem as follows:

$$Y_t^*\left(K_t, \{\omega_{i,t}\}_{i\in[0,1]}\right) = \max_{\{k_{i,t}\}_{i\in[0,1]}} \int_0^1 \mathbb{E}_{t-1}\left[\omega_{i,t} \cdot f\left(k_{i,t}; z_{i,t}\right) - \left(1-\omega_{i,t}\right) \cdot \left(\left(1-\delta\right) k_{i,t} - \phi_i\left(k_{i,t}\right)\right)\right] di$$

$$s.t.$$

$$K_t = \int_0^1 k_{i,t} di$$

This problem is now a special case of the environment in Hughes and Majerovitz (2025). We can use their main proposition to derive the cost of misallocation, up to a second-order approximation. Define

$$g_i\left(k_i\right) := \mathbb{E}_{t-1}\left[\omega_{i,t} \cdot f\left(k_{i,t}; z_{i,t}\right) - \left(1-\omega_{i,t}\right) \cdot \left(\left(1-\delta\right) k_{i,t} - \phi_i\left(k_{i,t}\right)\right)\right].$$

Proposition 2 shows the cost of intensive-margin misallocation.

Proposition 2 ((Special Case of Hughes and Majerovitz (2025))). *The cost of intensive-margin misallocation is given by*

$$\underbrace{\log Y_t^*\left(K_t, \left\{\omega_{i,t}\left(S^t\right)\right\}_{i\in[0,1]}\right) - \log Y_t}_{\textit{Cost of Intensive-Margin Misallocation}}$$
$$\approx \frac{1}{2} \cdot \underbrace{\mathbb{E}_{g_i(k_i)}\left[\mathcal{E}_i\right]}_{\textit{Sales-Weighted Elasticity}} \cdot \underbrace{Var_{g_i(k_i)\mathcal{E}_i}\left(\log\left(\frac{\partial}{\partial k_{i,t}} g_i\left(k_i\right)\right)\right)}_{\textit{Weighted Variance of Log Expected MPK}}$$

where $g_i(k_i)$ is the expected output of the firm as a function of k_i, $\mathcal{E}_i$ is the elasticity of expected output with respect to the cost of capital, $\mathbb{E}_{g_i(k_i)}[\cdot]$ denotes the weighted average, weighting by $g_i(k_i)$, $Var_{g_i(k_i)\mathcal{E}_i}(\cdot)$ denotes the weighted variance, weighting by $g_i(k_i)\mathcal{E}_i$. All moments are computed for the set of firms that are operating at time $t-1$. The formulas for the expected output of the firm and the elasticity of expected output with respect to the cost *of capital are given by:*

$$g_i(k_i) = \mathbb{E}_{t-1}\left[\omega_{i,t}\cdot f(k_{i,t};z_{i,t}) - (1-\omega_{i,t})\cdot((1-\delta)k_{i,t} - \phi_i(k_{i,t}))\right]$$

$$\mathcal{E}_i = -\frac{\left(\frac{\partial}{\partial k_i}g_i(k_i)\right)^2}{g_i(k_i)\cdot\frac{\partial^2}{(\partial k_i)^2}g_i(k_i)}$$

Note that in a Cobb-Douglas setting, with $f(k,z) = z\cdot k^\alpha$ and no default, the elasticity simplifies to $\mathcal{E} = \frac{\alpha}{1-\alpha}$. In our quantitative analysis, we will calibrate $\mathcal{E} = \frac{1}{2}$, consistent with $\alpha = \frac{1}{3}$. Moreover, note that although the proposition above provides a second-order approximation, it becomes exact in a setting where production is Cobb-Douglas and where productivity and distortions are jointly log-normal (the weights also fall out in that special case).

3.4 The Social Cost of Capital

We have already introduced the notion of the lender's discount rate, ρ, and the firm's cost of capital, r^{firm}. We now introduce the notion of the social cost of capital, r^{social}. This will reflect the social marginal product of capital at firm i. We define $r^{social}_{i,t}$ as the derivative of aggregate consumption $(Y_t - I_t)$ at time $t+1$ with respect to k_{it+1}, taking expectations at time t (when the investment decision is made).[6] We have:

$$\begin{aligned} r^{social}_{i,t} &:= \frac{\partial\mathbb{E}_t\left[Y_{t+1} - I_{t+1}\right]}{\partial k_{i,t+1}} \\ &= \mathbb{E}_t\left[\mathcal{P}_{i,t+1}\left(f_k(k_{i,t+1};z_{i,t+1}) + 1 - \delta\right)\right] + (1-\mathcal{P}_{i,t+1})\cdot\phi_i'(k_{i,t+1}) \end{aligned}$$

Combining this with the firm's first-order condition for investment in Equation (3) yields:

$$1 + r^{social}_{i,t} = \left(1 + r^{firm}_{i,t}\right)\mathcal{M}_{i,t} + (1-\mathcal{P}_{i,t+1})\cdot\phi_i'(k_{i,t+1}) \tag{7}$$

Note that $1 + r^{social}_{i,t-1} = \frac{\partial}{\partial k_i}g_i(k_i) + 1 - \delta$. This will allow us to use the distribution of r^{social} to measure the cost of misallocation. When we bring this result to the data, we will focus on measuring the variance of r^{social}, and use standard values to calibrate $\mathcal{E}$. Moreover, we will

[6]Proposition 2 is in terms of gross output, rather than consumption. Nevertheless, we define r^{social} in this way to parallel our definitions of ρ and r^{firm}.

make two further simplifying assumptions. First, we will focus on the unweighted variance, since the weights are difficult to observe in practice. Second, we will use the log-normal approximation $\text{Var}\left(\log\left(r^{social}_{i,t-1}+\delta\right)\right) \approx \log\left(1+\frac{\text{Var}\left(r^{social}_{i,t-1}+\delta\right)}{\mathbb{E}\left[r^{social}_{i,t-1}+\delta\right]^2}\right)$

To measure misallocation in our data, we combine this with our derivation of r^{social} to yield the following corollary:

Corollary 1. *Assume that* $\left(r^{social}_{i,t-1}+\delta\right)$ *is log-normally distributed, and also assume that weighted moments can be replaced with unweighted moments. The cost of intensive-margin misallocation is given by*

$$\begin{aligned}
&\log Y^*_t\left(K_t, \left\{\omega_{i,t}\left(S^t\right)\right\}_{i\in[0,1]}\right) - \log Y_t \\
&\approx \frac{1}{2}\cdot\mathcal{E}\cdot\log\left(1+\frac{Var\left(r^{social}_{i,t-1}\right)}{\mathbb{E}\left[r^{social}_{i,t-1}+\delta\right]^2}\right)
\end{aligned}$$

This corollary allows us to connect dispersion in r^{social}, an object that we will be able to measure in the microdata, with the cost of intensive-margin misallocation of capital. We next turn to how to measure ρ, r^{firm}, and r^{social} using credit registry data.

4 Empirical Methodology

This section describes the main data sources that we use, as well as the procedures we follow to map model objects to the data in order to estimate the three rates: the lender discount rate, ρ, the firm's cost of capital, r^{firm}, and the social cost of capital, r^{social}.

4.1 Data Sources

We rely on the FR Y-14Q dataset (Schedule H.1). This is a quarterly regulatory dataset maintained by the Federal Reserve for stress testing purposes, which contains information on individual loan facilities held in the books of the top 30 to 40 bank holding companies (BHCs) in the US. The Y-14 includes all loan facilities exceeding \$1 million and we consider data in the period ranging from 2014Q4 to 2023Q4. Importantly for the purposes of our analysis, the Y-14 contains detailed characteristics of credit facilities such as facility size, origination date and maturity, interest rate or spread, interest rate variability, and the type of loan. Additionally, the Y-14 also covers BHC's risk assessments for each borrower, which include estimates for the 1-year probability of default and loss given default. The probability of default is typically estimated using internal default models that have to be approved by regulators. While there is scope for some discretion in the assignment of these default probabilities (Plosser and Santos, 2018), these models are subject to standardized guidelines following Basel II (BCBS, 2001). We focus on term loans issued to non-governmental and nonfinancial companies based in the US. Our unit of observation is a loan origination. We do not include credit lines due to lack of information about the fee structure, which would be needed to price these facilities. Appendix B contains a detailed description of the data cleaning procedure and sample restrictions.

In terms of coverage, Faria-e-Castro et al. (2024) show that the FR Y-14Q Schedule H.1 accounts for 91% of Commercial & Industrial lending undertaken by the 25 largest banks in the US (FRED mnemonic: CIBOARD), and 55% of all Commercial & Industrial lending undertaken by all commercial banks in the US (FRED mnemonic: BUSLOANS). Our focus in term loans and relatively stringent cleaning procedures leave us with a total of 61,910 loans.

4.2 Mapping the Model to the Data

An important difference between the model and the data is the payment structure of loans. In the model, for tractability, we assume that firms borrow in long-term debt that is modeled as a perpetuity with geometrically decaying coupons. In the data, on the other hand, we focus our analysis on term loans with a fixed maturity. This section shows how we map model objects to the data, and how we exploit the Y-14 data to retrieve estimates of the lender's discount rate, ρ, the firm's cost of capital, r^{firm}, and the social cost of capital, r^{social}.

Consider a generic term loan with principal value B, maturity T, payment schedule $\{D_t\}_{t=1}^{T}$, repayment probability P assumed to be constant over time, and loss given default

LGD, also constant over time.[7] The break-even condition for a lender with discount rate ρ is given by:

$$B = \sum_{t=1}^{T} \left[\frac{P^t D_t + P^{t-1}(1-P)\,(1-LGD)\,B}{(1+\rho)^t} \right],$$

Assume now that the loan is a non-amortizing term loan, with each payment consisting of interest over the life of the loan, and the final payment consisting of a lump-sum principal repayment. Thus $D_t = r_t B$ for $t < T$ and $D_T = (1 + r_T)B$. The interest rate r_t is either a fixed interest rate, or a fixed spread over a floating benchmark rate. We can then rewrite the break-even condition at origination as:

$$1 = \sum_{t=1}^{T} \left[\frac{P^t \mathbb{E}_0\,(r_t) + P^{t-1}(1-P)\,(1-LGD)}{(1+\rho)^t} \right] + \frac{P^T}{(1+\rho)^T}, \tag{8}$$

This equation balances the present value of expected payments from the borrower against the lender's opportunity cost, ensuring that the lender breaks even. For a fixed-rate term loan, data on (P, LGD, T, r) allows us to solve this equation for the match-specific lender's discount rate ρ.

Floating Rate Loans. The data has loans with either fixed or floating rates. To estimate ρ for floating rate loans, it is necessary to obtain estimates of $\mathbb{E}_0\,(r_t)$, the expected interest rate. Floating rate loans typically charge a reference rate plus a spread. For our analysis, we use smoothed daily yield curve estimates provided by the Federal Reserve Board, based on the methodology described in Gürkaynak et al. (2007). Under the expectations hypothesis, long-term interest rates are assumed to reflect the market's expectations of future short-term rates. For each floating rate loan, we compute the sequence of forward short-term interest rates at the date of origination, and add the (fixed) loan spread to obtain a sequence of interest rates that are used to price that loan. Using this framework, we back out $\mathbb{E}_0\,(r_t)$

[7] To measure the amount borrowed, B, we note that banks record loans on their books using historical-cost accounting (Begenau et al., Forthcoming). When the loan is originated, the loan value recorded on the books is the amount of money that the bank gave the firm.

for each loan by combining the treasury forward rate with the loan's spread.[8] It is worth noting majority of floating rate loans in our sample are indexed to the LIBOR/SOFR rather than Treasury rates. However, for the period in analysis, the spread between the SOFR and short-term Treasury rates is negligible. In the absence of readily available forward curve estimates for the LIBOR or SOFR, we treat them as identical to the Treasury curve.

Lender's Discount Rate. Proposition 3 characterizes the lender's discount rate, ρ, in the context of fixed interest rate loans.

Proposition 3 (Lender's Discount Rate)**.** *For a fixed interest rate loan:*

$$1 + \rho_{i,t} = P_{i,t}\left(1 + r_{i,t}\right) + \left(1 - P_{i,t}\right)\left(1 - LGD_{i,t}\right).$$

Where $r_{i,t}$ is the fixed interest rate on the loan. This expression reflects the lender's return, accounting for repayment in non-default states and recovery in default states. A key result for fixed rate loans is that $\rho_{i,t}$ is independent of the loan's maturity $T_{i,t}$, which simplifies its calculation and interpretation.

For variable rate loans, however, the calculation of $\rho_{i,t}$ requires a numerical solution of the break-even condition presented in equation (8).

Firm's Cost of Capital. Returning to the firm's cost of capital of Proposition 1, we can estimate Λ for term loans, and then solve for r^{firm}. Proposition 4 provides an equation to estimate Λ directly from the data.

Proposition 4 (Firm's Cost of Capital)**.** *We can solve for $\Lambda_{i,t}$ as:*

$$\Lambda_{i,t} = \frac{\left(1 - P_{i,t}\right)\left(1 - LGD_{i,t}\right)}{1 + \rho_{i,t} - \left(1 - P_{i,t}\right)\left(1 - LGD_{i,t}\right)}.$$

[8]More specifically, the estimate for the reference rate n years ahead at time t is given by $f_t(n,0) = \beta_0 + \beta_1 \exp(-n/\tau_1) + \beta_2(n/\tau_1)\exp(-n/\tau_1) + \beta_3(n/\tau_2)\exp(-n/\tau_2)$ (equation 21 of their paper), where estimates for $(\beta_1, \beta_2, \beta_3\tau_1, \tau_2)$ are regularly updated by the Board of Governors and available at https://www.federalreserve.gov/data/nominal-yield-curve.htm for each date. We compute the sequence of forward rates at loan origination, and add the fixed spread to obtain an estimate for the interest rate at each repayment point in time.

This allows us to write the firm's cost of capital as:

$$1 + r_{i,t}^{firm} = \frac{1 + \rho_{i,t}}{1 + \Lambda_{i,t}} = (1 + \rho_{i,t}) - \underbrace{(1 - P_{i,t})(1 - LGD_{i,t})}_{Expected\ Recoveries}.$$

In this expression, $(1 - P_{i,t})(1 - LGD_{i,t})$ represents expected recoveries, capturing the key difference between $\rho_{i,t}$ and $r_{i,t}^{firm}$: lenders benefit from expected recoveries, but borrowers get zero profit in the default state, regardless of whether the lender recovers anything on the loan.

For fixed interest rate loans, we can use the expression derived in Proposition 3 to simplify the firm's cost of capital to:

$$1 + r_{i,t}^{firm} = (1 + r_{i,t})\, P_{i,t},$$

where $r_{i,t}$ is the fixed interest rate. This formula reflects how the borrower's cost adjusts based on the likelihood of repayment and default outcomes. As a result, we are able to measure the firm's cost of capital for each loan in the data at origination.

Social Cost of Capital. We can also use the data to estimate the social cost of capital, $r_{i,t}^{social}$. For measurement, we specialize and assume that the liquidation technology is linear, meaning it takes the form $\phi_i(k_i) = \phi_i \cdot k_i$. Combining with Equation (7), this yields a formula for $r_{i,t}^{social}$ in terms of objects in the data.

Proposition 5. *Assume a linear liquidation technology. The social cost of capital is then:*

$$\begin{aligned} 1 + r_{i,t}^{social} &= \left(1 + r_{i,t}^{firm}\right) \mathcal{M}_{i,t} + (1 - P_{i,t}) \cdot (1 - LGD_{i,t}) \cdot lev_{i,t} \\ &= (1 + \rho_{i,t})\, \mathcal{M}_{i,t} + (lev_{i,t} - \mathcal{M}_{i,t}) \cdot (1 - P_{i,t}) \cdot (1 - LGD_{i,t}) \end{aligned}$$

where $lev_{i,t} := \frac{Q_{i,t} \cdot b_{i,t+1}}{k_{i,t+1}}$ *is the firm's leverage ratio.*

In our empirical analysis, we will set the price feedback multiplier $\mathcal{M}_{i,t} = 1$.[9] Under that calibration, the social cost of capital simplifies further:

[9]We show that this is a good assumption in Appendix B.2, where we estimate this variable in the data and find a distribution extremely concentrated around 1.

$$1 + r_{i,t}^{social} = \underbrace{1 + \rho_{i,t}}_{\text{Matching efficiency}} + \underbrace{(\text{lev}_{i,t} - 1) \cdot (1 - P_{i,t}) \cdot (1 - LGD_{i,t})}_{\text{Financial Frictions}} \quad (9)$$

The social cost of capital is thus equal to the lender's discount rate (i.e., matching efficiency), plus a term that reflects financial frictions related to default and recovery. This latter term reflects the tension between lenders and borrowers. The presence of this term also implies that whether the social return on capital exceeds the lender's discount rate or not is a function of whether firm leverage exceeds 1 or not. This is due to the following: lenders care about average recovery per dollar of debt, $\phi_i(k_i)/b_i$, which is equal to $1 - LGD_i$ in the data. The planner, on the other hand, cares about the marginal recovery $\phi_i'(k_i)$, which is equal to $(1 - LGD_i) \cdot \text{lev}_i$ in the data. The two coincide when $\text{lev}_i = 1$, and so the social cost of capital equals the lender's discount rate in that case.

5 Empirical Results

5.1 Summary Statistics

We provide summary statistics for key variables in Table 1. Our unit of observation is a loan origination, and so all reported firm financials correspond to the financials of the quarter in which that origination took place. The average annual loan interest rate in our sample is 4.17%. These loans have an average expected default probability of 1.42% over the next year, and banks expect to lose, on average, 34.5% of the outstanding value of the loan in the event of default. As a result, the lender's discount rate, ρ, averages 3.75%.[10] The social cost of capital and firm's cost of capital are even lower, at 3.54% and 2.82% respectively.

Interest rates vary across loans, with a standard deviation of 1.7%, reflecting heterogeneity both within and across time. The lender's discount rate shows similar heterogeneity, with

[10] The negative covariance of r and P means that the average ρ is lower than we might have expected from the raw averages of P, r, and LGD. For fixed rate loans, $1 + \rho = P(1 + r) + (1 - P)(1 - LGD)$; the average value of $P(1 + r)$ will be brought down by the fact that P is low when $(1 + r)$ is high.

a standard deviation of 1.7%. In contrast, the social cost of capital and firm's cost of capital have higher heterogeneity, with standard deviations of 1.88% and 2.75%, respectively.

Why are the variances of the lender's discount rate and the contractual interest rate similar? To build intuition, we can focus on the formula for the lender's discount rate, ρ, for fixed rate loans. With some rearrangement, the lender's discount rate can be expressed as $\rho = r - (1 - P)(r + LGD)$. Since r is small at annual frequencies compared to LGD, we can use the approximation $\rho \approx r - (1 - P) \cdot LGD$. This yields the variance decomposition

$$\mathbb{V}[\rho] \approx \mathbb{V}[r] + \mathbb{V}[(1 - P) \cdot LGD] - 2 \cdot \mathbb{C}[r, (1 - P) \cdot LGD] \tag{10}$$

The variance of ρ is similar to the variance of r because the variance of expected losses,$(1 - P) \cdot LGD$, is offset by the covariance term: interest rates are higher when the lender's expected losses are high.

We view these results as a vindication for our method of estimating the cost of capital. If our observed measures of default probabilities and recovery rates were just noise, then the variance of the lender's discount rate would be substantially greater than the variance of interest rates: the covariance term would be zero, and the term $\mathbb{V}[(1 - P) \cdot LGD]$ would push the variance of ρ substantially above the variance of r. Instead, the variance of the cost of capital is similar to that of the interest rate, suggesting that default probabilities and recovery rates covary with interest rates in the way that we would expect in a financial market that is close to efficient.

5.2 Averages by Quarter of Origination

We begin by analyzing the time series of average values, by quarter of origination. The key inputs into our measures of the cost of capital are the interest rate, default probability, and loss given default. We first analyze the behavior of these averages over time, in Figure 1. We separate the interest rate time series into interest rates on fixed-rate loans and the spread for variable-rate loans. During the time period we study, interest rates fall and then rise concurrent with the movement of monetary policy; average spreads are very stable, ranging

Table 1: Summary Statistics

	mean	sd	p10	p50	p90
Interest rate	4.17	1.69	2.21	3.93	6.59
Maturity (yrs)	6.85	4.64	3.00	5.00	10.25
ρ (%)	3.75	1.69	2.05	3.69	5.88
r^{firm} (%)	2.82	2.75	0.87	3.04	5.26
r^{social} (%)	3.54	1.88	1.77	3.53	5.71
Prob. Default (%)	1.42	2.37	0.19	0.82	2.85
LGD (%)	34.50	13.20	16.00	36.00	50.00
Loan amount (M)	10.77	68.81	1.11	2.55	22.64
Sales (M)	1,254.73	5,923.53	2.17	58.80	1,556.58
Assets (M)	1,770.83	8,956.78	1.06	35.52	1,792.00
Leverage (%)	72.03	24.57	42.57	71.17	100.00
Return on assets (%)	22.61	29.05	4.68	15.56	44.04
N Loans	62687				
N Firms	38587				
N Fixed Rate	31540				
N Variable Rate	31147				

from 1.9% to 2.3%. Default probabilities show a modest upward secular trend, along with a temporary spike around the time of the COVID-19 pandemic. Expected losses given default fall around the onset of the pandemic, implying that banks expect larger recoveries in the event of default. Note, however, that the magnitude of the change in recoveries is sufficiently small that it has little effect on ρ, since this change is multiplied by the (small) probability of default.

Next, in Figure 2, we plot the lender's discount rate, ρ, the firm's cost of capital, r^{firm}, and the social cost of capital, r^{social}, against the five-year treasury rate. The average lender's discount rate is similar to the average r^{social}, and both rates covary strongly with the five-year treasury rate. There is an average spread of roughly 164 basis points between the lender's discount rate and the treasury rate, although it has a delayed reaction to movements in treasury rates: the spread is initially stable at 150 basis points, then rises above the average when treasury rates fall and falls below the average once treasury rates rise again. Note that the lender's discount rate is already adjusted for default risk, and so this cannot explain the

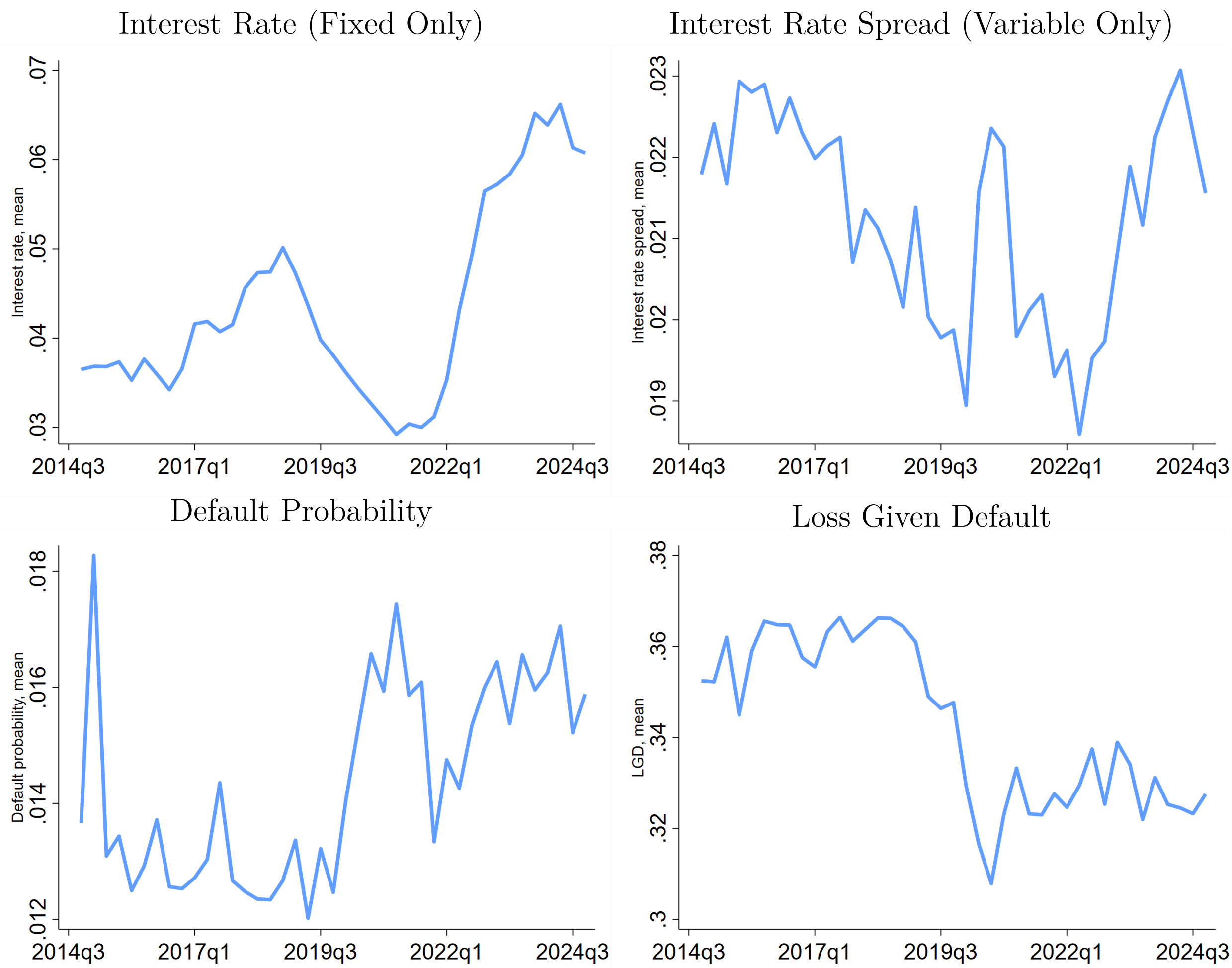

Figure 1: Average values of key inputs by quarter of origination: contractual interest rate (fixed rate loans only), interest rate spread (floating rate loans only), probability of default, and loss given default.

spread relative to treasuries. While the social cost of capital, r^{social} is quite close to ρ, the firm's cost of capital, r^{firm}, tracks the treasury rate more closely with a small spread.

While our analysis takes into account the maturity of the loan, there are potential concerns that loans of different maturities may face different rates, even if they reflect a constant spread on a (time-varying) risk-free rate. To mitigate these concerns, in Appendix B.5, we redo our analysis focusing on fixed-rate, five-year loans. This is the most common maturity for fixed rate loans. Focusing on fixed-rate loans is convenient because it is not sensitive to the term structure of the loan, nor to estimates of expected future rates derived from the

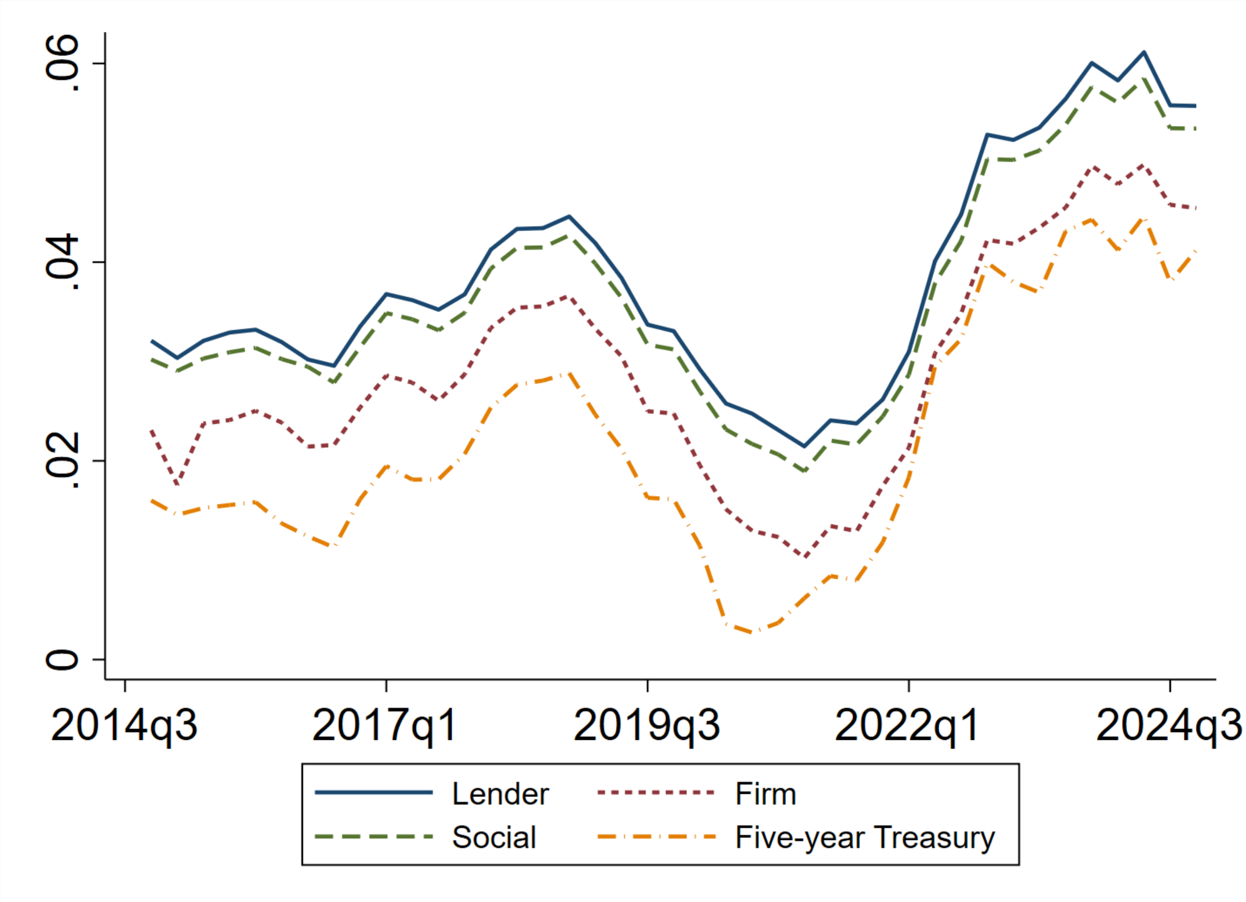

Figure 2: Averages by quarter of origination for the lender's discount rate (ρ), firm cost of capital (r_{firm}), social cost of capital (r_{social}), and time series for the five-year Treasury rate.

yield curve. Five-year loans are also convenient because they allow direct comparison to the five-year treasury rate. The average cost of capital for fixed-rate, five-year loans is very similar to the overall sample: there is a roughly 150 basis point spread relative to five-year treasuries, following the same dynamics as in the overall sample.

5.3 Cross-Sectional Heterogeneity

While the means display similar movements, they mask a substantial amount of cross-sectional heterogeneity in these measures. In this section, we show that this heterogeneity is substantial across measures even after controlling for time, lender and firm fixed effects. To this end, we follow the variance decomposition of Daruich and Kozlowski (2023). To ensure that we can estimate firm level fixed effects we subset our sample to the set of firms with five or more distinct loans. We then progressively add time, lender-time, and firm-lender-time fixed effects, building to the fixed-effects specification displayed in Equation 11 below, where i indexes firms, τ represents the quarter of origination, b indexes lenders, and l represents the particular loan.

$$r_{\tau bil} = \alpha_{\tau} + \gamma_{\tau b} + \delta_{\tau bi} + \varepsilon_{\tau bil} \tag{11}$$

The results are in Table 2, for the contractual interest rate, lender's discount rate, firm cost of capital, and social cost of capital. The time fixed effect explains 72% of the variance in interest rates and 62% of the variance in the lender's discount rate. For all four variables, adding in lender-time fixed effects explains a negligible share of the variance (at most 4.25% for r^{firm}), suggesting that heterogeneity across lenders is not an important source of heterogeneity in interest rates or the different measures of cost of capital. Adding in firm-lender-time fixed effects explains an additional 13% to 20% of the variance of these measures. Finally, notice that the loan-level variance, after controlling for firm-lender-time fixed effects, remains substantial, ranging from 13% for contractual interest rates up to 42% for the firm cost of capital. To the extent that our measure of misallocation depends on the variance of r^{social}, these results show that a significant share of that variance is loan-specific and cannot be accounted for by either time, lender, or borrower effects.

	Time	Bank	Firm	Loan
Contractual rate	71.88	1.63	13.45	13.04
Lender discount rate, ρ	61.94	3.08	14.02	20.96
Firm cost of capital, r^{firm}	33.23	4.25	20.12	42.4
Social cost of capital, r^{social}	53.84	3.87	16.21	26.08
N Firms	1681			
N Loans	14738			

Table 2: Variance decomposition for contractual interest rates and different measures of the cost of capital (ρ, r^{firm}, and r^{social}) using equation 11.

In Appendix B.3 we further explore the correlation between the cost of capital and firm-level covariates such as leverage, return on assets, and assets. Of the three covariates, the best predictor is the return on assets; interest rates and the cost of capital are consistently higher at firms with high return on assets. Although we cannot attach a causal interpretation to the estimated coefficients, this would be consistent with a model where causality runs from the cost of capital to firm decisions: firms with a higher cost of capital will demand a high return on their investments. Yet perhaps more notable is the very low R^2. The return on assets explains between 2 and 3% of the variance, depending on the measure of the cost of capital, with other covariates explaining less than 1%. Firm-level covariates explain approximately none of the variance in the cost of capital.

5.4 Misallocation

What does the heterogeneity in the cost of capital imply for the cost of misallocation? To answer this we use our approximate formula for misallocation from Corollary 1, calibrating $\delta = 0.06$.[11] We compute this statistic by quarter of origination, in order to focus on within-period misallocation. Our model does not contain aggregate shocks, and we would thus need a richer model to study misallocation across time. We interpret our results below as reflecting what misallocation would be for an economy that remained in the same steady state.

We find that the misallocation of capital resulting from heterogeneity in the cost of capital, plus the financial friction term, is small. We plot our estimates in Figure 3. In the period before the COVID pandemic the implied misallocation is flat and low: formally reallocating capital across firms would increase aggregate output by 0.54%. This number rises with the onset of the pandemic, averaging 1.11% during 2020 and 2021, before falling back to a somewhat elevated 0.76% starting in 2022.

Our model of misallocation studies an economy in steady-state, which complicates the

[11] Note that in steady state $\delta = \frac{I/Y}{K/Y}$. In the data, at the annual frequency, the capital-output ratio is about 3 while the investment-output ratio is about 0.18 (we measure capital as BEA Current-Cost Net Stock of Fixed Assets and investment is GDPI in FRED). Hence, at an annual frequency, $\delta = 0.06$. Additionally, recall that, as previously explained, we set $\mathcal{E} = 1/2$ and $\mathcal{M} = 1$.

interpretation of short-run changes in the distribution of the cost of capital. A temporary shock to the dispersion of r^{social} among newly originated loans will have only limited effects on the dispersion of r^{social} in the full population of firms. Moreover, a steady-state model with no aggregate shocks is not well suited to studying aggregate dynamics in response to a shock. Thus, we caution against over-interpreting the transitory rise in implied misallocation during the pandemic: if the increased dispersion of r^{social} were permanent, then misallocation would rise by 0.6% in steady state, but it is not obvious how much misallocation actually rose in response to the transitory shock. Instead, our main takeaway from the analysis is that in "normal times" (e.g. before the pandemic), heterogeneity in r^{social} implies a very small cost of misallocation in steady state.

5.5 Decomposing Misallocation

To further understand the drivers of misallocation, we decompose it into the component coming from heterogeneous cost of capital, ρ, and the component coming from heterogeneity in the financial frictions term. We perform two counterfactuals. In the first, we replace ρ with its average value for that quarter. This tells us how much misallocation arises from heterogeneity in the financial frictions term. In the second counterfactual we set the financial frictions term equal to its average value within the quarter, which allows us to measure the misallocation arising from heterogeneity in ρ. Note that neither counterfactual changes the within-quarter average of r^{social}, and thus the results are driven by changes in the variance of r^{social}.

We show the results of these decompositions in Figure 4. Misallocation is mostly driven by heterogeneity in the cost of capital. In the pre-pandemic period, if ρ is equalized across firms then the cost of misallocation falls to just 0.07% instead of 0.54%. In contrast, the cost of misallocation in the counterfactual with a constant financial friction is 0.39%. Note that there is an interaction between ρ and the financial frictions term as total misallocation exceeds the sum of the two counterfactuals. During the pandemic period (2020-2021), there is a significant increase in total misallocation, which more than doubles. While still small, the

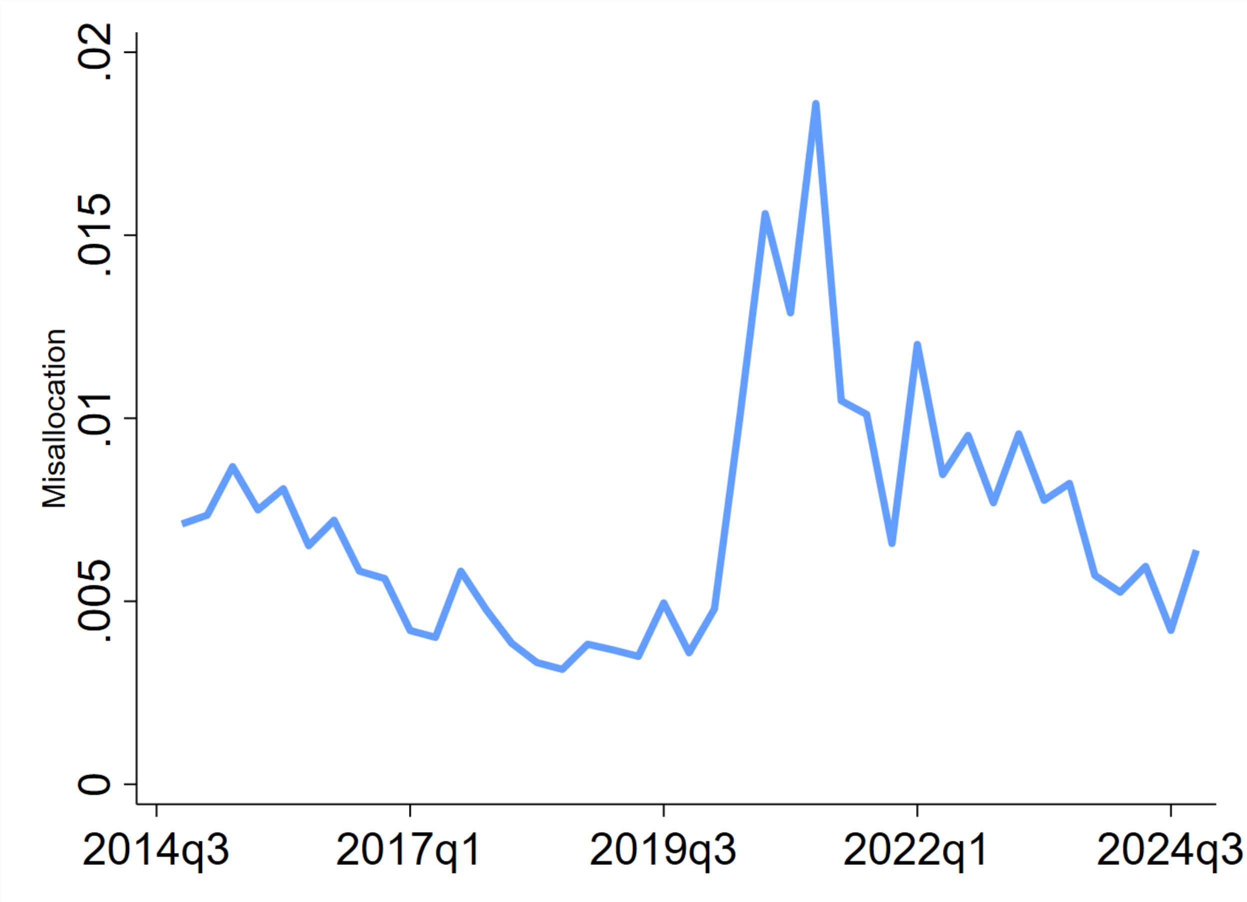

Figure 3: Time series for the cost of misallocation, using the formula in Corollary 1. The cost of misallocation is the percentage difference between actual output and output in a counterfactual economy where the planner is free to reallocate physical capital across firms, keeping exit decisions fixed.

financial frictions term also doubles in importance. Quantitatively, the large increase comes from an increased dispersion in ρ. In the post-pandemic period (2022-2024), misallocation remains 40% more elevated than in the pre-pandemic period. Once again, ρ plays the dominant role: misallocation would be only 0.11% in the constant ρ counterfactual.

In Appendix B.5, we repeat our misallocation analysis, focusing on five-year, fixed-rate loans. We find that the results are very similar to those of our main analysis.[12] This reinforces

[12] In the robustness sample, there is also a brief spike in the dispersion of r^{social} earlier in the period, but it only lasts for one quarter.

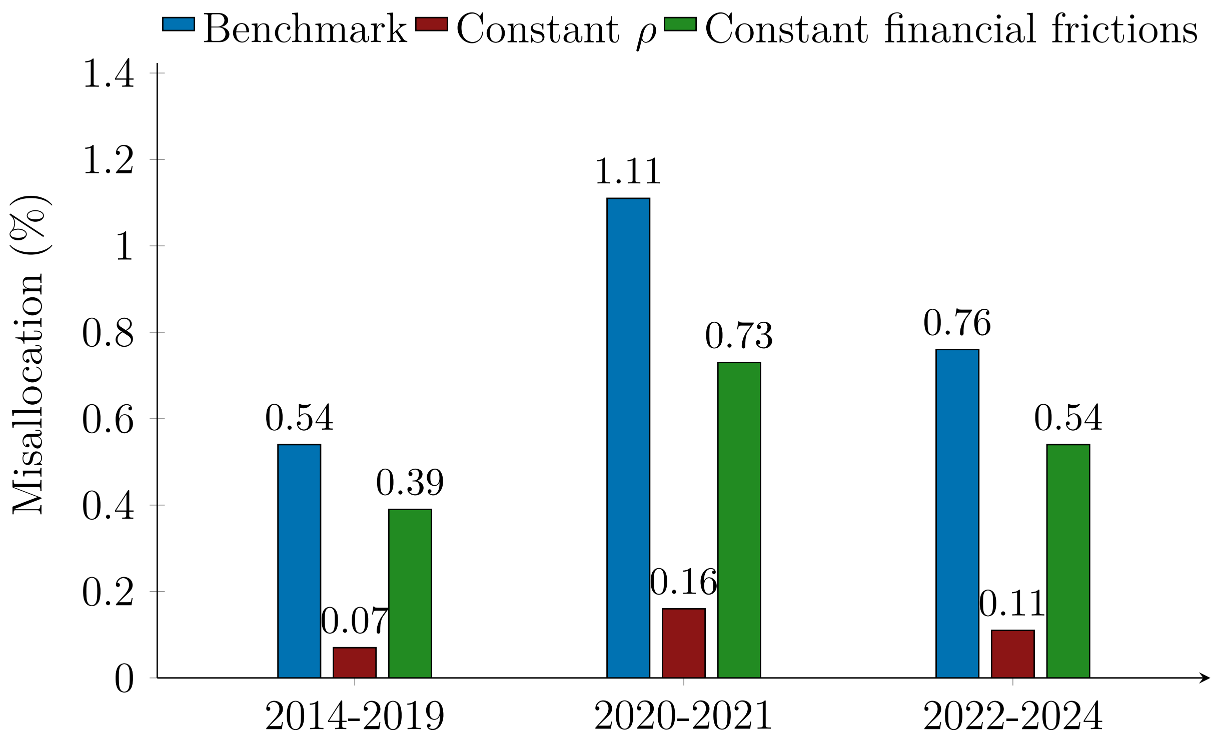

Figure 4: Decomposing the cost of misallocation. The "benchmark" columns correspond to the averages for the cost of misallocation in Figure 3 in the respective time period. "Constant ρ" is the average cost of misallocation in a counterfactual exercise where we set the estimated ρ for each newly originated loan to be equal to the average ρ for that quarter. "Constant financial frictions" repeats the exercise for a counterfactual where we keep the second term in equation 9 equal to its average value for each quarter, for all newly originated loans.

the robustness of our results, confirming that heterogeneity in the cost of capital across firms is not driven by differences in maturity or term structure.

5.6 The 2020-21 increase in misallocation

Our analysis in Figure 4 reveals that while both components of the social cost of capital contribute to an increase in misallocation in the 2020-21, the heterogeneity in ρ is quantitatively the most important factor. Figure 5 shows the mean and standard deviation of ρ. First, as already shown in Figure 2 the average ρ follows the risk-free interest rate. Hence, as risk free rates decreased during 2020-2021, the average ρ also decreased. Second, the standard deviation of ρ increases during this period. Both these movements contribute to an increase in the coefficient of variation for ρ. The increased dispersion in ρ translates into an increased dispersion of r^{social} which directly affects our measure of misallocation.

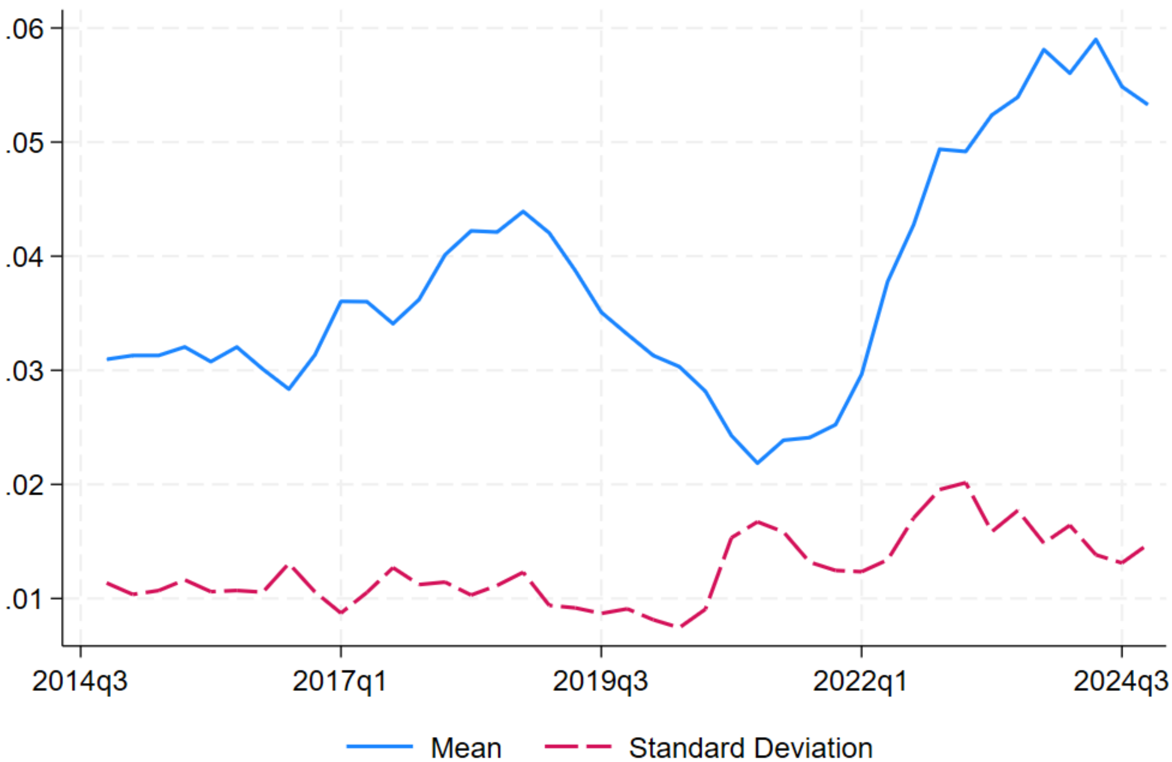

Figure 5: Mean and standard deviation of the lender's discount rate, ρ.

We can further understand the changes in the moments of ρ with the approximation in Equation (10), which allows us to approximate the coefficient of variation for ρ as a function of the means, variances, and covariance between the contractual interest rate r and the expected loss $(1-P)LGD$:

$$\frac{\mathbb{V}\left[\rho\right]^{0.5}}{\mathbb{E}\left[\rho\right]} \approxeq \frac{\left(\mathbb{V}\left[r\right]+\mathbb{V}\left[(1-P)LGD\right]-2\mathbb{COV}\left[r,(1-P)LGD\right]\right)^{0.5}}{\mathbb{E}\left[r\right]-\mathbb{E}\left[(1-P)LGD\right]}$$

We use this expression to consider three counterfactuals: one in which the mean and variance of contractual interest rates are kept constant at their sample averages, one in which the mean and variance of the expected loss are kept constant, and one in which the covariance between the two is kept constant. The results of this exercise are presented in Figure 6, where "Benchmark" refers to the actual coefficient of variation for ρ. While all factors contribute to the rise of the CV during the 2020-22 period, changes in the distribution of contractual interest rates are by far the most important factor.

Let us consider the counterfactual scenario of a constant contractual rate ("constant r"). In this exercise, the mean and variance of expected losses vary, but the mean and variance of contractual interest rates remain fixed. As shown by the red dashed line in Figure 6, under this counterfactual, the dispersion in ρ would have increased only modestly—from 0.27 to 0.46—representing a change of only 0.19, while in the benchmark the dispersion of ρ increased by 0.48. This finding implies that a substantial portion of the observed increase in the dispersion of ρ in the benchmark scenario is driven by increased dispersion in contractual rates.

What explains the rise in contractual rate dispersion during 2020–2021? Our inspection of the microdata reveals that this shift was primarily driven by very risky loans—those with high expected losses—being underpriced, i.e., offered with unusually favorable contractual rates. As a result, the implied ρ for these loans was quite low, contributing significantly to the overall increase in ρ dispersion.

We hypothesize that this pattern emerged due to broad-based fiscal and monetary interventions implemented during 2020–2021 to support the economy. In particular, many of these programs involved lending to or rescuing firms and sectors under financial distress (such as the Paycheck Protection Program, the Main Street Lending Program, Primary Market Corporate Credit Facility, or the Secondary Market Corporate Credit Facility, among others). Our hypothesis is that lenders inferred an implicit government guarantee for many of these loans: if risky borrowers were to default, lenders believed the government would likely intervene to cover the losses.

This perceived guarantee introduced a moral hazard problem, encouraging lenders to take on more risk than they otherwise would have. Ultimately, this behavior led to increased misallocation of credit. Absent such implicit guarantees, we argue that lenders would have priced risk more accurately—leading to greater allocative efficiency.

Risk premia and aggregate shocks. Alternatively, a plausible explanation for the rise in ρ is that it reflects risk premia, as lenders require an increased compensation for risk during an extremely uncertain and volatile period. Furthermore, since that different firms may be differentially exposed to aggregate shocks, the existence of heterogeneity in risk premia may not necessarily imply the presence of misallocation (David et al., 2022). Our framework is set in steady state and cannot accommodate aggregate shocks that would trigger increases in risk premia. An increase in risk premia for some firms, say firms whose cash flows were more exposed to COVID-19 disruptions, should raise the skewness of the distribution of ρ. What we find is the opposite: not only the average ρ falls from 3.6% (2014-19) to 2.7% (2020-21), but its skewness also becomes more negative (from -2.6 to -3.5). Thus, if anything, the "left tail" becomes more pronounced. This is at odds with an explanation related to the rise in risk premia and, if anything, suggests that risk premia seem to have fallen during this period, potentially due to explicit and implicit guarantees.

5.7 Cross-Country Comparison

Table 3: Cross-Country Comparison of Cost of Capital

	Pakistan 1980–1981	Pakistan 1996–2002	Brazil 2006–2016	Mexico 2003–2022	United States 2014–2024
Average contractual rate, %	78.7	14.1	83.0	16.8	3.9
St deviation of contractual rate, %	38.1	2.9	93.3	5.2	1.5
Default probability, %	2.7	16.9	4.0	8.9	1.4
Recovery rate (World Bank), %	42.8	42.8	18.2	63.9	81.0
Implied Misallocation, %	4.9	2.2	21.5	1.7	0.6

Notes: Data for Pakistan (1980–1981) are from Aleem (1990), and for 1996–2002 from Khwaja and Mian (2005). Brazilian data are from Cavalcanti et al. (2021), and Mexican data from Beraldi (2025). Recovery rates ($1 - \text{LGD}$) are country-level estimates obtained from the World Bank's Doing Business database.

Finally, in Table 3, we compare our results to related values for other countries. Our method for deriving the distribution of r_{social} incorporates the joint distribution of interest

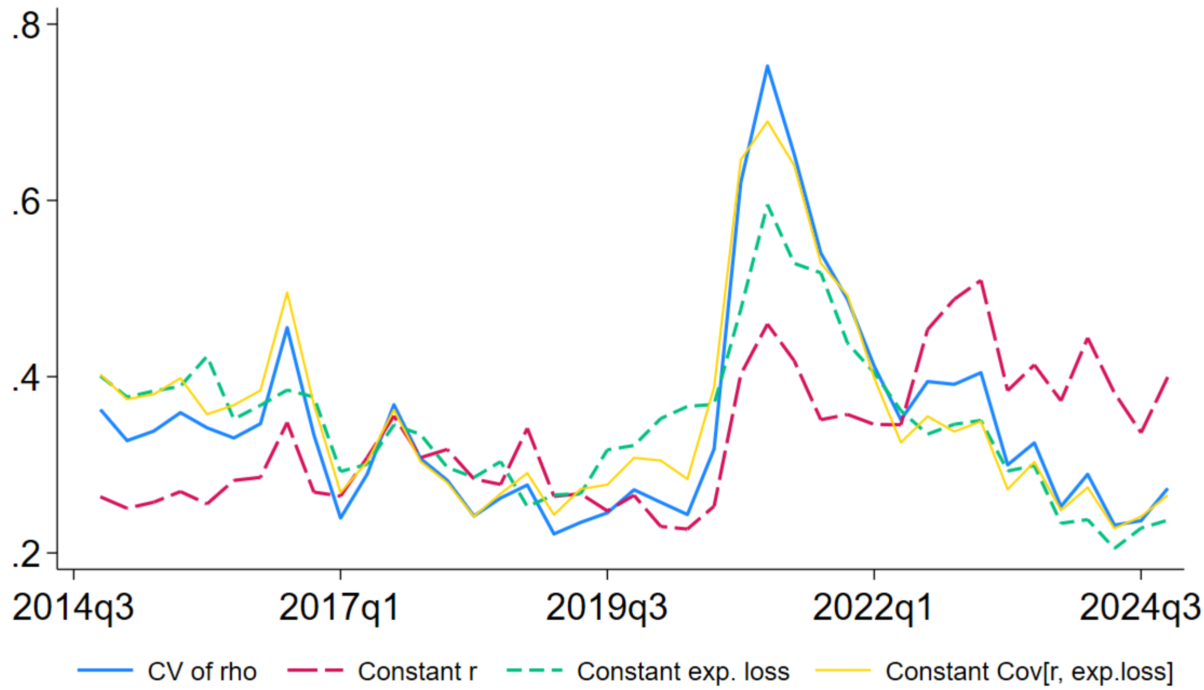

Figure 6: Decomposition of the coefficient of variation of ρ.

rates, default probabilities, expected losses given default, and leverage. Although prior work has not combined all of these variables in the ways necessary to apply our method we can still learn something from more commonly reported statistics. We focus on papers that report summary statistics for a representative sample of bank loans, or moneylenders in the case of Aleem (1990). We are able to find five such papers that report the mean and standard deviation of interest rates as well as the mean probability of default. For comparison, we also use our data to provide the same statistics for the United States.[13] Relative to less developed countries, the United States stands out for a low mean and low standard deviation of interest rates. However there is significant heterogeneity among the results for other countries, which

[13] Cavalcanti et al. (2021) report the spread relative to deposit rates; we compute the mean interest rate by adding the mean deposit rate, which they report in the paper, to the mean spread. Beraldi (2025) reports the spread relative to the Mexican 28-day interbank rate; we compute the mean interbank rate over this time period and add this to the mean spread. For both papers, we use the distribution of spreads to get the standard deviation of the spread. Beraldi (2025) reports quantiles instead of the standard deviation, so we use the 90/10 percentile range and a normal approximation to infer the standard deviation.

does not appear to simply track the level of development. Bank loans in Pakistan and Mexico have a standard deviation of interest rates that is only modestly higher than in the United States, while bank loans in Brazil and loans from moneylenders in Pakistan have an extremely high standard deviation of interest rates.

We can attempt to use these statistics to compute misallocation, at the cost of strong assumptions. We use recovery rates from the World Bank's Doing Business database to provide information on expected losses given default. Since we do not have information on firm leverage, we use the lender's cost of capital, ρ, in place of the social cost of capital, r_{social}. We then use the fixed rate formula for ρ and assume that the probability of default and the losses given default do not vary across firms. This allows us to compute a cost of misallocation, which we show in the last column. The cost of misallocation we compute for the United states is similar to the actual cost that we computed earlier, although this is no guarantee that the same is true for other countries.

There are two main takeaways from our misallocation analysis. First, the United states seems to have the most efficient credit markets of the countries in our table: misallocation is moderately higher for bank loans in Mexico and Pakistan, and substantially higher in Brazil and among Pakistani moneylenders. However the relationship between development and credit market efficiency varies across settings, even for countries like Brazil and Mexico that are at similar levels of economic development.

6 Conclusion

This paper develops a novel methodology to estimate the cost of capital using credit registry microdata, and examines the implications of dispersion in the cost of capital for misallocation. We show, in a dynamic corporate finance model, the connection between the lender's cost of capital, the firm's cost of capital, and the social cost of capital, and how to measure these objects in the data. We also show how the mean and variance of the social cost of capital can be used as sufficient statistics to measure the output losses from misallocation that arise from credit market imperfections.

After developing this general methodology, we apply it to credit registry data for the United States. We find that although the cost of capital varies across firms, the resulting misallocation is modest in normal times, resulting in output losses of only 0.5%. However, dispersion in the social cost of capital among newly originated loans rose dramatically during the COVID-19 pandemic, driven by a rise in the dispersion of lender discount rates. Understanding the causes of this rise in dispersion, as well as the consequences for aggregate productivity, is an important area for future research. Moreover, comparing the distribution of the cost of capital in the United States to the distribution in other economies, especially less developed economies, will help us better understand how financial markets contribute to development.

References

Irfan Aleem. Imperfect information, screening, and the costs of informal lending: a study of a rural credit market in pakistan. *The World Bank Economic Review*, 4(3):329–349, 1990.

Abhijit V. Banerjee and Esther Duflo. Chapter 7 growth theory through the lens of development economics. In *Handbook of Economic Growth*, pages 473–552. Elsevier, 2005. doi: 10.1016/s1574-0684(05)01007-5.

BCBS. The Internal Ratings-Based Approach. Consultative document, Basel Committee on Banking Supervision, January 2001.

Juliane Begenau, Saki Bigio, Jeremy Majerovitz, and Matias Vieyra. A q-theory of banks. *The Review of Economic Studies*, Forthcoming.

Francesco Beraldi. Banking relationships and loan pricing disconnect. Job Market Paper, Yale University, January 2025. URL https://francberaldi.github.io/MyPapers/FrancescoBeraldi_JMP.pdf.

Paul Carrillo, Dave Donaldson, Dina Pomeranz, and Monica Singhal. Misallocation in firm production: A nonparametric analysis using procurement lotteries. Technical report, jun 2023.

Tiago V Cavalcanti, Joseph P Kaboski, Bruno S Martins, and Cezar Santos. Dispersion in financing costs and development. Technical report, National Bureau of Economic Research, 2021.

Thomas F. Cooley and Vincenzo Quadrini. Financial markets and firm dynamics. *American Economic Review*, 91(5):1286–1310, December 2001. doi: 10.1257/aer.91.5.1286. URL https://www.aeaweb.org/articles?id=10.1257/aer.91.5.1286.

Diego Daruich and Julian Kozlowski. Macroeconomic implications of uniform pricing. *American Economic Journal: Macroeconomics*, 15(3):64–108, July 2023. ISSN 1945-7715. doi: 10.1257/mac.20210172.

Joel M. David, Lukas Schmid, and David Zeke. Risk-adjusted capital allocation and misallocation. *Journal of Financial Economics*, 145(3):684–705, 2022. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2022.06.001. URL https://www.sciencedirect.com/science/article/pii/S0304405X22001398.

Miguel Faria-e-Castro, Samuel Jordan-Wood, and Julian Kozlowski. An Empirical Analysis of the Cost of Borrowing. Working Papers 2024-016, Federal Reserve Bank of St. Louis, July 2024. URL https://ideas.repec.org/p/fip/fedlwp/98542.html.

Simon Gilchrist, Jae W. Sim, and Egon Zakrajsek. Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs. *Review of Economic Dynamics*, 16(1):159–176, January 2013. ISSN 1094-2025. doi: 10.1016/j.red.2012.11.001.

Niels Joachim Gormsen and Kilian Huber. *Corporate Discount Rates*. June 2023. doi: 10.3386/w31329.

Niels Joachim Gormsen and Kilian Huber. *Firms' Perceived Cost of Capital*. June 2024. doi: 10.3386/w32611.

Refet S. Gürkaynak, Brian Sack, and Jonathan H. Wright. The u.s. treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304, 2007. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2007.06.029. URL https://www.sciencedirect.com/science/article/pii/S0304393207000840.

John Haltiwanger, Robert Kulick, and Chad Syverson. Misallocation measures: The distortion that ate the residual. Technical report, jan 2018.

Hugo A. Hopenhayn. Firms, misallocation, and aggregate productivity: A review. *Annual Review of Economics*, 6(1):735–770, aug 2014. doi: 10.1146/annurev-economics-082912-110223.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics*, 124(4):1403–1448, November 2009. doi: 10.1162/qjec.2009.124.4.1403.

David Hughes and Jeremy Majerovitz. Measuring misallocation with experiments. Technical report, 2025.

Barış Kaymak and Immo Schott. Tax heterogeneity and misallocation. *International Finance Discussion Paper*, (1393), 2024.

Asim Ijaz Khwaja and Atif Mian. Do lenders favor politically connected firms? rent provision in an emerging financial market. *The quarterly journal of economics*, 120(4):1371–1411, 2005.

Matthew C Plosser and João A C Santos. Banks' Incentives and Inconsistent Risk Models. *The Review of Financial Studies*, 31(6):2080–2112, 2018. URL https://ideas.repec.org/a/oup/rfinst/v31y2018i6p2080-2112..html.

Diego Restuccia and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720, October 2008. doi: 10.1016/j.red.2008.05.002.

Diego Restuccia and Richard Rogerson. The causes and costs of misallocation. *Journal of Economic Perspectives*, 31(3):151–174, aug 2017. doi: 10.1257/jep.31.3.151.

Appendix

A Proofs

Proof of Proposition 1.

$$\mathbb{E}_t\left[\frac{\mathcal{P}_{t+1}\left(\theta+(1-\theta)\,Q_{t+1}\right)}{Q_t}\right] = (1+\rho)\frac{\mathbb{E}_t\left[\mathcal{P}_{t+1}\left(\theta+(1-\theta)\,Q_{t+1}\right)\right]}{\mathbb{E}_t\left[\mathcal{P}_{t+1}\left(\theta+(1-\theta)\,Q_{t+1}\right)\right]+\mathbb{E}_t\left[\left(1-\mathcal{P}_{t+1}\right)\phi(k')/b'\right]}$$
$$= (1+\rho)\left(1+\frac{\mathbb{E}_t\left[\left(1-\mathcal{P}_{t+1}\right)\phi(k')/b'\right]}{\mathbb{E}_t\left[\mathcal{P}_{t+1}\left(\theta+(1-\theta)\,Q_{t+1}\right)\right]}\right)^{-1}$$
$$= (1+\rho)\,(1+\Lambda)^{-1}$$

where

$$\Lambda \equiv \frac{\mathbb{E}_t\left[\left(1-\mathcal{P}_{t+1}\right)\phi(k')/b'\right]}{\mathbb{E}_t\left[\mathcal{P}_{t+1}\left(\theta+(1-\theta)\,Q_{t+1}\right)\right]}$$

□

Derivation of Equation 3. We combine the first-order conditions for capital and for debt. Recall the recursive formulation of the model.

$$V(k,b,z) = \max_{k',b'} \pi(k,b,z,k',b') + \beta\mathbb{E}\left[\max\left\{V(k',b',z'),0\right\} \mid z\right]$$
$$\pi(k,b,z,k',b') = f(k,z) + (1-\delta)\,k - k' - \theta b + Q(k',b',z,\rho)\,(b'-(1-\theta)\,b)$$

The firm's maximization yields the first-order conditions for tomorrow's capital, k', and tomorrow's debt, b'.

$$0 = \frac{\partial\pi(k,b,z,k',b')}{k'} + \beta\mathcal{P}(k',b',z)\,\mathbb{E}\left[\frac{\partial}{\partial k'}V(k',b',z') \mid z, V>0\right]$$
$$0 = \frac{\partial\pi(k,b,z,k',b')}{b'} + \beta\mathcal{P}(k',b',z)\,\mathbb{E}\left[\frac{\partial}{\partial b'}V(k',b',z') \mid z, V>0\right]$$

where $\mathcal{P}(k',b',z)$ is the probability of not defaulting, and $V>0$ indicates that the firm did not default.

We next use the Envelope Theorem to note that $\frac{\partial V(k',b',z')}{\partial k'} = \frac{\partial}{\partial k'}\pi(k',b',z',k'',b'')$ and similarly to note that $\frac{\partial V(k',b',z')}{\partial b'} = \frac{\partial}{\partial b'}\pi(k',b',z',k'',b'')$. Our first-order conditions become:

$$0 = \frac{\partial \pi(k,b,z,k',b')}{k'} + \beta \mathcal{P}(k',b',z)\,\mathbb{E}\left[\frac{\partial}{\partial k'}\pi(k',b',z',k'',b'') \mid z, V > 0\right]$$
$$0 = \frac{\partial \pi(k,b,z,k',b')}{b'} + \beta \mathcal{P}(k',b',z)\,\mathbb{E}\left[\frac{\partial}{\partial b'}\pi(k',b',z',k'',b'') \mid z, V > 0\right]$$

Next, we take derivatives of the profit function to plug into our first-order conditions. We have:

$$\frac{\partial \pi(k,b,z,k',b')}{k} = f_k(k,z) + (1-\delta)$$
$$\frac{\partial \pi(k,b,z,k',b')}{b} = -\theta - (1-\theta)\,Q(k',b',z)$$
$$\frac{\partial \pi(k,b,z,k',b')}{k'} = -1 + \frac{\partial Q(k',b',z)}{\partial k'}(b' - (1-\theta)\,b)$$
$$\frac{\partial \pi(k,b,z,k',b')}{b'} = Q(k',b',z) + \frac{\partial Q(k',b',z)}{\partial b'}(b' - (1-\theta)\,b)$$

Plugging these expressions in, our first-order conditions now become:

$$0 = -1 + \frac{\partial Q(k',b',z)}{\partial k'}(b' - (1-\theta)\,b) + \beta \mathcal{P}(k',b',z)\,\mathbb{E}\left[f_k(k',z') + (1-\delta) \mid z, V > 0\right]$$
$$0 = Q(k',b',z) + \frac{\partial Q(k',b',z)}{\partial b'}(b' - (1-\theta)\,b) + \beta \mathcal{P}(k',b',z)\,\mathbb{E}\left[-\theta - (1-\theta)\,Q(k'',b'',z') \mid z, V > 0\right]$$

We next combine these two first-order conditions. Rather than thinking about investment that is financed through earnings, we want to instead imagine that the firm is financing a marginal unit of capital through borrowing. To do this, we multiply the first-order condition for debt by

$$-\frac{1 - \frac{\partial Q(k',b',z)}{\partial k'}(b' - (1-\theta)\,b)}{Q(k',b',z) + \frac{\partial Q(k',b',z)}{\partial b'}(b' - (1-\theta)\,b)}$$

which reflects the amount of new debt needed to finance a marginal unit of capital. The denominator reflects the amount raised by selling a unit of debt, $Q(k',b',z)$, plus an adjustment factor, $\frac{\partial Q}{\partial b'}(b' - (1-\theta)\,b)$, that reflects how the change in the price of debt affects the cost of borrowing. Similarly, the numerator reflects how an increase in capital lowers is partly self-financing, because it lowers the cost of borrowing.

Combining the two equations then yields:

$$\beta\mathcal{P}\left(k',b',z\right)\mathbb{E}\left[f_k\left(k',z'\right)+\left(1-\delta\right)\mid z,V>0\right]=\frac{1-\frac{\partial Q\left(k',b',z\right)}{\partial k'}\left(b'-\left(1-\theta\right)b\right)}{Q\left(k',b',z\right)+\frac{\partial Q\left(k',b',z\right)}{\partial b'}\left(b'-\left(1-\theta\right)b\right)}$$
$$\times\beta\mathcal{P}\left(k',b',z\right)\mathbb{E}\left[\theta+\left(1-\theta\right)Q\left(k'',b'',z'\right)\mid z,V>0\right]$$

Further manipulation then yields:

$$\begin{aligned}\mathcal{P}\left(k',b',z\right)\mathbb{E}\left[f_k\left(k',z'\right)+\left(1-\delta\right)\mid z,V>0\right]&=\frac{1-\frac{\partial Q\left(k',b',z\right)}{\partial k'}\left(b'-\left(1-\theta\right)b\right)}{1+\frac{\partial Q\left(k',b',z\right)}{\partial b'}\cdot\left(b'-\left(1-\theta\right)b\right)/Q\left(k',b',z\right)}\\&\quad\times\mathcal{P}\left(k',b',z\right)\mathbb{E}\left[\frac{\theta+\left(1-\theta\right)Q\left(k'',b'',z'\right)}{Q\left(k',b',z\right)}\mid z,V>0\right]\\&=\mathcal{M}\cdot\left(1+r_t^{firm}\right)\end{aligned}$$

where $\mathcal{M}$ is given by the following formula:

$$\begin{aligned}\mathcal{M}&=\frac{1-\frac{\partial Q}{\partial k'}\left(b'-\left(1-\theta\right)b\right)}{1+\frac{\partial Q}{\partial b'}\cdot\left(b'-\left(1-\theta\right)b\right)/Q}\\&=\frac{1-\frac{\partial\log Q}{\partial\log k'}\cdot\frac{Q}{k'}\left(b'-\left(1-\theta\right)b\right)}{1+\frac{\partial\log Q}{\partial\log b'}\cdot\frac{Q}{b'}\left(b'-\left(1-\theta\right)b\right)/Q}\\&=\frac{1-\frac{\partial\log Q}{\partial\log k'}\cdot\frac{Q\cdot b'}{k'}\frac{\left(b'-\left(1-\theta\right)b\right)}{b'}}{1+\frac{\partial\log Q}{\partial\log b'}\cdot\frac{\left(b'-\left(1-\theta\right)b\right)}{b'}}\\&=\frac{1-\gamma\cdot\frac{Q\cdot b'}{k'}\cdot\frac{\partial\log Q}{\partial\log k'}}{1+\gamma\cdot\frac{\partial\log Q}{\partial\log b'}}\end{aligned}$$

where $\gamma:=\frac{\left(b'-\left(1-\theta\right)b\right)}{b'}$. This completes the proof. □

Proof of Proposition 3.

$$1=\sum_{t=1}^{T}\left(\frac{P}{1+\rho}\right)^t\left[r+\frac{\left(1-P\right)}{P}\left(1-LGD\right)\right]+\left(\frac{P}{1+\rho}\right)^T$$

Let $x=\frac{P}{1+\rho}$ so

$$1=\left(r+\frac{1-P}{P}\left(1-LGD\right)\right)\frac{x}{1-x}\left(1-x^T\right)+x^T$$

Guess that $1+\rho=(1+r)\,P+(1-P)\,(1-LGD)$

$$\frac{1-x}{x}=\frac{1}{x}-1=\frac{(1+r)\,P+(1-P)\,(1-LGD)}{P}-1=r+\frac{1-P}{P}\,(1-LGD)$$

And, therefore

$$1=1\left(1-x^{T}\right)+x^{T}$$

which validates the guess. □

Proof of Proposition 4. Rearranging Equations 1 and 2, we have

$$\begin{aligned}
1+\rho &= \frac{\mathbb{E}\left[\mathcal{P}_{t+1}\left(\theta+(1-\theta)Q_{t+1}\right)+\left(1-\mathcal{P}_{t+1}\right)\frac{\phi(k_{t+1})}{b_{t+1}}\middle| k_{t+1},b_{t+1},z_t\right]}{Q_t}\\
&= 1+r_t^{firm}+\mathbb{E}\left[\left(1-\mathcal{P}_{t+1}\right)\frac{\phi(k_{t+1})}{Q_t\cdot b_{t+1}}\middle| k_{t+1},b_{t+1},z_t\right]\\
&= 1+r_t^{firm}+(1-P)\cdot(1-LGD)
\end{aligned}$$

with the last line using our formula for LGD at origination.

□

Proof of Proposition 5. We start with Equation 7, then we use the fact that, by assumption, $\phi'\left(k_{t+1}\right)=\phi\left(k_{t+1}\right)/k_{t+1}$, and then we plug in the definitions of lev and LGD. This yields:

$$\begin{aligned}
1+r^{social} &= \left(1+r^{firm}\right)\mathcal{M}+(1-P)\cdot\phi'\left(k_{t+1}\right)\\
&= \left(1+r^{firm}\right)\mathcal{M}+(1-P)\cdot\frac{\phi\left(k_{t+1}\right)}{k_{t+1}}\cdot\frac{Q_t b_{t+1}}{Q_t b_{t+1}}\\
&= \left(1+r^{firm}\right)\mathcal{M}+(1-P)\cdot(1-LGD)\cdot lev
\end{aligned}$$

Plugging in our formula for r^{firm} from Proposition 4 yields

$$
\begin{aligned}
1 + r^{social} &= (1 + \rho - (1 - P)(1 - LGD))\,\mathcal{M} + (1 - P) \cdot (1 - LGD) \cdot lev \\
&= (1 + \rho)\,\mathcal{M} + (lev - \mathcal{M}) \cdot (1 - P)(1 - LGD)
\end{aligned}
$$

□

B Data

B.1 Details on Data Cleaning and Construction

While the FR Y-14Q Schedule H.1 data goes back to 2011, we keep only data from 2014Q4 due to data quality and consistency of reporting issues.

Borrowers. We drop all loans to borrowers without a Tax Identification Number. We keep only Commercial & Industrial loans to nonfinancial U.S. addresses, i.e. lines reported on FR Y-9C equal to 3, 4, 8, 9, and 10. We drop all borrowers with NAICS codes 52 (Finance and Insurance), 92 (Public Administration), 5312 (Offices of Real Estate Agents and Brokers), and 551111 (Offices of Bank Holding Companies), as some financial companies are classified under the later two NAICS codes in our sample.

Loans. We drop all loans with a negative committed exposure, or for which the utilized exposure exceeds the committed exposure as these are likely to be mistakes. We drop all observations for which the origination date exceeds the current date, and all those for which the maturity date precedes the current date.

We keep only "vanilla" term loans (Facility type equal to 7), and we thus exclude Type A, B, and C term loans, as well as bridge term loans. We keep only loans that are classified as fixed or variable rate, and drop mixed interest rate variability loans. We keep only loans with maturity between 1 and 10 years, thus excluding very short-term and very long-term loans. We keep only loans with interest rates in the 1st-99th percentiles for fixed rate loans, and spread in the 1st-99th percentiles for variable rate loans, as some of the very high and low rates/spreads are likely to be data errors. Additionally, we drop loans with interest rates higher than 50% at origination. We also drop loans for which the probability of default and the loss given default are either missing or outside of the $[0, 1]$ intervals. We also drop loans for which the probability of default is equal to 1, as that is an indicator that the loan is in default.

B.2 Estimating $\mathcal{M}$

In this section, we argue that our calibration of $\mathcal{M} = 1$ is a good approximation. To this end, we provide estimates of this object in the data. Recall that this object was defined as

$$\mathcal{M}_t := \frac{1 - \gamma \times \frac{Qb'}{k'} \times \frac{\partial \log Q}{\partial \log k'}}{1 + \gamma \times \times \frac{\partial \log Q}{\partial \log b'}}$$

Given estimates for the function Q, γ, and firm leverage Qb'/k' we can compute $\mathcal{M}$ for every observation (loan origination) in our data. The main challenge is to estimate Q as a function of firm borrowing and investment. This function can either be obtained by solving a calibrated version of our model, or estimated non-parametrically in the data. In this subsection, we present results for the latter approach.

First, we compute Q for every loan origination in the data. In a model setting such as ours, where loans are modeled as perpetuities that decay at a geometric rate θ, we can write Q as the present value of all future payments, discounted at the contractual interest rate r:

$$Q = \frac{\theta + (1 - \theta)Q}{1 + r} = \frac{\theta}{r + \theta}$$

r is directly observed in the data, and we can apply the common approximation that θ is equal to the inverse of the loan maturity, $\theta = 1/T$. This allows us to compute Q for every loan origination in the data.

The model establishes that Q is a function of firm investment k', firm borrowing b', as well as the current level of productivity z. Additionally, Q should also depend on the lender's cost of capital ρ. We therefore approximate (the log of) Q as a polynomial of these four variables. We measure firm investment as (the log of) tangible assets at loan origination, firm borrowing as (the log of) total debt owed by the firm at loan origination, firm productivity as the (the log of) sales over tangible assets (a measure of TFPR following Hsieh and Klenow (2009)). The lender's cost of capital ρ is measured as in the main text. We therefore estimate:

$$
\begin{aligned}
\log Q_i = \alpha + \beta_k \log k_i + \beta_b \log b_i + \beta_z \log z_i + \beta_\rho \rho_i \\
+\beta_{k,k}(\log k_i)^2 + \beta_{k,b} \log k_i \times \log b_i + \beta_{k,z} \log k_i \times \log z_i + \beta_{k,\rho} \log k_i \times \rho_i \\
+\beta_{b,b}(\log b_i)^2 \beta_{b,z} \log b_i \times \log z_i + \beta_{b,\rho} \log b_i \times \rho_i \\
+\beta_{z,z}(\log z_i)^2 \beta_{z,\rho} \log z_i \times \rho_i + \beta_{\rho,\rho}(\rho_i)^2 \epsilon_i
\end{aligned}
$$

The resulting estimates can be used to compute the partial derivatives of $\log Q$ with respect to investment and borrowing. Qb'/k' is measured in a consistent manner, as the sum of total liabilities plus new borrowings divided by total assets plus new borrowings. Finally, we take advantage of the fact that at the steady state, $\gamma = \theta = 1/T$.

Figure 7 presents the histogram for the estimated $\mathcal{M}_i$ in our sample. Clearly, the distribution is extremely concentrated around 1. The mean is equal to 0.996 and the median to 0.997, with a standard deviation of 0.006. Figure 8 replicates our measure of misallocation, when computed accounting for heterogeneity in $\mathcal{M}$, and compares it to our baseline, showing that the two measures are extremely similar, both in terms of magnitudes and dynamics. Taken together, these results suggest that our assumption that $\mathcal{M} = 1$ is a good one.

B.3 Cross-sectional Heterogeneity

We also explore the correlation between the cost of capital and firm-level covariates. We regress $\log(1+r)$ separately on log leverage, log return on assets, and log assets. We conduct this analysis for interest rates, ρ, r^{firm}, and r^{social}. The results are shown in Table 4. Of the three covariates, the best predictor is the return on assets; interest rates and the cost of capital are consistently higher at firms with high return on assets. Although we cannot attach a causal interpretation to the estimated coefficients, this would be consistent with a

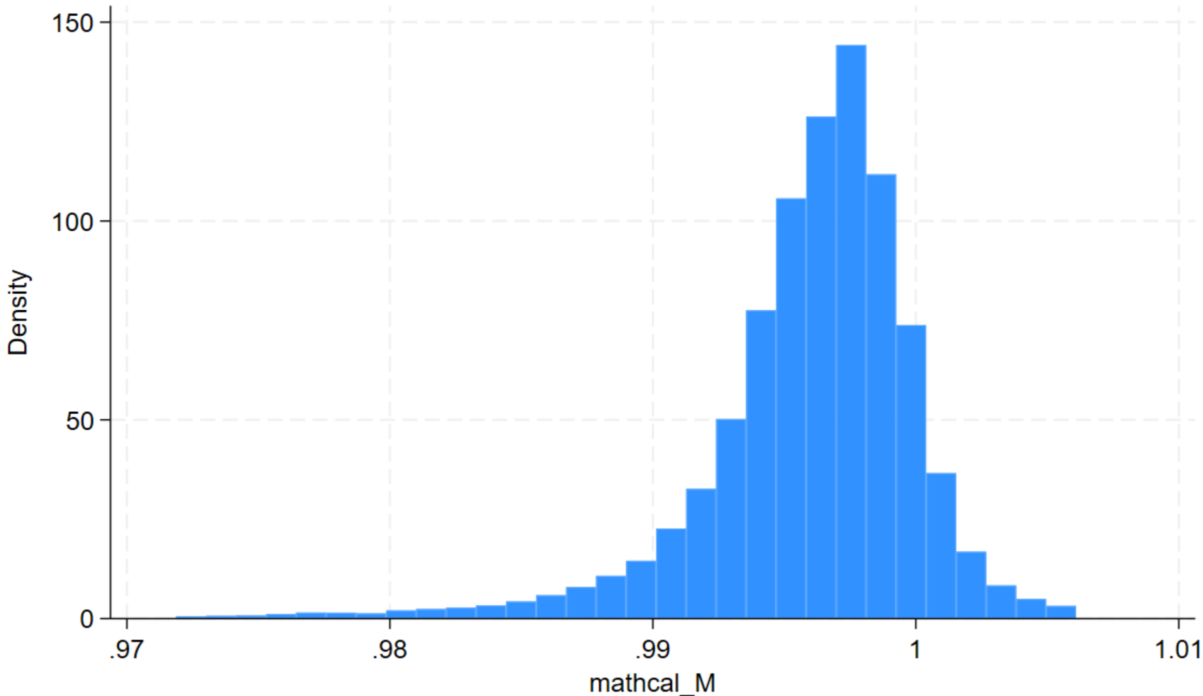

Figure 7: Histogram for estimated $\mathcal{M}_i$

model where causality runs from the cost of capital to firm decisions: firms with a higher cost of capital will demand a high return on their investments. Yet perhaps more notable is the very low R^2. The return on assets explains between 2 and 3% of the variance, depending on the measure of the cost of capital, with other covariates explaining less than 1%. Firm-level covariates explain approximately none of the variance in the cost of capital.

Table 4: Determinants of Capital Costs and Spreads

Panel A: Contractual Rate			
	(1)	(2)	(3)
log leverage	0.048***		
	(0.00)		
log roa		0.134***	
		(0.02)	
log assets			-0.094***
			(0.00)
Observations	62687	1665	60437
Adjusted R^2	0.002	0.017	0.009
Panel B: Lender Discount Rate			
	(1)	(2)	(3)
log leverage	-0.026***		
	(0.00)		
log roa		0.146***	
		(0.03)	
log assets			-0.007
			(0.00)
Observations	62687	1665	60437
Adjusted R^2	0.001	0.021	0.000
Panel C: Firm's Cost of Capital			
	(1)	(2)	(3)
log leverage	-0.086***		
	(0.00)		
log roa		0.174***	
		(0.02)	
log assets			0.071***
			(0.00)
Observations	62687	1665	60437
Adjusted R^2	0.007	0.030	0.005
Panel D: Social Cost of Capital			
	(1)	(2)	(3)
log leverage	0.073***		
	(0.01)		
log roa		0.146***	
		(0.03)	
log assets			-0.009**
			(0.00)
Observations	62687	1665	60437
Adjusted R^2	0.005	0.021	0.000

Standardized beta coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

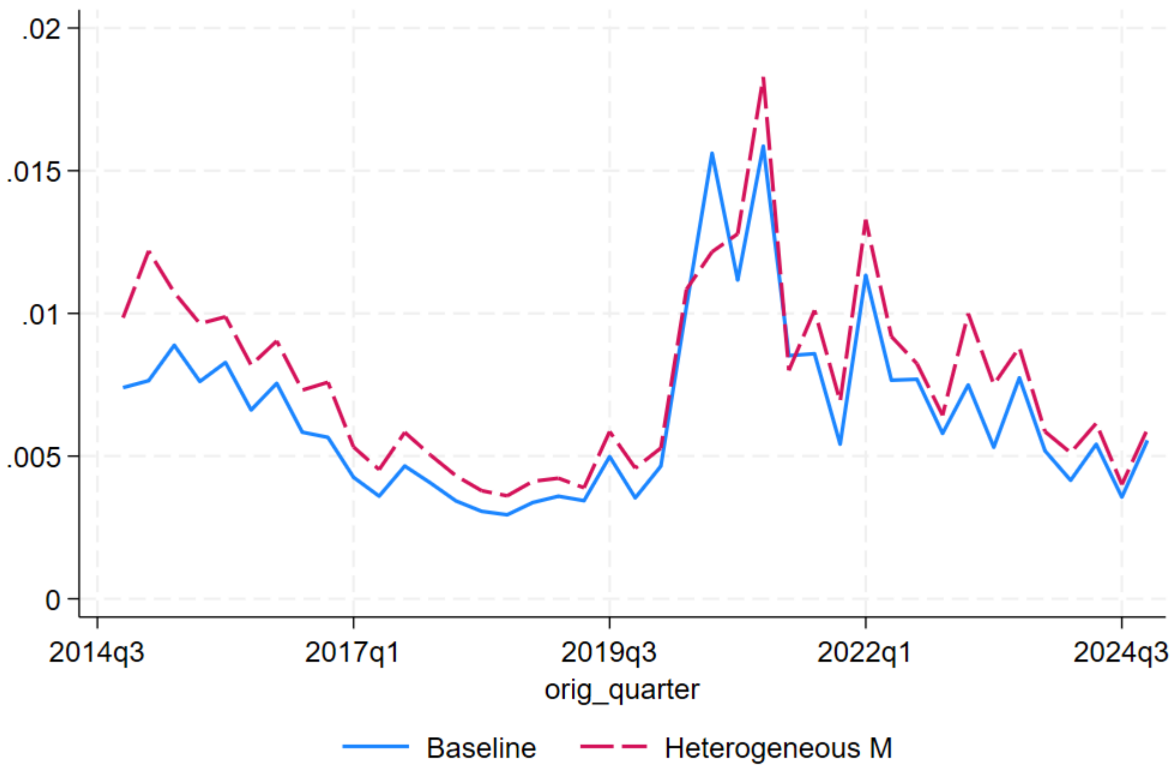

Figure 8: Misallocation measure with $\mathcal{M} = 1$ vs. estimated $\mathcal{M}_i$

B.4 Misallocation weighted by loan size

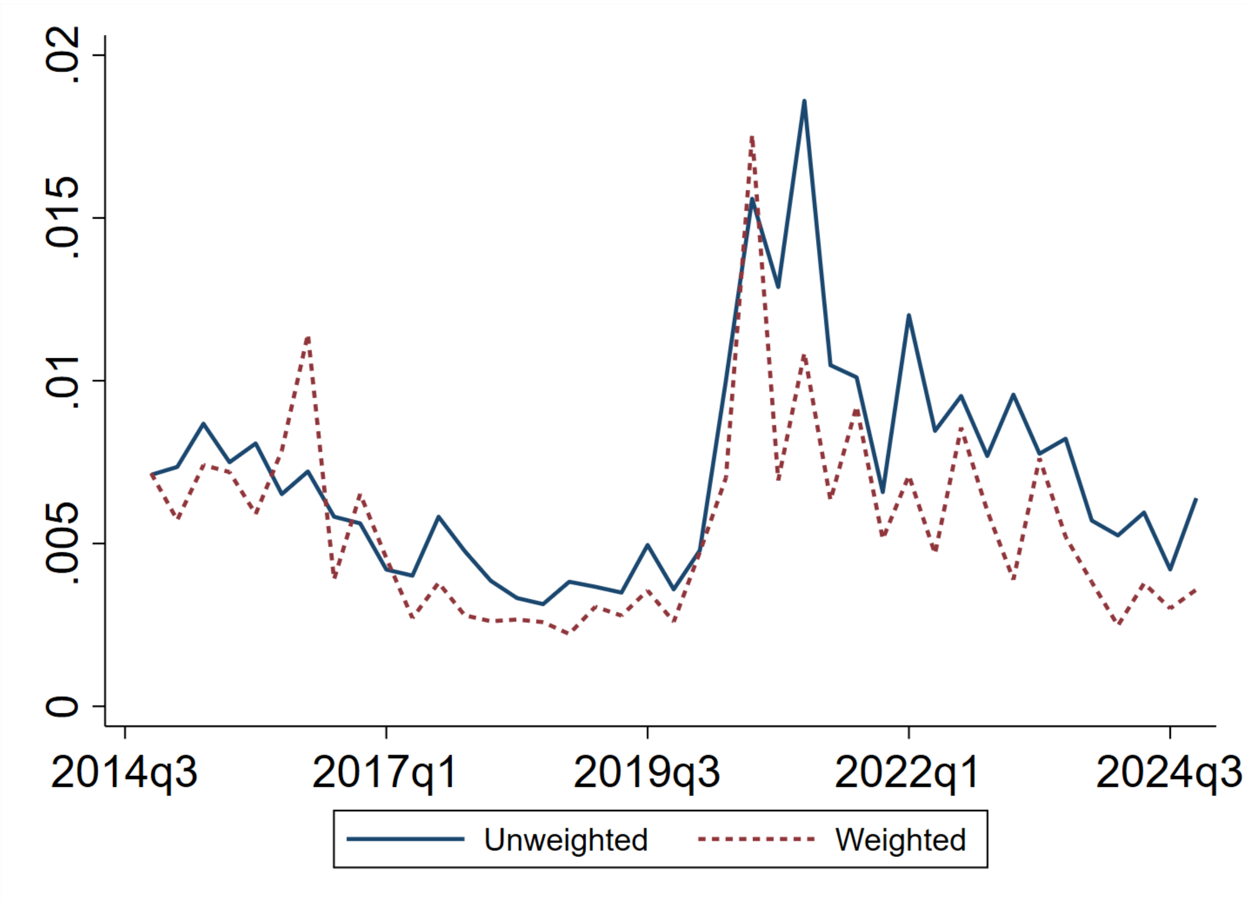

Figure 9: Misallocation, unweighted and weighted by loan size

B.5 Robustness: Results for Fixed-Rate Five-Year Loans

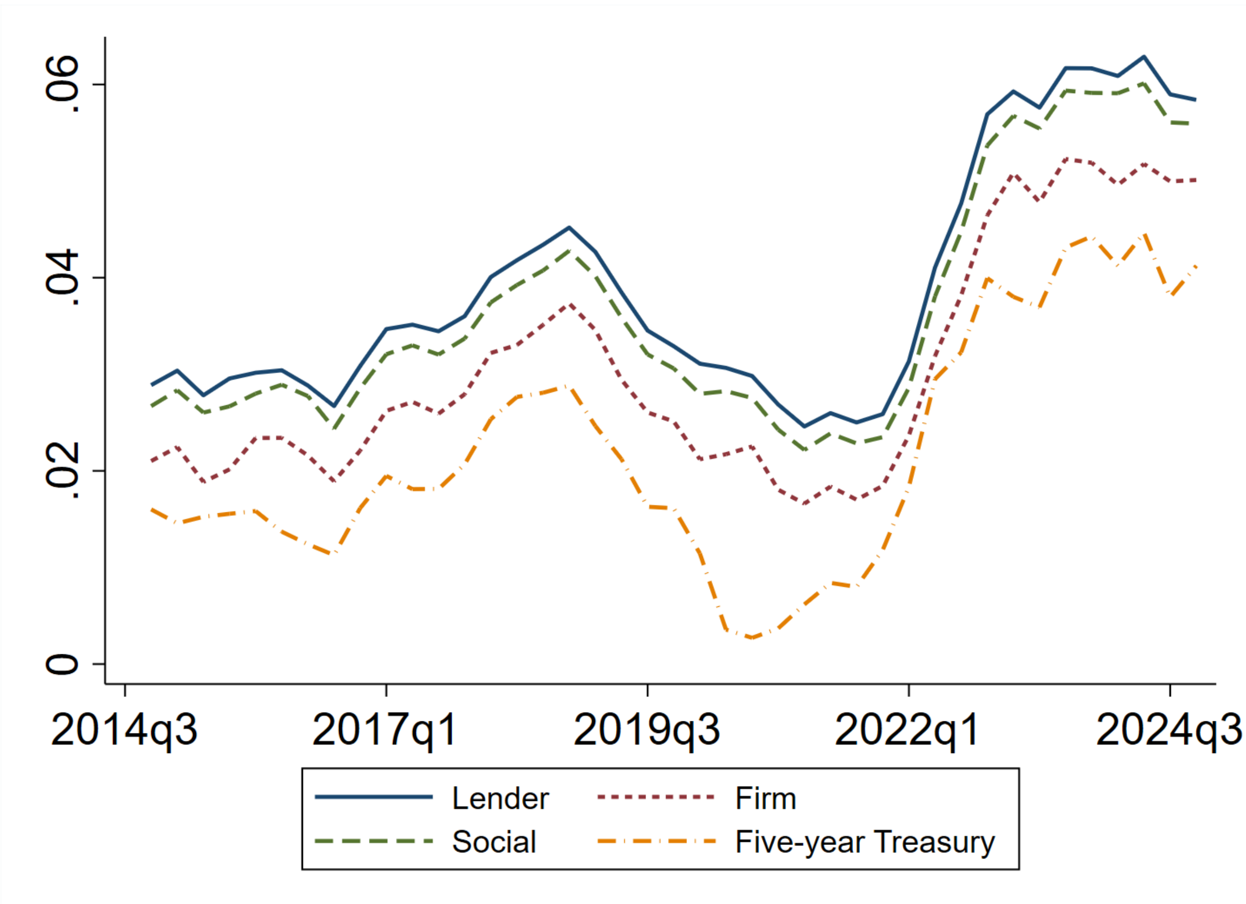

Figure 10: Averages by Quarter of Origination (Fixed-Rate Five-Year Sample)

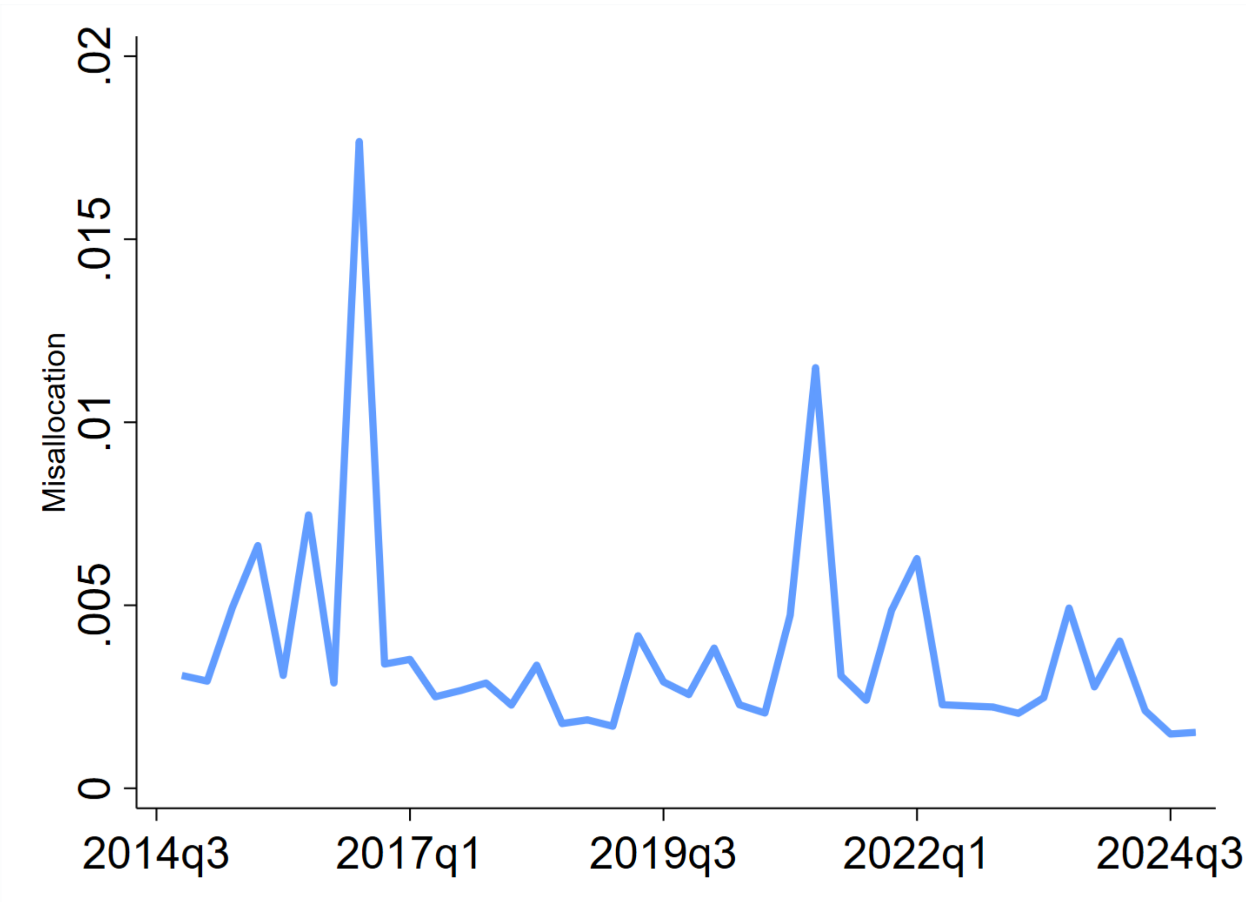

Figure 11: Cost of Misallocation (Fixed-Rate Five-Year Sample)

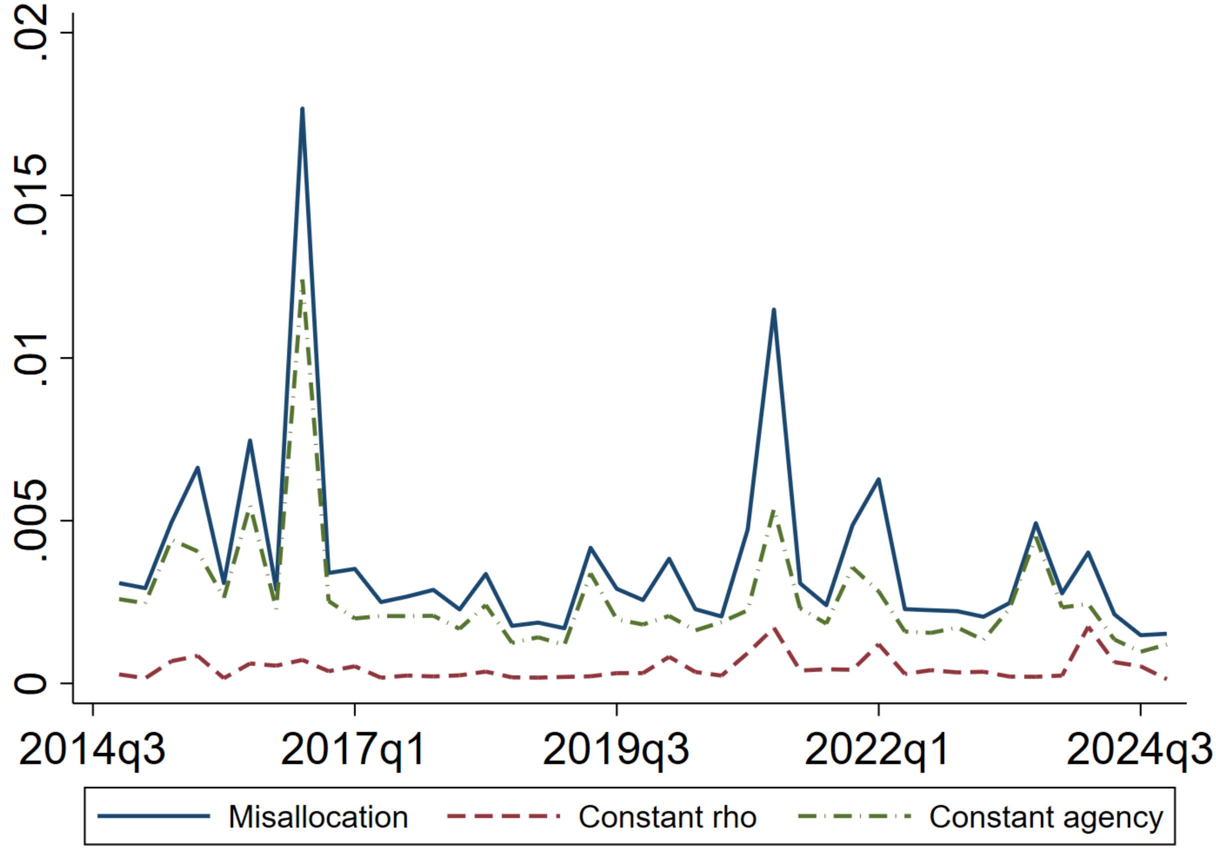

Figure 12: Decomposing Misallocation (Fixed-Rate Five-Year Sample)