The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXVI, NO. 5 • OCTOBER 2021

The Misallocation of Finance

TONI M. WHITED and JAKE ZHAO

ABSTRACT

We estimate real losses arising from the cross-sectional misallocation of financial liabilities. Extending a production-based framework of misallocation measurement to the liabilities side of the balance sheet and using manufacturing firm data from the United States and China, we find significant misallocation of debt and equity in China but not the United States. Reallocating liabilities of firms in China to mimic U.S. efficiency would produce gains of 51% to 69% in real value-added, with only 17% to 21% stemming from inefficient debt-equity combinations. For Chinese firms that are large or in developed cities, we estimate lower distortionary financing costs.

OVER A DECADE OF RESEARCH IN INDUSTRIAL organization, development, and macroeconomics has provided convincing evidence that cross-sectional misallocation of capital and labor is significant and can help explain why developing countries have lower total factor productivity (TFP).[1] Such pervasive evidence of factor misallocation begs the question of whether the financial liabilities that back the funding of capital goods and payroll are also misallocated. Are the right firms getting the right amount of finance, and is the mix of debt and equity optimal? In this paper, we tackle these two related questions, moving from the asset side of the balance sheet to the liability side to quantify the extent of the misallocation of finance.

Toni M. Whited is from the University of Michigan and the NBER. Jake Zhao is from Peking University HSBC Business School. We thank the editor; Amit Seru; an associate editor; three referees; Hengjie Ai; Joel David; Eduardo Dávila; Murray Frank; Arthur Korteweg; Xiaoji Lin; Andrey Malenko; Uday Rajan; Jiao Shi; Michael Song; Qi Sun; Stephen Terry; Xu Tian; James Weston; Yufeng Wu; Nan Xiong; Shaojun Zhang; Stefan Zeume; and participants at the Advances in Macro-Finance Tepper/LAEF Conference, the Washington University Corporate Finance Conference, the North American Summer Meetings of the Econometric Society, the China International Conference in Finance, the Fifth Symposium on Emerging Financial Markets, the Summer Institute of Finance Conference, the China Financial Research Conference, the Midwest Finance Association Meetings, and the American Finance Association Meetings for helpful comments. Both authors have read *The Journal of Finance* disclosure policy and have no conflicts of interest to disclose.

Correspondence: Toni M. Whited, Ross School of Business, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109; e-mail: twhited@umich.edu.

[1] Banerjee and Duflo (2005) offer an overview of the misallocation hypothesis in the development literature while Syverson (2011) and Restuccia and Rogerson (2013) survey the literature from an industrial organization and a macroeconomics perspective, respectively.

DOI: 10.1111/jofi.13031

To this end, we build on the empirical framework of Hsieh and Klenow (2009), so a brief outline of their approach helps clarify our own. They base their empirical work on a static model of cross-sectional factor allocation in which firms have constant returns-to-scale technology, are monopolistically competitive, and as a result face downward-sloping demand, which endogenously limits their size. At an optimal allocation, the marginal revenue products of each factor are equal across firms in an industry. Distortions in cross-sectional allocations break this equality and adversely affect TFP. The greater the dispersion in factor marginal revenue products within a sector, the greater the divergence of actual TFP from its potential level and thus the greater the potential gains from reallocating debt and equity to the firms that can use these funds most efficiently. Using establishment-level data on manufacturing sectors in the United States, China, and India, Hsieh and Klenow (2009) find that China and India could realize TFP gains of 30% to 50% and 40% to 60%, respectively, if these countries hypothetically reallocated their factors of production to achieve the U.S. level of efficiency.

Our model is analogous to the setup in Hsieh and Klenow (2009), with two differences. First, while they model the factor mix that leads to distortions in TFP, we model the financial liabilities that underlie these factors and potentially contribute to TFP distortions. In particular, we specify different types of financial liabilities as the primitive inputs into the production process. Although different forms of finance are not exactly equivalent to factors of production, our modeling strategy is reasonable in the sense that firms ultimately finance their purchases of factors of production using debt and equity. The proximate factors—capital, materials, labor, and energy—can then be thought of as unmodeled intermediate inputs.

Second, while Hsieh and Klenow (2009) specify capital and labor as imperfect substitutes, we extend their framework by allowing different forms of finance to be either perfect or imperfect substitutes. This extra flexibility in our model is important because it allows for a frictionless Modigliani-Miller (MM) world as a baseline. This flexibility also allows us to understand whether potential reallocation gains stem from rearranging the mix of financial liabilities across firms or from moving the gross flow of resources from less efficient to more efficient firms. This second avenue for reallocation is available even if debt and equity are perfectly substitutable, and as such likely captures the real misallocation in Hsieh and Klenow (2009). In addition, we show that within a class of well-specified dynamic capital structure models, the method we propose not only offers a good estimate of the misallocation of financial liabilities but also approximates the real misallocation of Hsieh and Klenow (2009).

In our framework, at an optimal social planner's allocation, the marginal contributions of debt and equity finance to nominal value-added are equal across firms in a sector. Empirically, we infer distortions by observing deviations from this first-best allocation. These deviations manifest as large differences (relative to our model) in debt-equity ratios across firms in a sector. Because distortions in these allocations lower productivity, these large differences imply poorly developed financial markets and large potential gains from the reallocation of finance.

Using data on manufacturing firms in the United States and China, we find significant misallocation of debt and equity. Although financial liabilities appear to be well allocated in the United States, this is not the case in China. If China's debt and equity markets were as developed as those in the United States, gains of 51% to 69% in real firm value-added would be available. Interestingly, 79% to 83% of these gains come from the misallocation of the total amount of finance, that is, from a suboptimal matching of firm productivity with resources. The remaining 17% to 21% of misallocation stems from the misallocation of the type of finance. This is the central result in our paper, as it implies that the problem of a mismatch between firm productivity and firm size is more important for observed misallocation than the problem of an inefficient mix of liability type.

We produce several further interesting results. First, our framework allows us to estimate the distorted costs of debt and equity for each firm. We analyze the cross-sectional patterns in these costs in China and find that larger firms and firms located in more developed cities face markedly lower costs. Second, we find that our results are not concentrated in the large number of extremely small Chinese firms in our sample. In particular, we find significant potential reallocation gains even when we compare Chinese and U.S. firms of similar size. Although we find smaller gains, the reduction in potential gains comes from eliminating extremely large U.S. firms from our sample and not from eliminating the extremely small Chinese firms. Third, our results are robust to including labor income in our model.

Because our strategy of modeling financial liabilities as factor inputs is unusual, it is worth outlining classes of models and types of market forces that might plausibly motivate our specification in which dispersion in leverage ratios reflects frictions or inefficiencies. One model that clearly fails to produce such an association is the frictionless world of Modigliani and Miller (1958), where leverage irrelevancy implies large dispersion in observed leverage ratios. However, a positive association between leverage dispersion and frictions arises in models in which frictions affect not only leverage but also firm value or real outcomes. For example, in contingent claims models such as Fischer, Heinkel, and Zechner (1989) or Goldstein, Ju, and Leland (2001), higher debt issuance costs induce wider optimal inaction bounds for leverage ratios and thus higher observed leverage variation. Similarly, we show that our method successfully detects frictions that are present in a dynamic model of optimal investment and leverage in the spirit of Hennessy and Whited (2005). Similar results are likely in the closely related set of dynamic contracting models based on limited commitment (Rampini and Viswanathan (2013), Li, Whited, and Wu (2016)).

These classes of models can incorporate various financial frictions that spill over to a firm's real decisions and that consequently lead to decreasing real marginal benefits of any particular type of finance. First, many capital structure models feature costs of financial distress that accompany too much debt, and many of these costs are real. For example, given limited liability, firms with too much debt have incentives to undertake excessively risky projects (Jensen and Meckling (1976)), as equity holders bear no downside risk. Also, excessive leverage can cause important employees to jump ship or deter potential workers from seeking employment (Brown and Matsa (2016)). Second, too much debt can result in debt overhang, whereby managers forgo profitable investment projects because too little of the project proceeds would accrue to shareholders (Myers (1977)). Third, monitoring frictions in the lending market can lead to covenants that suboptimally constrain real investment and acquisition decisions. Fourth, and in contrast, agency problems that lead managers to engage in wasteful spending can accompany too little debt (Jensen (1986)). Fifth, enforcement frictions lead to financial contracts in which external financing needs to be collateralized (Rampini and Viswanathan (2010)). Finally, adverse selection (Stiglitz and Weiss (1981), Myers and Majluf (1984)) can cause financial markets to break down. In this case, the strict rationing of credit would spill over into inefficient purchases of proximate factors of production.

Ultimately, these financial frictions deliver misallocation because they affect real outcomes. These frictions motivate two important insights: Different forms of finance are not equivalent, and any form of finance has a decreasing marginal benefit. An optimal mix of debt and equity mitigates these frictions to the greatest extent possible. Failure to minimize these frictions results in an inefficient allocation of financial liabilities, thereby reducing real output or firm value.

Given the unusual nature of our empirical strategy, we emphasize an important advantage of this approach. It offers a tractable alternative to dynamic equilibrium models for measuring the value to society of a well-functioning financial system. In the context of dynamic equilibrium models, this task would require specification of all quantitatively relevant financial frictions, which would result in an excessively complicated model that is difficult to interpret. Alternatively, if the goal is simply to measure the extent of misallocation, instead of to understand its various primitive sources, we think our simple approach has value because we do not have to take a stance on which frictions drive misallocation.

The literature on factor misallocation is extensive.[2] Within this body of work, our paper is most similar to Buera, Kaboski, and Shin (2011) and Midrigan

[2] Early work that provides theoretical underpinnings for misallocation includes Lucas (1978), Hopenhayn (1992), Hopenhayn and Rogerson (1993), Olley and Pakes (1996), and Cooley and Quadrini (2001). More recently, several studies use firm- or establishment-level data and heterogeneous firm models to investigate the quantitative importance of misallocation. A partial list of more recent papers includes Jeong and Townsend (2007), Foster, Haltiwanger, and Syverson (2008), Restuccia and Rogerson (2008), Alfaro, Charlton, and Kanczuk (2009), Hsieh and Klenow (2009), Banerjee and Moll (2010), Buera, Kaboski, and Shin (2011), Song, Storesletten, and Zilibotti (2011), Levinsohn and Petrin (2012), Bartelsman, Haltiwanger, and Scarpetta (2013), Chen and Song (2013), Midrigan and Xu (2014), Asker, Collard-Wexler, and Loecker (2014), Hsieh and Klenow (2014), Song and Wu (2015), Kehrig and Vincent (2016), Curtis (2016), Ai, Li, and Yang (2016), Bai, Lu, and Tian (2016), David and Venkateswaran (2016), and Wu (2018).

and Xu (2014), who also consider financial frictions. However, there are two substantive differences between our paper and these studies. First, while both Buera, Kaboski, and Shin (2011) and Midrigan and Xu (2014) emphasize the extensive margin of misallocation across sectors of the economy, we emphasize within-sector misallocation, in line with Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Because more developed financial markets can indeed cause new firms to enter, our analysis provides a lower bound on the extent of financial misallocation that a dynamic model with entry and exit might find.

Second, both Buera, Kaboski, and Shin (2011) and Midrigan and Xu (2014) feature calibrated dynamic models, whereas our approach is largely empirical. For example, Buera, Kaboski, and Shin (2011) study a model in which financial frictions affect the manufacturing sector primarily on the extensive margin, as these frictions prevent talented agents from entering this sector. In Midrigan and Xu (2014), financial frictions lead to little intensive misallocation but substantial misallocation across sectors because the more productive sector requires a cost of entry that is difficult to pay in the face of financial frictions.

In the finance literature, our work is related to Graham (2000), who also considers the cross-sectional allocation of debt and equity. However, there are again substantive differences between our work and his. Graham (2000) computes firm-level estimates of the point at which the marginal tax benefits of debt begin to decline. A firm that incurs interest deductions to the left of this "kink" has an inefficiently low level of debt. Estimates of this inefficiency imply large amounts of tax benefits left on the table by underleveraged firms. One notable feature of the framework in Graham (2000) is that he takes relative prices as given and interprets deviations from the optimal responses to these prices as suboptimal. In contrast, we assume that firms behave rationally and use our framework to back out the price distortions that lead to the financing decisions that we observe in the data. This alternative perspective seems reasonable in light of the finding in Blouin, Core, and Guay (2010) that the marginal tax rate estimates of Graham (2000) imply rational behavior when the kink points are derived from more accurate estimates of future taxable income.

The rest of the paper is organized as follows. Section I outlines the model and explains the empirical framework for measuring misallocation. Section II describes the U.S. and Chinese data. Section III describes the estimation of the model parameters. Section IV presents our baseline empirical results. Section V examines the robustness of our results to several assumptions in our baseline model, while Section VI considers alternative models. Section VII concludes. The Appendix contains a full derivation of the model.

I. Model

This section sketches our model. We closely follow Hsieh and Klenow (2009), who develop is a closed-economy version of the model in Melitz (2003). We start with a description of the environment and technology and a statement of the optimality conditions. We then discuss how to measure the benefits of reallocation. A full derivation of the model is in the Appendix.

A. Environment and Technology

Firms in our model are financed by debt and equity. In our model, we do not distinguish between external and internal equity. Given the rarity of seasoned equity offerings (DeAngelo, DeAngelo, and Stulz (2010)), and given that external equity constitutes a negligible source of funds over the last two decades in the Federal Reserve's Flow of Funds data, we view this simplification as innocuous for our purposes, which are primarily empirical.

Firms use the proceeds from issuing these financial assets to purchase proximate factors of production that generate a real benefit to shareholders. Because our economy is static in nature, in nearly all of our analysis we represent this benefit as one-period real value-added. However, to maintain generality, we use the term "real benefit," with the term "nominal benefit" then representing the real benefit times the corresponding price.[3]

We denote the total real benefit of finance (value-added) in the economy by F. We assume that the economy consists of S sectors, and hence for each sector s the real benefit is given by F_s. Using a Cobb-Douglas aggregator we obtain the real benefit across sectors as follows:

$$F = \prod_{s=1}^{S} F_s^{\theta_s}, \quad \text{with} \quad \sum_{s=1}^{S} \theta_s = 1. \tag{1}$$

The Cobb-Douglas aggregator in (1) implies that increasing the size of any particular sector while holding the others constant has a decreasing marginal benefit.

Next, we assume that the real benefit in each sector, F_s, comes from a constant elasticity of substitution (CES) aggregate of the real benefit created by I differentiated firms, that is,

$$F_s = \left(\sum_{i=1}^{I} F_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where F_{si} is the real benefit of firm i and σ is the elasticity of substitution of the real benefit between firms in a sector. As in Dixit and Stiglitz (1977), a finite elasticity of substitution implies monopolistically competitive behavior in the product market.

Finally, we assume that within an individual firm, debt and equity finance can be aggregated into the real benefit of finance using a CES function,

[3] Korteweg (2010) uses similar terminology.

according to

$$F_{si} = A_{si}\left(\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right)^{\frac{\gamma_s}{\gamma_s-1}}. \tag{3}$$

Here, D_{si} and E_{si} are the levels of debt and equity for firm i in sector s. In equation (3), $\alpha_s \in (0, 1)$ denotes the weight on the importance of debt in generating the real benefit, γ_s is the elasticity of substitution between debt and equity, and A_{si} denotes the total factor benefit (TFB), which is directly analogous to TFP. Note that certain variables are firm-specific, while others are sector-specific. For example, TFB, A_{si}, depends on both the sector and firm, while the weight, α_s, and the elasticity of substitution, γ_s, depend only on the sector. An important feature of (3) is imperfect substitutability between debt and equity, which results in decreasing marginal returns to the use of each of these forms of finance.

Our use of a CES aggregator is an important departure from Hsieh and Klenow (2009), who use a Cobb-Douglas production function. A CES functional form allows for perfect substitutability between different forms of finance and, as we show below, gives us the flexibility to distinguish between reallocation gains that come from the type of finance and those that come from the gross flow of finance. Nonetheless, equation (3) constitutes a strong functional-form assumption about how debt and equity are ultimately transformed into the real benefit of finance, as our primary measure of this benefit is value-added. Specifically, (3) describes the generation of value-added without explicitly showing how the proceeds raised through financing activities translate into capital, labor, or other factors of production. Thus, because firms ultimately finance their purchases of all factors of production using debt and equity, more immediate factors such as capital and labor can be thought of as unmodeled intermediate inputs.

B. Optimal Allocations

Next, we define the prices that enter the firm's optimization problem. We let r_{si} and λ_{si} be the costs associated with using debt and equity, respectively. The costs of debt and equity, r_{si} and λ_{si}, can vary both across sectors and across firms within a sector. These costs are also functions of the amounts of debt and equity. For our purposes, we do not have to specify a functional form for this relation and thus we omit this additional notation to be concise.

To the extent that financial market frictions distort these costs, we also need to define reduced-form cost distortions, which we refer to as distortions or wedges. Specifically, $\tau_{D_{si}}$ is a wedge embedded in the total cost of debt and $\tau_{E_{si}}$ is a wedge embedded in the total cost of equity. Positive values indicate that firms face additional costs of finance. As noted in the introduction, these costs can arise from frictions such as informational asymmetry or agency problems. Negative values, in contrast, suggest favorable financial relationships or government subsidies. Because $\tau_{D_{si}}$ and $\tau_{E_{si}}$ can vary across firms, we think of these wedges as measuring frictions that are idiosyncratic to the firm, such as political connections or manager-specific agency frictions. As our model is ultimately an empirical measurement framework, we do not model the explicit mechanisms behind these distortions and instead assume that they are well-encapsulated by $\tau_{D_{si}}$ and $\tau_{E_{si}}$.

Given these definitions, we now specify nominal profit, π_{si}, as

$$\pi_{si} = P_{si}F_{si} - \left[\left(1+\tau_{D_{si}}\right)r_{si}D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si}E_{si}\right]. \tag{4}$$

Because the term in square brackets on the right-hand side of equation (4) is the cost of capital, π_{si} can be interpreted as economic value-added (EVA), which is a sensible quantity to maximize in a static model. The firm maximizes π_{si} by choosing P_{si}, D_{si}, and E_{si}, taking $\tau_{D_{si}}$ and $\tau_{E_{si}}$ as given. Here. P_{si} is a choice variable because of the assumption of imperfect substitutability of the real benefit, F_{si}, across firms in a sector that is embedded in (2).

To maximize (4), the firm first chooses D_{si} and E_{si} to minimize the cost of capital, $(1+\tau_{D_{si}})r_{si}D_{si} + (1+\tau_{E_{si}})\lambda_{si}E_{si}$, subject to setting equation (3) equal to a fixed level, $\bar{F}_{si}$. Intuitively, the first-order conditions for optimal financing show that the marginal net benefits of debt and equity should be equal. As shown in the Appendix, this equality implies that the solution for the optimal ratio of debt to equity is given by

$$Z_{si} \equiv \frac{D_{si}}{E_{si}} = \left[\frac{\alpha_s}{1-\alpha_s}\frac{\frac{\partial}{\partial E_{si}}\left[\left(1+\tau_{D_{si}}\right)r_{si}D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si}E_{si}\right]}{\frac{\partial}{\partial D_{si}}\left[\left(1+\tau_{D_{si}}\right)r_{si}D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si}E_{si}\right]}\right]^{\gamma_s}. \tag{5}$$

Next, the firm chooses P_{si} to maximize π_{si}. As shown in the Appendix, the solution to this second problem is given by

$$P_{si} = \frac{\sigma}{\sigma-1}\left[\frac{1}{A_{si}}\left(\left(1+\tau_{D_{si}}\right)r_{si}\left(\alpha_s + (1-\alpha_s)Z_{si}^{-\frac{\gamma_s-1}{\gamma_s}}\right)^{-\frac{\gamma_s}{\gamma_s-1}} + \left(1+\tau_{E_{si}}\right)\lambda_{si}\left(\alpha_s Z_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)\right)^{-\frac{\gamma_s}{\gamma_s-1}}\right)\right]. \tag{6}$$

Because $Z_{si} \equiv D_{si}/E_{si}$ is an optimal allocation, by comparing equation (4) to the term in square brackets in equation (6), it is clear that this latter term is the minimized marginal cost of providing one unit of the marginal benefit of finance. Thus, equation (6) naturally shows that price is a markup over marginal cost, with the term $\sigma/(\sigma-1)$ being the markup.

Next, we solve for the sector price P_s as a function of the firm price P_{si} by defining P_s to be the minimum price of acquiring a unit of the sector real benefit. The solution is

$$P_s = \left(\sum_{i=1}^{I} P_{si}^{-(\sigma-1)}\right)^{-\frac{1}{\sigma-1}}. \tag{7}$$

Finally, cost minimization of the Cobb-Douglas aggregator across sectors gives

$$P = \prod_{s=1}^{S} \left(\frac{P_s}{\theta_s} \right)^{\theta_s}, \tag{8}$$

where θ_s is the weight on industry s and P is the minimum price of acquiring a unit of the aggregate real benefit. We assume that the nominal benefit satisfies value additivity at both the sector and the firm levels, so

$$\sum_{s=1}^{S} P_s F_s = PF$$

and

$$\sum_{i=1}^{I} P_{si} F_{si} = P_s F_s.$$

Ultimately, from the derivation of P, the industry weights θ_s are found to be the fractions of the economy allocated to each industry, that is,

$$P_s F_s = \theta_s PF. \tag{9}$$

We close with one comment about preferences. Although we do not model preferences explicitly, in a more explicit general equilibrium setting the implicit preferences that produce the above results are CES preferences over the benefit of finance from firms in a sector and Cobb-Douglas preferences over the benefit of finance from sectors in the economy.

C. Reallocation

We now calculate the real gains from reallocation using the framework above. One obstacle to overcome in making this calculation is that the real benefit of finance is unobservable. Although the nominal benefit, $P_{si}F_{si}$, is in principle observable, the real benefit is not because of the lack of observable prices. While many output price indices exist at the national or industry level, firm-level prices, which are necessary for calculating reallocation gains, are difficult to measure with any accuracy. This measurement difficulty implies that to calculate the gains from reallocation, we need to rely on the structure of the model. Specifically, we first plug the optimal allocations of debt and equity into the firm-level CES aggregate of debt and equity, (3), to obtain the optimal, first-best, firm-level real benefit of finance. Next, we calculate an estimate of the actual firm-level real benefit, which we obtain from plugging actual observed debt and equity into (3). We then calculate the economy-wide real benefit by aggregating the firm-level real benefits into sectors and the sectors into the aggregate economy. Finally, we compare the optimal with the actual aggregated real benefit to measure aggregate gains.

Before such aggregation can take place, inspection of (3) shows that as a first step toward calculating the firm-level real benefit, we need to find an expression for A_{si}. Here we rely on the observability of our main measure of the nominal benefit, $P_{si}F_{si}$, which is nominal value-added. We can therefore write A_{si} in terms of $P_{si}F_{si}$. In the Appendix, we show that we can express A_{si} as

$$A_{si} = \eta_s \frac{(P_{si}F_{si})^{\frac{\sigma}{\sigma-1}}}{\left(\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right)^{\frac{\gamma_s}{\gamma_s-1}}}, \quad \text{where} \quad \eta_s = \frac{1}{P_s(P_sF_s)^{\frac{1}{\sigma-1}}}. \tag{10}$$

All of the variables on the right-hand side of (10) are observable except for η_s. However, as shown in the Appendix, the reallocation gains do not depend on η_s, as it does not vary by firm, and can thus be divided out of the problem. We therefore normalize η_s to one.

The final ingredients necessary for the calculation of reallocation gains are the efficient levels of debt $\hat{D}_{si}$ and equity $\hat{E}_{si}$, where a hat above a variable indicates the efficient level after reallocation. Under this efficient allocation, total debt and total equity in a sector are held constant, but debt and equity are reallocated across firms in a sector to maximize the sector-level real benefit. To derive the efficient allocation, we define $\hat{D}_{si}$ and $\hat{E}_{si}$ as the solution to the problem faced by a social planner trying to maximize firm net benefits, subject to the total amount of finance allocated to the firm. Because the goal of a social planner's problem is the optimal allocation of quantities, the prices of debt and equity are not a part of the solution. As shown in the Appendix, from the first-order conditions for this planner's problem, we find that the optimal debt-equity ratio after reallocation is the same across firms at the sector level:

$$\frac{\hat{D}_{si}}{\hat{E}_{si}} = \frac{D_s}{E_s}. \tag{11}$$

This result of a constant debt-to-equity ratio within a sector relies on the assumption that the elasticity of substitution between debt and equity, γ_s, is constant within a sector. While we recognize that this assumption is strong, we provide justification for it in Section VI.B. below.

Given this optimal ratio, the total amount of finance is then reallocated among firms in the sector. The optimal total amount of finance of each firm can be found from the first-order conditions obtained from differentiating the expression for the sector-level real benefit, (2), with respect to the firm-level real benefits. These optimality conditions are given by

$$\hat{D}_{si} = \frac{A_{si}^{\sigma-1}}{\sum_j A_{sj}^{\sigma-1}} D_s, \tag{12}$$

$$\hat{E}_{si} = \frac{A_{si}^{\sigma-1}}{\sum_j A_{sj}^{\sigma-1}} E_s. \tag{13}$$

These two conditions express the intuitive result that the most productive firms in the economy, that is, those with the highest A_{si}, receive the most finance at an optimal allocation.

Once we determine optimal debt and equity, we can write the optimal real benefit for a firm, a sector, and the economy, respectively, as

$$\hat{F}_{si} = A_{si}\left(\alpha_s \hat{D}_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)\hat{E}_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right)^{\frac{\gamma_s}{\gamma_s-1}}, \tag{14}$$

$$\hat{F}_s = \left(\sum_i \hat{F}_{si}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{15}$$

$$\hat{F} = \prod_s \hat{F}_s^{\theta_s}. \tag{16}$$

As explained above, the original, prior to reallocation, real benefit can be computed by replacing $\hat{D}_{si}$ and $\hat{E}_{si}$ by the actual observable debt, D_{si}, and equity, E_{si}, in equation (14). We can therefore quantify the potential reallocation gains by calculating the observed allocation as a fraction of the efficient allocation. Letting F denote the observed real benefit, these gains are given simply by $F/\hat{F}$, which we refer to throughout as the fractional benefit.

D. MM Limit

We now consider the case in which $\sigma \to \infty$, $\gamma_s \to \infty$, and $\tau_{D_{si}} = \tau_{E_{si}} = 0$. We refer to this case as the MM benchmark or limit. In particular, as the elasticity of substitution between debt and equity goes to infinity, debt and equity become perfect substitutes. In general, a nontrivial elasticity of substitution between debt and equity can be interpreted as the result of some not-explicitly modeled financial frictions.

In the Appendix, we confirm that in this limit, the nominal profit of finance converges to zero, as follows:

$$\pi_{si} = P_{si}F_{si} - (r_{si}D_{si} + \lambda_{si}E_{si}) = 0. \tag{17}$$

Next, as shown in the Appendix, from a dual problem in which the total amount of finance is the firm's sole financial liability, we can derive a similar expression

$$P_{si}F_{si}/u_{si} = D_{si} + E_{si}, \tag{18}$$

where u_{si} is the average cost of capital for firm i in sector s. The relation in equation (18) is essentially MM Proposition I. That is, the nominal benefit of finance of any firm is independent of its capital structure, but rather depends only on the sum of its liabilities.

Combining equations (17) and (18) produces a standard expression for the weighted average cost of capital:

$$u_{si} = r_{si}\frac{D_{si}}{D_{si}+E_{si}} + \lambda_{si}\frac{E_{si}}{D_{si}+E_{si}}, \tag{19}$$

where u_{si} can also be interpreted as the fixed return on unlevered equity. As is standard, one implication of equation (19) is that the costs of debt, r_{si}, and equity, λ_{si}, adjust to satisfy MM Proposition II, which is a simple rearrangement of (19):

$$\lambda_{si} = u_{si} + \frac{D_{si}}{E_{si}}(u_{si} - r_{si}). \tag{20}$$

Recall that r_{si} and λ_{si} are functions of the debt-to-equity ratio, D_{si}/E_{si}. However, we do not assume a specific functional form, as many possible forms of r_{si} and λ_{si} can satisfy MM Proposition II as long as u_{si} is fixed at the firm level.

E. Elasticities

With the MM benchmark in mind, it is worth discussing the role of the parameters σ and γ_s in the quantification of reallocation gains. We first consider σ, the elasticity of substitution of the real benefit between firms in a sector. The extent of misallocation and thus the magnitude of potential reallocation gains depends positively on σ. To see this, consider a case in which firms in a sector are all the same size but have wide dispersion in TFB, A_{si}, implying inefficient allocations of debt and equity. In this case, if σ is high, moving to the efficient allocation would result in large dispersion in firm size. For example, as the more productive firms receive more finance, high product substitutability implies that their prices fall only slightly, so efficient reallocation implies in turn that their size increases substantially. Conversely, a low value for σ implies that reallocating debt and equity efficiently would result in the most productive firms getting only a modest amount of finance. In this latter case, larger size implies movement down a steep demand curve, which limits any reallocation gains. Overall, reallocation gains are greater when σ is higher.

Turning to γ_s, the elasticity of substitution between debt and equity, it is intuitive to see that when γ_s approaches infinity, debt and equity are perfect substitutes, and thus the potential gains from changing the debt-equity mix are zero. It follows that, even at an efficient allocation, any level of dispersion in debt-equity ratios could be associated with efficiency. More precisely, we show in the Appendix that if $\tau_{D_{si}} = \tau_{E_{si}} = 0$ and $\gamma_s \to \infty$, as in a classic MM world, our model is dual to a problem in which any debt-equity ratio can exist.

The reallocation gains are decreasing in γ_s. For example, if γ_s is large but still finite, then after reallocation, firms' debt-equity ratios are still equalized within a sector. However, high substitutability between debt and equity means that even large deviations from this optimum produce small efficiency losses, so the reallocation gains are also small. Conversely, if γ_s is small, then even a tight, but not degenerate, distribution of debt-equity ratios before reallocation could result in large reallocation gains.

Importantly, reallocation gains can exist even if debt and equity are perfect substitutes as long as σ is finite. Equations (12) and (13) imply that if the allocation of total resources to each firm is not proportional to its productivity, inefficiencies can exist. We call this source of misallocation the misallocation of scale, which could manifest as the real misallocation from Hsieh and Klenow (2009). For example, a firm that obtains too much finance given its optimal size and that does not disperse these excess funds to shareholders might use the funds to purchase too many assets or to hire too many workers in possibly suboptimal proportions. Conversely, a firm that cannot obtain enough finance operates at an inefficient scale. It might also possibly use an inefficient capital-labor mix. Finally, purely real distortions, such as labor subsidies, also show up in the misallocation of scale. Of course, while our model is not sufficiently rich to measure these different sources of the misallocation of scale, the balance sheet identity that assets equal liabilities implies that misallocation of real assets in Hsieh and Klenow (2009) should be closely related to our misallocation of financial liabilities.

We next discuss the different roles of the elasticity γ_s and the wedges $\tau_{D_{si}}$ and $\tau_{E_{si}}$ in quantifying misallocation. In a world with no financial frictions, $\gamma_s \to \infty$ and $\tau_{D_{si}} = \tau_{E_{si}} = 0$. Thus, relative to this baseline, inefficiency can stem from idiosyncratic factors or from sector-wide factors that do not vary at the firm level. Any sector-wide factors are, by definition, incorporated into the parameter γ_s. In Section VI.B below, we show in the context of a dynamic model of investment and finance that γ_s is primarily a function of parameters governing technology and uncertainty, which are likely to vary at the industry level. The wedges $\tau_{D_{si}}$ and $\tau_{E_{si}}$ then capture the idiosyncratic components of any frictions. This interpretation of γ_s, $\tau_{D_{si}}$, and $\tau_{E_{si}}$ means that any reallocation gains calculated by setting the price wedges to zero capture industry-wide imperfect substitutability between debt and equity. Thus, by comparing reallocation gains from a model with finite γ_s to an otherwise identical model with $\gamma_s \to \infty$, we can isolate the gains that stem from this substitutability.

II. Data

We first present our data from China; we then repeat the analysis for our U.S. data. The data set for China comes from the National Bureau of Statistics (NBS) of China and contains a panel of firms from 1999 to 2007. During this time period, firms with more than 5 million Chinese yuan (CNY), or approximately 600,000 U.S. dollars (USD), in sales are required to provide detailed financial information. This information includes statistics such as employment, income statement items, balance sheet items, and, after 2004, cash flow items.

This data set contains firms only from the mining, manufacturing, and utilities sectors. We focus on the manufacturing sector because the mining sector is relatively small and operationally different from manufacturing, while utilities are highly regulated in China. In addition, we remove state-owned and collective corporations, which are also known as Township and Village Enterprises (TVEs). Each firm-year observation is classified as a private corporation operating-year if the total state and collective paid-in capital is less than 50%.[4] We drop firms with negative and missing value-added, total liabilities, and shareholders' equity. We also drop firms with less than 5 million 1999 CNY in sales because the lack of reporting requirements for these firms likely results in significant selection bias and undersampling. After applying these screens, we are left with 1,248,729 firm-year observations.

We use total liabilities as our measure of D_{si} and shareholders' equity as our measure of E_{si}. This measure of equity is the stock of book equity and thus includes both external equity finance and retained earnings. As such, it reflects cumulative payout policy. As our measure of the nominal benefit of finance, $P_{si}F_{si}$, we use nominal value-added, which we compute as the sum of profits, indirect taxes, wages, and depreciation. These variables are all directly available in the NBS data. Note that we use total liabilities rather than debt, for two important reasons. First, debt is not an item in the NBS data, and second, using total liabilities can offer more robust estimation because there are almost no firms with zero liabilities. For a CES function with a finite elasticity of substitution, the marginal benefit of a factor input is unbounded at zero, and this property of the CES aggregator would present omitted observation problems in the estimation. These choices for debt and equity imply that the sum of the two liabilities, debt and equity, equals total assets.

Summary statistics for the sample of Chinese firms are presented in Table I. Panel A reports various statistics in the sample stratified by size, with cumulative density breakpoints of 5%, 15%, 30%, 50%, 70%, 85%, 95%, and 100%. This partition is useful because the mean of total assets roughly doubles for each size group, with the exception of the largest size group, which contains extremely large firms, and reflects the well-known right skewness of the firm-size distribution in China.

In Panel A of Table I, we find two patterns of interest in the data. First, the ratio of liabilities to assets increases with size. Because most external finance of any type is debt in this sample of primarily private firms during this period in China, the positive dependence of leverage on size generally indicates better access to finance for the larger Chinese firms. Second, it is clear from comparing the value-added and assets columns that the smaller firms use their assets far more efficiently in producing value-added. This pattern, juxtaposed with the higher leverage across size classes, points strongly to potential misallocation of debt and equity, as more productive firms should have more access to finance.

Panel B of Table I presents summary statistics by year. We find that the ratio of liabilities to assets is little changed over the sample period. Interestingly, average firm size shrinks somewhat from the beginning to the end of the

[4] This type of classification is often used because official corporate ownership registrations can lag several years behind actual ownership changes. See Guariglia, Liu, and Song (2011) for a similar approach.

Table I

Chinese Firm Summary Statistics

Calculations are based on a sample of Chinese firms from the annual survey conducted by the National Bureau of Statistics of China (NBS) from 1999 to 2007. All firms are in the manufacturing sector and all have more than 5 million Chinese yuan (CNY) in sales. There are a total of 1,248,729 firm-year observations, and all variables are reported in millions of 2005 CNY. Panel A presents summary statistics by firm-size percentile. Panel B presents summary statistics by year.

Panel A: Summary Statistics by Size

Size percentile	Observations	Assets	Liabilities	Equity	Liabilities/ Assets	Value-added
0–5	62,497	1.9	1.0	1.0	0.483	1.6
5–15	124,849	3.8	2.1	1.8	0.536	2.0
15–30	187,328	6.4	3.5	2.9	0.549	2.5
30–50	249,697	11.2	6.2	5.0	0.555	3.5
50–70	249,750	21.6	11.9	9.6	0.552	5.4
70–85	187,304	46.3	25.5	20.8	0.551	9.7
85–95	124,870	120.3	66.6	53.7	0.553	22.2
95–100	62,434	859.3	493.1	365.8	0.569	138.8

Panel B: Summary Statistics by Year

Year	Observations	Assets	Liabilities	Equity	Liabilities/ Assets	Value-added
1999	51,646	75.6	42.3	33.4	0.566	11.3
2000	62,822	78.1	43.6	34.4	0.566	12.4
2001	78,893	73.1	39.9	33.1	0.554	11.8
2002	95,520	70.9	38.9	32.0	0.551	11.9
2003	119,292	72.5	40.8	31.7	0.548	12.6
2004	181,692	59.5	34.3	25.2	0.564	10.7
2005	190,022	68.8	39.3	29.4	0.545	12.9
2006	217,242	70.8	40.3	30.5	0.539	14.0
2007	251,600	71.9	40.9	31.0	0.535	15.7

sample period. However, these slightly smaller firms create 40% more value-added at the end of the sample period than at the beginning.

For the United States, we use Compustat data. To make the data comparable to the Chinese data, we only keep the years from 1999 through 2007 and also only keep firms in the manufacturing sector with Standard Industrial Classification (SIC) codes between 2000 and 3999. We subsequently drop firms with missing data.

We compute value-added as in Imrohoroglu and Tuzel (2014). First, we estimate labor costs from the National Bureau of Economic Research and U.S. Census Bureau's Center for Economic Studies (NBER-CES) Manufacturing Industry Database by computing the mean wage per employee by three-digit SIC industry. We then multiply these wage figures by the number of employees in Compustat (EMP) to obtain a firm-level imputation of the wage bill. Value-added is then calculated as operating income before depreciation (OIBDP) plus

Table II

U.S. Firm Summary Statistics

Calculations are based on a sample of manufacturing firms (SIC 2000 to 3999) from Compustat. The sample period is 1999 to 2007 and includes 14,158 firm-year observations. All variables are reported in millions of 2005 USD. Value-added is operating income before depreciation (OIBDP) plus imputed wages. Imputed wages are calculated by multiplying the employment of each firm with the mean wage per employee in the appropriate three-digit SIC industry. Panel A presents summary statistics by firm-size percentile. Panel B presents summary statistics by year.

Panel A: Summary Statistics by Size

Size percentile	Observations	Assets	Liabilities	Equity	Liabilities/ Assets	Value-added
0–5	713	9.9	3.3	6.6	0.330	4.7
5–15	1,416	25.6	9.7	15.9	0.378	11.0
15–30	2,123	63.5	22.9	40.6	0.369	25.2
30–50	2,829	170.1	62.7	107.4	0.367	62.1
50–70	2,833	507.8	226.8	281.0	0.438	188.0
70–85	2,124	1,480.5	796.7	683.8	0.540	518.0
85–95	1,416	4,707.0	2,802.9	1,904.1	0.591	1,543.1
95–100	704	28,128.4	16,154.0	11,974.5	0.596	8,169.6

Panel B: Summary Statistics by Year

Year	Observations	Assets	Liabilities	Equity	Liabilities/ Assets	Value-added
1999	1,936	1,479.1	883.7	595.4	0.466	507.6
2000	1,782	1,719.4	1,000.2	719.2	0.460	597.7
2001	1,612	2,013.2	1,172.5	840.7	0.450	614.3
2002	1,541	2,131.0	1,257.1	873.9	0.448	622.1
2003	1,544	2,263.7	1,286.4	977.3	0.430	663.4
2004	1,517	2,469.3	1,365.8	1,103.5	0.418	750.5
2005	1,470	2,578.9	1,411.5	1,167.4	0.424	811.2
2006	1,415	2,860.4	1,549.0	1,311.4	0.426	877.7
2007	1,341	3,112.7	1,695.0	1,417.7	0.431	912.2

the imputed wages. For D_{si} and E_{si}, we use total liabilities (LT) and shareholders' equity (AT − LT).

We partition the firms by size according to the same densities as for China. Although the average firm is much larger in the United States than in China, the observed patterns by firm size can still be informative. Table II provides summary statistics by firm size in Panel A and by year in Panel B. In Panel A, we present eight firm-size categories, with the last three Chinese firm-size categories approximately equivalent to the first three U.S. firm-size categories. As in the Chinese data, we find a positive relation between size and leverage for U.S. firms. However, in contrast to Chinese firms, both small and large U.S. firms have approximately the same ratio of value-added to assets. Panel B reveals two more differences between Chinese and U.S. firms. In particular, U.S. firms grow from the beginning to the end of our sample, and they become less leveraged over time.

III. Estimation

Before we present our results, we explain how we estimate several parameters. First and most important, we estimate γ_s separately at the sector level using an extension of the method in Kmenta (1967), which is designed for the estimation of CES production functions of capital and labor. We do not impose the structure of the model in Section I on this estimation because in our framework, a world without financial frictions is one in which both $\gamma_s \to \infty$ and the wedges $\tau_{D_{si}} = \tau_{E_{si}} = 0$. Thus, estimating γ_s in this way allows us to use a source of data variation that differs from the source used to pin down the estimates of the wedges. Because γ_s is, by definition, a sector-level parameter in our model, this estimate captures any sector-wide average variation between leverage ratios and productivity. For example, if Chinese or U.S. firms were free of financial frictions, we would uncover no relation between leverage and productivity (Modigliani and Miller (1958)), which the method in Kmenta (1967) would translate into an infinite elasticity. In contrast, the presence of financial frictions produces a nonzero relation, which translates into a finite elasticity. Thus, estimating γ_s confers the advantage of being able to quantify the extent to which frictions are not firm-specific, even though we cannot identify individual frictions.

As described in Section III in the Internet Appendix,[5] in our main analysis we use ordinary least squares (OLS) estimation with fixed effects on the regression produced by the method in Kmenta (1967), which in our context is a nonlinear regression of value-added on debt and equity. The error in this regression is a function of A_{si}, and because A_{si} is naturally correlated with both D_{si} and E_{si}, the error is correlated with the regressors. Thus, the identifying assumption in our specification with fixed effects is that A_{si} varies only at the firm level and not over time.

We also consider an alternative identification strategy based on the assumption that A_{si} follows an autoregressive process. This strategy allows us to quasi-difference our estimating equation and use lagged instruments, as in Blundell and Bond (2000) and Davis, Fisher, and Whited (2014).[6] Because this generalized method of moments (GMM) estimation method can be poorly behaved in small samples, and because many of our two-digit sectors contain few observations, this method does not always produce sensible or precise estimates at the two-digit level. Therefore, in most of our analysis we use fixed-effects estimation at the two-digit sector level, but we provide additional analysis on a country-level basis using GMM estimation.

To apply either of these estimation strategies, we need to construct a measure of real value-added. In the U.S. data, we do so by deflating nominal value-added by the shipments price deflator from the NBER productivity database. For our Chinese data, no such price deflator is available, but for the period 1999

[5] The Internet Appendix may be found in the online version of this article.

[6] Other production function estimation methods, such as those in Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2015), require data on variables such as input prices or intermediate inputs that are either unavailable or ill-defined in our context.

Table III
Elasticity of Substitution Summary Statistics

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. This table presents summary statistics from the two-digit sector-level estimation of the elasticity of substitution between debt and equity, γ_s. The last two rows show the whole economy γ estimate using the Kmenta (1967) method with fixed effects and using GMM, respectively.

	United States	China
Mininum two-digit γ_s	1.28	1.38
Maximum two-digit γ_s	3.38	2.05
Mean two-digit γ_s	1.72	1.52
Weighted mean two-digit γ_s	1.60	1.52
Whole economy γ (Kmenta)	1.57	1.50
Whole economy γ (GMM)	1.52	1.58

to 2003, industrial output is reported at both current and constant prices. We therefore build a firm-specific deflator that can be used to compute real value-added from nominal value-added.

Table III summarizes the sector-level estimates of γ_s; the detailed estimates are provided in Table IA.I in the Internet Appendix. We find that the sector-level estimates range between 1.28 and 3.38 in the United States and between 1.38 and 2.05 in China, with means of 1.72 and 1.52, respectively. Table III also compares the whole-economy estimates from the Kmenta (1967) method with whole-economy GMM estimates. We find that the GMM approach produces corresponding estimates nearly identical to those from the Kmenta (1967) method with fixed effects.

Next, we set the elasticity of substitution for real value-added between firms in an industry to $\sigma = 1.77$. This choice equates the standard deviations of the observed and efficient size distributions in the United States where observed size is given by $\hat{D}_{si} + \hat{E}_{si}$ and actual size is given by $D_{si} + E_{si}$. Calibrating σ to match this moment makes sense if one believes that the U.S. size distribution is efficient because there is a monotonic mapping between σ and the efficient size standard deviation.

Ultimately, the estimates of the two parameters σ and γ_s allow us to quantify the relative importance of the debt-equity mix and firm scale for misallocation. With these parameters fixed at their estimated values, the model serves as a lens to filter the cross-sectional distributions of firm size and debt-equity ratios, with the relative size of the parameters, σ and γ_s, dictating the relative importance of scale and mix.

IV. Results

With the above parameters in hand, we now use the framework developed in Section I to quantify the extent of the misallocation of finance. Specifically, we use (10) to (13) to compute the hypothetically efficient levels of debt and equity for each firm. We then compare value-added computed with these efficient levels to value-added computed with actual levels, thus obtaining the reallocation gains. We conduct this comparison by sector and then aggregate sectors, where we define sectors by three-digit industry classifications from the Chinese NBS and three-digit SIC industries from Compustat.

Table IV contains year-by-year estimates of the potential gains from the reallocation of finance across firms in a sector.[7] Column (1) shows the observed U.S. real benefit of finance as a fraction of the optimal real benefit, $F_{\mathrm{US}}/\hat{F}_{\mathrm{US}}$. Column (2) shows the corresponding hypothetical percentage gain from moving from the observed to the optimal allocation of debt and equity. We find that the percent discrepancy between the optimal and observed allocation of finance for U.S. firms is 10.8% to 12.3%. In other words, these firms would stand to gain 10.8% to 12.3% in terms of value-added if they were to move to an optimal allocation.

In columns (3) and (4), we present analogous calculations for the Chinese firms. We find that the potential reallocation gains appear enormous, as value-added could increase by over 70% if the Chinese firms were to move to an efficient allocation of debt and equity. Although these figures seem large, they are of the same order of magnitude as the estimated gains found in Hsieh and Klenow (2009) regarding capital and labor allocations. Somewhat surprisingly, we also find, that the efficiency of the allocation of debt and equity falls over our sample period in China. This phenomenon can be attributed largely to the expansion of the NBS survey in the 2004 Industrial Census, as discussed in Brandt, Van Biesebroeck, and Zhang (2014). When we restrict our sample to firms in the NBS survey that are present before and after the 2004 Industrial Census, the pattern of increasing misallocation diminishes greatly.[8]

To put these results into perspective, we express the estimates of misallocation in China relative to the estimates of misallocation in the United States. This comparison is motivated by the observation in Hsieh and Klenow (2009) that because a simple static model based on the framework in Melitz (2003) is likely to be misspecified, a researcher may infer positive potential gains even when allocations are efficient. For example, potential sources of misspecification include uncertainty or lags in raising finance. This observation is particularly applicable in our context of financial misallocation because U.S. financial markets are highly developed. Thus, comparison of Chinese gains and U.S.

[7] We calculate standard errors in two ways. First, we calculate asymptotic standard errors of the gains by stacking the influence functions for the individual components of F and $\hat{F}$ and then using the delta method. Second, we use a simple bootstrap. The asymptotic standard errors are an order of magnitude smaller than bootstrapped standard errors, so we report the latter to be conservative. Even in this conservative case, all of the standard errors are quite small. This result makes sense inasmuch as the figures that we present are all functions of means, which can be estimated with a great deal of precision with several thousand data points.

[8] Table IA.II in the Internet Appendix presents results with γ estimated at the country level using the GMM technique described in Section III of the Internet Appendix. The results are quantitatively similar.

Table IV
Reallocation Gains by Year

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. This table presents potential reallocation gains when the substitutability between debt and equity varies at the two-digit sector level, and the elasticity of substitution between the real benefit of firms in a sector is $\sigma = 1.77$. Column (1) presents the observed U.S. allocation of real value-added as a fraction of the optimal U.S. allocation, $F_{\mathrm{US}}/\hat{F}_{\mathrm{US}}$. Column (2) presents the corresponding percentage gain from moving from the observed to the optimal allocation. Columns (3) and (4) present analogous calculations for Chinese firms. Columns (5) and (6) show the Chinese efficiency ratio as a fraction of the U.S. efficiency ratio, $(F_{\mathrm{China}}/\hat{F}_{\mathrm{China}})/(F_{\mathrm{US}}/\hat{F}_{\mathrm{US}})$, and the corresponding percentage gains. Columns (7) and (8) provide a breakdown of total misallocation into the misallocation due to the scale of finance and due to the mix of debt and equity, holding scale fixed. Standard errors are in parentheses below each estimate.

	United States		China		United States vs. China			
Year	(1) Fractional Benefit	(2) Percent Gain	(3) Fractional Benefit	(4) Percent Gain	(5) Fractional Benefit	(6) Percent Gain	(7) Percent Scale	(8) Percent Type
1999	0.901	11.0	0.581	72.3	0.644	55.2	45.0	10.2
	(0.006)	(0.7)	(0.008)	(2.4)	(0.010)	(2.3)	(1.9)	(0.7)
2000	0.891	12.3	0.573	74.6	0.643	55.5	46.3	9.2
	(0.009)	(1.1)	(0.013)	(3.8)	(0.016)	(3.7)	(3.2)	(0.8)
2001	0.897	11.4	0.587	70.3	0.654	52.8	43.4	9.4
	(0.009)	(1.1)	(0.007)	(2.1)	(0.011)	(2.6)	(2.1)	(0.7)
2002	0.891	12.2	0.579	72.7	0.650	53.9	44.5	9.4
	(0.008)	(1.0)	(0.011)	(3.3)	(0.013)	(3.1)	(2.7)	(0.8)
2003	0.892	12.2	0.591	69.3	0.663	50.9	41.3	9.6
	(0.009)	(1.1)	(0.007)	(2.0)	(0.009)	(2.2)	(1.5)	(0.8)
2004	0.902	10.9	0.556	80.0	0.616	62.3	50.7	11.6
	(0.008)	(1.0)	(0.005)	(1.7)	(0.008)	(2.2)	(1.7)	(0.7)
2005	0.902	10.8	0.542	84.4	0.601	66.4	53.3	13.2
	(0.009)	(1.1)	(0.005)	(1.8)	(0.009)	(2.4)	(1.8)	(0.8)
2006	0.897	11.5	0.546	83.2	0.608	64.4	50.8	13.5
	(0.009)	(1.0)	(0.004)	(1.5)	(0.007)	(1.9)	(1.5)	(0.7)
2007	0.896	11.6	0.529	88.9	0.591	69.2	54.4	14.8
	(0.008)	(1.0)	(0.004)	(1.3)	(0.007)	(2.0)	(1.4)	(0.8)

gains isolates the potential gains in China relative to an assumed efficient allocation. The results are presented in columns (5) and (6) of Table IV, which report the ratio $(F_{\text{China}}/\hat{F}_{\text{China}})/(F_{\text{US}}/\hat{F}_{\text{US}})$ and the corresponding percentage gain. This normalization delivers results that are more modest. We find potential gains of 51% to 56% before the expansion of the NBS survey and of 62% to 69% after the expansion.

To understand whether the reallocation gains come from the amount of finance available to Chinese firms or to the type of finance, we next compare the relative fractional benefit in column (5) to the case in which $\gamma_s = \infty$. If $\gamma_s = \infty$, then the type of finance does not matter for aggregate value-added because debt and equity are perfect substitutes. This exercise produces an interesting result. As can be seen in columns (7) and (8), we find that most of the potential reallocation gains come from the misallocation of scale. Before the expansion of the NBS survey in 2004, gains of only 9.2% to 10.2% could be realized by reallocating the type of finance. After the expansion, this figure rises to 11.6% to 14.8%. These results mean that the mix of debt and equity finance is important, but what matters more is the efficiency of the allocation of total firm resources to firms with respect to their productivities. Thus, as discussed above, because the misallocation of scale is difficult to distinguish from the real misallocation in Hsieh and Klenow (2009), our main finding is the small size of the reallocation gains stemming from the type of finance.

We also examine the implications of our estimates for the cross-sectional distribution of firm size. Recall that because of downward-sloping demand, each firm has a well-defined optimal size, with an optimal financing mix. Deviations of the financing amount and mix from the optimal allocation therefore affect firm size, and thus comparing the distributions of firm size under the actual and efficient allocations is a useful way to quantify misallocation. Figure 1 illustrates this idea with plots of the observed and efficient firm-size distributions for the United States and China. We compute observed firm size as $\log(D_{si} + E_{si})$ and efficient firm size as $\log(\hat{D}_{si} + \hat{E}_{si})$. Panel A shows that the efficient U.S. firm-size distribution exhibits approximately as much dispersion as the actual distribution. Of course, this result is to be expected, given our calibration of σ, with the slight discrepancy in the distributions stemming from the log transformation. In Panel B, in contrast, we see that the efficient firm-size distribution for China has significantly fatter tails, with many Chinese firms being either too small or too large. These size distortions stem from both the misallocation of the scale and the mix of finance documented in Table IV.

Although the plots in Figure 1 show the firm-size distributions before and after reallocation, they do not illustrate the individual changes in firm size that accompany reallocation. Figure 2 shows these potential movements via heat maps. Panel A is a heat map of a three-dimensional histogram in which the observed U.S. firm-size distribution is on the x-axis and the efficient U.S. firm-size distribution is on the y-axis. The legend for the heat map z-axis is located on the right of the map and represents the number of observations in each bin. Similarly, Panel B is a heat map for China. In Panel A, we find that U.S. firms are concentrated along the 45-degree line, where firm size before and

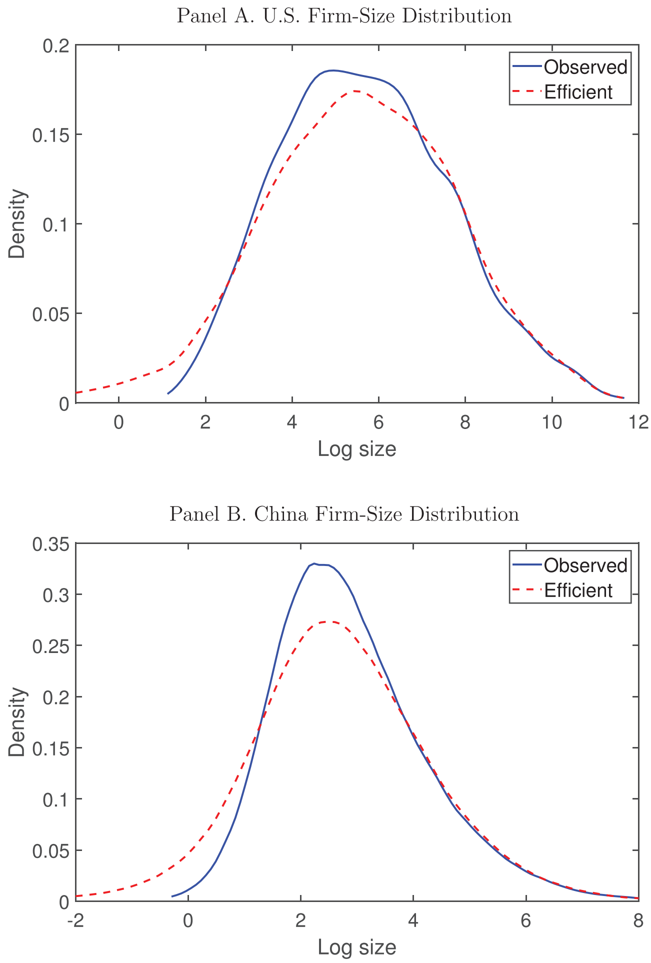
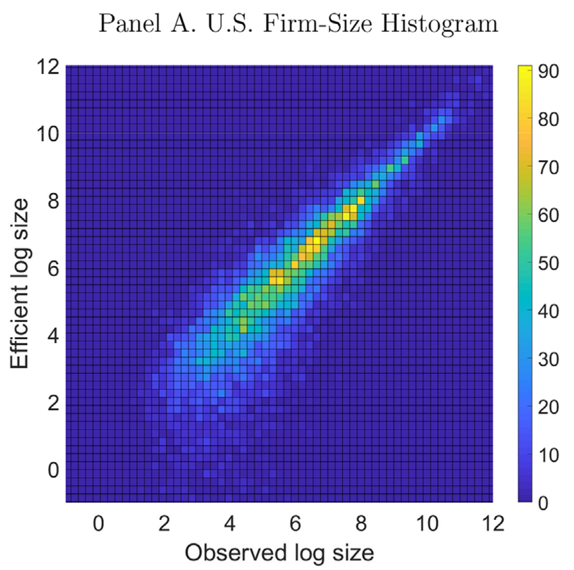


Figure 1. Efficient versus actual size densities. Panel A compares the U.S. observed and efficient firm-size distributions using a kernel density estimator. Observed firm size is computed as $\log(D_{si} + E_{si})$, and efficient firm size is computed as $\log(\hat{D}_{si} + \hat{E}_{si})$, where D_{si}, E_{si}, $\hat{D}_{si}$, and $\hat{E}_{si}$ are measured in millions of 2005 USD. Panel B similarly compares the observed and efficient firm-size distributions in China. Firm size is computed in the same manner, but D_{si}, E_{si}, $\hat{D}_{si}$, and $\hat{E}_{si}$ are measured in millions of 2005 CNY. (Color figure can be viewed at wileyonlinelibrary.com)

after reallocation is the same. In contrast, Chinese firms are much more spread out, reflecting the substantial efficiency gains available from reallocation. Both heat maps are more concentrated toward the top right than toward the bottom left, with this contrast being somewhat more pronounced for U.S. firms. This pattern indicates that small firms are more likely to suffer from financial misallocation than large firms, and relatively more so in the United States.

Next, we present estimates of the costs of debt and equity, as given by the right-hand sides of (A37) and (A38) in the Appendix. These marginal costs



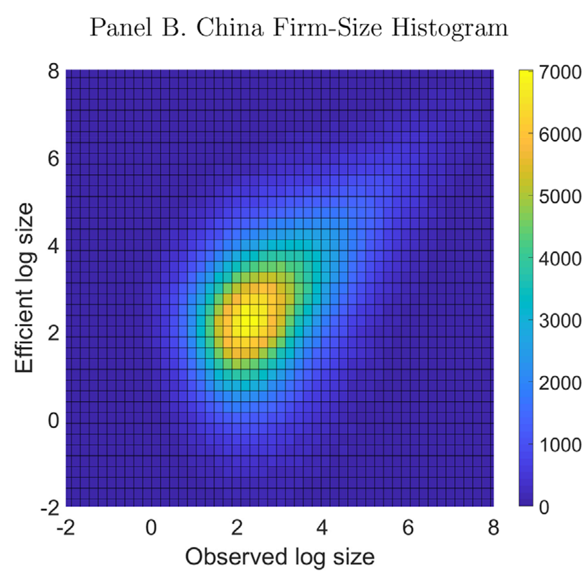


Figure 2. Efficient versus actual size histograms. Panel A contains the heat map of a three-dimensional histogram where the observed U.S. firm-size distribution is on the x-axis and the efficient U.S. firm-size distribution is on the y-axis. The legend for the z-axis heat map is located at the right of the plot and represents the number of observations in each bin. Observed firm size is computed as $\log(D_{si} + E_{si})$, and efficient firm size is computed as $\log(\hat{D}_{si} + \hat{E}_{si})$, where D_{si}, E_{si}, $\hat{D}_{si}$, and $\hat{E}_{si}$ are measured in millions of 2005 USD. Panel B similarly compares the observed and efficient firm-size distributions in China. Firm size is computed in the same manner, but D_{si}, E_{si}, $\hat{D}_{si}$, and $\hat{E}_{si}$ are measured in millions of 2005 CNY. (Color figure can be viewed at wileyonlinelibrary.com)

include the debt and equity wedges, $\tau_{D_{si}}$ and $\tau_{E_{si}}$. Table V summarizes these postdistortion costs, again under the assumptions that γ_s varies at the two-digit sector level and $\sigma = 1.77$. Panel A contains means and Panel B contains medians. In Panel A, we find that the distortion-inclusive costs of debt and equity decline over the sample period in the United States while these costs rise in China over the same time period. This pattern reinforces the results in Table IV that indicate greater misallocation after 2004, when the NBS survey samples more firms. These extra firms exhibit more misallocation and consequently greater costs of debt and equity. Finally, the figures in Panel B are uniformly much smaller than those in Panel A, especially for the Chinese firms. In line with extreme right skewness in the distribution of the cost of finance, this result implies that some firms are effectively barred from financial markets.

Table VI presents the distortion-inclusive costs of debt and equity when we stratify the sample by size instead of year. In both countries, the average costs of both debt and equity decrease with size, but these costs fall more sharply in China.

Making sense of the magnitudes of these costs in terms of standard rates of returns on securities or loans is difficult because our costs embed frictions related to the misallocation of both scale and mix. Nonetheless, one question we can answer is whether overleveraged firms have a higher weighted average cost of finance than underleveraged firms. This calculation speaks to the result in Korteweg (2010) that the costs of being over- or underleveraged are asymmetric. In Table IA.IV in the Internet Appendix, we present the average unit cost of finance, defined as $[(1+\tau_{D_{si}})r_{si}D_{si} + (1+\tau_{E_{si}})\lambda_{si}E_{si}]/(D_{si}+E_{si})$, for firms stratified by size in both the United States and China. We define overleveraged as a situation in which $D_{si}/(D_{si}+E_{si}) > D_s/(D_s+E_s)$. This definition makes sense in our framework because at an efficient allocation, all firms in a sector have leverage ratios equal to $D_s/(D_s+E_s)$. As in Korteweg (2010), we find that for U.S. firms, it is more costly to be overleveraged than underleveraged. Interestingly, the converse holds in China, with underleveraged firms having a higher average unit cost of finance.

We finish our analysis by examining how these costs vary with firm characteristics. Specifically, using our sample of Chinese firms, we regress the distortion-inclusive costs of debt and equity on location, state investment, firm size, time, and firm age. The first of these variables, *Location*, is a dummy variable that equals one if a firm is located in Beijing, Shanghai, Shenzhen, or Guangzhou and zero otherwise. These four Chinese cities are also known as first-tier cities and are the most developed in China. The second variable, *State investment*, is a dummy variable that equals one if a firm has a nonzero percentage of paid-in-capital from state sources and zero otherwise. The remaining variables are defined as follows: *Size* is the log of total assets measured in 2005 CNY, *Time* is a simple linear time trend, and *Young* is a dummy variable that equals one if the firm is three or fewer years old and zero otherwise.

Table VII presents the results. We find that the costs of debt and equity are significantly lower for larger firms in China. This result confirms our cross-sectional sorts by firm size in Table VI. Firms operating in first-tier Chinese

Table V

Costs of Debt and Equity by Year

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. This table displays the estimated mean (Panel A) and median (Panel B) distortion-inclusive marginal costs of debt and equity in the United States and China by year. All parameter settings are as in Table IV. Standard errors are in parentheses below each estimate.

	United States		China	
Year	Debt Cost	Equity Cost	Debt Cost	Equity Cost
	Panel A: Mean Costs			
1999	0.232	0.189	0.191	0.164
	(0.0042)	(0.0117)	(0.0032)	(0.0039)
2000	0.227	0.185	0.205	0.167
	(0.0044)	(0.0043)	(0.0083)	(0.0030)
2001	0.214	0.172	0.203	0.173
	(0.0046)	(0.0045)	(0.0036)	(0.0034)
2002	0.217	0.176	0.202	0.179
	(0.0047)	(0.0062)	(0.0030)	(0.0027)
2003	0.220	0.173	0.270	0.173
	(0.0046)	(0.0053)	(0.0083)	(0.0014)
2004	0.222	0.169	0.226	0.182
	(0.0048)	(0.0039)	(0.0026)	(0.0017)
2005	0.218	0.171	0.288	0.202
	(0.0037)	(0.0040)	(0.0057)	(0.0022)
2006	0.213	0.167	0.314	0.204
	(0.0044)	(0.0044)	(0.0149)	(0.0012)
2007	0.206	0.166	0.355	0.228
	(0.0063)	(0.0060)	(0.0046)	(0.0017)
	Panel B: Median Costs			
1999	0.188	0.154	0.085	0.096
	(0.0034)	(0.0022)	(0.0004)	(0.0005)
2000	0.184	0.153	0.091	0.102
	(0.0041)	(0.0024)	(0.0004)	(0.0004)
2001	0.168	0.137	0.095	0.106
	(0.0033)	(0.0029)	(0.0004)	(0.0004)
2002	0.171	0.138	0.098	0.107
	(0.0035)	(0.0029)	(0.0004)	(0.0004)
2003	0.171	0.134	0.102	0.109
	(0.0035)	(0.0030)	(0.0003)	(0.0004)
2004	0.177	0.138	0.101	0.112
	(0.0035)	(0.0027)	(0.0003)	(0.0003)
2005	0.177	0.141	0.112	0.117
	(0.0039)	(0.0027)	(0.0003)	(0.0002)
2006	0.169	0.139	0.117	0.122
	(0.0041)	(0.0025)	(0.0003)	(0.0003)
2007	0.162	0.130	0.127	0.131
	(0.0041)	(0.0030)	(0.0003)	(0.0003)

Table VI

Costs of Debt and Equity by Firm Size

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. This table displays the estimated mean (Panel A) and median (Panel B) distortion-inclusive marginal costs of debt and equity in the United States and China by firm size. All parameter settings are as in Table IV. Standard errors are in parentheses below each estimate.

	United States		China	
Year	Debt Cost	Equity Cost	Debt Cost	Equity Cost
Panel A: Mean Costs				
0–5	0.355	0.202	0.832	0.468
	(0.0159)	(0.0127)	(0.0172)	(0.0045)
5–15	0.289	0.198	0.438	0.307
	(0.0090)	(0.0082)	(0.0068)	(0.0027)
15–30	0.249	0.181	0.334	0.240
	(0.0036)	(0.0043)	(0.0054)	(0.0014)
30–50	0.246	0.170	0.255	0.194
	(0.0027)	(0.0071)	(0.0034)	(0.0011)
50–70	0.214	0.166	0.210	0.158
	(0.0029)	(0.0025)	(0.0029)	(0.0017)
70–85	0.166	0.179	0.173	0.130
	(0.0018)	(0.0036)	(0.0030)	(0.0023)
85–95	0.146	0.169	0.131	0.105
	(0.0017)	(0.0032)	(0.0040)	(0.0011)
95–100	0.135	0.142	0.152	0.089
	(0.0021)	(0.0023)	(0.0557)	(0.0009)
Panel B: Median Costs				
0–5	0.266	0.151	0.338	0.286
	(0.0104)	(0.0063)	(0.0019)	(0.0011)
5–15	0.221	0.149	0.196	0.199
	(0.0050)	(0.0035)	(0.0007)	(0.0005)
15–30	0.207	0.140	0.143	0.156
	(0.0040)	(0.0026)	(0.0003)	(0.0003)
30–50	0.210	0.132	0.111	0.124
	(0.0031)	(0.0015)	(0.0003)	(0.0002)
50–70	0.176	0.138	0.090	0.099
	(0.0023)	(0.0020)	(0.0002)	(0.0002)
70–85	0.153	0.147	0.077	0.083
	(0.0019)	(0.0019)	(0.0002)	(0.0002)
85–95	0.136	0.151	0.068	0.074
	(0.0019)	(0.0024)	(0.0002)	(0.0002)
95–100	0.127	0.137	0.062	0.068
	(0.0039)	(0.0034)	(0.0003)	(0.0002)

Table VII

Firm Characteristics and Debt and Equity Costs

Calculations are based on a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. This table presents two OLS regressions, in which the dependent variables are the distortion-inclusive marginal costs of debt and equity, respectively. The regressors are location, state investment, firm size, time, and firm age. *Location* is a dummy variable equal to one if a firm is located in Beijing, Shanghai, Shenzhen, or Guangzhou and zero otherwise. *State investment* is a dummy variable equal to one if a firm has a nonzero percentage of paid-in-capital from state sources and zero otherwise. *Foreign investment* is a dummy variable equal to one if a firm has a nonzero percentage of paid-in-capital from foreign sources and zero otherwise. *Size* is log total assets measured in 2005 CNY, *Time* is a linear time trend, and *Young* is a dummy variable equal to one if the firm is three or fewer years old and zero otherwise. t-statistics are in parentheses.

	Debt Cost	Equity Cost
Location	−0.054	−0.017
	(−5.8)	(−7.9)
State investment	0.019	0.011
	(0.9)	(2.2)
Foreign investment	0.088	−0.002
	(11.8)	(−1.3)
Size	−0.098	−0.057
	(−40.8)	(−104.6)
Time	0.022	0.007
	(16.5)	(24.7)
Young	0.015	−0.006
	(2.2)	(−3.7)

cities also face lower costs. Surprisingly, firms with nonzero state investment face slightly higher costs on average. It is important to note that this result is conditional on firm size. If we break down the full set of firms into those with and without state paid-in-capital, we find that firms with state paid-in-capital have lower costs. However, these firms are also significantly larger, so the effect of state investment on costs reverses after we control for firm size. Foreign investment, on the other hand, is associated with higher costs of debt but lower costs of equity, although this second effect is insignificant. We also find a positive coefficient on the time trend, which reflects the increasing costs also evident in Table V. Finally, we find that young firms face slightly highs of debt but slightly lower costs of equity.

V. Robustness

In this section, we consider several extensions of our analysis. First, we examine the robustness of the results in Table IV to the calibration of the parameters γ_s and σ, which are the elasticity of substitution between debt and equity and the elasticity of substitution between firms in a sector, respectively. The results are in Table VIII. The benchmark γ_s is estimated at the two-digit sector level, but we vary γ at the economy-wide level in Table IV for ease of

Table VIII

Reallocation Gains by Elasticities of Substitution

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. This table presents potential reallocation gains averaged across all years when we allow the elasticities of substitution, γ and σ, to vary. When γ is varied, σ is set to 1.77, and when σ is varied, γ_s is set to the benchmark two-digit sector values. Column (1) shows the observed U.S. allocation of real value-added as a fraction of the optimal U.S. allocation, $F_{\mathrm{US}}/\hat{F}_{\mathrm{US}}$. Column (2) shows the corresponding percentage gain from moving from the observed to the optimal allocation. Columns (3) and (4) present analogous calculations for Chinese firms. Columns (5) and (6) show the Chinese efficiency ratio as a fraction of the U.S. efficiency ratio, $(F_{\mathrm{China}}/\hat{F}_{\mathrm{China}})/(F_{\mathrm{US}}/\hat{F}_{\mathrm{US}})$, and the corresponding percentage gains. Columns (7) and (8) provide a breakdown of total misallocation into the misallocations due to the scale of finance and due to the mix of debt and equity, holding scale fixed. Standard errors are in parentheses below each estimate.

	United States		China		United States vs. China			
Parameter	(1) Fractional Benefit	(2) Percent Gain	(3) Fractional Benefit	(4) Percent Gain	(5) Fractional Benefit	(6) Percent Gain	(7) Percent Scale	(8) Percent Type
$\gamma = 1$	0.877	14.0	0.515	94.1	0.588	70.2	47.6	22.6
	(0.003)	(0.4)	(0.003)	(0.9)	(0.003)	(0.9)	(0.6)	(0.5)
γ_s	0.897	11.5	0.565	77.0	0.630	58.7	47.6	11.1
	(0.003)	(0.3)	(0.002)	(0.7)	(0.003)	(0.8)	(0.6)	(0.2)
$\gamma = 2$	0.903	10.8	0.583	71.6	0.646	54.9	47.6	7.2
	(0.002)	(0.3)	(0.002)	(0.7)	(0.003)	(0.7)	(0.6)	(0.2)
$\gamma = 5$	0.916	9.2	0.610	63.8	0.667	50.0	47.6	2.4
	(0.002)	(0.3)	(0.002)	(0.6)	(0.003)	(0.6)	(0.6)	(0.1)
$\gamma = 10$	0.920	8.7	0.618	61.7	0.672	48.7	47.6	1.1
	(0.002)	(0.3)	(0.002)	(0.6)	(0.003)	(0.6)	(0.6)	(0.0)
$\sigma = 1.5$	0.909	10.1	0.613	63.1	0.675	48.2	38.9	9.3
	(0.002)	(0.3)	(0.002)	(0.6)	(0.003)	(0.6)	(0.5)	(0.2)
$\sigma = 1.77$	0.897	11.5	0.565	77.0	0.630	58.7	47.6	11.1
	(0.003)	(0.3)	(0.002)	(0.7)	(0.003)	(0.8)	(0.6)	(0.2)
$\sigma = 2$	0.886	12.8	0.524	90.9	0.591	69.1	56.0	13.2
	(0.003)	(0.4)	(0.003)	(0.9)	(0.003)	(0.9)	(0.7)	(0.3)
$\sigma = 2.5$	0.862	16.0	0.435	129.8	0.505	98.1	78.3	19.8
	(0.004)	(0.5)	(0.003)	(1.6)	(0.004)	(1.5)	(1.1)	(0.5)
$\sigma = 3$	0.836	19.6	0.353	183.7	0.422	137.2	107.8	29.4
	(0.004)	(0.6)	(0.003)	(2.8)	(0.004)	(2.4)	(1.8)	(0.8)

juxtaposition. We find that assigning different values for γ has a negligible effect on our estimates of the percent gains in the United States. Although we find substantial sensitivity of the percent gains to γ in China, we can still conclude that the vast majority of potential gains stems from the misallocation of scale, as the gains from reallocating the type of finance run between 22.6% and 1.1%.

While varying γ_s has only a modest effect on our estimated reallocation gains, Table VIII shows that the estimated gains increase sharply with σ. When we vary σ, we set the parameters γ_s to their baseline two-digit sector values. As explained in Section I.E above, a high value of σ implies small price responses to reallocation, so the gains are larger. Our calibration of σ on the low end of this range therefore means that our baseline results are conservative. If σ were lower, the efficient size distribution would be more compressed than the observed distribution, and the reallocation gains would be small. However, when σ rises above 1.77, the efficient size distribution expands substantially and the reallocation gains become quite large. Moreover, our calibration of $\sigma = 1.77$ is conservative in light of the evidence in Broda and Weinstein (2006), who estimate the elasticity of substitution between goods within various three-digit industries using import data—our value of 1.77 is at the low end of their range of estimates.

A natural concern is the small overlap between the distributions of the sizes of Chinese and U.S. firms. To understand whether this feature of our data drives our results, each year we limit the sample to the size intersection between the Chinese and U.S. firms, where size is total assets in USD adjusted by the nominal exchange rate. In computing this intersection, we drop the 10 smallest U.S. firms and the 10 largest Chinese firms from the sample to minimize the effects of extreme outliers. We then recompute the potential reallocation gains. The results reported in Table IX indicate that the relative percentage reallocation gains for Chinese firms are approximately halved.

Interestingly, and surprisingly, the factor driving most of the difference is the removal of the large U.S. firms, not the removal of the small Chinese firms. In other words, a nonoverlapping size distribution appears to mask observed inefficiencies in the U.S. firms rather than exacerbate observed inefficiencies in the Chinese firms. To understand this issue, note that we remove approximately the largest one-sixth of the U.S. firms after taking the size intersection, and these firms are close to operating at an optimal allocation. Thus, the removal of these big U.S. firms increases the percentage reallocation gain in the United States from approximately 10% to over 20%. This result implies that smaller U.S. firms could make substantial gains from reallocation, but these gains appear insignificant when we consider all of the firms in Compustat, which include the top one-sixth of firms that are large, productive, and already at an efficient size and financing mix. In contrast, productive Chinese firms can be of any size, so potential reallocation gains do not come only from tiny firms. This pattern in the Chinese data is also evident in the heat map in Figure 2, where large deviations between original and efficient sizes can be seen for both large and small firms.

Table IX

Reallocation Gains for the Firm-Size Intersection

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. In each year, we limit the sample to the size intersection between U.S. and Chinese firms. Column (1) shows the observed U.S. allocation of real value-added as a fraction of the optimal U.S. allocation, $F_{\text{US}}/\hat{F}_{\text{US}}$. Column (2) shows the corresponding percentage gain from moving from the observed to the optimal allocation. The next two columns present analogous calculations for Chinese firms. The final two columns show the Chinese efficiency ratio as a fraction of the U.S. efficiency ratio, $(F_{\text{China}}/\hat{F}_{\text{China}})/(F_{\text{US}}/\hat{F}_{\text{US}})$, and the corresponding percentage gains. Standard errors are in parentheses below each estimate.

	United States		China		United States vs. China	
Year	(1) Fractional Benefit	(2) Percent Gain	(3) Fractional Benefit	(4) Percent Gain	(5) Relative Fractional Benefit	(6) Percent Gain
1999	0.817	22.4	0.616	62.3	0.754	32.6
	(0.011)	(1.5)	(0.010)	(2.7)	(0.016)	(2.7)
2000	0.792	26.2	0.643	55.4	0.812	23.2
	(0.014)	(2.1)	(0.007)	(1.6)	(0.016)	(2.4)
2001	0.798	25.4	0.639	56.4	0.801	24.8
	(0.011)	(1.6)	(0.007)	(1.7)	(0.013)	(2.1)
2002	0.797	25.5	0.629	59.1	0.789	26.8
	(0.012)	(1.8)	(0.012)	(3.0)	(0.019)	(3.1)
2003	0.792	26.3	0.638	56.8	0.805	24.2
	(0.012)	(1.8)	(0.006)	(1.5)	(0.015)	(2.3)
2004	0.808	23.8	0.613	63.2	0.759	31.8
	(0.012)	(1.8)	(0.006)	(1.5)	(0.014)	(2.5)
2005	0.817	22.4	0.625	59.9	0.765	30.7
	(0.013)	(1.8)	(0.006)	(1.5)	(0.014)	(2.4)
2006	0.825	21.3	0.628	59.3	0.761	31.3
	(0.011)	(1.6)	(0.005)	(1.3)	(0.012)	(2.1)
2007	0.819	22.1	0.616	62.2	0.753	32.9
	(0.012)	(1.8)	(0.005)	(1.4)	(0.012)	(2.2)

Next, we investigate whether our benchmark results in Table IV depend on the exclusion of state-owned firms from our sample of Chinese firms. This issue is potentially important because in terms of aggregate size, state-owned firms make up a large fraction of the Chinese economy. Table X reports the potential reallocation gains for state-owned Chinese firms alone and for the entire sample of Chinese firms. For reference, column (1) of Table X shows the fractional reallocation benefit by year for the entire sample of U.S. firms and thus repeats column (1) in Table IV. Columns (3) and (4) show that, somewhat surprisingly, state-owned firms stand to gain less in percentage terms than non–state-owned firms, with the fractional reallocation gains ranging from 36% to 53%, figures slightly lower than the corresponding values for the sample of non–state-owned Chinese firms. Two forces are at work here. On the one hand, the public status of the state-owned Chinese firms confers upon them much better access to external financial markets. On the other hand, these state-owned firms do not necessarily seek to maximize profits when planning their activities, so they are likely to be less efficient. Our evidence suggests that the first force dominates the second. Note also that the average state-owned firm is nearly twice as large as the average non–state-owned firm. Finally, the percentage of available reallocation gains is slightly smaller for the pooled sample of state-owned and private Chinese firms. However, because the measured economy is larger with the inclusion of the state-owned firms, the total gross amount of potential reallocation gains is still substantially larger.

Finally, we check to see whether our results are robust to using net debt instead of gross debt in our analysis. This issue is important because financial frictions can influence firms' choices of cash balances as well as their choices of debt. However, as can be seen in Table IA.III in the Internet Appendix, our results are largely similar when we use net debt instead of gross debt in our analysis.

Table XI examines whether our results depend on our measure of the nominal benefit of finance. This issue arises because measures of value-added likely contain substantial measurement error (Bils, Klenow, and Ruane (2017)). Instead of value-added, a natural alternative measure is the sum of the market values of debt and equity. We cannot use this measure in our sample of Chinese firms, as most of these firms are not publicly traded, but we can examine this alternative measure in our sample of U.S. firms. We find that overall reallocation gains are similar to those in Table IV. One exception occurs during the dot-com boom, in which we find more misallocation, indicating that the easy access to public equity markets in this period did not produce an efficient allocation of resources.

VI. Alternative Models

A. Labor Income

In this section, we consider whether our choice to model both capital and labor as intermediate inputs matters for our results. On the one hand, if

Table X
State-Owned Firms

Calculations are based on two samples of firms. One is a sample of U.S. firms from Compustat, and the other is a sample of Chinese firms from the National Bureau of Statistics of China. The sample period is 1999 to 2007. Column (1) shows the observed U.S. allocation of real value-added as a fraction of the optimal U.S. allocation, $F_{\mathrm{US}}/\hat{F}_{\mathrm{US}}$. Columns (2) to (4) report the Chinese state-owned firm efficiency ratio, $F_{\mathrm{China}}/\hat{F}_{\mathrm{China}}$, the Chinese state-owned firm efficiency ratio as a fraction of the U.S. efficiency ratio, $(F_{\mathrm{China}}/\hat{F}_{\mathrm{China}})/(F_{\mathrm{US}}/\hat{F}_{\mathrm{US}})$, and the corresponding percentage gains. Columns (5) to (7) repeat columns (2) to (4), except that calculations are based on a sample of all Chinese firms, private and state-owned. Standard errors are in parentheses below each estimate.

	United States	China State-Owned Firms			China All Firms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Year	Fractional Benefit	Fractional Benefit	Relative Fractional Benefit	Percent Gain	Fractional Benefit	Relative Fractional Benefit	Percent Gain
1999	0.901	0.614	0.681	46.8	0.587	0.652	53.5
	(0.006)	(0.010)	(0.012)	(2.7)	(0.007)	(0.009)	(2.0)
2000	0.891	0.611	0.686	45.8	0.577	0.647	54.5
	(0.009)	(0.013)	(0.016)	(3.5)	(0.010)	(0.013)	(3.2)
2001	0.897	0.587	0.654	52.8	0.569	0.634	57.8
	(0.009)	(0.015)	(0.019)	(4.3)	(0.010)	(0.013)	(3.3)
2002	0.891	0.635	0.712	40.4	0.590	0.662	51.1
	(0.008)	(0.011)	(0.014)	(2.7)	(0.008)	(0.011)	(2.5)
2003	0.892	0.654	0.733	36.4	0.602	0.675	48.1
	(0.009)	(0.011)	(0.014)	(2.5)	(0.007)	(0.010)	(2.2)
2004	0.902	0.649	0.720	39.0	0.574	0.637	57.0
	(0.008)	(0.014)	(0.017)	(3.2)	(0.005)	(0.008)	(2.0)
2005	0.902	0.640	0.709	41.0	0.556	0.616	62.3
	(0.009)	(0.018)	(0.020)	(4.0)	(0.007)	(0.009)	(2.5)
2006	0.897	0.654	0.729	37.1	0.561	0.626	59.8
	(0.009)	(0.017)	(0.019)	(3.5)	(0.005)	(0.008)	(2.0)
2007	0.896	0.651	0.727	37.6	0.543	0.607	64.8
	(0.008)	(0.018)	(0.022)	(4.1)	(0.005)	(0.008)	(2.3)

Table XI

U.S. Reallocation Gains with Market Value Benefit

Calculations are based on a sample of U.S. firms from Compustat. The sample period is 1999 to 2007. The nominal benefit of finance is measured as the market value of debt plus the market value of equity, instead of value-added. This table presents potential reallocation gains when the substitutability between debt and equity varies at the two-digit sector level and the elasticity of substitution between the real benefit of firms in a sector is $\sigma = 1.77$. Column (1) shows the observed U.S. allocation of the real benefit of finance as a fraction of the optimal U.S. allocation, $F_{\mathrm{US}}/\hat{F}_{\mathrm{US}}$. Column (2) shows the corresponding percentage gain from moving from the observed to the optimal allocation. Standard errors are in parentheses below each estimate.

	United States	
Year	Fractional Benefit	Percent Gain
1999	0.728	37.4
	(0.026)	(4.8)
2000	0.739	35.4
	(0.024)	(4.2)
2001	0.853	17.3
	(0.012)	(1.6)
2002	0.879	13.8
	(0.011)	(1.4)
2003	0.878	13.9
	(0.012)	(1.5)
2004	0.877	14.1
	(0.012)	(1.6)
2005	0.874	14.4
	(0.013)	(1.7)
2006	0.885	13.0
	(0.009)	(1.1)
2007	0.867	15.4
	(0.011)	(1.4)

liabilities only fund fixed asset accumulation, then because labor income represents over 60% of U.S. national income, our results can only be interpreted narrowly in terms of capital income, not in terms of value-added. On the other hand, liabilities can be viewed as the accumulated internal and external funding that finances all factors of production. In this case, our representation of the productive process via the CES aggregator in (3) is sensible in that all proximate factors of production such as capital and labor can be thought of as unmodeled intermediate inputs. More generally, if labor is perfectly flexible, then Ejarque (2002) shows that labor is self-financing, so it is sensible to view debt and equity as financing only fixed assets. However, if labor is a quasi-fixed factor, then funds raised via either debt or equity can finance either capital or labor. A useful anecdote in support of this view comes from DeAngelo and Roll (2015), who note that in 1966, IBM used a seasoned equity offering to fund a payroll shortfall.

In Section I in the Internet Appendix, we show that taking a stand on either of these views matters little for our quantitative measures of financial misallocation. Briefly, we assume that the real benefit of finance for firm i in sector s is given by

$$F_{si} = A_{si}\left(\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right)^{\frac{\beta_s\gamma_s}{\gamma_s-1}} L_{si}^{1-\beta_s}, \tag{21}$$

where $1-\beta_s$ is labor's share. This functional form captures the notion that debt and equity are combined to finance capital, which in turn is combined with labor via a Cobb-Douglas aggregator to generate the real benefit of finance. Next, we assume that the firm minimizes the cost of production, given by

$$\min_{\{D_{si},E_{si},L_{si}\}} \{(1+\tau_{D_{si}})r_{si}D_{si} + (1+\tau_{E_{si}})\lambda_{si}E_{si} + (1+\tau_{L_{si}})wL_{si}\}, \tag{22}$$

where w is the real wage and $\tau_{L_{si}}$ is the wedge associated with the real wage. This minimization is subject to a certain level of output, $F_{si} = \bar{F}_{si}$.

We show that the optimal debt-equity ratio in this alternative model is still given by (5). This result makes sense given that our Cobb-Douglas assumption implies that the first-order conditions for debt and equity contain an identical multiplicative term in labor that divides out of the expression for the optimal debt-equity ratio. We also show that after reallocation, an individual firm's debt, equity, and labor are all proportional to the firm's productivity, just as in the baseline model of Section I.

We find that the results using the model extension to calculate reallocation gains are quantitatively similar to those in Table IV. To calculate these gains, we assume that the capital share of income is one-third for the United States and one-half for China, in line with aggregate data. We report our results in Table IA.V in the Internet Appendix. We find that the percentage gains for the United States and China are slightly smaller and slightly larger, respectively, than the baseline results from Table IV. The relative percentage gains are therefore slightly larger, ranging from 68.2% to 77.3%, in contrast to the 50.9% to 69.2% figures from Table IV. Regardless of whether we use our baseline model from Section I or this augmented model, we can conclude that the overall reallocation gains are substantial.

B. Dynamic Model of the Firm

Finally, we ask whether our unconventional framework from Section I can detect financial frictions generated by a more standard model of capital structure. We turn to a model with a production framework that is closely related to the one in Hennessy and Whited (2005). While we lay out the exact assumptions of the model of Section II in the Internet Appendix, the basic structure can be understood as follows. Time is discrete. The equilibrium setting consists of a single representative consumer and a unit continuum of infinitely lived firms, each of which combines tangible capital, intangible capital, and labor to produce output using a stochastic technology that is subject to persistent shocks. Each firm chooses two types of capital, labor, and net debt to maximize the expected present value of distributions, which are defined as after-tax profits minus factor expenditures plus net debt issuance. Negative distributions are equivalent to injections of funds and thus can be thought of as equity issuance. While labor can be adjusted costlessly, both types of capital incur smooth adjustment costs. In addition, both types of capital depreciate, while labor does not.

The model contains three financial frictions. The first is a tax benefit, which makes the firm behave impatiently relative to its discount rate. By itself, this friction would incentivize the firm to use an unlimited amount of debt if it were not for the next two frictions. The second friction is a collateral constraint such that firms can only borrow up to a fraction of their tangible capital. With only these two frictions, the model has a simple solution for optimal debt. In each period, to satisfy its impatience, the firm borrows up to the limit given by the collateral constraint. Any extra funds needed for its optimal hiring and investment policies come from deep-pocket shareholders. An interior solution for optimal debt comes from the third friction, which is an equity issuance cost. In this case, facing the possibility of having to raise costly external equity, the firm preserves debt capacity, choosing optimal debt at a level less than that given by the collateral constraint.

We solve the model, use the solution to simulate data, and then estimate the percentage gains from reallocation in this economy. We consider two versions of the model: one with no equity issuance costs and one with different levels of this cost. Although the first version does contain frictions, it can be considered closer to a frictionless world, as the firm is never constrained by the availability or cost of funds in implementing its optimal labor and capital investment policies. While in principle it would be interesting to calculate reallocation gains from a model with no financial frictions, in this case optimal debt policy is undefined.

Panel A of Figure IA.1 in the Internet Appendix plots three measures of misallocation: the percentage gains, the percent of the gains that come from the scale of finance, and the percent of the gains that come from the type of finance. We plot these three measures as a function of the cost of equity issuance by solving and simulating the model 20 times, with each solution corresponding to a different value for the equity issuance cost. Intuitively, we find that the reallocation gains are increasing in the issuance cost. We also find that the percent type is increasing in the issuance cost, while the percent scale is decreasing. This later result is also intuitive in that leverage hugs the collateral constraint for near-zero issuance costs, so the type of finance varies little. This variation increases as the issuance cost rises and firms optimally choose a positive and time-varying amount of free debt capacity relative to the collateral constraint.

Five features of this basic result are of interest. First, and most importantly, we can conclude that our CES production function of debt and equity can detect financial frictions, as our measure of misallocation is monotonically increasing in the cost of equity issuance. While this conclusion is confined to the laboratory of dynamic models of finance and investment, as surveyed in

Strebulaev and Whited (2012) and Bazdresch, Kahn, and Whited (2018), these models have proven useful for quantitatively matching many features of firm-level data.

Second, the no-issuance-cost version of the model produces efficiency losses. The reason is simple model misspecification, as our static model from Section I is only an approximation to the dynamic model in the Internet Appendix. The largest difference between these two models is that in the dynamic model, the existence of a stochastic shock implies that the connection between debt-equity ratios and value-added varies in a way that is inconsistent with a simple static framework. However, our experiment does tell us that this misspecification induces only a modest amount of apparent misallocation.

Third, the increase in reallocation gains that accompanies the introduction of issuance costs arises because the firm issues less equity and uses less debt. As such, the constrained optimal choices of debt and equity leave the firm with insufficient funds to implement its first-best investment and hiring policies. Moreover, the collateral constraint induces the firm to utilize more tangible capital than it would in the absence of financial frictions. Finally, the issuance cost implies that the firm chooses debt to avoid a binding collateral constraint, so, as seen in Panel B of Figure IA.1, the dispersion in debt-equity ratios rises with financial frictions. More generally, we observe misallocation in the cross-section because current leverage affects future capital stock choices and thus future productivity. At the same time, financial frictions induce path dependence in leverage (Hennessy and Whited (2005)), so current leverage also affects future leverage choices. The model therefore produces cross-sectional correlations between leverage and measures of value or productivity. Note that we generate cross-sectional dispersion in leverage ratios without explicit heterogeneity in the cost of external finance. Such an additional source of heterogeneity would mechanically augment this dispersion.

Fourth, the decrease in our measure of fractional benefit is comparable in magnitude to two other measures of misallocation that capture the effects of the financial friction. The first is the average firm market-to-book ratio, and the second is aggregate output, which we normalize to one when the issuance cost is zero. This second measure directly captures the real misallocation of Hsieh and Klenow (2009).

Panel A of Figure IA.2 shows that as the issuance cost rises from 0 to 0.12,[9] the fractional benefit falls by 16%, while the market-to-book ratio falls by 21% and aggregate output falls by 15%. These measures are comparable in magnitude to one another. They are also comparable to the decrease in the measure of misallocation of type in China in the later years of our sample, as can be seen in Table IV. Type misallocation is the relevant empirical quantity for comparison because in this comparative static, we only change the equity issuance cost.

Fifth, we use our simulated data to estimate the elasticity of substitution between debt and equity, that is, the parameter γ_s in our static model from

[9] This second figure is close to the estimated issuance cost for small firms from Hennessy and Whited (2007).

Section I. Interestingly, the R^2 from the regression is 0.89, indicating that although our dynamic model does not collapse exactly into our static measurement framework from Section I, the approximation is good. As can be seen in Panel B of Figure IA.2, we find that the estimate of γ_s falls as financial frictions rise.

While this result is intuitive, because γ_s is a function of all model parameters, we further examine which of these parameters influence γ_s the most. To this end, we solve and simulate the dynamic model 400 times, with each time corresponding to a randomly chosen set of parameter values. The upper and lower bounds for the parameter values are in Panel A of Table IA.VI in the Internet Appendix. For each set of parameter values, we then estimate γ_s using the simulated stocks of debt, equity, and value-added. Next, we regress the 400 estimates of γ_s on the model parameters that correspond to each estimate. The results are presented in the first two columns of Table IA.VI, Panel B. Here, we find significant coefficients on many of the parameters that describe the firm's technology, as well as the collateral constraint. In the third column we present the best set of least absolute shrinkage and selection operator (LASSO) predictors, calculated with 10-fold cross-validation. Here we find similar results, with the parameters that carry significant OLS coefficients also providing the most explanatory power.

The most important message from this exercise is that the parameter γ_s is likely to be a function of deep technological parameters, as well as the tangibility of capital. While these purely technological characteristics are clearly not constant across the individual firms in our sample, they are likely to vary more across industries than across individual firms within an industry. Thus, the results in Table IA.VI lend credence to our assumption that γ_s varies across industries but not across firms in an industry.

VII. Conclusion

This paper entertains the possibility that finance may be misallocated in the cross-section of firms. We explore this hypothesis using a tractable model of differentiated firms based on Hsieh and Klenow (2009). Our empirical investigation is based on the intuitive result that in our framework, the optimal allocation of debt and equity equates their marginal benefits across firms within an industry. Thus, any observed dispersion in the marginal benefits of debt or equity is symptomatic of misallocation.

Our evidence from U.S. data points to only modest misallocation distortions or, equivalently, modest potential reallocation gains. These firms stand to gain only 10.8% to 12.3% in terms of aggregate real firm value-added if they were to move to an efficient allocation. Our results are much more dramatic for China, where we estimate that if China were able to achieve the more reasonable U.S. level of efficiency, gains of 51% to 69% would still be possible. When we break this figure down by the amount of finance versus the type of finance, we find that approximately four-fifths of the potential gains can be attributed to the amount of finance, with the rest attributed the mix of securities used to fund a firm's operations.

Our work thus introduces a new methodology for quantifying the extent to which an efficient allocation of financial resources across the economy can affect both aggregate output and the TFP of an economy. While we find that these effects are potentially large, the simplicity of our measurement framework implies that more work needs to be done before any such results could be confidently used for policy analysis. We examine several extensions of our framework and find that our results are largely robust, but we can imagine many more extensions, such as making the model dynamic, including a more granular set of financial instruments, or incorporating risk. It remains to be seen whether these natural extensions impact the quantitative importance of the misallocation of finance.

Initial submission: May 27, 2017; Accepted: October 12, 2020
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

Appendix

Aggregate Price

We begin by solving for the aggregate price P as a function of the sector price P_s, where P is defined as the minimum price of acquiring a unit of the aggregate benefit. Formally, the minimization problem is

$$\min_{F_s}\left\{\sum_s P_s F_s\right\}, \tag{A1}$$

subject to

$$\prod_s F_s^{\theta_s} = \bar{F}. \tag{A2}$$

The Lagrangian is

$$\mathcal{L} = -\sum_s P_s F_s + M\left[\prod_s F_s^{\theta_s} - \bar{F}\right], \tag{A3}$$

where M is the Lagrange multiplier. The first-order condition with respect to F_s gives

$$P_s = M\theta_s \frac{\prod_s F_s^{\theta_s}}{F_s}. \tag{A4}$$

Because $M = P$, (A4) simplifies to

$$\frac{P_s F_s}{\theta_s} = PF. \tag{A5}$$

By aggregating the sectors in the economy, we can write the aggregate price as a function of sector price

$$P = \prod_s \left(\frac{P_s}{\theta_s}\right)^{\theta_s}. \tag{A6}$$

Sector Price

In a similar fashion, we can solve for the sector price P_s as a function of the firm price P_{si}, where P_s is defined as the minimum price of acquiring a unit of the sector benefit. Formally, the minimization problem is

$$\min_{F_{si}} \left\{ \sum_i P_{si} F_{si} \right\}, \tag{A7}$$

subject to

$$\left(\sum_i F_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} = \bar{F}_s. \tag{A8}$$

The Lagrangian is

$$\mathcal{L}_s = -\sum_i P_{si} F_{si} + M_s \left[\left(\sum_i F_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} - \bar{F}_s \right], \tag{A9}$$

where M_s is the Lagrange multiplier. The first-order condition with respect to F_{si} gives

$$P_{si} = M_s \left(\sum_i F_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1}{\sigma-1}} F_{si}^{-\frac{1}{\sigma}}. \tag{A10}$$

Because $M_s = P_s$, (A10) simplifies to

$$P_{si}^{\sigma} F_{si} = P_s^{\sigma} F_s. \tag{A11}$$

By aggregating the firms in a sector, we can write the sector price as a function of firm price:

$$P_s = \left(\sum_i P_{si}^{-(\sigma-1)} \right)^{-\frac{1}{\sigma-1}}. \tag{A12}$$

Firm's Problem

A firm i in sector s chooses its price P_{si}, debt D_{si}, and equity E_{si} to maximize the nominal net benefit of finance, π_{si}. The debt and equity decision aims to minimize the total cost of finance for a given level of real benefit, $\bar{F}_{si}$, and can be separated from the price decision. Formally, the minimization problem is

$$\min_{D_{si},E_{si}} \left\{ \left(1+\tau_{D_{si}}\right) r_{si} D_{si} + \left(1+\tau_{E_{si}}\right) \lambda_{si} E_{si} \right\}, \tag{A13}$$

subject to

$$A_{si}\left[\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s) E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{\gamma_s}{\gamma_s-1}} = \bar{F}_{si}, \tag{A14}$$

where the weight on debt, α_s, is set to $D_s^{1/\gamma_s}/(D_s^{1/\gamma_s}+E_s^{1/\gamma_s})$, in which D_s and E_s are the sector debt and equity levels. The prices of debt and equity, r_{si} and λ_{si}, vary by both firm and sector and are functions of D_{si}/E_{si}. Because we do not need to specify the functional form, we omit this additional notation to be concise. After setting up the Lagrangian, taking the first-order conditions with respect to D_{si} and E_{si}, and rearranging, we arrive at the following optimal debt-equity ratio:

$$\frac{D_{si}}{E_{si}} = \left[\frac{\alpha_s}{1-\alpha_s}\frac{\frac{\partial}{\partial E_{si}}\left[\left(1+\tau_{D_{si}}\right) r_{si} D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si} E_{si}\right]}{\frac{\partial}{\partial D_{si}}\left[\left(1+\tau_{D_{si}}\right) r_{si} D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si} E_{si}\right]}\right]^{\gamma_s}, \tag{A15}$$

which expands out to

$$\frac{D_{si}}{E_{si}} = \left[\frac{\alpha_s}{1-\alpha_s}\frac{-\left(1+\tau_{D_{si}}\right)\frac{\partial r_{si}}{\partial E_{si}}\left(\frac{D_{si}}{E_{si}}\right)^2 - \left(1+\tau_{E_{si}}\right)\left(\frac{\partial \lambda_{si}}{\partial E_{si}}\frac{D_{si}}{E_{si}} - \lambda_{si}\right)}{\left(1+\tau_{D_{si}}\right)\left(\frac{\partial r_{si}}{\partial D_{si}}\frac{D_{si}}{E_{si}} + r_{si}\right) + \left(1+\tau_{E_{si}}\right)\frac{\partial \lambda_{si}}{\partial D_{si}}}\right]^{\gamma_s}. \tag{A16}$$

To simplify notation, let Z_{si} denote the optimal ratio:

$$Z_{si} \equiv \left(\frac{D_{si}}{E_{si}}\right)^{*}. \tag{A17}$$

Debt and equity can thus be expressed as linear functions of the real benefit, as follows:

$$\begin{aligned} D_{si} &= \frac{\bar{F}_{si}}{A_{si}}\left[\alpha_s + (1-\alpha_s) Z_{si}^{-\frac{\gamma_s-1}{\gamma_s}}\right]^{-\frac{\gamma_s}{\gamma_s-1}}, \\ E_{si} &= \frac{\bar{F}_{si}}{A_{si}}\left[\alpha_s Z_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)\right]^{-\frac{\gamma_s}{\gamma_s-1}}. \end{aligned} \tag{A18}$$

Using the above expressions for debt and equity, the minimum cost function becomes a function of the fixed real benefit, $\bar{F}_{si}$:

$$\begin{aligned} C(\bar{F}_{si}) &= \left(1+\tau_{D_{si}}\right) r_{si} D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si} E_{si} \\ &= C_{si}\bar{F}_{si}, \end{aligned} \tag{A19}$$

where

$$\begin{aligned} C_{si} = \frac{1}{A_{si}} \Bigg\{ &\left(1+\tau_{D_{si}}\right) r_{si} \left[\alpha_s + (1-\alpha_s) Z_{si}^{-\frac{\gamma_s - 1}{\gamma_s}}\right]^{-\frac{\gamma_s}{\gamma_s - 1}} \\ &+ \left(1+\tau_{E_{si}}\right)\lambda_{si} \left[\alpha_s Z_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)\right]^{-\frac{\gamma_s}{\gamma_s-1}} \Bigg\}. \end{aligned} \tag{A20}$$

Next, we choose P_{si} to maximize the nominal net benefit of finance, that is,

$$\max_{P_{si}}\{\pi_{si}\} = \max_{P_{si}}\left\{P_{si}F_{si} - C_{si}F_{si}\right\}. \tag{A21}$$

Recall from the derivation of the sector price that firm-level real benefit is a function of the sector price, firm price, and sector real benefit, as follows: $F_{si} = (\frac{P_s}{P_{si}})^{\sigma} F_s$. Therefore, because the optimal debt-equity ratio does not depend on the price, the debt-equity ratio can be maximized out of the problem of the optimal determination of the price, leaving the firm's real benefit as just a function of price. Because the firm faces a downward-sloping demand curve, the maximization problem is bounded even though the firm has constant returns to scale. From the first-order condition for the price, we find

$$P_{si} = \frac{\sigma}{\sigma - 1} C_{si}. \tag{A22}$$

Note that the price is a fixed markup over marginal cost and a higher elasticity of substitution between firms in a sector reduces the price the firm can charge for the real benefit it is generating.

MM Limit

Recall that the nominal net benefit of finance is

$$\pi_{si} = P_{si}F_{si} - \left[\left(1+\tau_{D_{si}}\right) r_{si} D_{si} + \left(1+\tau_{E_{si}}\right)\lambda_{si} E_{si}\right]. \tag{A23}$$

To characterize a frictionless economy in the spirit of (Modigliani and Miller, 1958, MM hereafter), we assume $\sigma \to \infty$, $\gamma_s \to \infty$, and $\tau_{D_{si}} = \tau_{E_{si}} = 0$. First, we show that under these conditions, $\pi_{si} = 0$. As an initial step, we show that the firm-level price under the MM limit is

$$\begin{aligned} P_{si} &= \lim_{\sigma\to\infty,\gamma_s\to\infty} \frac{\sigma}{\sigma-1}\frac{1}{A_{si}}\Bigg\{(1+\tau_{D_{si}})r_{si}\left[\alpha_s+(1-\alpha_s)\left(\frac{D_{si}}{E_{si}}\right)^{-\frac{\gamma_s-1}{\gamma_s}}\right]^{-\frac{\gamma_s}{\gamma_s-1}} \\ &\quad + (1+\tau_{E_{si}})\lambda_{si}\left[\alpha_s\left(\frac{D_{si}}{E_{si}}\right)^{\frac{\gamma_s-1}{\gamma_s}}+(1-\alpha_s)\right]^{-\frac{\gamma_s-1}{\gamma_s}}\Bigg\} \\ &= \frac{2}{A_{si}}\left[r_{si}\left(1+\frac{E_{si}}{D_{si}}\right)^{-1}+\lambda_{si}\left(\frac{D_{si}}{E_{si}}+1\right)^{-1}\right] \\ &= \left(\frac{2}{A_{si}}\right)\left(\frac{r_{si}D_{si}+\lambda_{si}E_{si}}{D_{si}+E_{si}}\right). \end{aligned} \tag{A24}$$

Similarly, the real benefit of finance is

$$\begin{aligned} F_{si} &= \lim_{\sigma\to\infty,\gamma_s\to\infty} A_{si}\left[\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}}+(1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{\gamma_s}{\gamma_s-1}} \\ &= \left(\frac{A_{si}}{2}\right)(D_{si}+E_{si}). \end{aligned} \tag{A25}$$

Finally, we combine (A24) and (A25) to confirm that π_{si} under the MM limit is

$$\pi_{si} = \lim_{\sigma\to\infty,\gamma_s\to\infty}\{P_{si}F_{si}-[(1+\tau_{D_{si}})r_{si}D_{si}+(1+\tau_{E_{si}})\lambda_{si}E_{si}]\}=0. \tag{A26}$$

Dual Problem

In this subsection, we show that if $\gamma_s\to\infty$ in the original model, nominal value-added equals an average cost of total finance times the quantity of total finance, without regard for the composition of finance. For this alternative problem, let total debt and equity be defined as $T_{si}=D_{si}+E_{si}$. Next, let the cost of finance be $(1+\tau_{T_{si}})u_{si}T_{si}$, where u_{si} is the base cost of finance without any distortions and $\tau_{T_{si}}$ is a wedge attached to the total cost of finance. Then, similar to (A25), we can express the real benefit as

$$F_{si}=A_{si}T_{si}. \tag{A27}$$

The cost minimization problem subject to a given level of the real benefit, $\bar{F}_{si}$, yields the trivial solution $T_{si}=\bar{F}_{si}/A_{si}$. Thus, the minimum cost function can be written as a function of a fixed level of the real benefit:

$$C(\bar{F}_{si})=(1+\tau_{T_{si}})u_{si}T_{si}=C_{si}\bar{F}_{si}, \tag{A28}$$

where

$$C_{si} = \frac{1}{A_{si}}(1+\tau_{T_{si}})u_{si}. \tag{A29}$$

Next, the firm chooses P_{si} to maximize the nominal benefit of finance, as follows:

$$\max_{P_{si}}\{\pi_{si}\} = \max_{P_{si}}\{P_{si}F_{si} - C_{si}F_{si}\}. \tag{A30}$$

As before, price is a markup over marginal cost:

$$P_{si} = \frac{\sigma}{\sigma-1}C_{si}. \tag{A31}$$

The MM limit when debt and equity are already perfectly substitutable assumes $\sigma \to \infty$ and $\tau_{T_{si}} = 0$. Therefore, we have

$$\pi_{si} = \lim_{\sigma\to\infty,\gamma_s\to\infty}\{P_{si}F_{si} - u_{si}T_{si}\} = 0. \tag{A32}$$

MM Propositions

Equation (A32) is essentially MM Proposition I. That is, the nominal benefit of finance of any firm is independent of its capital structure. Next, combining (A26) and (A32), we can produce the weighted average cost of capital formula,

$$u_{si} = r_{si}\frac{D_{si}}{D_{si}+E_{si}} + \lambda_{si}\frac{E_{si}}{D_{si}+E_{si}}, \tag{A33}$$

where u_{si} can also be interpreted as the fixed return on unlevered equity. This expression implies that the prices of debt, r_{si}, and equity, λ_{si}, adjust to satisfy MM Proposition II. Rearranging (A33), we have

$$\lambda_{si} = u_{si} + \frac{D_{si}}{E_{si}}(u_{si} - r_{si}), \tag{A34}$$

which is the standard formulation of MM Proposition II.

Wedges

To solve for the debt and equity wedges, we first substitute (A11) into $P_{si}F_{si}$ to write the nominal benefit of finance as

$$P_{si}F_{si} = P_sF_s^{\frac{1}{\sigma}}F_{si}^{\frac{\sigma-1}{\sigma}}. \tag{A35}$$

The marginal nominal benefit of debt must equal the marginal nominal cost of debt for the maximizing firm, so the first-order condition with respect to D_{si}

gives

$$\frac{\sigma-1}{\sigma}P_sF_s^{\frac{1}{\sigma}}F_{si}^{-\frac{1}{\sigma}}A_{si}\left[\alpha_sD_{si}^{\frac{\gamma_s-1}{\gamma_s}}+(1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{1}{\gamma_s-1}}\alpha_sD_{si}^{-\frac{1}{\gamma_s}} = \frac{\partial}{\partial D_{si}}\left[(1+\tau_{D_{si}})r_{si}D_{si}+\left(1+\tau_{E_{si}}\right)\lambda_{si}E_{si}\right], \tag{A36}$$

which simplifies to

$$\alpha_s\frac{\sigma-1}{\sigma}\frac{P_{si}F_{si}}{\alpha_sD_{si}+(1-\alpha_s)D_{si}^{\frac{1}{\gamma_s}}E_{si}^{\frac{\gamma_s-1}{\gamma_s}}} = \frac{\partial}{\partial D_{si}}\left[(1+\tau_{D_{si}})r_{si}D_{si}+\left(1+\tau_{E_{si}}\right)\lambda_{si}E_{si}\right]. \tag{A37}$$

Similarly, the first-order condition with respect to E_{si} simplifies to

$$(1-\alpha_s)\frac{\sigma-1}{\sigma}\frac{P_{si}F_{si}}{\alpha_sD_{si}^{\frac{\gamma_s-1}{\gamma_s}}E_{si}^{\frac{1}{\gamma_s}}+(1-\alpha_s)E_{si}} = \frac{\partial}{\partial E_{si}}\left[(1+\tau_{D_{si}})r_{si}D_{si}+\left(1+\tau_{E_{si}}\right)\lambda_{si}E_{si}\right]. \tag{A38}$$

Without the wedges, the marginal cost of debt and equity are just the marginal costs from the MM benchmark.

Reallocation

We now turn to the derivation of the efficient allocation in a sector. Under the efficient allocation, total debt and total equity in a sector are held constant, but debt and equity are reallocated across firms in a sector to maximize sector real benefit. Because this problem is essentially one of a social planner, it is not necessary to consider the prices of output or either debt or equity. Instead, the social planner first chooses debt and equity to maximize the real benefit of finance of a firm, subject to the total amount of finance that it is allocated. That is, the planner chooses $\hat{D}_{si}$ and $\hat{E}_{si}$ to maximize

$$F_{si} = A_{si}\left[\alpha_s\hat{D}_{si}^{\frac{\gamma_s-1}{\gamma_s}}+(1-\alpha_s)\hat{E}_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{\gamma_s}{\gamma_s-1}} \tag{A39}$$

subject to

$$\hat{D}_{si}+\hat{E}_{si} = \bar{T}_{si}, \tag{A40}$$

where $\bar{T}_{si}$ is the fixed total amount of finance.

From the firm first-order conditions, we find that this optimal debt-equity ratio after reallocation is

$$\frac{\hat{D}_{si}}{\hat{E}_{si}} = \left(\frac{\alpha_s}{1-\alpha_s}\right)^{\gamma_s}. \tag{A41}$$

Because $\alpha_s = D_s^{1/\gamma_s}/(D_s^{1/\gamma_s} + E_s^{1/\gamma_s})$, the optimal ratio simplifies to

$$Z_s \equiv \left(\frac{D_s}{E_s}\right)^*. \tag{A42}$$

Thus, the debt-equity ratio is the same for all firms i in sector s when debt and equity are reallocated to achieve efficiency. The real benefit can then be written as a function of $\hat{D}_{si}$:

$$\hat{F}_{si} = \left[\alpha_s + (1-\alpha_s)Z_s^{-\frac{\gamma_s-1}{\gamma_s}}\right]^{-\frac{\gamma_s}{\gamma_s-1}} A_{si}\hat{D}_{si}, \tag{A43}$$

where a hat above a variable indicates the level after reallocation. The Lagrangian for maximizing sector financial benefit is

$$\hat{\mathcal{L}}_s = \left\{\sum_i \left[\left[\alpha_s + (1-\alpha_s)Z_s^{-\frac{\gamma_s-1}{\gamma_s}}\right]^{-\frac{\gamma_s}{\gamma_s-1}} A_{si}\hat{D}_{si}\right]^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}} + \hat{M}_s\left[\sum_i \hat{D}_{si} - D_s\right], \tag{A44}$$

where $\hat{M}_s$ is the Lagrange multiplier. The first-order condition with respect to $\hat{D}_{si}$ and $\hat{D}_{sj}$, for firms i and j, respectively, rearranges to

$$\left(\frac{\hat{D}_{si}}{\hat{D}_{sj}}\right)^{-\frac{1}{\sigma}} = \left(\frac{A_{sj}}{A_{si}}\right)^{\frac{\sigma-1}{\sigma}}. \tag{A45}$$

After aggregation, the expression above can be simplified to

$$\hat{D}_{si} = \frac{A_{si}^{\sigma-1}}{\sum_j A_{sj}^{\sigma-1}} D_s. \tag{A46}$$

The optimal equity allocation can be similarly derived as

$$\hat{E}_{si} = \frac{A_{si}^{\sigma-1}}{\sum_j A_{sj}^{\sigma-1}} E_s. \tag{A47}$$

The real benefit F_{si} is assumed to be unobservable. However, A_{si} can be expressed in terms of observable variables, such as the nominal benefit, $P_{si}F_{si}$, that is,

$$A_{si} = \eta_s \frac{(P_{si}F_{si})^{\frac{\sigma}{\sigma-1}}}{\left[\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{\gamma_s}{\gamma_s-1}}}, \tag{A48}$$

where

$$\eta_s = \frac{1}{P_s(P_sF_s)^{\frac{1}{\sigma-1}}}, \tag{A49}$$

because

$$F_{si}P_s(P_sF_s)^{\frac{1}{\sigma-1}} = (P_{si}F_{si})^{\frac{\sigma}{\sigma-1}}. \tag{A50}$$

Reallocation gains are not affected if η_s is normalized to one for all sectors s. Because η_s does not vary across firms in a sector, it divides out of our measure of reallocation gains.

Aggregation

The ultimate goal is to find the ratio of the aggregate real benefit computed from data to the aggregate real benefit that would occur if the allocation were efficient. The real benefit computed from data is given by

$$\begin{aligned} F_{si} &= A_{si}\left[\alpha_s D_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)E_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{\gamma_s}{\gamma_s-1}}, \\ F_s &= \left(\sum_i F_{si}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \\ F &= \prod_s F_s^{\theta_s}, \end{aligned} \tag{A51}$$

while the efficient allocation is given by

$$\begin{aligned} \hat{F}_{si} &= A_{si}\left[\alpha_s \hat{D}_{si}^{\frac{\gamma_s-1}{\gamma_s}} + (1-\alpha_s)\hat{E}_{si}^{\frac{\gamma_s-1}{\gamma_s}}\right]^{\frac{\gamma_s}{\gamma_s-1}}, \\ \hat{F}_s &= \left(\sum_i \hat{F}_{si}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \\ \hat{F} &= \prod_s \hat{F}_s^{\theta_s}. \end{aligned} \tag{A52}$$

Therefore, the ratio is $F/\hat{F}$.

REFERENCES

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, 2015, Identification properties of recent production function estimators, *Econometrica* 83, 2411–2451.

Ai, Hengjie, Kai Li, and Fang Yang, 2016, Financial intermediation and capital reallocation, Working paper, University of Minnesota.

Alfaro, Laura, Andrew Charlton, and Fabio Kanczuk, 2009, Plant-size distribution and cross-country income differences, in Jeffrey A. Frankel, and Christopher Pissarides, eds., NBER International Seminar on Macroeconomics (National Bureau of Economic Research, Cambridge, MA).

Asker, John, Allan Collard-Wexler, and Jan De Loecker, 2014, Dynamic inputs and resource (mis)allocation, *Journal of Political Economy* 122, 1013–1063.

Bai, Yan, Dan Lu, and Xu Tian, 2016, Do financial frictions explain Chinese firms' saving and misallocation? Working paper, University of Rochester.

Banerjee, Abhijit V., and Esther Duflo, 2005, Growth theory through the lens of development economics, in Philippe Aghion, and Steven N. Durlauf, eds., Handbook of Economic Growth (Elsevier, Amsterdam).

Banerjee, Abhijit V., and Benjamin Moll, 2010, Why does misallocation persist? *American Economic Journal: Macroeconomics* 2, 189–206.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, 2013, Cross-country differences in productivity: The role of allocation and selection, *American Economic Review* 103, 305–334.

Bazdresch, Santiago, R. Jay Kahn, and Toni M. Whited, 2018, Estimating and testing dynamic corporate finance models, *Review of Financial Studies* 31, 322–361.

Bils, Mark, Peter J. Klenow, and Cian Ruane, 2017, Misallocation or mismeasurement? Working paper, Stanford University.

Blouin, Jennifer, John E. Core, and Wayne Guay, 2010, Have the tax benefits of debt been overestimated? *Journal of Financial Economics* 98, 195–213.

Blundell, Richard, and Stephen Bond, 2000, GMM estimation with persistent panel data: An application to production functions, *Econometric Reviews* 19, 321–340.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, 2014, Challenges of working with the Chinese NBS firm-level data, *China Economic Review* 30, 339–352.

Broda, Christian, and David E. Weinstein, 2006, Globalization and the gains from variety, *Quarterly Journal of Economics* 121, 541–585.

Brown, Jennifer, and David A. Matsa, 2016, Boarding a sinking ship? An investigation of job applications to distressed firms, *Journal of Finance* 71, 507–550.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, 2011, Finance and development: A tale of two sectors, *American Economic Review* 101, 1964–2002.

Chen, Kaiji, and Zheng Song, 2013, Financial frictions on capital allocation: A transmission mechanism of TFP fluctuations, *Journal of Monetary Economics* 60, 683–703.

Cooley, Thomas F., and Vincenzo Quadrini, 2001, Financial markets and firm dynamics, *American Economic Review* 91, 1286–1310.

Curtis, Chadwick C., 2016, Economic reforms and the evolution of China's total factor productivity, *Review of Economic Dynamics* 21, 225–245.

David, Joel M., and Venky Venkateswaran, 2016, Capital misallocation: Frictions or distortions? Working paper, USC.

Davis, Morris A., Jonas D. M. Fisher, and Toni M. Whited, 2014, Macroeconomic implications of agglomeration, *Econometrica* 82, 731–764.

DeAngelo, Harry, Linda DeAngelo, and Réne M. Stulz, 2010, Seasoned equity offerings, market timing, and the corporate lifecycle, *Journal of Financial Economics* 95, 275–295.

DeAngelo, Harry, and Richard Roll, 2015, How stable are corporate capital structures? *Journal of Finance* 70, 373–418.

Dixit, Avinash K., and Joseph E. Stiglitz, 1977, Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297–308.

Ejarque, Joao, 2002, Do financial market imperfections affect the cyclicality of employment, Working paper, University of Copenhagen.

Fischer, Edwin O., Robert Heinkel, and Josef Zechner, 1989, Dynamic capital structure choice: Theory and tests, *Journal of Finance* 44, 19–40.

Foster, Lucia, John Haltiwanger, and Chad Syverson, 2008, Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98, 394–425.

Goldstein, Robert, Nengjiu Ju, and Hayne Leland, 2001, An EBIT-based model of dynamic capital structure, *Journal of Business* 74, 483–512.

Graham, John R., 2000, How big are the tax benefits of debt? *Journal of Finance* 55, 1901–1941.

Guariglia, Alessandra, Xiaoxuan Liu, and Lina Song, 2011, Internal finance and growth: Microeconometric evidence on Chinese firms, *Journal of Development Economics* 96, 79–94.

Hennessy, Christopher A., and Toni M. Whited, 2005, Debt dynamics, *Journal of Finance* 60, 1129–1165.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing?, Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Hopenhayn, Hugo A., 1992, Entry, exit, and firm dynamics in long run equilibrium, *Econometrica* 60, 1127–1150.

Hopenhayn, Hugo, and Richard Rogerson, 1993, Job turnover and policy evaluation: A general equilibrium analysis, *Journal of Political Economy* 101, 915–938.

Hsieh, Chang-Tai, and Peter J. Klenow, 2009, Misallocation and manufacturing TFP in China and India, *Quarterly Journal of Economics* 124, 1403–1448.

Hsieh, Chang-Tai, and Peter J. Klenow, 2014, The life cycle of plants in India and Mexico, *Quarterly Journal of Economics* 129, 1035–1084.

Imrohoroglu, Ayse, and Selale Tuzel, 2014, Firm-level productivity, risk, and return, *Management Science* 60, 2073–2090.

Jensen, Michael C., 1986, The agency costs of free cash flow: Corporate finance and takeovers, *American Economic Review* 76, 323–329.

Jensen, Michael, and William H. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3, 305–360.

Jeong, Hyeok, and Robert M. Townsend, 2007, Sources of TFP growth: Occupational choice and financial deepening, *Economic Theory* 32, 179–221.

Kehrig, Matthias, and Nicolas Vincent, 2016, Do firms mitigate or magnify capital misallocation? Evidence from plant-level data, Working paper, Duke University.

Kmenta, Jan, 1967, On estimation of the CES production function, *International Economic Review* 8, 180–189.

Korteweg, Arthur G., 2010, The net benefits to leverage, *Journal of Finance* 65, 2137–2170.

Levinsohn, James, and Amil Petrin, 2003, Estimating production functions using inputs to control for unobservables, *Review of Economic Studies* 70, 317–341.

Levinsohn, James, and Amil Petrin, 2012, Measuring aggregate productivity growth using plant-level data, *RAND Journal of Economics* 43, 705–725.

Li, Shaojin, Toni M. Whited, and Yufeng Wu, 2016, Collateral, taxes, and leverage, *Review of Financial Studies* 29, 1453–1500.

Lucas, Robert E., Jr, 1978, On the size distribution of business firms, *Bell Journal of Economics* 9, 508–523.

Melitz, Marc J., 2003, The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica* 71, 1695–1725.

Midrigan, Virgiliu, and Daniel Yi Xu, 2014, Finance and misallocation: Evidence from plant-level data, *American Economic Review* 104, 422–458.

Modigliani, Franco, and Merton H. Miller, 1958, The cost of capital, corporation finance and the theory of investment, *American Economic Review* 48, 261–297.

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.

Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187–221.

Olley, G. Steven, and Ariel Pakes, 1996, The dynamics of productivity in the telecommunications equipment industry, *Econometrica* 64, 1263–1297.

Rampini, Adriano A., and S. Viswanathan, 2010, Collateral, risk management, and the distribution of debt capacity, *Journal of Finance* 65, 2293–2322.

Rampini, Adriano A., and S. Viswanathan, 2013, Collateral and capital structure, *Journal of Financial Economics* 109, 466–492.

Restuccia, Diego, and Richard Rogerson, 2008, Policy distortions and aggregate productivity with heterogeneous establishments, *Review of Economic Dynamics* 11, 707–720.

Restuccia, Diego, and Richard Rogerson, 2013, Misallocation and productivity, *Review of Economic Dynamics* 16, 1–10.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, 2011, Growing like China, *American Economic Review* 101, 196–233.

Song, Zheng, and Guiying Laura Wu, 2015, Identifying capital misallocation, Working paper, Nanyang Technological University.

Stiglitz, Joseph E., and Andrew Weiss, 1981, Credit rationing in markets with imperfect information, *American Economic Review* 71, 393–410.

Strebulaev, Ilya A., and Toni M. Whited, 2012, Dynamic models and structural estimation in corporate finance, *Foundations and Trends in Finance* 6, 1–163.

Syverson, Chad, 2011, What determines productivity? *Journal of Economic Literature* 49, 326–365.

Wu, Guiying Laura, 2018, Capital misallocation in China: Financial frictions or policy distortions? *Journal of Development Economics* 130, 203–223.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.