NBER WORKING PAPER SERIES

GOOD DISPERSION, BAD DISPERSION

Matthias Kehrig
Nicolas Vincent

Working Paper 25923
http://www.nber.org/papers/w25923

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2019, Revised March 2024

This paper previously circulated under the title "Do Firms Mitigate or Magnify Misallocation? Evidence from Plant-level Data." We thank Andrew Abel, Rudi Bachmann, Christian Bayer, David Berger, Ilan Cooper, Russ Cooper, Andrea Eisfeldt, Xavier Giroud, Jo˜ao Gomes, John Haltiwanger, Hugo Hopenhayn, Pete Klenow, Virgiliu Midrigan, Ben Moll, Gordon Phillips, Vincenzo Quadrini, Esteban Rossi-Hansberg, Cian Ruane, Yongseok Shin, Michèle Tertilt and Venky Venkateswaran as well as seminar and conference participants at many places for helpful comments and discussions. Vincent gratefully acknowledges financial support from the Fondation HEC Montréal. Any views expressed are those of the authors and not those of the U.S. Census Bureau or of the National Bureau of Economic Research. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2210.(CBDRB-FY24-P2210-R11148) The latest version of this paper can be downloaded at https://sites.google.com/site/matthiaskehrig/research.

Good Dispersion, Bad Dispersion
Matthias Kehrig and Nicolas Vincent
NBER Working Paper No. 25923
June 2019, Revised March 2024
JEL No. E2,G3,L2,O4

ABSTRACT

We document that most dispersion in marginal revenue products of inputs occurs across plants within firms rather than between firms. This is commonly thought to reflect misallocation: dispersion is “bad.” However, we show that eliminating frictions hampering internal capital markets in a multi-plant firm model may in fact increase productivity dispersion and raise output: dispersion can be “good.” This arises as firms optimally stagger investment activity across their plants over time to avoid raising costly external finance, instead relying on reallocating internal funds. The staggering in turn generates dispersion in marginal revenue products. We use U.S. Census data on multi-plant manufacturing firms to provide empirical evidence for the model mechanism and show a quantitatively important role for good dispersion. Since there is less scope for good dispersion in emerging economies, the difference in the degree of misallocation between emerging and developed economies looks more pronounced than previously thought.

Matthias Kehrig
Department of Economics
Duke University
237 Social Sciences, Box 90097
Durham, NC 27708
and NBER
matthias.kehrig@duke.edu

Nicolas Vincent
HEC Montreal
Institute of Applied Economics
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec), H3T 2A7
Canada
nicolas.vincent@hec.ca

1 Introduction

A considerable body of recent research has documented a large, persistent and ubiquitous degree of productivity dispersion across production units, leading to a revival of interest in the causes and consequences of resource misallocation. In their seminal work, Hsieh and Klenow (2009) argue that resource reallocation in China or India that would render dispersion similar to that in the United States would yield substantial aggregate output gains. Multiple papers have since expanded and refined their methodology. Yet, the common view in this literature remains that a high level of productivity dispersion is a sign of resource misallocation and therefore reduced welfare; i.e., dispersion is "bad."

We challenge this interpretation: we show that the relationship between frictions and dispersion is non-monotonic. Specifically, in the context of a simple conceptual model, we highlight situations in which *relaxing a friction* or constraint leads to increased output through a more efficient allocation of resources, despite generating *higher dispersion* of marginal products of inputs. In this context, more dispersion is "good." As a consequence, the level of fundamental frictions is non-monotonically related to empirically observable productivity dispersion, rendering the latter a potentially misleading indicator of misallocation. While this insight is broad and could be applied in other contexts, we focus on the role of multi-plant firms in shaping economy-wide productivity dispersion. This setting is particularly relevant for two reasons.

First, even if the literature has historically made little distinction between them, firms and plants are fundamentally different institutions. While firms compete for resources in markets, they act as planners in allocating these resources across their plants. This latter allocation activity is economically predominant: For instance, multi-plant firms account for most of aggregate output and investment in the U.S. manufacturing sector. As such, work studying the sources of productivity dispersion should take into account how firm-internal decisions differ from those across firms.

Second, using plant-level data from the U.S. Economic Census, we document that almost two-thirds of the overall dispersion in marginal revenue products of capital originates across plants operated by the same firm rather than between firms. This novel empirical finding highlights the importance of paying attention to the role played by the allocative decisions of firms. It also poses a theoretical challenge: Why do firms tolerate surprisingly high levels of dispersion in marginal revenue products across their plants and forgo seemingly large output gains? Why do they not reallocate resources across their plants in order to reduce within-firm productivity dispersion? Does the firm allocate resources less efficiently than markets? Understanding how firms work differently from markets may therefore shed new light on the causes of resource (mis)allocation.

To answer these questions, we build a quantitative model of capital allocation for a multi-plant firm that faces various constraints. Beyond its multi-plant nature, our framework is standard and embeds many frictions and imperfections that have been suggested in models of investment (see Caballero (1999) for an overview). It includes "technological" frictions, such as a fixed investment adjustment cost (see, among others, Abel and Eberly (1994), Caballero et al. (1995), Doms and Dunne (1998), Cooper and Haltiwanger (2006), Gourio and Kashyap (2007) and Baley and Blanco

(2023)), which are more relevant at the level of the plant, as well as external financing constraints (see, among others, Fazzari et al. (1988) or Gilchrist and Himmelberg (1995)), which affect the firm as a whole.[1] In our model, the firm organizes internal and costly external financing across plants that face both fixed and convex costs of adjustment. The optimal allocation of financial funds to investment projects is shaped by the interaction of frictions at the plant level, the firm level and across plants within the firm. The model is calibrated using moments from the Annual Survey of Manufactures collected by the U.S. Census Bureau and standard parameter values in the literature.

Using our model simulations, we then address the key question of our study: Can higher dispersion be associated with a more efficient allocation? In our setting, the answer is unambiguously "yes." To assess the potential role of this *good dispersion*, we compare the dispersion of marginal products of capital as well as aggregate quantities in our model economy as we vary the ability of the firm to pool and reallocate funds across its plants. As expected, we find that shutting down internal capital markets leads to lower aggregate capital and output, to the order of 12% and 8% respectively. More surprisingly, this less efficient allocation of resources is accompanied by a variance of the marginal products of capital that is about 40% *lower* than in the economy with with fully functional internal capital markets.

While this non-standard relationship between economic activity and dispersion may seem surprising, the forces behind it are intuitive: When firms are constrained in their access to external funds, they leverage internal capital markets and focus investment on only a few plants *even if the expected rate of returns of all plants are identical*. In the next period, the firm will concentrate its internal financial resources on another set of plants. Such an internal reallocation of funds will continue until the firm has carried out investment projects in all its plants. This "staggered investment" policy therefore leads to a rise in the dispersion of both investment rates and marginal revenue products of capital within the firm. In the stochastic equilibrium with frequent shocks, this *good dispersion* is persistent and never vanishes. In addition, these firms exhibit less correlated investment rates across their plants than their unconstrained counterparts.

Note that, while we illustrate our mechanism in the context of constraints in external financial markets, it can arise due to any firm-level factor in scarce supply: for instance, headquarter support such as engineers and implementation managers might be limited to only a few plants in a given year, or costly information acquisition may force the marketing team to focus on one market segment at a time, while remaining at first rationally inattentive to others. In Section 2.2 we illustrate a large set of factors that could result in the staggered investment policy and generate *good dispersion*.

Our quantitative exercises suggest two things. First, the interaction of plant- and firm-level frictions is quite powerful in generating substantial within-firm heterogeneity and dispersion in general. Second, the dispersion of marginal revenue products of capital is not necessarily an indicator of misallocation or inefficient investment policies. In our context, *good dispersion* arises because the firm manages to partially overcome external financial frictions and reallocate resources to a limited

[1] Gomes (2001), Khan and Thomas (2013) and Eisfeldt and Muir (2016) are examples of papers that combine real and financial frictions in a unified model of a firm operating a single plant.

extent. These constraint-efficient reallocation decisions result in a second-best allocation of the firm's resources, which increases output and welfare. Hence, we view our results as a cautionary tale about the risks of interpreting higher productivity dispersion as a sign of resource misallocation.

We then turn to plant-level data from the U.S. Annual Survey of Manufactures and find empirical support for the model predictions. To proxy for frictions to internal capital markets, we rely on the work of Giroud (2013) and Gumpert et al. (2022). We interpret the geographic distance between plants within the firm as impacting a headquarter's ability to gather information on production units, evaluate their capital needs, efficiently assess the optimal allocation of funds and handle the managerial implications of reshuffling cash flows across plants. We then show that as within-firm distance increases and the firm faces more significant frictions to its internal capital market, the level and dispersion of revenue products, the need for external funds and the cross-plant investment dynamics evolve empirically in line with what our model predicts.

Our findings have potentially important implications for the literature on resource allocation. Starting with Hsieh and Klenow (2009), a large body of work has studied productivity dispersion across plants. This literature typically envisions frictions that by their nature increase dispersion in marginal revenue products, decrease output and therefore fall under our label of *bad dispersion*; it does not explicitly entertain the presence of *good dispersion* that we described above. Note that this distinction does not depend on the type of friction. Even if they are technological in nature, such as fixed investment cost and time-to-build (Asker et al. (2014)) or partial irreversibilities (Baley and Blanco (2023)), a tightening of the frictions leads to more dispersion and lower output. In our setting, we show instead that tightening frictions may lead to less dispersion and vice versa. Bayer et al. (2015) distinguish the long-term from the temporary components of dispersion, while Buera et al. (2011), Khan and Thomas (2013) and Moll (2014) have studied the impact of financial frictions on misallocation, usually ignoring their effects within firms that operate several plants. In related work, Midrigan and Xu (2014) study the effects of financial frictions on firm entry and factor misallocation across firms. While we abstract from the entry channel, a modified version of the latter effect is present in our analysis, albeit in a framework in which financially constrained firms operate several plants and can overcome external financial frictions by internally reallocating financial resources.

Building on these insights, we argue that Hsieh and Klenow (2009) may in fact have *under-estimated* the gains from reallocation in emerging economies, such as India or China. When we make the U.S. economy more comparable to that of emerging economies by populating it with single-plant instead of multi-plant firms, dispersion in the U.S. is even smaller relative to that of India and China. This, in turn, implies that the distortions inferred to explain the observed dispersion and thus output losses from a distorted resource allocation are arguably higher than initially estimated. A quantitative exercise suggests that previous work may have missed between one-tenth and one-half of the output benefits from reallocation because it ignored the beneficial effects of reallocation within firms. This efficiency gain is distinct from comparable studies, such as Bau and Matray (2023), in which a market liberalization reduced *bad dispersion*.

Ultimately, we also see our project as a first step toward modelling how the organizational structure of a firm may impact the micro-level adjustment of capital, as well as understanding the role of firms for efficiency. While most research ignores the within-firm dimension of decision making, some theoretical research has been done on the efficiency of internal versus external capital markets: Stein (1997) and Malenko (2016) study mostly principal-agent problems between a firm's owner and manager in a single-plant setup. Gertner et al. (1994); Scharfstein and Stein (2000) show that division managers may exploit imperfect monitoring by firm headquarters to build up "inefficient empires," resulting in lower firm values Rajan et al. (2000). Doerr et al. (2021) show that the competitive pressures from international trades reduce misallocation within multi-unit firms. Through the lens of these frameworks, diversified firms with complex internal capital markets will suffer from more acute agency problems, further misallocation and, consequentially, more *bad dispersion* within firms. While all these factors may be present, we are the first to explore the potential role of *good dispersion* in this context.

Eisfeldt and Papanikolaou (2013) stress the importance of organizational or intangible capital at the firm level in order to understand a firm's productivity, albeit without the multi-plant dimension we are interested in. With the exception of Lamont (1997), Schoar (2002), Giroud (2013), Matvos and Seru (2014), Giroud and Mueller (2015), Doerr et al. (2021) and Kabir et al. (2023), empirical research on within-firm dynamics is scarce and often limited to studying major business divisions of conglomerates. An exception is Giroud and Mueller (2019), whose work is closely related to ours. They show empirically how local shocks propagate through the firm's internal organization, and that the reaction of other establishments is only significant if the parent is financially constrained. Finally, while we take the organizational structure of the firm as given, Ševčík (2015) considers the endogenous formation of multi-plant firms (which he calls "business groups").

Our paper is organized as follows. In Section 2, we show evidence on the importance of the within-firm dimension for the dispersion of marginal revenue products of capital and investment; then we illustrate theoretically how relaxing frictions within a firm may increase rather than decrease dispersion. Section 3 describes our multi-plant model of a firm that faces an external financing constraint. Section 4 conducts various quantitative exercises geared toward understanding the nature of productivity dispersion and provides supporting evidence. Section 5 presents an application to the setting in Hsieh and Klenow (2009), and Section 6 concludes.

2 Motivation

In this section, we motivate both empirically and theoretically our subsequent quantitative work. As discussed in the introduction, many studies have documented the ubiquitous presence of a large and persistent dispersion of marginal revenue products of inputs across production units. We first show empirically that in U.S. manufacturing, the majority of the dispersion in both (log) marginal revenue products of capital ($mrpk$) and investment rates (i/k) occurs across plants within firms rather than across firms. In the standard economic models generally used in the literature on misallocation, reallocating capital from low-$mrpk$ to high-$mrpk$ plants through investment activity reduces $mrpk$ dispersion and increases aggregate output.[2] Yet as we show with the help of a simple framework in Section 2.2, the opposite may be true: Relaxing frictions within firms may lead to more dispersion of the marginal revenue products of capital. In the quantitative model of Section 3, we show that our empirical finding could be interpreted as the outcome of an improved allocation rather than as evidence of a suboptimal allocation of resources within the firm.

2.1 Empirical motivation: Dispersion within and across firms

Data sources and variables of interest Our data source is collected by the U.S. Census Bureau in the Annual Survey of Manufactures (ASM), which is an annual dataset covering manufacturing businesses described in detail in Appendix A.2. The Census Bureau collects its manufacturing data at the level of an "establishment," which is defined as a physical business unit at a single location for which the primary activity is production. In this paper, we generally refer to establishments as "plants." Each plant also carries information about its parent firm, which is defined by Census as a collection of plants under common ownership or control.

Following the literature, we assume a Cobb-Douglas production technology, which is common for all plants in a 4-digit NAICS industry, and approximate the marginal revenue product of plant n in year t with its real value added per unit of capital.[3] We study the variance of its logarithm, $V_t\left(mrpk_{nt}\right) \equiv V_t\left(\log MRPK_{nt}\right) = V_t\left(\log(y_{nt}/k_{nt})\right)$, within a 4-digit NAICS industry and aggregate industries using value-added weights, as detailed in Appendix A.5.

Dispersion in U.S. manufacturing First, overall dispersion in marginal revenue products across plants is large, as shown in Table 1. In the average industry and year, the standard deviation of its logarithm is 0.9. This means that a plant that is one standard deviation above the mean produces $e^{0.9} \approx 2.5$ times the value added as the average plant with the same capital stock; the difference between the plant at the 90th percentile and that at the 10th percentile even implies an $e^{2.12} \approx 8.3$-fold value-added difference given the same capital stock.

Interestingly, Table 1 also indicates that the cross-sectional distribution of $mrpk$ is positively skewed. The standard coefficient of skewness is 0.634, while the quantile-based Kelley skewness measure is 0.128 on average.[4] The latter moment implies that the top half of the inter-decile range, $mrpk^{90} - mrpk^{50}$, is about 29% more spread out than the bottom half, $mrpk^{50} - mrpk^{10}$. As we will later argue, this evidence is supportive of some of our modeling assumptions.

[2] For details on the assumptions underlying this, see Appendix A.1. Based on Hsieh and Klenow (2009), the misallocation literature usually postulates equalizing revenue total factor productivity ($TFPR$). Like Asker et al. (2014), we focus on the capital allocation problem and hence equalizing $mrpk$. In that context, investment should not necessarily flow toward units with the highest $TFPR$ if they already operate a large capital stock; it should flow to units with the highest expected capital return.

[3] Though we study average rather than marginal revenue products of capital, we consider their difference in Appendix A.6.2.

[4] Following Kelley (1947), p. 250, we define the Kelley skewness as $\gamma^{Kelley} = \frac{mrpk^{90}+mrpk^{10}-2mrpk^{50}}{mrpk^{90}-mrpk^{10}}$.

Table 1: Cross-sectional moments of capital and investment

Variable	*Cross-sectional moments*					
	Mean	StD	IDR	Skew-ness	Kelley Skewn.	Excess Kurtosis
mrpk		0.905	2.120	0.634	0.128	1.978
		(0.013)	(0.032)	(0.028)	(0.010)	(0.085)
i/k	0.112	0.362	0.175	6.113	0.479	57.204
	(0.015)	(0.093)	(0.008)	(0.099)	(0.008)	(2.378)

Note: Data consist of our benchmark panel comprising annual plant-level data from the ASM 1972-2009. Moments are computed in a given year and 4-digit NAICS industry first before being aggregated by industry and then averaged across years. For details see Appendix A.4.

Investment rates also differ substantially across plants, which implies that the allocative activity of capital differs greatly across units within a typical year and industry. The cross-sectional standard deviation of 36% is large, given that the average plant in the economy has an investment rate of 11.2% – an indication of the well-known lumpy nature of investment. This also makes investment rates highly leptokurtic, which is reported in the last column of Table 1.

Under the standard interpretation of *mrpk* dispersion as evidence of misallocation, reallocating capital to high-*mrpk* plants in the same industry could hence result in a considerable boost in aggregate output. We show next that the majority of this dispersion occurs across plants within firms rather than across firms.

Dispersion within and across firms We decompose the total variance of marginal revenue products of capital, denoted by V_t, into two components: the variance between firms, denoted by V_t^B, and the average variance between plants within firms, denoted by V_t^W. To compare sufficiently similar units, we perform our analysis within 4-digit NAICS industries. This means we break up diversified conglomerates along industry lines, thus reducing the scope of actual within-firm dispersion. Our results should thus be regarded as a lower bound on within-firm dispersion. In a given 4-digit NAICS industry, our variance decomposition is then:

$$V_t\left(mrpk_{nt}\right) \equiv V_t = \underbrace{\sum_j \omega_{jt}\left(mrpk_{jt} - mrpk_t\right)^2}_{V_t^B \text{ average between-firm}} + \underbrace{\sum_j \omega_{jt} \underbrace{\sum_{n\in j}^{N_j} \omega_{nt}^j \left(mrpk_{njt} - mrpk_{jt}\right)^2}_{V_{jt}^W \text{ within firm } j}}_{V_t^W \text{ average within-firm}}. \tag{1}$$

The variable $mrpk_{njt}$ denotes the logarithm of the marginal revenue product of capital of plant n belonging to firm j in year t; $mrpk_{jt}$ the average *mrpk* in firm j in an industry; and $mrpk_t$ the average *mrpk* in a given industry. ω_{njt} is the weight of plant n at time t, ω_{jt} that of firm j and $\omega_{nt}^j = \omega_{njt}/\omega_{jt}$ that of plant n just inside firm j. While unweighted dispersion is our benchmark, we also consider capital-weighted dispersion to account for economic relevance. In the former case, we have $\omega_{njt} = 1/N_t$ (where N_t is the number of observations), while in the latter $\omega_{njt} = k_{njt}/k_t$ and accordingly for ω_{jt} and ω^j_{nt}. More details about this decomposition can be found in Table A3 in the appendix, while the main results are displayed in Table 2.

Table 2: Share of variance within firms in moments of capital and investment

Variable	*A. Full MUF panel*	*B. Single-product firms*
mrpk	0.601	0.559
i/k	0.679	0.689

Note: The data underlying Panel A are our benchmark panel comprising annual plant-level data from the ASM 1972-2009. Moments in Panel B are based on a subsample of firms in which all plants produce the same and a single at the 7-digit SIC product level as defined by Foster et al. (2008). Moments are computed for each industry and years first before being aggregated by industry and then averaged across years. For details, see Appendix A.4.

The main takeaway from our accounting exercise is that for the full sample (Panel A), about 60% of the dispersion of marginal revenue products of capital and 68% of the dispersion in investment rates in a typical industry occur within firms, with the remainder accounted for by between-firm variations.[5] The within-firm dispersion of *mrpk* is economically large: A plant that is one standard deviation ($0.702 = \sqrt{0.601 \times 0.905^2}$) above the firm's average produces twice the value added with the same capital stock as a plant that would reflect the firm average.

One might worry that this result is driven by residual product heterogeneity within 4-digit NAICS industries. To alleviate this concern, we repeat the decomposition, but this time focus on plants that produce only one good. Ideally, one would find physically perfectly homogeneous goods to eliminate the possibility of heterogeneous production technologies driving dispersion. We make a step in that direction and follow Foster et al. (2008) to consider industries that produce physically more homogeneous goods, such as cement, sugar or carbon black, etc.[6] Even if we focus solely on these more homogeneous 7-digit SIC industries, the within-firm share of dispersion in marginal revenue products of capital and investment rates displayed in Panel B amounts to 56% and 69%, respectively. This means that the high within-firm share of dispersion does not reflect mechanical aggregation across heterogeneous products within 4-digit NAICS industry firms. This exercise, along with additional robustness checks, can be found in Appendix A.7.

In light of the misallocation literature (see Hsieh and Klenow (2009)), our finding that most dispersion occurs within firms may appear surprising: It seems to suggest the presence of particularly large frictions within firms, rendering them an inferior allocation mechanism. In the next

[5] As Appendix A.3 shows, while multi-plant firms operate only 28% of all plants, they account for roughly 80% of aggregate economic activity.

[6] More specifically, these "industries" are defined by the following SIC product codes: Sugar (2061011), Block and Processed Ice (2097011 and 2097051), Gasoline (2911131), Hardwood Flooring (2426111), Concrete (3273000), Whole Bean and Ground Coffee (2095111 and 2095117 & 2095118 – later merged into 2095115 – and 2095121), Carbon Black (2895011 and 2895000), Bread (2051111, later split into 2051121 and 2051122) and Plywood (2435100, later split into 2435101, 2435105, 2435107 and 2435147).

subsection, however, we present a simple conceptual framework that shows how the opposite might be true: Relaxing frictions within the firm can increase the dispersion of marginal revenue products.

2.2 Theoretical motivation: *Bad dispersion* vs. *good dispersion*

Hsieh and Klenow's work on distortions and misallocation has been highly influential, spawning a myriad of studies on both the empirical and modeling fronts. Some have tried to map abstract distortions into empirically measurable market imperfections, often with the objective of quantifying potential output gains from eliminating specific imperfections. Others have attempted to clarify the distinction between imperfections and technological constraints (see, among others, Asker et al. (2014), David and Venkateswaran (2019), Baley and Blanco (2021) and Haltiwanger et al. (2018)). What is common among all these papers is that reducing frictions spurs beneficial reallocation, brings down the dispersion of factor revenue products, and increases aggregate output. Crucially, dispersion is always assumed to be inversely related to aggregate output: Economies with higher dispersion in factor revenue products are thought to be worse off (see Syverson (2011) for a summary of the academic consensus and Dabla-Norris et al. (2015), Cirera and Maloney (2017) and Cusolito and Maloney (2018) for the importance of that view in global policy making). In other words, dispersion is "bad" because it is a symptom of misallocation.

In this section, we show that the opposite can be true: While reducing frictions always improves factor allocation, it may *increase* dispersion rather than reducing it.[7] In this context, dispersion is "good" because it reflects better resource allocation. In what follows, we lay out the key aspects of our framework and analyze the relationship between dispersion and misallocation.

Framework. Consider a firm that invests in N plants subject to two constraints. First, each plant provides only a limited amount of funds x to finance investment generated by, for example, past profits. Second, a fixed investment adjustment cost implies that τ units of funds are lost for every plant the firm invests in. Production in each plant is given by

$$y = \begin{cases} k^{\alpha} & \text{if no investment} \\ (k+i-\tau)^{\alpha} & \text{if investment} \end{cases} \tag{2}$$

where k is the existing capital stock in a plant that can be augmented by investment.[8] Returns to capital are positive and decreasing, which is reflected in $0 < \alpha < 1$. The firm uses the total funds available, Nx, for investment activity across its plants in order to maximize the sum of output subject to the fixed adjustment cost and the technology in Equation (2).

We now use this simple framework to illustrate the complex relationship between frictions, misallocation and dispersion. First, let us define our two concepts of dispersion:

[7] Indeed, Bai et al. (2018) present a related finding: Reducing trade barriers in China has *worsened* misallocation.

[8] Though we assume the same k across all plants, our logic holds if we assume heterogeneous k.

Definition Dispersion in marginal revenue products, $V(mrpk)$, depends on the level of friction and is defined to be

- *good dispersion* if it decreases in the level of frictions,
- *bad dispersion* if it increases in the level of frictions.

Note that, while *good dispersion* and *bad dispersion* are local concepts, aggregate output is unambiguously decreasing in the level of frictions.

Bad dispersion. Let us first focus on the role of the fixed investment adjustment cost. If $\tau = 0$, the firm effectively incurs no penalty from investing only small amounts in each plant. As a result, the optimal course of action is to invest equally across all plants to equate their marginal revenue products of capital. By definition, dispersion of *mrpk* is therefore nil. As the friction tightens (τ rises), the firm trades off the concavity of returns (pushing it to equalize investment across its plants) against the fixed adjustment cost (pushing it to concentrate on a few plants). The optimal action is to pick a share of plants, denoted by $n^* < 1$, in which the firm will invest equal amounts $i^* = x/n^*$. As a result, the variance of *mrpk* increases to

$$n^*(1 - n^*) \left[\alpha \log(1 + i^*/k)\right]^2 > 0, \tag{3}$$

and firm output is lower.

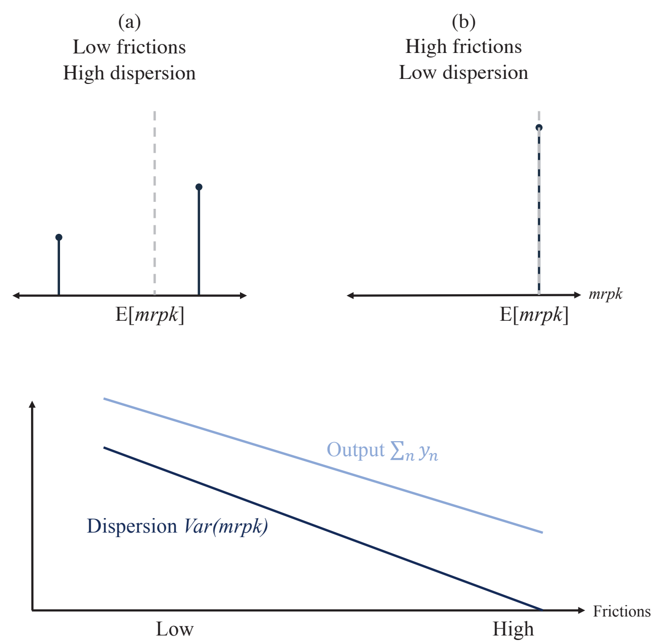
In sum, this example displays the standard relationship between dispersion and misallocation: As the friction is tightened, resource allocation moves further away from the unconstrained optimum and the dispersion of marginal revenue products of capital rises while output falls. In other words, more dispersion is *bad dispersion*. Note that it does not matter if τ represents a distortion as in Hsieh and Klenow (2009) or a fixed adjustment cost as in Asker et al. (2014).

Good dispersion. We now show that the interpretation of *mrpk* dispersion is, in reality, more complex and subtle than the impression given by the previous, standard argument. First, let us continue with the example of the fixed investment adjustment cost. Consider a starting level of τ so high that the firm decides to not invest at all, i.e., $n^* = 0$; the implication is that dispersion is again equal to zero, as was the case without any cost ($\tau = 0$). As τ is lowered, investing in a positive fraction of plants $n^* > 0$ becomes optimal. The outcome is a higher level of dispersion of *mrpk*, as can be seen from Equation (3), *as well as* more economic activity and output. In this case, as the friction is relaxed, resource allocation moves closer to the unconstrained optimum, and the dispersion of marginal revenue products of capital *mrpk* rises along with output. In other words, more dispersion is good.

Next, we turn our attention to the role of internal capital markets, holding $\tau > 0$ fixed. If frictions within the firm choke its ability to shift financial resources across its plants, effectively shutting down its internal capital market, only up to x funds can be used for investment in each plant. To simplify exposition, let's assume that $x < \tau < Nx$: The implication is that without the ability to pool funds, the firm cannot invest in any of its plants. Hence, without internal capital markets, the variance of (log) marginal revenue products is zero. This extreme case is depicted in Panel (b) of Figure 1.

What happens when we allow the firm to pool all Nx plant-level funds and reallocate them freely? Equipped with a functional internal capital market, the firm again trades off the concavity of returns (pushing it to equalize investment across its plants) against the fixed adjustment cost (pushing it to concentrate on a few plants). The optimal action is to pick a share of plants $n^* > 0$ in which the firm will invest equal amounts $i^* = x/n^*$. As a result, the variance of marginal revenue products increases to a positive value (see Equation (3)). The ex ante homogeneous plants, hit by the same shock, are now heterogeneous ex post *even though the allocation is more efficient*. This case is depicted in Panel (a) of Figure 1.

Figure 1: Frictions, output and dispersion



In this example, relaxing a friction and allowing for internal capital markets leads to a *higher* dispersion of marginal products of capital. In this sense, the nature of dispersion is again good.

More generally, we have shown that the relationship between frictions, misallocation and dispersion is non-monotonic.

Broader applications Even if our conceptual exercise were performed in an investment framework for a multi-plant firm, the underlying logic carries through in a broad range of settings. Hence, while the productive units in our example were plants, they could alternatively be business divisions, teams or even workers. The resources being pooled across productive units were financial funds, yet they could be interpreted as any firm-wide resource such as time, cognitive attention, technical knowledge or even managerial skills. Finally, the activity and its friction, in our case capital investment and the fixed adjustment cost, could instead be applied to a wide array of contexts:

- The introduction of new products subject to non-convex frictions, such as clinical trials for new drugs (see DiMasi et al. (2003)) or mandatory emissions regulations for new vehicles.
- Some innovation process that implies overhead costs of research and development as in Cohen and Klepper (1992, 1996) and Aw et al. (2011) or fixed start-up costs for research labs.
- The decision to enter a new export market, subject to export rules and regulations that may render exporting small amounts of goods unprofitable (see Melitz (2003), Das et al. (2007) and Creusen et al. (2011)).
- Business restructuring that requires a minimal fixed amount of attention or time from managers (Caliendo et al. (2015)).
- Hiring activity subject to non-convex frictions τ such as costs related to job postings and interview procedures (Mortensen and Pissarides (1994)).

In sum, any tax, transaction cost, trade barrier, cost of doing business, or menu cost that is less than proportionally[9] related to an input will have the effect we described.

It should also be noted that while all of the examples above rely on some non-convex frictions, they are not a necessary ingredient for our result: the key to generating *good dispersion* is that the firm finds it optimal to focus the activity on a subset of units. For example, in our original multi-plant firm setting, it would suffice to assume locally increasing marginal returns to capital – say, up to a capacity constraint. In this case, eliminating internal capital market frictions would push the firm to pool resources across its plants and redistribute them toward a few units for investment purposes, generating more dispersion of *mrpk* in the process. This means that the scope of our mechanism is broad, and so is the potential for *good dispersion*.

In the next section, we build a quantitative model of an economy in which firms operate several units and allocate capital across them. Capital is often pinned as the most distorted production factor (see Gopinath et al. (2017), among others) in explaining aggregate misallocation. We focus on multi-plant firms, as they account for the lion's share of economic activity in the U.S. economy. Importantly, that model features a rich set of frictions, as in Cooper and Haltiwanger (2006) and David and Venkateswaran (2019), which gives rise to the non-monotonic relationship between distortions and dispersion in marginal revenue products of capital.[10] Matching the rich model to establishment-level data from the ASM, we show that the introduction of frictions to internal capital markets leads to a *decrease* in the dispersion of marginal revenue products of capital.

[9] Proportional distortions, as in Hsieh and Klenow (2009), can be added without changing our argument.

[10] Brown et al. (2016) make a similar point, in which the non-convex distortion affects the discrete entry decision.

3 A model of finance and investment in multi-plant firms

In this section and the next, we describe, solve, simulate and analyze a simple model of a multi-plant firm. Our focus is on the role of within-firm frictions – more specifically, on those that regulate the functioning of internal capital markets – in shaping investment decisions and the dispersion of marginal revenue products of capital across plants within the firm. At one extreme, these frictions are so excessive that firms are merely a collection of disconnected production units: Decisions are made on a plant-by-plant basis. At the other extreme, firm management can fully use its frictionless internal capital market to mitigate or offset the other frictions and constraints that it faces. In particular, we show that as frictions to internal capital markets are eliminated, the firm optimally alters the size and timing of plant-level investment projects, which in turn generates more dispersion across its plants. While our focus in this application is on financing frictions, recall from the end of Section 2.2 that many other mechanisms would lead to similar results.

3.1 The problem of the firm

We focus on the basic problem of a firm that operates two plants n, where $n = A, B$. We limit our model to only two plants in an effort to keep the numerical analysis of our model, which features non-differentiable investment policies, computationally feasible. A larger number of plants would exponentially increase the size of the state vector of the firm, which must include the capital stock and technology level of each of its plants, without adding insight into the underlying fundamental economic mechanisms. We start by describing the technology and frictions at the level of the individual plant, before discussing the financing decision and analyzing the problem of the firm. In what follows, lowercase letters refer to plant-level variables, uppercase letters to firm-level variables and bold uppercase letters to vectors of a firm's plant variables.

3.2 Technology and frictions at the plant level

The plant operates a Cobb-Douglas production function that combines the beginning-of-period capital stock, k_{nt}, and other variable inputs in order to produce output, y_{nt}. While capital is fixed throughout the period, we assume that plants can freely choose any other variable inputs in perfectly competitive markets.[11] This means we can substitute out any static first-order condition for variable inputs and write revenue net of variable factor costs for plant n as

$$y_{nt} = e^{z_{nt}} k_{nt}^{\alpha}. \tag{4}$$

z_{nt} contains plant (log) total factor productivity and prices of other statically chosen production factors, while α is the revenue elasticity of capital, which embeds the production elasticity of

[11] Given our Cobb-Douglas production function, flexible factor markets will result in revenue products of flexible inputs that are completely equalized across plants and firms in the economy. This will not be the case, however, for marginal revenue products of capital since capital is chosen one period in advance and because of decreasing returns to scale as well as fixed investment adjustment cost.

capital, those of the static inputs, as well as the curvature of the demand function the firm faces. The productivity level of plant n in firm j consists of a component common to both plants in the firm and an idiosyncratic plant component: $z_{nt} = z_{njt} + z_{jt}$. The two components evolve as follows:

$$z_{njt} = \rho_p z_{njt-1} + \eta_{njt} \tag{5}$$

$$z_{jt} = \rho_f z_{jt-1} + \eta_{jt}. \tag{6}$$

where η_{njt} and η_{jt} are both *iid*, mean zero and have variances σ_p^2 and σ_f^2, respectively.

The capital stock of plant n depreciates every period at rate δ and grows with investment i_{nt}. This implies that its evolution over time corresponds to the conventional expression

$$k_{nt+1} = (1-\delta)k_{nt} + i_{nt} \tag{7}$$

As documented in a number of studies (see Cooper and Haltiwanger (1993), Cooper et al. (1999), Doms and Dunne (1998) and Caballero and Engel (1999), among others), investment dynamics at the plant level are characterized by lumpiness: multiple periods of inactivity (no or only small amounts of maintenance investment) are followed by "investment spikes." The traditional modeling feature used to reproduce this stylized fact is to introduce a fixed cost of investing: the firm must pay a certain cost, $\psi_1 k_{nt}$, if investment is non-zero. Such costs can arise because investment activity – no matter how small or large – has a disruptive effect on production activities in the short run, for example. The parameter ψ_1 regulates how much revenue is forgone when the plant needs to shut down production in order to install new capital. As a result of aggregation, firm-level investment activity will be less lumpy, as documented by Eberly et al. (2012). Baley and Blanco (2021) show that absent further frictions, fixed costs must differ for upsizing and downsizing and must significantly vary over time in order to explain capital dynamics across firms in Chile. We abstract from these additional margins as we incorporate more adjustment costs that shape to those investment dynamics.

In addition to this non-convex adjustment cost, we include a traditional quadratic adjustment cost. This convex adjustment cost captures the notion that larger investment projects become increasingly disruptive with size.[12] The parameter ψ_2 below captures the importance of this margin.

To summarize, frictions at the plant level are expressed as:

$$\theta\left(i_{nt}, k_{nt}\right) = \left[\psi_1 \mathbb{I}\left\{\frac{i_{nt}}{k_{nt}} > \vartheta\right\} + \frac{\psi_2}{2}\left(\frac{i_{nt}}{k_{nt}}\right)^2\right] k_{nt} \tag{8}$$

where $\mathbb{I}$ is an indicator function equal to 1 if the plant investment rate is above ϑ[13]; ψ_1 is a parameter regulating the forgone sales due to production disruptions if the plant undergoes any investment; and ψ_2 regulates the impact of the quadratic adjustment cost. Everything is scaled by the plant's

[12] This formulation is similar to assuming lower profitability during large capital adjustments, which has been documented by Power (1998) and Sakellaris (2004).

[13] In the simulations, ϑ is set to 1%.

capital stock k_{nt} in order to eliminate any mechanical effect from size.

Combining equations (4) and (8) above, plant cash flow is given by

$$\pi_{nt} = z_{nt} k_{nt}^{\alpha} - \theta(i_{nt}, k_{nt}). \tag{9}$$

3.3 Financing frictions and the internal capital market

While the benefits a firm provides to its plants are likely numerous, our focus in this application is on its ability to organize and allocate finance. As is often assumed in models of corporate financing, the firm can raise funds in external financial markets by issuing equity. Alternatively, it can pool the cash flows generated by its plants and reallocate these funds to their most productive use. Hence, reallocating funds via internal capital markets allows the firm to relax its external financing constraint. However, as we describe later, the actions of raising external finance and reallocating internal funds are both hampered by frictions.

All production and investment activities physically take place at the level of the individual plant. We assume, however, that it is the firm that coordinates investment projects across all of its plants, organizes financing of investment through either internal cash flow or external finance, and allocates funds to plants in line with planned investment projects. Our assumption that only the firm can organize external finance is realistic and sensible: while large and complex corporations such as General Electric operate hundreds of plants, only the firm issues bonds, borrows from banks or raises equity.

A firm's need for external financing, however, is indirectly determined by how functional its internal capital market is. In other words, how much financing a firm needs to raise externally is shaped by its ability to optimally reshuffle funds that are generated internally. With this in mind, we begin by characterizing two extreme cases. In the first scenario, the firm enjoys a frictionless internal capital market and can freely reallocate cash flows across its plants. In the other polar case, internal capital markets are nonexistent and investment in a given plant can only be financed by the cash flow it generates. Later, we allow for a situation that lies between these two extremes.

Frictionless internal capital market A firm with a frictionless internal capital market collects the cash flow from all of its plants and then decides freely how to allocate funds to finance investment projects across its plants. In this "Full Internal Capital Market" ($FullICM$) case, firm cash flow is defined as

$$\Pi_t = \pi_{At} + \pi_{Bt}. \tag{10}$$

In the event that desired firm-wide investment exceeds firm cash flow, the firm needs to raise external funds, denoted by ϕ_t:

$$i_{At} + i_{Bt} = I_t \leq \Pi_t + \phi_t. \tag{11}$$

Organizing external finance, however, is a costly and imperfect process. More specifically, we assume that the firm faces a cost of raising new equity, captured by the following function:

$$\Phi_t^{FullICM} = \left[\xi_1 \mathbb{I}\left\{\phi_t > 0\right\} + \xi_2 \max\{0, \phi_t\}\right] K_t. \tag{12}$$

This parameterization, which incorporates both a fixed and proportional cost of equity issuance, is in line with the one used in the empirical study of Bolton et al. (2021). Hennessy and Whited (2007) estimate a similar reduced-form specification to capture the role of underwriting fees and adverse selection costs that arise from issuing equity. The presence of the parameter ξ_1 is informed by the fact that equity issuance is lumpy, potentially reflecting fixed costs linked to the process.[14] The parameter ξ_2 reflects the notion that placing larger amounts of equity (relative to the firms asset size) gets more and more costly. This may result from a lower price of equity the firm can generate when it tries to sell larger amounts or higher commission fees for financial intermediaries placing that new equity.

Note that ultimately, investment activity at the plant level generates both direct and indirect costs. In addition to the amount of capital needed, there are both fixed and quadratic adjustment costs from investment activity but also potential costs associated with the issuance of new equity, in the event that the firm's internal cash flow do not suffice for its investment. Finally, investment activity at a specific plant may indirectly raise the cost of investing at the other plant as it depletes internal funds and imposes a financial constraint shared by the entire firm. In our model, the firm, as the organizer of funding and investment activity, internalizes this externality.

No internal capital market The specification described above represents the specific, extreme case of a firm with full, frictionless internal capital markets: the firm is able to costlessly shift funds generated by a specific plant to finance investment activity at the other plant. It can also freely allocate any new equity issued. There are various frictions and constraints that explain why this might not be the case, however. Stein (1997) and Scharfstein and Stein (2000), for example, emphasize how individual divisions within a firm compete for corporate resources in order to build "local empires," while Giroud (2013) has quantified the impact of imperfect information flows within a firm on the investment efficiency at the level of individual plants.

When these frictions become prohibitive, internal capital markets may cease to function altogether. In such extreme case, the firm effectively operates its plants as standalone units: financing investment activity at one plant by using the funds generated by another is not anymore an option, as funds are retained at the level of the plant. In other words, while plants may be generating positive cash flows in the aggregate, the firm might still need to raise external funds to finance investment at a specific cash-strapped plant. Hence, in this "No Internal Capital Market" ($NoICM$) extreme, the need for external funds is specific to plant n and defined as:

$$i_{nt} \leq \pi_{nt} + \phi_{nt} \quad \forall n \in \{A, B\} \tag{13}$$

[14] Hennessey and Whited (2007) also allow for convex costs, but find them not to be quantitatively important.

and the firm's cost of external financing is given by:

$$\Phi_t^{NoICM} = [\xi_1 \mathbb{I}\{\phi_{At} \text{ or } \phi_{Bt}\} > 0\} + \xi_2 \max\{0, \phi_{At}\} + \xi_2 \max\{0, \phi_{Bt}\}] K_t \tag{14}$$

While similar to the case of "Full Internal Capital Market" described in equation (12), the cost of raising external financing is now a function of the separate needs for funds at the plant level (ϕ_{At} and ϕ_{Bt}) instead of the aggregate need for funds at the firm level ($\phi_t = \phi_{At} + \phi_{Bt}$). Equation (14) captures the extreme case in which dysfunctional internal capital markets fully impede the firm's ability to shift financial resources across its plants, and Φ_t^{NoICM} is always at least as large as $\Phi_t^{FullICM}$.

Note that leaving aside the fixed financing cost which is paid only once, the firm in this scenario effectively operates its plants as standalone units: maximizing firm-level profits then boils down to separately maximizing the value of each plant in isolation.

Imperfect internal capital market Arguably, the situation of most multi-unit firms lies between these two extremes. In fact, even if a firm's headquarter was in theory able to pool and redistribute the cash flows generated by its plants, the process might be hampered by imperfect information or made costly due to limited management efficiency or the need for corporate oversight. In our quantitative analysis, we therefore allow for the firm to face partial internal capital markets by adopting a specification that weighs these two polar cases:

$$\Phi_t(\phi_{At}, \phi_{Bt}, \gamma) = (1 - \gamma)\Phi_t^{FullICM} + \gamma\Phi_t^{NoICM}. \tag{15}$$

We think of the parameter γ as capturing internal frictions within the firm that regulate how efficiently it can direct financing flows across its plants. Higher values of γ reflect a tighter level of internal capital market frictions: $\gamma = 0$ and $\gamma = 1$ correspond, respectively, to the extreme cases of full and no internal capital markets. When $0 < \gamma < 1$, the firm is uncertain about the internal capital markets frictions it will ultimately face, and therefore acts as if moving funds across its plants is possible yet costly.

More specifically, at the beginning of the period, the firm observes the productivity realizations z_{At} and z_{Bt}, and therefore learns the cash flow that will be generated by each plant. Based on this information, it chooses the level of investment at each of its units, i_{At} and i_{Bt}. However, while investment plans are made early on, for instance due to the presence of time-to-build constraints, financing activity only occurs later in the period. As a result, investment decisions must be made as a function of the *expected* cost of internal financing, according to equations (12), (14) and (15). That is, the firm recognizes that when funds need to be gathered to finance investment later in the period, there is a probability γ that it might be unable to pool cash flows from its two plants. It must therefore choose its plant-level investment based on the expected cost of external financing, given by a weighted average of the costs under the full and no internal capital markets, as given by equation (15).

3.4 Firm value and policy

We now describe the full problem of the firm. We define the vectors of technology levels and capital stocks within the firm as $\mathbf{Z_t} = \{z_{At}, z_{Bt}\}$ and $\mathbf{K_t} = \{k_{At}, k_{Bt}\}$, respectively. Given the plant-level fixed adjustment cost, the firm's state consists of the distribution of capital stocks, $\mathbf{K_t}$, and technology levels, $\mathbf{Z_t}$, across plants within the firm. The firm chooses investment in plants A and B in order to maximize firm value, which corresponds to the net present value of discounted future gross profits net of investment and financing costs. When deciding the investment level of each plant, the firm takes into account the various adjustment costs and whether external funds are required to finance the desired level of investment, conditional on the expected cost of external financing. The firm's problem can be written in recursive form as

$$V(\mathbf{Z_t}, \mathbf{K_t}) = \max_{i_{At}, i_{Bt}} \left\{ \Pi_t - I_t - \Phi_t + \beta \mathbb{E} V(\mathbf{Z_{t+1}}, \mathbf{K_{t+1}}) \right\}$$

subject to the revenue function (4); the shock processes (5) and (6); the law of motion of capital (7); plant n's investment adjustment cost $\theta(i_{nt}, k_{nt})$, defined in equation (8); the definition of the need and cost of funds under the $FullICM$ case, given by equations (11) and (12); the equivalent for the $NoICM$ case, given by equations (13) and (14); and the expression for the overall expected cost of external financing, equation (15).

Before moving to the quantitative analysis, it is worth highlighting a key trade-off faced by the firm in the model. This trade-off arises from the interplay between the plant-level adjustment costs, the firm-level financial constraints and the internal reallocation constraint γ. On its own, the fixed investment adjustment cost at the plant level nudges the firm towards lumping its investment activity through infrequent investment spikes. This lumpy investment policy, however, is in conflict with the firm's optimal financing policy: because external financing is costly, the firm would prefer to avoid raising external funds, which is harder to achieve in the presence of investment spikes which are likely correlated across plants due to the common firm component in TFP. The firm can resolve this tension, at least partly, by leveraging its internal capital market and adjusting the *timing* of its plant-level investment policies. By pooling the cash flows generated by its plants and by staggering investment spikes over time, the firm can limit the need to raise external funds. This means that frictions at the plant level (investment adjustment cost), the firm level (costly external financing) and within the firm (imperfect internal capital markets) interact to shape the timing and size of investment plans as well as the size and composition of internal and external finance.

In the next section, we study quantitatively the consequences of tightening the internal capital markets frictions faced by the firm, which is captured by raising the parameter γ in constraint (15).

4 Quantitative analysis

In this section, we first perform a numerical analysis of the model of the previous section to illustrate quantitatively our main insight: that eliminating a friction – in this case, the constraints to leveraging internal capital markets – can lead to both a more efficient allocation of resources *and* a rise in the dispersion of marginal products. Second, we provide empirical evidence that supports the main model mechanism.

4.1 Calibration

Table 3 summarizes the parameter values used to calibrate our model for the quantitative analysis. We strive to match as many of parameters in our unified multi-plant firm model as possible to the ASM data we are using to motivate our work. Most of the commonly used values are in line with calibrated parameters generally used in the investment literature.

Table 3: Model Calibration

Model param. & meaning		Value	Target/Source
β	Discount rate	0.95	Long-run real interest rate
α	Revenue elasticity	0.636	ASM: estimated for multi-unit firms
δ	Depreciation rate	0.089	ASM: long-run average i/k
ρ^p	TFP persistence plant	0.60	ASM: serial corr. $\log(mrpk)_p$: 0.25
ρ^f	TFP persistence firm	0.85	ASM: serial corr. of $\log(mrpk)_f$: 0.31
σ^p	TFP shock plant	0.25	ASM: volat. of $\log(mrpk)_p$: 0.33
σ^f	TFP shock firm	0.24	ASM: volat. of $\log(mrpk)_f$: 0.26
ψ_1	Fixed inv. adj. cost	0.039	ASM: Cooper/Haltiwanger (2006)
ψ_2	Quadr. inv. adj. cost	0.049	ASM: Cooper/Haltiwanger (2006)
ξ_1	Fixed equity issuance cost	0.00154	ASM: Frequency of equity issuance of SUF
ξ_2	Prop. equity issuance cost	0.11	ASM: Average size of ϕ in SUF
γ	Internal friction	[0,1]	vary between $FullICM$ and $NoICM$

Technology, shocks, depreciation and adjustment costs. In line with previous research, we set β at 0.95, which corresponds to an annual real interest rate of roughly 5%. To inform us about the production function elasticity, α, we extend the structural framework of Cooper and Haltiwanger (2006) to accommodate multi-plant firms and re-estimate plant-level revenue functions. Our GMM estimate puts α at 0.627, which is fairly close to the value they find. We reiterate that α is the revenue elasticity of capital, which embeds the production elasticities of capital and the static inputs as well as the curvature of the demand function. The parameters governing persistence, ρ^p and ρ^f, and volatility, σ^p and σ^f, of the plant and firm TFP shock processes are chosen to match the persistence and volatility of $mrpk$ at the plant and firm levels for the median firm among the sample of two-plant firms in the ASM. The depreciation rate δ is set to match the establishment-level average investment rate over time for equipment in our ASM data. For the fixed and convex investment adjustment cost parameters, ψ_1 and ψ_2, we rely on the structural estimates of Cooper and Haltiwanger (2006), which are somewhat smaller than the analogous values estimated by Asker et al. (2014).

Equity issuance. The literature provides limited guidance for the calibration of ξ_1 and ξ_2, which respectively capture the fixed and proportional costs of issuing new equity. While Hennessy and Whited (2007) estimate similar parameters using Simulated Method of Moments, their model also incorporates additional external financing frictions, making comparisons difficult. In addition, for the reasons discussed in Section 3.3, one can expect interactions between ξ_1, ξ_2 and the frictions to internal capital markets, captured by γ. To rule out interactions with γ, we find values of ξ_1 and ξ_2 that allow us the model to match the frequency and median size of equity issuance ϕ *for the sample of single-plant firms.* Focusing on single-plant firms allows us to make abstraction of frictions to the flow of financial resources within the firm, set $\gamma = 0$ and isolate the role of the costs of equity issuance in shaping the moments of the object ϕ. The parameter values we obtain can be found in Table 3.

4.2 Internal frictions, resource allocation and *good dispersion*

Before analyzing the forces at play inside the model, let us go straight to the main questions of the paper: Does the model generate a quantitatively relevant increase in *both* dispersion *and* output once the firm is allowed to pool the cash flows and capacities for external finance (capital stock) from its two plants? Can relaxing a friction be welfare-improving, yet at the same time generate *more* dispersion in marginal revenue products?

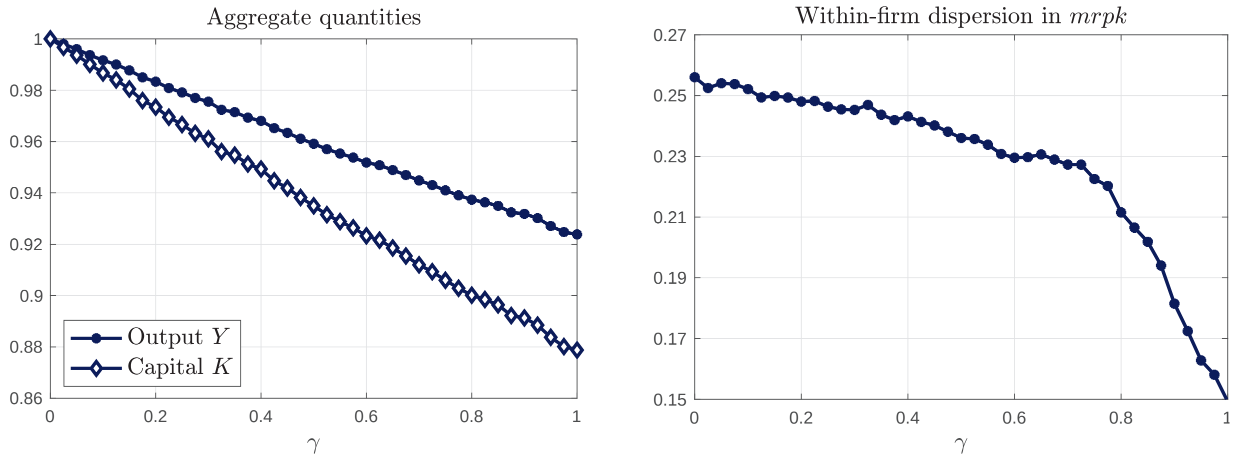
To answer these questions, we compare the quantitative predictions of our model for various levels of the internal friction parameter γ. Recall from Section 3.3 that in the model, this parameter captures the probability that the firm will not be able to take advantage of internal capital markets by pooling and redistributing firm-level cash flows, and instead need to raise costly external funds. This, in turn, will influence the firm's optimal investment policy at its two plants. Specifically, we vary γ from zero to unity.

For each value of γ, the problem of a representative firm is solved using a value function iteration procedure that is described in detail in Appendix B. We then simulate an economy composed of 1,000 two-plant firms for 1,000 periods to investigate two questions and analyze the impact of the internal capital markets friction on the interaction between the dispersion of marginal revenue products and aggregate misallocation in this model economy. Naturally, two-plant firms with functional internal capital markets cannot do worse than the $NoICM$ economy: they convexify the choice set and can always reproduce the allocation of an economy in which the pooling of resources within the firm is not possible. Our objective is to determine whether (1) these gains can be substantial and (2) dispersion of marginal revenue products is an indicator of misallocation.

The left panel of Figure 2 displays aggregate output and capital as a function of the parameter γ, while the right panel plots the average within-firm dispersion of expected marginal revenue products. To facilitate the presentation and comparisons, we normalize the values of aggregate output and capital to unity for the economy with frictionless internal capital markets $(\gamma = 0)$.

Focusing first on the left panel, we see an expected result: imposing tighter frictions on internal capital markets by raising γ leads to lower aggregate levels of capital and output. Specifically, when

Figure 2: Aggregate quantitative effects of internal capital market frictions



Note: This figure displays the quantitative effects of the internal capital market friction γ on dispersion and aggregates. Aggregate magnitudes are normalized to unity for the $FullICM$ economy ($\gamma = 0$). Without ICM ($\gamma = 1$), aggregate capital and output respectively drop by 12% and 8% compared to the frictionless benchmark. At the same time, despite the additional friction, firms make investment choices that generate *lower* planned *mrpk* dispersion across their plants than in an economy populated by firms with fully-functioning ICMs.

firms cannot shuffle funds across their plants ($\gamma = 1$), aggregate capital is about 12% lower than under the frictionless $FullICM$ case ($\gamma = 0$), while output falls by almost 8%. In other words, eliminating this internal friction is, as expected, welfare improving.

More surprising, however, is the result in the right panel, which displays the dispersion of marginal products as a function of γ. It indicates that tightening the frictions on internal capital markets (raising γ) not only depresses aggregate output, but also generates a *lower* dispersion of the marginal revenue product of capital. More specifically, an economy without internal capital markets ($\gamma = 1$) displays a within-firm dispersion of expected *mrpk* over 40% lower relative to the frictionless ICM case ($\gamma = 0$).[15] Note that we focus on the dispersion of *expected mrpk* to emphasize the fact that the higher dispersion of *mrpk* is a result of the firms' optimal choice of capital, and not only ex post plant-level TFP shocks as in Asker et al. (2014).

In sum, we have shown that a *more efficient* allocation can be accompanied by a *higher dispersion* of marginal revenue products. This *good dispersion* is a result of the firm's optimal actions in the face of a relaxation of the frictions it faces, and not a sign of misallocation. Next, we delve into the mechanism behind this *good dispersion*.

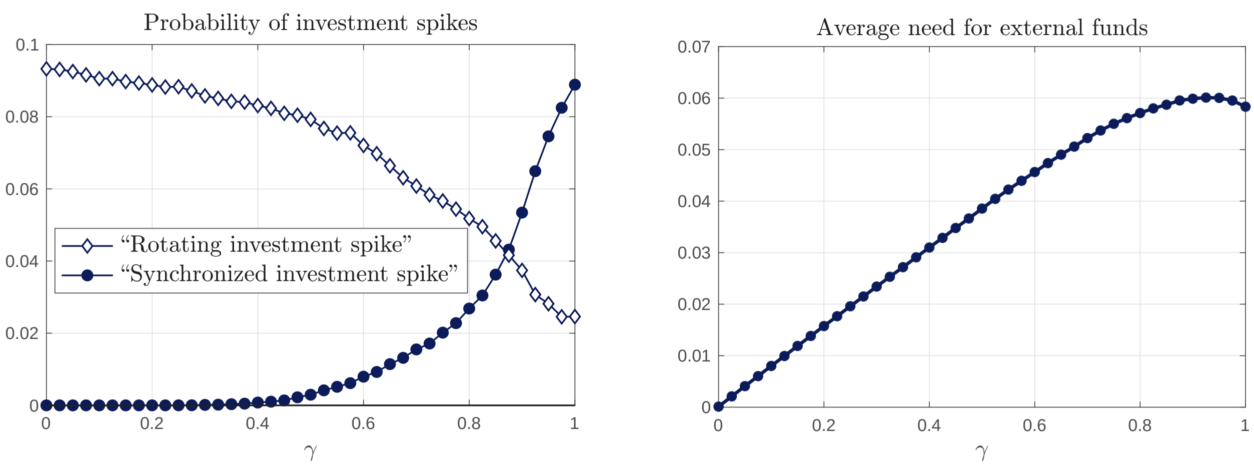
4.3 Why do firms create more dispersion? The mechanism

Our objective in this section is to show how the firm's optimal allocation of capital across its plants can generate *lower* dispersion of marginal products in the face of tighter frictions. To do so, we continue to solve and simulate our model for different values of γ.

[15] Total dispersion in our model economy also falls as γ increases.

Due to the presence of a fixed adjustment cost of investing, investment at the plant level is characterized by periods of inactivity interspersed with "investment spikes," a pattern that has been heavily documented in the empirical investment literature. These dynamics, in turn, have implications for financing activity: investment spikes are generally accompanied by an acute need for funds, which are costly when raised externally. Hence, while fixed adjustment costs encourage lumpy investment, costly external finance puts a brake on the desire for investment spikes. This is particularly true when these spikes are correlated across plants, say in response to a firm-wide productivity shock. In theory, the presence of an internal capital market within the firm offers an opportunity to reconcile these opposing forces through the appropriate timing of plant-level investment spikes.

Figure 3: Quantitative effects of financial constraints on the multi-plant firm economy



Note: This figure plots various moments of investment and the need for external funds and thir associated costs as a function of the internal friction γ.

To understand why, we display key moments from the simulations as a function of the ICM friction γ in Figure 3. The left panel focuses on the joint dynamics of investment across plants within the firm, while the right panel plots the expected need for external funds of the typical firm.

Let us start from the $NoICM$ scenario at the right end of the plot, when $\gamma = 1$ and firms cannot leverage internal capital markets. In this economy, funds generated in one plant cannot be used to finance investment at the other plant. In other words, the firm basically manages its plants as standalone units.

As can be seen in the left panel of Figure 3, under $\gamma = 1$, the probability of observing synchronized investment spikes within the firm (spikes occurring at both plants simultaneously) is close to 9%. This is a direct consequence of the fact that plant-level TFP has a firm-wide component. Hence, following a positive firm-level shock, the firm finds it optimal to invest in both plants: TFP, and therefore investment, are positively correlated across plants. Due to the presence of fixed adjustment costs, a corollary is that we can expect investment spikes (defined in our simulations as an investment rate of 15% or more) to be synchronized to some degree.[16]

We now turn to the financing needs of the firm that are implied by the optimal investment policy, illustrated in the right panel of Figure 3. When $\gamma = 1$, firms cannot pool the funds generated by their plants. As a result, both investment and financing decisions are essentially taken at the level of the plant: investment activity in one plant has no bearing on the optimal choice of capital and external financing at the other. Because plant TFP shocks are correlated, so are investment activities resulting in synchronized external financing needs, which are costly but impossible to avoid. In this scenario, the average need for funds is equal to about 6% of the firm's capital stock. As a result, in this $NoICM$ scenario, the average cost of raising external finance is equal to 3.5% of capital.

Next, we relax the ICM friction by lowering γ (i.e. moving towards the left in each plot of Figure 3). When the ICM frictions fall (γ declines), the firm gains the ability to pool and internally redistribute the funds generated by its plants and avoid costly external finance. However, to take full advantage of functional internal capital markets, the firm must strategically adjust the timing of investment spikes: when hit by a firm-wide productivity shock, the firm reacts by pooling cash flows and channels them to finance investment in only one plant at a time. Then, since shocks are persistent, it is likely to turn to financing an investment spike at the other plant in the next period. Indeed, as can be seen in the top-left panel of Figure 3, this "rotating investment spikes" pattern is much more prevalent when the firm faces less constrained internal capital markets (γ is low).[17] Conversely, as γ falls, the firm now finds it optimal to avoid costly external finance by avoiding synchronized investment spikes: the probability of observing synchronized investment spikes within the firm drops from 9% to 0% as γ goes from 1 to 0. This process of staggering investment spikes over time, in turn, creates higher within-firm and economy-wide dispersion of not only investment rates but also $mrpk$, as we saw in Figure 2.[18] Note that the firm has no incentive to strategically stagger investment activity when it has no access to internal capital markets: in this case, the firm sees no gain in desynchronizing investment and, as a byproduct, generating dispersion.[19]

As explained earlier, internal capital markets are valuable to the firm because they limit its need to access costly external financing. From the right panel of Figure 3, we can see that the firm almost never needs external funds in the $FullICM$ case ($\gamma = 0$), against an expected need equal to 6% of capital in the $NoICM$ scenario. [20] As a result, the average cost of external financing when $\gamma = 0$ is negligible at less than 0.5% of capital, compared to 3.5% when $\gamma = 1$. In sum, the presence of a functional internal capital market dramatically relaxes a firm's need for external funds and

[16] The fact that the probability is not higher is because (1) most periods see no spikes due to the lumpy nature of investment and (2) investment is not perfectly positively correlated due to the idiosyncratic plant-level TFP shock.

[17] Formally, a rotating investment spike is defined as a period in which only one of the two plants spikes, followed by a spike at the other plant in the subsequent period.

[18] In an economy in which firms and plants are constantly buffeted by shocks, this effect occurs frequently: We find a very similar pattern if we instead compute the dispersion of 5-year moving averages of logged expected $mrpk$.

[19] Besley et al. (1993) study Rotating Savings and Credit Associations (ROSCA), a financial institution mostly found in developing economies that relies on a similar mechanism.

[20] Moreover, when the firm does raise funds, the average size of external financing is less than 5% of capital in the $FullICM$ case, against 31% when ICM is not functional.

lowers its financing costs, freeing resources that can be reinvested in additional capital.

To summarize, we have shown that alleviating a friction, in this case to firms' internal capital markets, can result in a more efficient resource allocation *and* a higher degree of dispersion in marginal revenue products across producers. In other words, higher *mrpk* dispersion is not necessarily a sign of more severe misallocation: the increased dispersion is the optimal response of the firm to more relaxed constraints it faces rather than a consequence of inefficient distortions. The presence of such *good dispersion* is due to the interaction of the internal capital markets with two crucial frictions: the fixed investment adjustment cost ψ_1 and costly external finance embedded in ξ_1 and ξ_2. For example, in the absence of a cost to external finance, the firm would have unlimited access to financial resources and would invest to equalize *mrpk* across its plants, irrespective of ICM frictions: any dispersion in *mrpk* would be *bad dispersion* and merely reflect the presence of the technological fixed investment adjustment cost ψ_1, as in Asker et al. (2014).[21] Conversely, without fixed adjustment costs, a financially constrained firm would invest suboptimally low amounts, but would still allocate investment to equalize marginal revenue products across its plants.

Multi-plant firms and investment flow adjustment costs. Finally, it is worth briefly discussing the link between the mechanism at the core of our multi-plant firm model and the investment flow adjustment cost of the form $\kappa(I_t/I_{t-1})^2$ that is commonly used in DSGE macro models. While this reduced-form friction has been found to fit the aggregate dynamics of investment much better than other types of adjustment costs (see Christiano et al. (2005)), the literature offers little in the way of micro-foundations to rationalize its use (an exception is Lucca (2007)). Our model of a multi-plant firm provides a natural interpretation: The optimal capital allocation it generates can reconcile the lumpy investment dynamics at the plant level (Cooper and Haltiwanger (2006)) with the finding of Eberly et al. (2012) that investment for larger firms is smooth and compatible with an investment flow adjustment cost.[22]

4.4 Good dispersion: empirical evidence for the mechanism

Next, we document empirically that a number of predictions of the mechanism described in the previous section are borne out in the data.

Validation of calibration. As discussed in Section 4.1, our model is calibrated to target specific micro moments from the Annual Survey of Manufactures. The within-firm dispersion of marginal products, a concept central to the reallocation literature and our motivation, was not among the targeted moments. As a first step, we verify how well the model can fit its empirical counterpart, the dispersion of *mrpk* within two-plant firms.

[21] Similar results could be obtained in the case of a borrowing constraint that would see the interest rate paid by the firm increase with the size of its borrowing. Under such conditions, the firm would again wish to stagger investment projects in order to minimize the cost of borrowing.

[22] In our model, investment at the firm level is smoother than at the plant level. Hence, owners of multi-plant firms will not be exposed to volatile consumption, thus alleviating the concern of Thomas (2002) that general equilibrium forces would render plant-level investment spikes irrelevant for aggregate dynamics.

First, it is important to note that our model does not capture factors that could generate permanent or highly persistent productivity differences across plants, such as industry specificities, firm life-cycle dynamics, heterogeneity in production functions (e.g. heterogeneous α's) or markups (e.g. heterogeneity in the stochastic processes), etc. Therefore, we start by taking out plant-level fixed effects from *mrpk* to purge it of structural factors that may generate permanent dispersion in *mrpk* across plants within firms. We then recompute the variance of demeaned *mrpk* across plants within 2-plant firms. We find an average empirical within-firm variance of 0.214 across firms (against 0.492 for the non-demeaned data).[23] In our model, the comparable within-firm variance ranges from 0.149 to 0.255 depending on the level of internal capital market friction γ (details below).

Measuring the internal reallocation friction. In order to compare the model predictions to the data, we need a metric that can plausibly be viewed as a proxy for the degree of internal capital market friction. Yet, unlike the case of more conventional adjustment costs and financial frictions, there is very limited guidance on how to empirically measure the type of within-firm friction that we capture through the parameter γ in our framework

Ideally, a researcher would have access to firm-level information that could be used to proxy dimensions such as the information flow within firms, the effectiveness of corporate oversight and/or the extent of financial flows between establishments. Such information, however, is difficult to come by. Instead, we argue that the distance between plants within the firm is plausibly related to frictions to internal capital markets. This is not new: empirical work by Giroud (2013) and Gumpert et al. (2022) has suggested that distance may play a crucial role in all of these firm functions. Intuitively, the time it takes corporate headquarters to visit individual production sites, assess their productivity and evaluate their capital needs entails a cost to the firm, which is conceptually similar to our internal friction. In the spirit of this work, we therefore interpret geographic distance as a proxy for γ.

More specifically, we focus on how spread out a firm's plants are across the U.S. Our assumption is that if establishments are located close to each other, the plant-specific information relevant to the optimal allocation of resources is on average easier to acquire by headquarters, all else equal, and frictions to internal capital markets less prevalent. On the other hand, managing firms that are more spread out geographically involves additional travel and effort to acquire the information and accurately target financial flows to the right investment projects. For example, consider two multi-plant firms. The first one, with all its establishments in Saint Louis, MO, has its center point in that city and an average distance of 0. The other has its center in Saint Louis as well but is characterized by a within-firm distance of 780 km: half of its establishments in Kansas City, MO, and the other half in Louisville, KY. Our notion is that the first firm concentrated in Saint Louis is more likely to have a well-functioning internal capital market than the second one.

To measure distance between plants of a specific firm, we rely on the information contained

[23] In Table 2 we showed this corresponds to 69% of overall dispersion.

in our dataset (since 2007) about a plant's latitude and longitude. This measurement leaves us with about 140,000 firm-year observations.[24] For each firm we use the latitude and longitude of all the establishments to compute gradient vectors from the earth's center to establishment locations. Then, we compute the value added weighted average vector from the earth's center to a location on the earth's surface, which represents the center of the firm. Finally, we compute the value added weighted distance from each establishment to that center point to obtain the average distance within the firm. We choose this procedure because it is not only easy to implement for firms with any number of plants.

For our analysis, we take the logarithm of 1 plus the average distance between establishments and the firm center point as our proxy for internal frictions. This measure varies from zero, which reflects a firm with establishments in the same location, to a maximum of about 8.4, which corresponds to establishments that are about 4,400 km from each other. This would correspond to a firm with plants on both the West and the East Coast. Table 4 reports summary statistics of geographic spread of multi-units firms. The average firm features a distance of 4.7, which corresponds to plants being about 55 km away from the firm center. A firm with one standard deviation above the average has its plants 571 km away from its center. There is some mass point at zero leading to a slightly negatively skewed distribution.

Table 4: Summary statistics of within-firm distance

	Moment
Mean	4.70
St. Dev.	2.34
Skewness	−0.96

Note: Source: Authors' calculations of the average distance across plants within firms in the ASM panel.

With an empirical proxy for γ at hand, we assess how the friction affects various firm-level moments in both the model and the data. Specifically, we study how the variation of within-firm dispersion of average revenue products of capital; dynamic investment patterns consistent with either synchronized or rotating investment spikes; the associated dynamic patterns of changes in average revenue products; and the reliance on external funds for investment vary as a function of γ in the model and distance in the data. In the data, we limit our attention to the sample of firms operating exactly two plants. That way, we can compare our model more closely to its empirical analogue in the real world.[25] Empirically, all these moments are constructed at the firm-level and

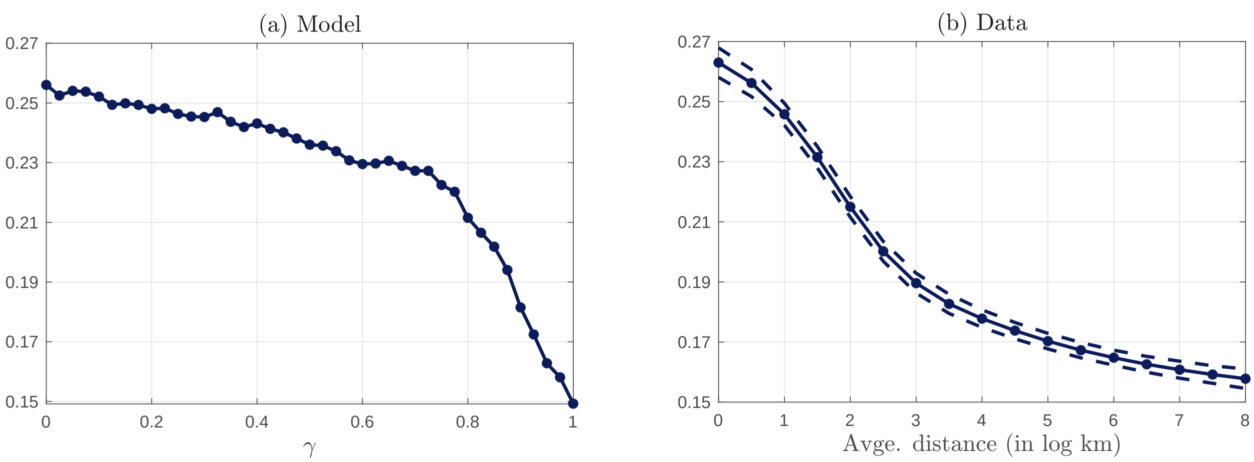
[24] More precisely, there are about 50,000 firm-year observations for which we directly observe the precise geolocation of a firm's plants. Because for a specific plant there is sometimes missing geolocation information in some years, we then use the stable plant identifier to roll that information backward and forward in time, as long as a plant is observed in the data. This greatly expands our sample of firms with precise geolocation codes, leaving us with about 140,000 firm-year observations. In addition, to corroborate our empirical findings below, we also used the economic center of a county of a plant's location as a proxy, which provides geolocation information for all 214,000 multi-unit firms across time. The results are all qualitatively similar if only slightly weaker quantitatively.

[25] We have also constructed the same set of moments for the universe of all multi-plant firms, which show similar patterns as the ones described in this section. We display those figures in Appendix C.

then purged in the same way as described above for the within-firm dispersion of *mrpk* by regressing them on industry, time and size dummies to control for structural factors that are not captured in our model. The residual variation is then projected on our distance measure. We center each moment at its unconditional mean.

Dispersion of *mrpk*. We start with the impact of the friction on the dispersion of marginal revenue products of capital across plants within the firm. The model predicts that a firm that faces external financial frictions but flexible internal capital markets will exhibit *more* dispersion of marginal products than an analogue firm with a limited ability to shuffle funds across its plants. As explained earlier, this is because firms find it optimal to substitute costly external finance with internal finance, even if it means staggering investment projects across its plans over time and hence creating dispersion of marginal products. We characterize this behavior as giving rise to *good dispersion*: in this case, higher dispersion is not a sign of misallocation, as it is associated with higher output.

Figure 4: Dispersion of *mrpk* within firms

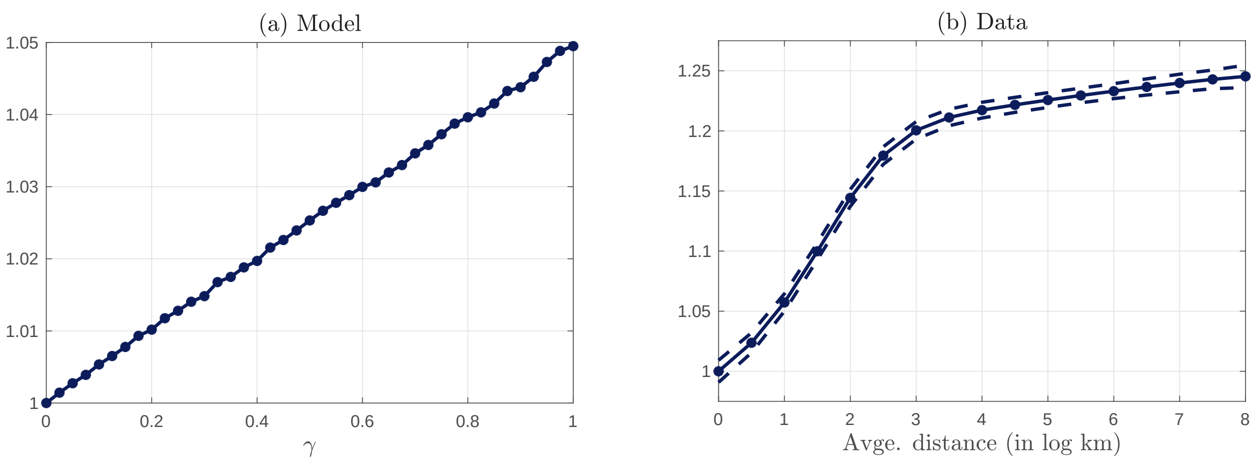


Note: Left panel: Dispersion of expected *mrpk* within firms as in Figure 2. The right panel plots its best empirical analogue, the variance of average revenue products of capital across plants within firms vis-a-vis geographic distance of the plants from the firm's center point. Underlying sample comprises the set of 2-plant firms in the ASM.

Figure 4 displays in the left panel the model-implied dispersion of *mrpk* as a function of γ: as shown and explained earlier, dispersion falls as the internal capital markets friction is tightened. In the right panel, we turn to the data and plot the relationship between within-firm dispersion and our distance measure. One may *a priori* expect that plants of more geographically concentrated firms would be more likely to be hit by similar shocks or share similar characteristics, and therefore exhibit lower dispersion of revenue products. Instead, we find empirically that the average variance of revenue products of capital within the firm varies *negatively* with distance: firms that are more concentrated geographically, and for which we therefore expect corporate oversight to be easier and more efficient based on the work of Giroud (2013), exhibit higher dispersion than firms that are geographically spread out, as our model would predict. The difference is quite significant, with the variance in geographically concentrated firms being about ten log points larger than their more spatially spread-out counterparts. Not only does the model predictions qualitatively match the pattern found in the data, but the variations are also quantitatively similar: at $[0.15, 0.25]$ the range in dispersion in the model (as a function of γ) is comparable to that found in the data (as a function of distance).

Level of firm *mrpk*. A higher level of *mrpk* is commonly interpreted as a symptom of a distorted firm (see Hsieh and Klenow (2009) among many others). In our model, γ acts a friction that raises the cost of internal finance. Therefore, as γ falls, the level of *mrpk* declines even while the dispersion of *mrpk* rises. In that sense, *good dispersion* in *mrpk* is accompanied by lower levels of *mrpk* as the firm is less distorted as a whole. *Bad dispersion*, on the other hand, would associate higher variance in *mrpk* with higher levels or *mrpk*. We show next that this relationship holds in the data.

Figure 5: Firm-level *mrpk*



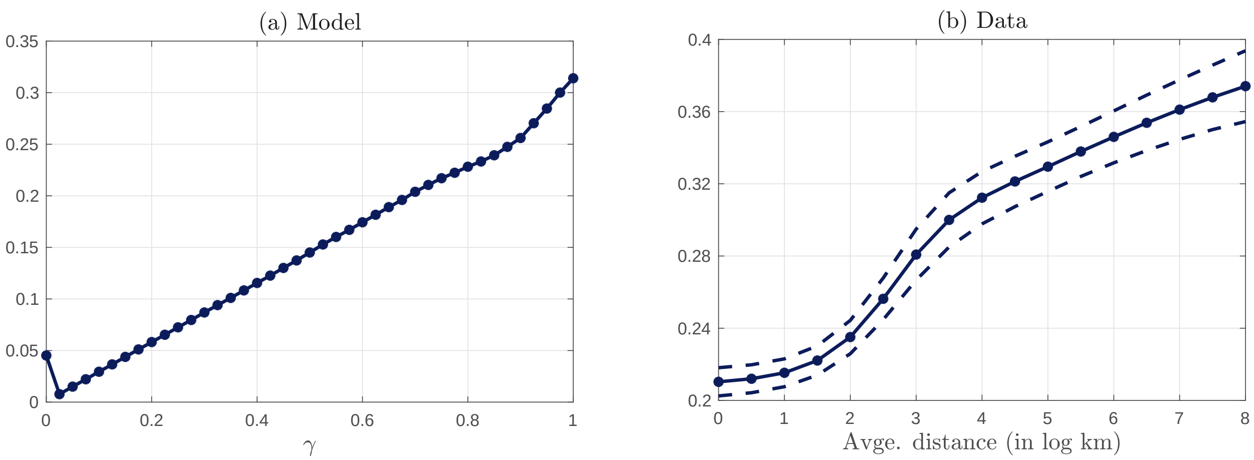
Note: See notes to Figure 4.

Figure 5 displays firm-level *mrpk* as a function of γ (model, left panel) or distance (empirically, right panel). For both the model and the data, we normalize the average firm-level *mrpk* at the left end to unity.[26] In the model, as one would expect, the level of *mrpk* increases as the firm finds it harder to reshuffle funds across its plants (γ rises). In other words, the dispersion in marginal products that is generated as γ is lowered and the friction is eliminated (see Figure 4) corresponds to *good dispersion*: it is accompanied by more capital and a lower level of *mrpk*. Empirically, we find a very similar relationship, both qualitatively and quantitatively: a higher level of intra-firm friction (higher distance) is associated not only with higher dispersion (as shown in Figure 4), but also a higher, not lower, level of *mrpk*.

[26] Any arbitrary measurement of output or capital will impact the log level of *mrpk*. Note that because measurements are additive in the log level or *mrpk*, they do not impact the dispersion measure.

External finance. Next, we examine the impact of internal capital markets on external financing activity by the firm. Because external finance is costly, a firm with functioning internal capital markets (low γ) will try avoid raising external funds if it can. On the other hand, a firm with dysfunctional internal capital markets (high γ) will be forced to access external finance more intensively if it wants to finance valuable investment projects in its plants, despite the higher financing cost. Recall that in the model, the amounts of funds raised externally by the firm is captured by the variable $\phi \geq 0$, which measures the amount by which firm investment exceeds firm cash flow (i.e. external funding required to finance firm investment). This variable, too, can be measured in the data. We focus on the volume of external finance (which may be zero or strictly positive) rather than the expected need for external funds. While we can easily compute the latter moment in the model, it is much more challenging in the data, where we may observe a given firm only for a few periods. Note how the left panel of Figure 6 is a monotone transformation of the expected average need for external funds displayed in the right panel of Figure 3.

Figure 6: Volume of external financing conditional on external finance



Note: See notes to Figure 4.

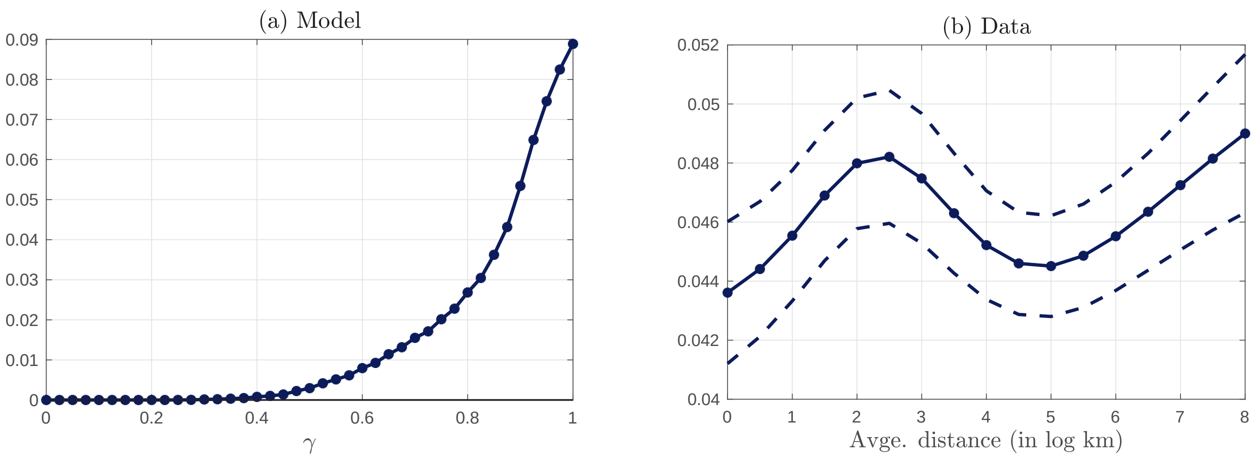
As the left panel of Figure 6 shows, the need for external funds is extremely limited when the firm can fully exploit internal capital markets ($\gamma = 0$). However, things are very different when the firm does not have the ability to shuffle funds around its plants: when $\gamma = 1$ and funds cannot be pooled across plants, the typical firm can expect to raise external funds equal to about 30% of its capital stock (conditional on needing funds).

The right panel in Figure 6 shows that this pattern is similar in the data: conditional on needing external funds, the size of external funding is clearly increasing in the average distance between plants within the firm. The increase in the data is not quite as dramatic as in the model, but still shows a strong and significant increase.

Investment activity. Finally, we turn our attention to the dynamics of investment within the firm. As noted in our earlier discussion of the model mechanism, when firms cannot pool funds and have no other option than to finance investment externally, they have no reason to stagger investment activity across their plants. Instead, driven by the common firm-specific component of TFP, they tend to synchronize investment spikes. Conversely, firms with well-functioning internal capital markets prefer to stagger investment activity in order to save on the external financing cost, i.e. choose investment plans that are asynchronized.

We assess these dynamic patterns first by comparing the relationship between synchronized investments within multi-unit firms, the share of plants which experience and investment spike, i.e. investment rates in excess of 15%. Then, we will look at the synchronization of the changes of investment rates as well as the associated synchronization of changes of *mrpk*s across plants within firms. If investment plans tend to be synchronized, the variance should be lower and vice versa for staggered investment policies.

Figure 7: Probability of synchronized investment spikes

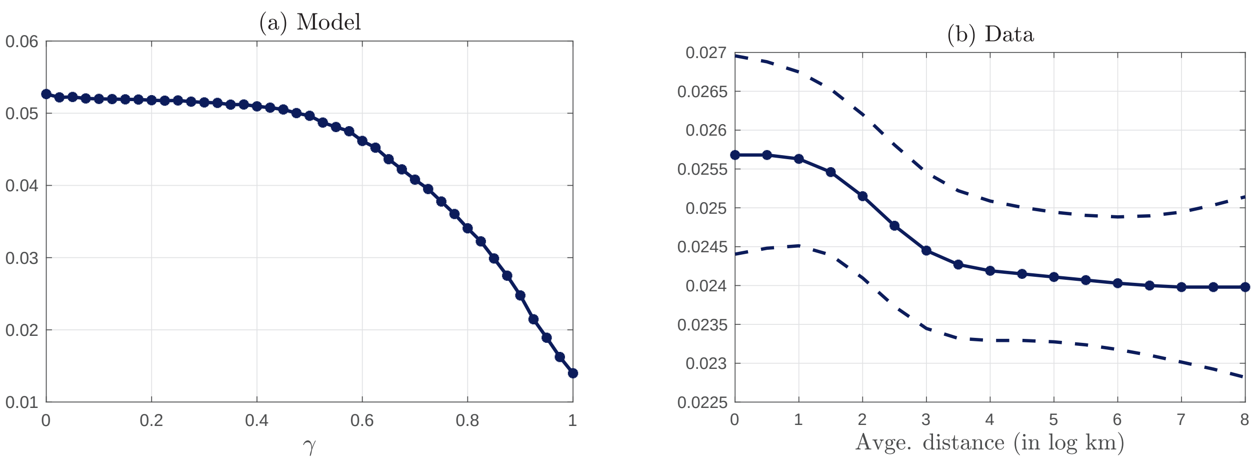


Note: See notes to Figure 4.

Figure 7 plots for the typical firm the probability of observing investment spikes at both its plants in a given period, both in the model (left) and data (right), where an investment spike is defined as an investment rate of 15% or more. As we already know, the probability of observing a joint spike in the model is higher when the firm faces dysfunctional internal capital markets and has no choice but to raise funds externally (γ high). The right panel shows that, while the relationship between distance and joint spikes is not as striking and unequivocal in the data, it is qualitatively comparable. Despite a degree of non-monotonicity in the relationship, the overall pattern is consistent with the model in the sense that the point estimate at low values of γ is outside the 95%-confidence interval at the upper end and vice versa. Firms that have plants further apart tend to experience synchronized spikes a tenth more frequently than those that are geographically concentrated.[27]

An alternative way of comparing within-firm investment dynamics between the model and the data is to compute the average variance of changes in the investment rate i/k across plants within the firm. This moment is more general as it does not focus solely on investment spikes and encompasses any investment activity. As can be seen in the left panel of Figure 8, the model predicts that this variance is higher when the firm's internal capital market is closer to frictionless (γ is low): when the firm has the ability to leverage internal capital markets, it tries to avoid raising costly external funds by staggering investment projects across plants over time, as was illustrated in Figure 3. This, naturally, tends to generate increases and decreases in plant investment rates that are asynchronized activity across plants within the firm, hence higher $Var(\Delta i/k)$ at lower values of γ.

Figure 8: Within-firm co-movement of investment $Var(\Delta i/k)$



Note: See notes to Figure 4.

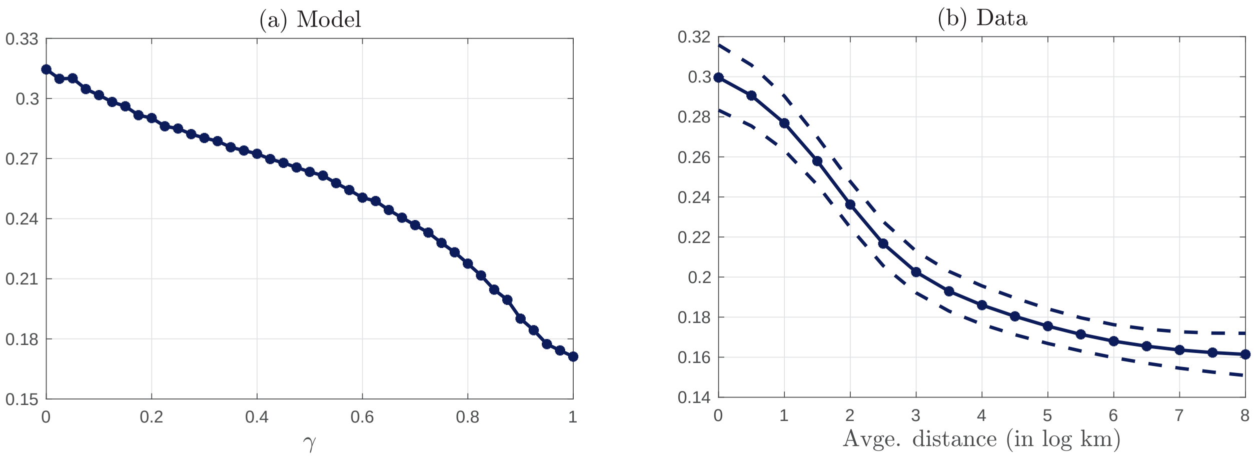
The right panel of Figure 8 indicates a similar pattern in the data: the variance of investment rate changes is higher for firms that are regionally concentrated and for which, as we discussed earlier, we would expect internal capital markets to be less subject to frictions. This is consistent with the presence of staggered investment patterns as described in Section 4.3. While the variance is declining in distance, the magnitude of the variation is significantly less pronounced in the data. Still, the point estimates at the extremes are outside the error bands at the other end of the geographic distance distribution.

Finally, as a complement to our evidence on investment activity, we also look at the dynamics of revenue products across plants within the firm. In theory, we expect the within-firm dynamics of investment and revenue products to be closely related: even if the levels of revenue products

[27] One specific limitation that we face in the data is that, since time is continuous, any definition of synchronized or staggered investment activity based on a yearly window is prone to measurement error. For example, a plant might experience a 12-month period of intense investment activity, yet this period straddles two yearly windows as defined by the ASM. This issue is not present in the model, where shocks occur discretely over time.

within the firm were to be different due to factors not modelled, the *changes* of $mrpk$ across plants within the firm should be more closely aligned when investment plans are synchronized. This is confirmed in the left panel of Figure 9: the variance of the changes in $mrpk$ is higher when the firm is able to leverage internal capital markets, as was the case for changes in investment rates.[28] The right panel shows that this pattern is borne out in the data: firms that are spatially concentrated display a variance of revenue products that is almost twice as large than for their most spread out counterparts.

Figure 9: Within-firm co-movement of $mrpk$: $Var(\Delta mrpk)$



Note: See notes to Figure 4.

Overall, we find that the empirical evidence provides convincing support for the model mechanism of internal capital markets, both qualitatively and quantitatively.

5 Revisiting Hsieh/Klenow with internal capital markets

Next, we investigate the potential implications of *good dispersion* for quantifying aggregate losses from misallocation in emerging economies. In their seminal work, Hsieh and Klenow (2009) estimate that lowering distortions in the Chinese and Indian manufacturing sectors to match dispersion in revenue total factor productivity found in the U.S. would lead to aggregate TFP gains of 39% for China and 47% for India.[29] While these are already sizable numbers, our findings suggest that they may in fact be lower bounds on potential gains.

The reason lies in differences in the organizational complexity of firms in developed economies such as the United States relative to emerging economies such as China and India. The vast majority of economic activity in the former group is accounted for by multi-plant firms (see Table A1). As we showed with the help of our model, internal capital markets in these multi-plant firms

[28]Note that this is not the product of the shock process, which is kept unchanged as we vary γ.

[29]To compute this number, we take the average gain across the three years analyzed by Hsieh and Klenow (2009). See their Table VI for details.

can generate *good dispersion* that should not be interpreted as a sign of misallocation. Developing economies, on the other hand, are largely populated by single-plant firms.[30] With a limited role for internal capital markets, dispersion in marginal revenue products is predominantly caused by distortions or other frictions and thus reflects *bad dispersion*. Therefore, the differences in marginal revenue product dispersion due to distortions (*bad dispersion*) between developed and emerging economies should in fact be even greater than those found by Hsieh and Klenow (2009).

To summarize, the potential gains from eliminating misallocation through resource reallocation are twofold: First, reducing distortions increases output as shown by Hsieh and Klenow (2009); and second, introducing multi-plant firms with internal capital markets also renders the economy more efficient, even if it generates more dispersion. This second efficiency gain has been overlooked so far. In the preliminary exercise that follows, we show that its magnitude is likely not trivial.

Hsieh and Klenow (2009) use differences in $TFPR$ dispersion to infer potential aggregate TFP and output gains from eliminating inefficiencies. To maintain consistency with that framework, we adopt their main assumptions of constant returns to scale, monopolistic competition, and a joint log-normal distribution of physical and revenue total factor productivity, denoted by $TFPQ$ and $TFPR$, respectively. Although our model was silent on the potential causes for dispersion in marginal revenue products of labor, we assume further that the variance of $mrpl$ in the two-plant-firm economy relative to the one-plant-firm economy behaves similarly to that of $mrpk$, and that its covariance with $mrpk$ is unchanged.[31] Under these assumptions, aggregate (or sectoral) TFP can be decomposed into an efficiency and dispersion term, as in Equation (16) of Hsieh and Klenow (2009). Absent further sectoral information on the two terms and the capital share (α in their notation), as well as the sectoral weights in aggregate production (θ_s in their notation), we abstract from differences in these terms and parameters that are specific to the country and sector. Using their Equation (16), we can then write the conventionally computed output gain from reducing misallocation in emerging economies (EM) down to the level of the U.S. economy (US) as

$$\log(1+\text{conv. output gain}) = \frac{1}{\sigma-1}\log\left(\frac{\sum_i A_{n,US}^{\sigma-1}}{\sum_n A_{n,EM}^{\sigma-1}}\right) - \frac{\sigma}{2}\left[V(tfpr_{n,US}) - V(tfpr_{n,EM})\right] \tag{16}$$

where "conv. output gain" is the output gain in percent; $\sigma = 3$ the elasticity of substitution between product varieties within industries; A_n is the technological efficiency ($TFPQ$) of plant n; and $V(tfpr_n)$ is the variance of log revenue total factor productivity. In the next two subsections, we generate the counterfactual variances of $tfpr$ that would arise if emerging economies had the same multi-plant firm structure as the United States. By plugging this information into Equation (16), we can ultimately quantify additional aggregate TFP gains from internal capital markets. We start with the case of India.

[30]For example, very few plants in the Indian manufacturing data report another plant in the same firm. We are grateful to Pete Klenow and Cian Ruane for making this information available to us.

[31]As a conservative robustness check, we consider below the quantitative implications of assuming no change in $mrpl$ dispersion between the two economies.

5.1 Quantifying the gains from internal capital markets in India

In the Indian data, we have some information about the prevalence of multi-unit and single-unit firms. This makes a quantitative exercise possible, in which we increase the share of multi-unit firms to U.S. levels by merging single-unit firms into two-plant firms. This exercise is akin to moving from a two-plant firm without functioning internal capital markets to one with perfect internal capital markets. As seen in the previous section, this counterfactual merge will produce an increase in *good dispersion*, and allow us to study the aggregate implications. To account for the coexistence of multi- and single-plant firms, we rewrite the total dispersion of $tfpr$ across Indian plants as follows:

$$\begin{aligned} V(tfpr_n) &= \sum_n \omega_n (tfpr_n - \overline{tfpr})^2 \\ &= \omega^M V(tfpr_n^M) + (1-\omega^M) V(tfpr_n^S) + \omega^M (1-\omega^M) \left(\overline{tfpr}^M - \overline{tfpr}^S\right)^2 \end{aligned} \tag{17}$$

where the superscript index M denotes the set of plants belonging to multi-plant firms; ω^M their share in the economy; $\overline{tfpr}^M$ their average level of revenue total factor productivity; and $V(tfpr^M)$ their dispersion. The superscript index S indicates the analogue variables for standalone plants.

Our counterfactual exercise consists in computing $V(tfpr_n)$ in India under the assumption that the fraction of plants belonging to multi-plant firms, ω^M, which is 8.9% in India, is the same as in the U.S., at 21.9%. We leave plants that are already part of a multi-unit firm and their dispersion untouched, so the first term in Equation (17) is unchanged.

Next, we randomly merge $21.9\% - 8.9\% = 13\%$ of Indian plants that are standalone firms into two-plant firms, so 21.9% of all plants in India are part of a multi-unit firm, as in the U.S. The dispersion among those newly merged plants would increase by up to 71%, which reflects *good dispersion*. The second term will therefore be higher than before by $0.13 \times 0.71 \times V(tfpr^S)$. Hsieh and Klenow (2009) do not separately report the variance across plants that are standalone firms vs. the overall variance, but we use their data to verify that it is essentially the same. We therefore set $V(tfpr^S) = 0.458$, the average value across years (see their Table II). We leave the variance of those plants that remain single-unit firms in our counterfactual exercise unchanged.

Last, we consider the change in the third term in Equation (17), which increases because ω_M is now closer to 1/2. Again using data from Hsieh and Klenow (2009), we find that the term $\left(\overline{tfpr}^M - \overline{tfpr}^S\right)^2 \approx 0.083$ averaged across years and industries. This will allow us to compute the increase in the third term.

Merging standalone plants into two-plants firms would give those new firms internal capital markets and higher *good dispersion*, which amounts to:

$$\begin{aligned}\Delta V(tfpr_{n,IN}) &= (\omega_{US}^{M} - \omega_{IN}^{M}) \times 71\% \times V(tfpr_{IN}^{S}) \\ &\quad + \left[\omega_{US}^{M}(1-\omega_{US}^{M}) - \omega_{IN}^{M}(1-\omega_{IN}^{M})\right]\left(\overline{tfpr}^{M} - \overline{tfpr}^{S}\right)^{2} \\ &= (0.219 - 0.089) \times 0.71 \times 0.458 \\ &\quad + [0.219 \times (1-0.219) - 0.089 \times (1-0.089)] \times 0.083 \\ &= 0.05.\end{aligned}$$

Substituting this increased variance into Equation (16), we obtain the true output gain from internal capital markets, which exceeds the conventional output gain. Quantitatively, this is given by:

$$\begin{aligned}\log(1{+}\text{true output gain}) &= \frac{1}{\sigma-1}\log\left(\frac{\sum_i A_{n,US}^{\sigma-1}}{\sum_j A_{n,IN}^{\sigma-1}}\right) + \frac{\sigma}{2}\left[V(tfpr_{n,IN}) + \boldsymbol{\Delta V(tfpr_{n,IN})} - V(tfpr_{n,US})\right] \\ &= \log(1{+}\text{conv. output gain}) + \frac{\sigma}{2} \times 0.05 \\ &= 47\% + 7.5\% = 54.5\%.\end{aligned}$$

The implication is that internal capital markets raise the aggregate TFP and output gains by an additional sixth of what Hsieh and Klenow (2009) computed. Note that if the typical Indian multi-plant firm operates fewer plants than its U.S. counterpart, then this number should be interpreted as a lower bound on the gains of internal capital markets. Additionally, we note that our quantitative analysis was computed in a framework with fixed adjustment costs; this mitigates the argument raised by Asker et al. (2014) that fixed adjustment costs can explain a large portion of the *bad dispersion* documented by Hsieh and Klenow (2009).

5.2 Quantifying the gains from internal capital markets in China

We cannot employ the same methodology to compute the benefits from *good dispersion* in China, because the data are sampled at the firm level. We can, however, still compute the aggregate gains from a comparable measure of dispersion by focusing on the portion that occurs *between* firms j. While this means that we must omit the within-firm portion, it allows us to carry out the analysis of Hsieh and Klenow (2009) at a comparable level of aggregation. In short, unlike the India exercise, we will be comparing *bad dispersion* in the U.S. to *bad dispersion* in China.

We showed in Section 2.1 that between-firm dispersion was 39.9% of overall dispersion in U.S.

manufacturing. Consequently, we adjust the U.S. variance in Equation (16) as follows:

$$\begin{aligned}\log(1+\text{true output gain}) &= \frac{1}{\sigma-1}\log\left(\frac{\sum_i A_{j,US}^{\sigma-1}}{\sum_j A_{j,CHI}^{\sigma-1}}\right) - \frac{\sigma}{2}\left[V(tfpr_{j,US}) - V(tfpr_{j,CHI})\right] \\ &= \log(1+\text{conv. output gain}) + \frac{\sigma}{2} \times 0.601 \times V(tfpr_{n,US}) \\ &= 39\% + 18.4\% = 57.4\%\end{aligned}$$

where use $V(tfpr_{n,US}) = 0.2035$, the average dispersion in the U.S. from Hsieh and Klenow (2009). This calculation implies that the aggregate TFP gains in China are almost one-half larger than previously thought.

While the above exercises provide a useful starting point to think about the role of *good dispersion*, they remain preliminary, in part due to data limitations. For example, a more detailed analysis would consider the exact empirical distribution of plants per firm in each economy, differences in production functions, the number of firms and plants in each sector, and how both $TFPR$ and $TFPQ$ are distributed. In addition, one strong assumption we made is that $V(mrpl)$ behaves similarly to $V(mrpk)$ in the two-plant-firm and single-plant-firm economies. While they likely exist in the real world, fixed adjustment costs of labor are arguably of lesser importance than those impeding the allocation of capital. As a result, our assumption that $V(mrpl)$ is affected similarly to $V(mrpk)$ by the presence of firm-level internal markets may arguably be too strong, and the extra output gains relative to Hsieh and Klenow (2009) may not be as large.[32] Yet despite its limitations, our development accounting exercise highlights the importance of taking into account the within-firm dimension for aggregate outcomes.

6 Conclusion

This paper shows that dispersion in marginal revenue products of capital need not indicate distortions. Motivated by evidence that dispersion mostly occurs within firms rather than across firms, we build a model of a firm operating several plants. Such firms have at their disposal an internal capital market that helps ease external financial constraints, support investment activity and generate extra output. Most importantly, economies with multi-plant firms may well exhibit more dispersion in marginal revenue products of capital than economies with single-plant firms, but still produce more aggregate output with the same technologies. An implication is that output gains from capital reallocation may be higher than previously thought in emerging economies, where single-plant firms are relatively more prevalent.

[32] According to the expression of $TFPR$ on p.1410 in Hsieh and Klenow (2009), our additional output gains would only be an α_s portion of what we computed in the benchmark case above. This still remains a significant adjustment, highlighting the potential impact of our mechanism on exercises that compute the gains from reducing distortions in emerging economies.

References

Andrew B. Abel and Janice C. Eberly. A unified model of investment under uncertainty. *American Economic Review*, 84(5):1369–1384, December 1994. (Cited on page 2.)

John Asker, Allan Collard-Wexler, and Jan De Loecker. Dynamic inputs and resource (mis)allocation. *Journal of Political Economy*, 122(5), October 2014. (Cited on pages 4, 6, 9, 10, 19, 21, 24, and 35.)

Bee Yan Aw, Mark J. Roberts, and Daniel Y. Xu. R&D investment, exporting, and productivity dynamics. *American Economic Review*, 102(4):1312–1344, June 2011. (Cited on page 12.)

Rüdiger Bachmann and Christian Bayer. Investment dispersion and the business cycle. *American Economic Review*, 104(4):1392–1416, April 2014. (Cited on page 51.)

Yan Bai, Keyu Jin, and Dan Lu. Misallocation under trade liberalization. *Working Paper*, 2018. (Cited on page 9.)

Isaac Baley and Andrés Blanco. Aggregate dynamics in lumpy economies. *Econometrica*, 89(3): 1235–1264, May 2021. (Cited on pages 9 and 14.)

Isaac Baley and Andrés Blanco. The macroeconomics of partial irreversibility. *Working Paper*, 2023. (Cited on pages 2 and 4.)

Eric J. Bartelsmann, John C. Haltiwanger, and Stefano Scarpetta. Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1):305–334, February 2013. (Cited on page 47.)

Natalie Bau and Adrien Matray. Misallocation and capital market integration: Evidence from India. *Econometrica*, 91(1):67–106, January 2023. (Cited on page 4.)

Christian Bayer, Ariel M. Mecikovsky, and Matthias Meier. Productivity dispersions: Could it simply be technology choice? *Working Paper*, 2015. (Cited on page 4.)

Timothy Besley, Stephen Coate, and Glenn Joury. The economics of rotating savings and credit associations. *American Economic Review*, 83(4):792–810, September 1993. (Cited on page 23.)

Nicholas Bloom. The impact of uncertainty shocks. *Econometrica*, 77(3):623–685, May 2009. (Cited on page 52.)

Nicholas Bloom, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry. Really uncertain business cycles. *Econometrica*, 86(3):1031–1065, May 2018. (Cited on pages 51 and 52.)

Patrick Bolton, Neng Wang, and Jinqiang Yang. Leverage dynamics under costly equity issuance. *NBER Working Paper No. 26802*, June 2021. (Cited on page 16.)

J. David Brown, Emin Dinlersoz, and John S. Earle. Does higher productivity dispersion imply greater misallocation? a theoretical and empirical analysis. *Census Discussion Paper CES-WP-16-42*, September 2016. (Cited on page 12.)

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002, August 2011. (Cited on page 4.)

Ricardo J. Caballero. Aggregate investment. In John B. Taylor and Michael Woodford, editors, *Handbook of Macroeconomics*, volume 1, Part B, chapter 12, pages 813–862. Elsevier, Amsterdam, December 1999. (Cited on page 2.)

Ricardo J. Caballero and Eduardo M. R. A. Engel. Explaining investment dynamics in U.S. manufacturing: A generalized (*S,s*) approach. *Econometrica*, 67(4):783–826, July 1999. (Cited on page 14.)

Ricardo J. Caballero, Eduardo M. R. A. Engel, and John C. Haltiwanger. Plant-level adjustment and aggregate dynamics. *Brookings Papers on Economic Activity*, 1995(2):1–54, 1995. (Cited on page 2.)

Lorenzo Caliendo, Ferdinando Monte, and Esteban Rossi-Hansberg. The anatomy of French production hierarchies. *Journal of Political Economy*, 123(4):809–852, August 2015. (Cited on page 12.)

Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, February 2005. (Cited on page 24.)

Lawrence J. Christiano, Roberto Motto, and Massimo Rostagno. Risk shocks. *American Economic Review*, 104(1):27–65, January 2014. (Cited on page 52.)

Xavier Cirera and William F. Maloney. *The Innovation Paradox: Developing-Country Capabilities and the Unrealized Promise of Technological Catch-Up*. The World Bank, Washington, DC, 2017. (Cited on page 9.)

Wesley M. Cohen and Steven Klepper. The anatomy of industry R&D intensity distributions. *American Economic Review*, 82(4):773–799, September 1992. (Cited on page 12.)

Wesley M. Cohen and Steven Klepper. A reprise of size and R&D. *Economic Journal*, 106(437): 925–951, July 1996. (Cited on page 12.)

Russell Cooper and John C. Haltiwanger. The aggregate implications of machine replacement: Theory and evidence. *American Economic Review*, 83(3):181–186, June 1993. (Cited on page 14.)

Russell Cooper and John C. Haltiwanger. On the nature of capital adjustment costs. *Review of Economic Studies*, 73(3):611–633, July 2006. (Cited on pages 2, 12, 19, and 24.)

Russell Cooper, John C. Haltiwanger, and Laura Power. Machine replacement and the business cycle: Lumps and bumps. *American Economic Review*, 89(4):921–946, September 1999. (Cited on page 14.)

Russell W. Cooper and Immo Schott. Capital reallocation and aggregate productivity. *Working Paper*, August 2018. (Cited on page 51.)

Harold Creusen, Henk Kox, Arjan Lejour, and Roger Smeets. Exploring the margins of Dutch exports: A firm-level analysis. *De Economist*, 159(4):413–434, December 2011. (Cited on page 12.)

Ana Paula Cusolito and William F. Maloney. *Productivity Revisited: Shifting Paradigms in Analysis and Policy*. World Bank, Washington, D.C., 2018. (Cited on page 9.)

Era Dabla-Norris, Si Guo, Vikram Haksar, Minsuk Kim, Kalpana Kochhar, Kevin Wiseman, and Aleksandra Zdzienicka. The new normal: A sector-level perspective on growth and productivity trends in advanced economies. *IMS Staff Discussion Note*, March 2015. (Cited on page 9.)

Sanghamitra Das, Mark J. Roberts, and James R. Tybout. Market entry costs, producer heterogeneity, and export dynamics. *Econometrica*, 75(3):837–873, May 2007. (Cited on page 12.)

Joel M. David and Venky Venkateswaran. The sources of capital misallocation. *American Economic Review*, 109(7):2531–2567, July 2019. (Cited on pages 9 and 12.)

Joseph A. DiMasi, Ronald W. Hansen, and Henry G. Grabowski. The price of innovation: New estimates of drug development costs. *Journal of Health Economics*, 22(2):151–185, March 2003. (Cited on page 12.)

Sebastian Doerr, Dalia Marin, Davide Suverato, and Thierry Verdier. Mis-allocation within firms: Internal finance and international trade. *CESifo Working Paper No. 9426*, 2021. (Cited on page 5.)

Mark Doms and Timothy Dunne. Capital adjustment patterns in manufacturing plants. *Review of Economic Dynamics*, 1(2):409–429, April 1998. (Cited on pages 2 and 14.)

Janice Eberly, Sergio Rebelo, and Nicolas Vincent. What explains the lagged investment effect? *Journal of Monetary Economics*, 59(4):370–380, May 2012. (Cited on pages 14 and 24.)

Andrea L. Eisfeldt and Tyler Muir. Aggregate external financing and savings waves. *Journal of Monetary Economics*, 84:116–133, December 2016. (Cited on page 3.)

Andrea L. Eisfeldt and Dimitris Papanikolaou. Organization capital and the cross-section of expected returns. *Journal of Finance*, 68(4):1365–1406, August 2013. (Cited on page 5.)

Andrea L. Eisfeldt and Adriano A. Rampini. Capital reallocation and liquidity. *Journal of Monetary Economics*, 53(3):369–399, April 2006. (Cited on pages 51 and 52.)

Andrea L. Eisfeldt and Yu Shi. Capital reallocation. *Annual Review of Financial Economics*, 10: 361–386, September 2018. (Cited on page 51.)

Steven M. Fazzari, R. Glenn Hubbard, and Bruce C. Petersen. Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1988(1):141–206, 1988. (Cited on page 3.)

Lucia Foster, John C. Haltiwanger, and Chad Syverson. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425, March 2008. (Cited on pages 8 and 49.)

Xavier Gabaix. The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772, May 2011. (Cited on page 52.)

Robert Gertner, David Scharfstein, and Jeremy Stein. Internal versus external capital markets. *Quarterly Journal of Economics*, 109(4):1211–1230, November 1994. (Cited on page 5.)

Simon Gilchrist and Charles P. Himmelberg. The role of cash flow in reduced-form investment equations. *Journal of Monetary Economics*, 36(3):541–572, December 1995. (Cited on page 3.)

Xavier Giroud. Proximity and investment: Evidence from plant-level data. *Quarterly Journal of Economics*, 128(2):861–915, May 2013. (Cited on pages 4, 5, 16, 25, and 27.)

Xavier Giroud and Holger M. Mueller. Capital and labor reallocation within firms. *Journal of Finance*, 70(4):1767–1804, August 2015. (Cited on page 5.)

Xavier Giroud and Holger M. Mueller. Firms' internal networks and local economic shocks. *American Economic Review*, 2019. (Cited on page 5.)

João F. Gomes. Financing investment. *American Economic Review*, 91(5):1263–1285, December 2001. (Cited on page 3.)

Gita Gopinath, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. Capital allocation and productivity in South Europe. *Quarterly Journal of Economics*, 132(4): 1915–1967, November 2017. (Cited on pages 12, 44, and 51.)

François Gourio and Anil K. Kashyap. Investment spikes: New facts and a general equilibrium exploration. *Journal of Monetary Economics*, 54:1–22, September 2007. (Cited on page 2.)

Anna Gumpert, Henrike Steimer, and Manfred Antoni. Firm organization with multiple establishments. *Quarterly Journal of Economics*, 137(2):1091–1138, May 2022. (Cited on pages 4 and 25.)

John Haltiwanger, Robert Kulick, and Chad Syverson. Misallocation measures: The distortion that ate the residual. *NBER Working Paper No. 24199*, October 2018. (Cited on page 9.)

Christopher A. Hennessy and Toni M. Whited. How costly is external financing? evidence from a structural estimation. *Journal of Finance*, 62(4):1705–1745, August 2007. (Cited on page 16.)

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4), November 2009. (Cited on pages 2, 4, 5, 6, 8, 10, 12, 28, 32, 33, 34, 35, and 36.)

Poorya Kabir, Eugene Tan, and Ia Vardishvili. Does marginal product dispersion imply productivity losses? the case of maintenance flexibility and endogenous capital user costs. *Working Paper*, 2023. (Cited on page 5.)

Matthias Kehrig. The cyclical nature of the productivity distribution. *Working Paper*, 2015. (Cited on pages 43, 44, 51, and 52.)

Truman L. Kelley. *Fundamentals of Statistics*. Harvard University Press, Cambridge, Massachusetts, 1947. (Cited on pages 6 and 44.)

Maurice G. Kendall and Alan Stuart. *Advanced Theory of Statistics*, volume 1. Oxford University Press, New York, 5th edition, 1987. (Cited on page 44.)

Aubhik Khan and Julia K. Thomas. Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, 121(6):1055–1107, December 2013. (Cited on pages 3 and 4.)

Owen A. Lamont. Cash flow and investment: Evidence from internal capital markets. *Journal of Finance*, 52(1):83–109, March 1997. (Cited on page 5.)

Andrea Lanteri. The market for used capital: Endogenous irreversibility and reallocation over the business cycle. *American Economic Review*, 108(9):2383–2419, September 2018. (Cited on pages 51 and 52.)

David O. Lucca. Resuscitating time-to-build. *Working Paper*, 2007. (Cited on page 24.)

Andrey Malenko. Optimal design of internal capital markets. *Working Paper*, February 2016. (Cited on page 5.)

Gregor Matvos and Amit Seru. Resource allocation within firms and financial market dislocation: Evidence from diversified conglomerates. *Review of Financial Studies*, 27(4):1143–1189, April 2014. (Cited on page 5.)

Marc J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, November 2003. (Cited on page 12.)

Virgiliu Midrigan and Daniel Y. Xu. Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–458, February 2014. (Cited on page 4.)

Benjamin Moll. Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221, October 2014. (Cited on page 4.)

Dale T. Mortensen and Christopher A. Pissarides. Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, 61(3):397–415, July 1994. (Cited on page 12.)

Ezra Oberfield and Devesh Raval. Micro data and macro technology. *Econometrica*, 89(2):703–732, March 2021. (Cited on page 48.)

Laura Power. The missing link: Technology, investment, and productivity. *Review of Economics and Statistics*, 80(2):300–313, May 1998. (Cited on page 14.)

Raghuram Rajan, Henri Servaes, and Luigi Zingales. The cost of diversity: The diversification discount and inefficient investment. *Journal of Finance*, 55(1):35–80, February 2000. (Cited on page 5.)

Plutarchos Sakellaris. Patterns of plant adjustment. *Journal of Monetary Economics*, 51(2):425–450, March 2004. (Cited on page 14.)

David S. Scharfstein and Jeremy C. Stein. The dark side of internal capital markets: Divisional rent-seeking and inefficient investment. *Journal of Finance*, 55(6):2537–2564, December 2000. (Cited on pages 5 and 16.)

Antoinette Schoar. Effects of corporate diversification on productivity. *Journal of Finance*, 57(6): 2379–2403, 2002. (Cited on page 5.)

Jae Song, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. Firming up inequality. *Quarterly Journal of Economics*, 134(1):1–50, February 2019. (Cited on pages 43 and 46.)

Jeremy C. Stein. Internal capital markets and the competition for corporate resources. *Journal of Finance*, 52(1):111–133, March 1997. (Cited on pages 5 and 16.)

Chad Syverson. What determines productivity? *Journal of Economic Literature*, 49(2):326–365, June 2011. (Cited on page 9.)

Julia K. Thomas. Is lumpy investment relevant for the business cycle? *Journal of Political Economy*, 110(3):508–534, June 2002. (Cited on page 24.)

Pavel Ševčík. Financial frictions, internal capital markets, and the organization of production. *Review of Economic Dynamics*, 18(3):502–522, July 2015. (Cited on page 5.)

Appendix

— for online publication —

A Additional empirical findings

A.1 Investment and equalizing the marginal revenue product of capital

The assumption that investment should ideally be undertaken to equalize marginal revenue products of capital rests on some assumptions that we work out in this appendix. In a standard frictionless economy with decreasing returns to scale, agents should choose investment to equate the expected capital return. To understand how equating capital returns relates to equating marginal revenue products, consider the expression of the expected capital return. This return is defined as the proceeds of one unit of capital at the end of next period – the value of undepreciated capital plus its marginal revenue product – divided by the cost of next period's capital in the current period. Defining industry output as the numéraire, we denote the price of next period's capital k' in terms of this period's numéraire by $P_t^{k'}$, the industry-wide depreciation rate by δ_t, the marginal revenue product of capital in plant n and year t by $MRPK_{nt}$, real value added by y_{nt}, and the capital stock by k_{nt}. Then the expected gross return, $\mathbb{E}\mathcal{R}_{nt+1}$, in a given year and industry is

$$\mathbb{E}\mathcal{R}_{nt+1} = \mathbb{E}\frac{P_{t+1}^{k}(1-\delta_t) + MRPK_{nt+1}}{P_t^{k'}}.$$

We assume all units in an industry face the same price of capital, $P_t^{k'}$, and the same depreciation rate δ_t. Then the only source of heterogeneity in returns stems from differences in expected marginal revenue products of capital, $MRPK_{nt+1}$. In a large set of models with Cobb-Douglas technology, this object is proportional to the expected average product of capital, $\mathbb{E}\frac{y_{t+1}}{k_{t+1}}$. Since we do not measure the expected marginal revenue product of capital, we approximate it by the realized marginal revenue product of capital. This is a good approximation if capital is chosen one period in advance, all other inputs are chosen statically, and total factor productivity is sufficiently persistent. Only unexpected innovations to profitability will then render the realized and the expected marginal revenue product of capital different. All of these assumptions are plausible and widely used in the macroeconomic and investment literature. From now on we study the logarithm of marginal revenue products of capital which is denoted by lower-case letters: $mrpk_{nt} \equiv \log(MRPK_{nt})$. Given our above assumptions, we measure dispersion in marginal revenue products of capital as $V_t\left(mrpk_{nt}\right) = V_t\left(\log(y_{nt}/k_{nt})\right)$.

A.2 Data

We mainly use confidential data on manufacturing establishments collected by the U.S. Census Bureau that comprise the 1972-2009 Annual Survey of Manufactures (ASM), the Census of Manufactures (CMF) from 1972-2007, and the Longitudinal Business Database (LBD) from 1976-2009. These data inform us about age, output, capital stocks, investment expenditures, and other inputs at the level of the individual establishment. In the manufacturing sector, the CMF defines an "establishment" as a business location where the principal activity is production; we hence think of an "establishment" as a production plant. The CMF also contains information about the ownership of each plant (denoted by the variable `FIRMID`), which allows us to construct the hierarchical plant structure of "firms" necessary for our main object of interest, the within-firm and between-firm component of heterogeneity in returns, productivity and reallocation.

From the CMF and the ASM, we construct a large dataset of plants in the U.S. manufacturing sector. In order to obtain a consistent longitudinal panel, we limit attention to the ASM and the ASM portion of the CMF data (identified by establishment type `ET`=0). We prefer the ASM over the CMF as our benchmark dataset because we want to test the dynamic implications of our model of investment in multi-plant firms at the highest possible frequency. Many aspects of our mechanism would disappear at the quinquennial frequency of the CMF. By focusing on the ASM portion in all years, we automatically eliminate all administrative observations (identified by `AR`=1), which are imputed mainly off industry means and would thus corrupt moments of the distribution we are interested in. Our resulting panel spans the years 1972-2009, which allows us to study the long-run features of the dispersion of marginal revenue products of capital and reallocation. Every year, we observe about 55,000 plants which total to 2.1 million observations.

We combine the Census data with industry-level data from several publicly available sources: input and output price deflators from the NBER-CES Manufacturing Industry Database (NBER-CES), various asset data from the the Capital Tables published by the Bureau of Labor Statistics (BLS), and the Fixed Asset Tables published by the Bureau of Economic Analysis (BEA). Unless otherwise noted, all datasets are at annual frequency. Most of the information contained in the non-Census datasets (BEA, BLS, NBER-CES) other than manufacturing data, are only needed to estimate productivity and the replacement value of capital at current market conditions.

For each plant in these data, we construct real value added, the real capital stock, and real investment. To obtain real value added, y_{nt}, we first compute nominal value added as sales less intermediate and energy inputs, correct for inventory changes and resales,[33] and deflate the resulting measure by the 6-digit NAICS shipment price deflator from the NBER-CES manufacturing database. The real capital stock, k_{nt}, is the sum of structure and equipment capital each of which are expressed as real replacement values at current market conditions. These replacement values are computed individually for structure and equipment capital with the perpetual inventory method, using investment expenditures and depreciation rates. When a plant is observed for the first time, we initialize its capital stock at its book value, which is transformed as follows. First, we convert nominal book values into nominal market values and then deflate this measure using the BLS's price deflators for capital goods at the 3-digit NAICS industry level.[34] Like capital, we compute real investment, i_{nt}, as the sum of real structure and equipment investment by deflating the respective nominal investment expenditures by the 3-digit NAICS industry investment price deflators from the BLS. Our capital measure denotes beginning-of-year stock values, while our investment and value added measures refer to flow values during the year. To avoid outliers driving our results about dispersion and the investment-productivity link, we drop the 1% tails of the productivity and investment rate distributions in a given 4-digit NAICS industry.

A firm is defined as all manufacturing plants within the same `FIRMID`[35] in a given year and 4-digit NAICS industry. The `FIRMID` defines the collection of plants under common ownership or control. All plants of subsidiary firms are included as part of the owning or controlling firm. If the same firm is active in several industries, we define each subset of plants belonging to the

[33] Resales are goods purchased from another producer and resold in an unchanged condition. Correcting for them means we assess productivity of the plant as a producer rather than its productivity as a trader.

[34] For more details about the primary data and the transformation needed to obtain measures of the real capital stock and estimate productivity, see the description in the appendix to Kehrig (2015).

[35] Song et al. (2019) identify firms off the `EIN`, the employer identification number, which comes from tax records. Since we are interested in organizational control rather than tax liability and because the same `FIRMID` may operate hundreds of `EIN`s for tax purposes, we prefer `FIRMID` to indicate firms.

same industry as separate firms. Our within-firm dispersion measures are hence an understatement because we ignore the between-industry component of within-firm dispersion.

A.3 The economic importance of multi-plant firms

Our between-firm/within-firm analysis is economically relevant if a significant portion of aggregate economic activity is accounted for by multi-plant firms. Table A1 shows that while single-plant firms dominate in numbers, multi-plant firms operate the majority of the capital stock, produce most output, and generate most investment. In fact, firms that consist of 20 or more plants operate almost one-half of all the capital stock in U.S. manufacturing.

Table A1: Economic activity by firm type in U.S. manufacturing

	Share of ...			
	plants	value added	capital stock	investment
Single-plant firms	0.719	0.220	0.178	0.215
Multi-plant firms	0.281	0.780	0.822	0.785
Firms with at least...				
... 10 plants	0.131	0.513	0.602	0.548
... 20 plants	0.095	0.398	0.470	0.421
... 40 plants	0.060	0.252	0.296	0.261

Note: The sample underlying this table comprises all establishments in the Census of Manufactures 1972-2007 less administrative records. The share of each variable in multi-plant vs. single-plant firms is computed for each Census year and then averaged across Census years. Non-manufacturing operations of firms are ignored.

A.4 Empirics of cross-sectional moments

This appendix details how the cross-sectional moments underlying Table 2, Panel B were computed. First, we compute cross-plant moments $\mathcal{M}_{it}$ and their standard errors in a given industry i and year t. $\mathcal{M}_{it}$ stands for the cross-sectional standard deviation, inter-decile range, skewness, Kelley skewness, and excess kurtosis. We adopt the formulae for the first four moments, the inter-quantile range and their standard errors from Kendall and Stuart (1987). Kelley skewness is a quantile-based measure of skewness whose predecessor was proposed by Kelley (1947).

Every cross-sectional moment is computed by industry and by year. To get long-run industry-specific moments, we first aggregate over years in order to exclude any industry-specific trends. As do Kehrig (2015); Gopinath et al. (2017), we note an upward trend in dispersion and – to a lesser extent – in skewness and a downward trend in kurtosis. Notice that the cross-plant standard deviation increases about 10 log points per decade; both between-firm and within-firm dispersion increases evenly, so there is no discernible trend in the within-firm share of the overall industry variance. The cross-plant skewness becomes more positive over time: Kelley skewness increases from around zero (unskewed) to 0.25 (right tail about 1.66 times as wide at the bottom tail) in 2007. We compute the typical cross-sectional moment in a given NAICS-4 industry in 1990 that corresponds to the middle of our sample.

Then, we aggregate across industries using that industry's average share in value added: $\mathcal{M}_t = \sum_i \omega_{it}\mathcal{M}_{it}$. Standard errors are computed according to this aggregation: $SE_{\mathcal{M}_t} = \sqrt{\sum_i (\omega_{it} SE_{\mathcal{M}_{it}})^2}$.

This yields the moments within the average industry in the middle of our sample.

A.5 Empirics of between-firm and within-firm moments

In this section, we detail how we compute the within-firm and between-firm dispersion in marginal revenue products of capital and capital reallocation that underlie Table 2 and the robustness exercises in Section A.7.

First, we decompose the overall variance in marginal revenue products of capital into three components: one between industries (reflecting differences in measurement and the definition of capital and value added), one between firms in a given industry, and one across plants within a firm and industry. We define firms that operate plants in separate industries as different firms, thus biasing the true within-firm component of dispersion downward.

$$
\begin{aligned}
V_t &= \sum_n \omega_{njit} \left(mrpk_{njit} - mrpk_t\right)^2 \\
&= \underbrace{\sum_i \omega_{it} \left(mrpk_{it} - mrpk_t\right)^2}_{V_t^{Ind} \text{ between-industry}} + \underbrace{\sum_i \omega_{it} \underbrace{\sum_{j \in i} \omega_{jt}^i \left(mrpk_{jit} - mrpk_{it}\right)^2}_{V_{it}^B \text{ between firms within ind. } i}}_{V_t^B \text{ average between-firm}} + \underbrace{\sum_i \omega_{it} \sum_{j \in i} \omega_{jt}^i \underbrace{\sum_{n \in j,i}^{N_j} \omega_{nt}^{ji} \left(mrpk_{njit} - mrpk_{jit}\right)^2}_{V_{jit}^W \text{ within firm } j \text{ and industry } i}}_{V_t^W \text{ average within-firm}}
\end{aligned}
\tag{A1}
$$

where n indicates the plant, j the firm, i the 4-digit NAICS industry and t the year. $mrpk_{njit}$ denotes the marginal revenue product of capital of plant n belonging to firm j and industry i in year t, $mrpk_{jit}$ the average return in firm j in industry i, $mrpk_{it}$ the average return in industry i, and $mrpk_t$ the average level of returns in the economy.

An industry's level of marginal revenue product of capital is determined by the level of P_t^k and the asset bundle it typically reflects in that industry. This and other industry specificities in measurement will artificially drive V^{Ind} – an object we ignore for its lack of economic meaning. In our empirical analysis in Section 2.1, we focus only on the V_i^B and V_i^W of firms with at least two plants, because it is meaningful to compare them and how much of the dispersion in marginal revenue products of capital within an industry originates within firms as opposed to between firms in that same industry: $\mathcal{W}_i \equiv \frac{V_i^W}{V^W{}_i + V_i^B}$. When computing an "aggregate" number for $\mathcal{W}$, we compute the average of industry ratios, which is weighted by ω_i, i.e., that industry's share in plants or capital, depending on whether we are looking at unweighted or capital-weighted dispersion.

Although investment rates do not suffer from industry-specific measurement issues, such as the marginal revenue product of capital, we proceed in a similar way to assess between-firm and within-firm investment-rate dispersion.

A.6 Robustness

A.6.1 Accounting for measurement error

In Section A.7 we will deal with some measurement error. If plant-level variables are measured with noise, then firm-level averages will be measured more precisely and artificially inflate the within-firm variance. Time aggregation should filter out this type of measurement error. Because time aggregation cannot deal with persistent measurement error, we now consider that type. To do that, we consider marginal revenue products of capital that are computed using separate measures of capital and values added. Our alternative measures come from different datasets or are separately measured variables in our baseline dataset. We have:

- K^{TAB} — we use appropriately deflated values of variable `TAB` instead of the perpetual inventory method;
- Y^{IRS} — we use administrative data on sales from the IRS instead of `TVS` from CMF/ASM;
- Y^{PCU} — we use collected data on actual production from the Plant Capacity Utilization Survey (PCU) instead of `TVS`.

These alternative measures should be correlated with our original measures of K and Y in the ASM (since they measure the same underlying object), but they should still be different due to different coverage or handling by the statistical agency. We recompute marginal revenue products of capital using the three alternative measures and redo the cross-sectional within-firm between-firm decomposition on these alternative measures. If the dominance of within-firm share is true, then this should show up in all of these measures. Song et al. (2019) follow a similar procedure.

Table A2: Accounting for measurement error

	Alt. Measure	$Corr\left(\log\left(\frac{y}{k}\right)^{bench}, \log\left(\frac{y}{k}\right)^{alt}\right)$	$\left(\frac{V^W}{V^W+V^B}\right)^{bench}$	$\left(\frac{V^W}{V^W+V^B}\right)^{alt}$
I: CMF 1972-2007	K^{TAB}	0.979	0.563	0.556
			(0.008)	(0.005)
II: CMF 2002-2007	Y^{IRS}	0.990	0.538	0.542
			(0.005)	(0.005)
III: ASM 1974-2007	Y^{PCU}	0.494	0.581	0.626
			(0.015)	(0.017)

Note: This table displays the within-firm share of overall dispersion for alternative measures of value added Y – collected either from tax records or separately measured in the Plant Capacity Utilization Survey (PCU) – and capital K (real replacement value at current market prices directly computed from book values instead of from the perpetual inventory method). Correlation of the computed marginal revenue products of capital measures are positive, some are high, and the within-firm share of overall marginal revenue products of capital dispersion is not statistically different at the 95% level except when using value added from the PCU, which yields an even higher within-firm share. Error bands constructed from averaging across 86 NAICS-4 industries.

Since using these alternative measures limits our sample at times, we also recompute the within-firm/between-firm decomposition using our original data so that we are comparing moments for the same underlying sample, for which we have both our benchmark measure as well as the alternative. It turns out that the differences in the within-firm share are marginal and almost always lie in the 95% error bands of the other measure. Only when using value added from the PCU does the benchmark differ from the alternative, which yields an even higher within-firm share. Error bands constructed from averaging across 86 NAICS-4 industries. We conclude from this exercise that our main result of the within-firm dispersion accounting for the largest portion in overall dispersion does not go away when using alternative measures of output and capital.

A.6.2 Marginal vs. average revenue products

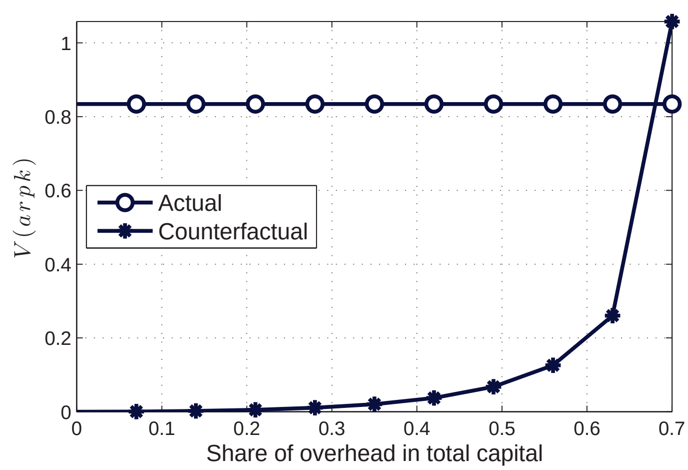
Our empirical work in Section 2.1 aimed at measuring marginal revenue products of capital, which are the relevant measure of what should be equalized across production units. But in the data, one can measure only average revenue products of capital; as in the literature, we approximate the dispersion of marginal revenue products with average revenue products. This approximation is usually justified with a Cobb-Douglas production function, since in this framework average revenue products are proportional to marginal ones. But they are not if technology is not multiplicative or if it is Cobb-Douglas with overhead inputs. Bartelsmann et al. (2013) document that overheads in production are quite powerful in explaining differences between micro production units, and they find significant aggregate consequences. Most overhead inputs are likely at the headquarters level of a firm rather than the plant level. Though this casts some doubt that all our results could be driven by constant inputs, we cannot dismiss this possibility.

Overheads Any constant input requirements at the plant level are hard to identify empirically. We therefore carry out a quantitative exercise to examine how large overheads would have to be in order to explain all or a portion of the empirically observed dispersion. Suppose the true technology is $y_n = z_n(k_n - \bar{k})^{\alpha}$. In that case, the average revenue product of capital can be written as $arpk_n = mrpk_n - \log\alpha + \log\left(1 - \bar{k}/k_n\right)$. Further suppose that marginal revenue products – which we cannot measure – are completely equalized. Then the entire variance of average revenue products would reflect the differential share of overheads across firms of different capital size:

$$V(arpk_n) = V\left(\log\left(1 - \frac{\bar{k}}{k_n}\right)\right). \tag{A2}$$

We simulate a firm-size distribution realistically, assuming that capital – unlike employment – is distributed log-normally. We consider how large the right-hand-side variance in Equation (A2) is for different levels of $\bar{k}$.

Figure A1: How much overheads is necessary to explain the observed $V(arpk)$?



Note: Simulation of the right-hand-side of Equation (A2) against the share of overhead in total inputs.

Figure A1 plots the RHS of Equation (A2) as a function of $\mathbb{E}\left[\bar{k}/k_n\right]$. Naturally, when $\bar{k} = 0$, this variance will be zero and the observed variance of average revenue products must be caused by the variance of marginal revenue products. This does not change much for low and moderate levels of overheads. Even if half of all inputs are overhead, less than one-tenth of the empirically observed variance in average revenue products can be explained by overhead. Clearly, this amount of overhead inputs at the level of the plant is unreasonable. Only if the average share of overhead in total inputs approaches 70% can the observed variance be explained by overhead. We conclude that overhead may only play a limited role in explaining the long-run dispersion of average revenue products of capital.

Non-unitary elasticity of substitution Another empirically plausible alternative to a simple Cobb-Douglas production function would be a constant elasticity of substitution production function. Suppose $y_n = \left[\alpha k_n^{\frac{\sigma-1}{\sigma}} + (1-\alpha) x_n^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$ where σ is the elasticity of substitution and x_n the variable inputs of plant n. In that case, the average revenue product of capital can be written as $arpk_n = \sigma[mrpk_n - \log\alpha]$ and the variance of average and marginal revenue products as

$$V(arpk_n) = \sigma^2 V(mrpk_n). \tag{A3}$$

The true dispersion of marginal revenue products could then be much lower if the elasticity of substitution is larger than unity. However, Oberfield and Raval (2021), who estimate the elasticity of substitution at the plant level, put that number significantly smaller than 1, suggesting that the empirically measured dispersion of average revenue products would be a lower bound on that of marginal revenue products. We conclude that a non-unitary elasticity of substitution would most plausibly measure only a portion of the true dispersion of marginal revenue products.

A.7 Further dimensions of within-firm dispersion

In the previous section, we documented that most dispersion in marginal revenue products of capital and investment rates originates within firms rather than between firms. We now study the between-firm/within-firm decomposition for a number of subsamples. The objective is to confirm the robustness of our main empirical result and identify possible causes behind the importance of within-firm dispersion. Table A3 provides an overview of our robustness checks.

The first row in Table A3 reiterates the baseline result of within-firm vs. between-firm dispersion of both marginal revenue products of capital and investment rates: About 60% of the variance of $mrpk$ and 68% of the variance in i/k arise within firms.

Ruling out ASM sampling specificities We use the ASM as our benchmark panel. This panel is known to overrepresent large plants, which in turn are more likely to be part of a multi-plant firm. Since we do not know whether within-firm dispersion is larger in firms with few or many plants, we repeat our between-firm/within-firm decomposition for the entire Census of Manufactures (every five years). This allows us to study the full sample of manufacturing plants in the economy. Row (2) illustrates our findings: While the within-firm share of dispersion is slightly lower, it still remains dominant at 57% and 66% for $mrpk$ and i/k, respectively.

Ruling out life-cycle dynamics Next, we examine whether our result is driven by entry, exit, or other life-cycle dynamics. Young, presumably more productive plants will be characterized by higher revenue products and will hence attract higher investment rates. The opposite may be true of older plants the firm keeps operational until the capital stock depreciates away. We therefore redo the decomposition using only "mid-age firms," which we define as plants that are at least three years old and at least three years away from exit. As Row (3) shows, the within-firm share of dispersion in both variables is almost unchanged, and the same is true when we consider a five-year distance to entry and exit. Next, we turn to a more drastic exercise and consider only a strongly

Table A3: Dispersion of *mrpk* within and between firms

	Sample	Share of $V(mrpk)$		Share of $V(i/k)$	
		b/w plants within firms	b/w firms	b/w plants within firms	b/w firms
(1)	Full panel	0.601	0.399	0.679	0.321
(2)	Census sample	0.566	0.434	0.663	0.337
(3)	Mid-age plants	0.595	0.405	0.685	0.315
(4)	Balanced panel	0.809	0.191	0.895	0.105
(5)	Permanent components	0.599	0.401	0.715	0.285
(6)	Transitory components	0.692	0.308	0.736	0.264
(7)	y is physical output	0.625	0.376		
(8)	K-weighted	0.540	0.460	0.680	0.320
(9)	Equipment	0.591	0.409	0.631	0.369
(10)	Private firms	0.713	0.287	0.795	0.205
(11)	5-plant firms	0.668	0.332	0.786	0.214
(12)	Counterfactual firms	0.482	0.518		
		(0.030)	(0.043)		

balanced panel. This will filter out any life cycle dynamics at some point in the sample and may. Row (4) shows that the share of within-firm dispersion is even larger at 81%.

Permanent versus transitory characteristics In our next robustness check, we want to study how much of the dispersion patterns are a permanent versus transitory phenomenon. Indeed, transitory plant-level shocks would "wash out" at the firm level, and thus make the within-firm share of dispersion larger at higher frequencies. We therefore decompose a plant's *mrpk* into its long-run average and transitory movements around it. We perform a variance decomposition on both components, display them in Rows (5) and (6) of the table and find the within-firm share of dispersion to be higher for the permanent component in both mrpk (59.9%) and i/k (71.5%). After taking out permanent components, the transitory components exhibit a within-firm dispersion that is even higher at 69.2% and 73.6%, for *mrpk* and i/k respectively. This suggests that the importance of within-firm dispersion is a phenomenon that is structural feature of the economy.

Ruling out markups Limiting our attention to homogeneous goods has another advantage: It allows us to derive real value added in two ways. In addition to the standard approach of deflating sales, we can also use the measured physical quantity of production, a meaningful object for these homogeneous product groups. This makes it possible to study how much of dispersion in capital revenue products reflects price differences – due to differential markups or transfer prices – rather than physical productivity differences. Row (7) shows that the within-firm share of marginal revenue products of capital is slightly higher, at 63%, when using physical output to compute y rather than deflated sales. This suggests that if anything, prices impact the within-firm and between-firm variances in a way that stacks the odds against our main empirical finding. It is also consistent with the fact that plant-level prices and physical productivity are negatively correlated, as documented by Foster et al. (2008).

Demonstrating economic relevance Next, we wish to confirm that our findings are of economic relevance. Instead of decomposing the unweighted variance, we now consider capital weights for the ω's in Equation (1) and redo the decomposition of the dispersion in $mrpk$ and i/k. Row (8) shows that while the within-firm share of capital-weighted dispersion is slightly lower, it still remains dominant at 54% and 68%, respectively.

Examining different capital types In Row (9), we display our decomposition results when focusing only on equipment capital when computing both revenue products and investment rates. Arguably, equipment can be more easily reallocated across production units than structures, which would lower dispersion. Again, results of the unweighted between-firm/within-firm decomposition are almost unchanged at 59% and 63%, respectively.

Ruling out mechanical aggregation If firms were mere random collections of plants, one would expect the within-firm variance of $mrpk$ to merely reflect random noise and not contain any meaningful economic information. To check how much dispersion within firms would arise from such randomness, we construct counterfactual firms in the following way: Within a 3-digit NAICS industry, we maintain the firms and how many plants each firm operates. Then, we randomly assign plants to all firms and repeat the between-within decomposition on this set of counterfactual firms. We repeat this exercise 1,000 times. Row (12) in Table A3 shows the average and standard deviations of these draws. The within-firm share of the variance drops from about 0.6 in the actual full panel to 0.48 in the bootstrapped sample of counterfactual firms. This difference is statistically significant, as the standard error of 0.03 makes clear.

Another way to look at the effects of mechanical aggregation is to study only samples of firms with the same number of plants per firm. In a world composed of many plants owned and operated by a single firm, the within-firm share of dispersion would be 100% by definition. As such, one could worry that the large share of dispersion occurring within the firm is driven by large entities. This is unlikely, since in our benchmark decomposition each firm receives equal weight, irrespective of its size. Because in our sample small 2-plant and 3-plant firms are much more numerous than large, complex firms, this bias probably does not play a large role. Still, in order to determine whether this mechanical aggregation could be an issue, we recompute the between-/within-firm decomposition for the set of firms that operate exactly five plants.[36] Row (11) shows that for this set of firms, the within-firm share of dispersion is indeed higher than for the whole manufacturing sector (67% for $mrpk$ and 79% for i/k), but not dramatically so. This suggests that our main result is not merely driven by the statistical importance of large firms.

These exercises have confirmed that the importance of the within-firm share of dispersion in revenue products of capital and investment rates is robust to changing the sampling frame in order to account for measurement and aggregation problems, life-cycle dynamics, multi-product firms, markups and transfer prices, the predominance of multi-plant firms with little capital, the type of capital, or the sampling of the ASM. In many cases, the within-firm share of overall dispersion is even higher, suggesting that our baseline results may in fact represent lower bounds on the actual within-firm share of dispersion.

A.8 Cyclicality of dispersion between and within firms

So far, we have focused on time-series averages of between-firm and within-firm dispersion in marginal revenue products and investment rates. At the aggregate level, cyclical movements in

[36]According to our more restrictive definition of a firm as all plants operated by the same firm within a 4-digit NAICS industry, half of the capital stock is operated by firms with five plants or more.

either dispersion measure are well known: The countercyclical nature of productivity dispersion has been documented empirically using marginal revenue products of capital in Compustat data by Eisfeldt and Rampini (2006), in used assets by Lanteri (2018), TFP levels by Kehrig (2015), and TFP innovations by Bloom et al. (2018), while Bachmann and Bayer (2014) have shown that the dispersion in investment rates is procyclical. Eisfeldt and Shi (2018) provide a good summary of the literature of the cyclicality of capital reallocation and the various frictions governing that process. These findings have important implications for the literatures on Schumpeterian creative destruction, misallocation, development or uncertainty-driven business cycles. For example, Cooper and Schott (2018) study the effects of cyclical capital reallocation on aggregate productivity. Yet to our knowledge, no one has investigated separately the cyclicality of dispersion within firms. We close this gap by studying the time-series properties of the various components of dispersion. We first compute detrended measures of the between-firm and within-firm variance,[37] which we then use for time-series analysis. In addition, we study the cyclical properties of the lower and upper portions of the distribution as measured by the distance between the 90th percentile and the median as well as that between the median and the 10th percentile. Table A4 displays properties of the long-run averages, autocorrelations, and time-series standard deviations for each measure.

Table A4: Dynamic properties of the $mrpk$ and i/k distributions

	$V^B(mrpk)$	$V^W(mrpk)$	$V^B(i/k)$	$V^W(i/k)$
A.1 Dispersion: Time-series moments				
Average	0.315	0.476	0.016	0.037
Autocorrelation	0.732	0.657	0.667	0.679
Volatility	0.042	0.084	0.006	0.016
A.2 Dispersion: Cyclicality				
$Corr(\Delta Y_{t+1}^{mfg}, ...)$	-0.252	-0.147	0.306	0.330
$Corr(\Delta Y_{t}^{mfg}, ...)$	-0.582	-0.304	0.389	0.394
$Corr(\Delta Y_{t-1}^{mfg}, ...)$	-0.413	-0.229	0.222	0.241
	$mrpk^{50} - mrpk^{10}$	$mrpk^{90} - mrpk^{50}$	$i/k^{50} - i/k^{10}$	$i/k^{90} - i/k^{50}$
B. Skewness: Cyclicality				
$Corr(\Delta Y_{t+1}^{mfg}, ...)$	-0.026	-0.071	-0.155	0.079
$Corr(\Delta Y_{t}^{mfg}, ...)$	-0.293	-0.152	0.148	0.197
$Corr(\Delta Y_{t-1}^{mfg}, ...)$	-0.226	-0.109	0.225	0.244

Note: The table reports time-series moments of between-firm and within-firm variance for both marginal revenue products of capital and capital reallocation. "Average" denotes the long-run average of each variance term; "Autocorrelation" the annual persistence, $Corr(V_t, V_{t-1})$; "Volatility" the time-series standard deviation, $StD(V_t)$; and cyclicality is with respect to the growth rate of aggregate manufacturing value added, denoted by ΔY_t^{mfg}.

Consistent with the evidence from Section 2.1, Panel A.1 in Table A4 shows that the within-firm portion of the variance in $mrpk$ is larger than that between firms. When studying fluctuations of the two variances over time, we find that the volatility of the within-firm portion is also twice as strong as that of the between-firm portion. This is true even if one compares the time-series coefficient of variations instead of the time-series standard deviation, and similar patterns are observed in the between-firm and within-firm dispersion of investment rates.

[37] As do Kehrig (2015) and Gopinath et al. (2017), we find an upward trend in cross-sectional dispersion.

As can be seen from Panel A.2. of Table A4, both $V^B(mrpk)$ and $V^W(mrpk)$ are countercyclical. This result could have important implications for the uncertainty literature. If one interprets dispersion as a cause of cycles, as do Bloom et al. (2018) and Christiano et al. (2014), then our within-firm result suggests that cycles manifest themselves within as well as between firms. This means that looking at the granular level of the firm, as Gabaix (2011) and Eisfeldt and Rampini (2006) do, may underestimate the role of dispersion for aggregate fluctuations. Since fluctuations in heterogeneity at the subgranular level of plants within firms are larger and countercyclical, this suggests tha one should not discard the dynamics inside multi-unit firms when studying business cycles.

Lastly, we study which tail of the marginal revenue product and investment rate distribution is more cyclically sensitive. This is motivated by recent work on cyclical capital reallocation and the various frictions governing that process. Lanteri (2018), for example, develops an endogenous process of capital reallocation based on capital resales that posits that the left, less productive tail of the distribution of marginal revenue products is more cyclically sensitive. Panel B in Table A4 confirms this hypothesis empirically: The distance between the median and the 10th percentile is more countercyclical than its corresponding portion in the upper tail of the marginal revenue product of capital distribution. Such a cyclical pattern of the lower tail of the productivity distribution has been shown to hold as well for TFP levels by Kehrig (2015).

A.9 Discussion

More generally, our findings in this section imply that welfare gains from a more efficient allocation of resources would not only stem from reallocation across firms, but also within. This highlights the importance of developing a better understanding of the factors that impede capital from flowing to its most productive use inside the firm. As such, our findings have implications for micro-founded macroeconomic models and their calibration. In much of the literature, the concepts of plants and firms are used interchangeably, with little discussion of their respective roles and constraints. For example, in the empirical uncertainty literature, plants are almost always interpreted as independent decision makers facing various frictions that impede the reallocation of productive capital. Arguably, some frictions, such as technological ones as in Bloom (2009) and Bloom et al. (2018), are indeed most relevant at the level of the plant. Yet others are more likely to impact the decisions of firms. This is, for example, the case of external financing constraints, which affect interactions between firms and their lenders, as in Christiano et al. (2014).

What is the link between firm-level financial frictions and *mrpk* dispersion? To investigate this issue, we repeat our between-firm/within-firm decomposition on the sample of privately held firms only. We find that the share of within-firm dispersion for private firms is *higher* than for the whole sample, at 71% and 80% for *mrpk* and i/k, respectively. This is despite the fact that privately held firms tend to operate fewer plants than their publicly traded counterparts, which, as discussed earlier, leaves less room for within-firm dispersion in the first place.

This result may appear surprising. After all, it could be expected that by impeding the efficient allocation of capital, financial frictions would increase the dispersion of *mrpk* across firms. Yet our results indicate that firm-level borrowing constraints may in fact shape the allocation of capital across plants *within* the firm. This suggests that the internal capital market of a multi-plant firm could play an important role in overcoming external financial frictions. In the next section, we aim to gain insight into this channel by building a model of a multi-plant firm facing various types of frictions, including financial ones.

B Numerical solution of model

To solve the model, we discretize plant-level capital stock using an N_k-point grid, where $N_k = 100$ to produce the results in this paper. This implies that the two-plant capital grid contains a total of 10,000 points. In addition, the shock process is approximated by an 8-point grid, from the combination of plant-specific and firm-specific Markov chain processes.

We use a hybrid iterative method to solve the model. First, we iterate and maximize over the (k'_A, k'_B) pairs of plant-specific capital until convergence of the policy function. Then, we continue iterating until changes in the value of the firm between iteration steps is below a given threshold for all states.[38] We then ensure that the policy function is indeed stable. This method, while not particularly computationally efficient, allows us to handle the numerous non-convexities of our model. We also tested to verify that lowering or increasing N_k did not have any meaningful impact on our results.

Next, we simulate a single two-plant firm over 100,500 periods, throw out the first 500 observations to allow for burn-in, and then create a panel of 500 two-plant firms with the simulations that were kept. This approach is appropriate because there are no aggregate shocks in our setup: With uncorrelated firm-level shocks, we are not required to simulate a panel of firms period-by-period. Simulated moments are computed on this firm panel.

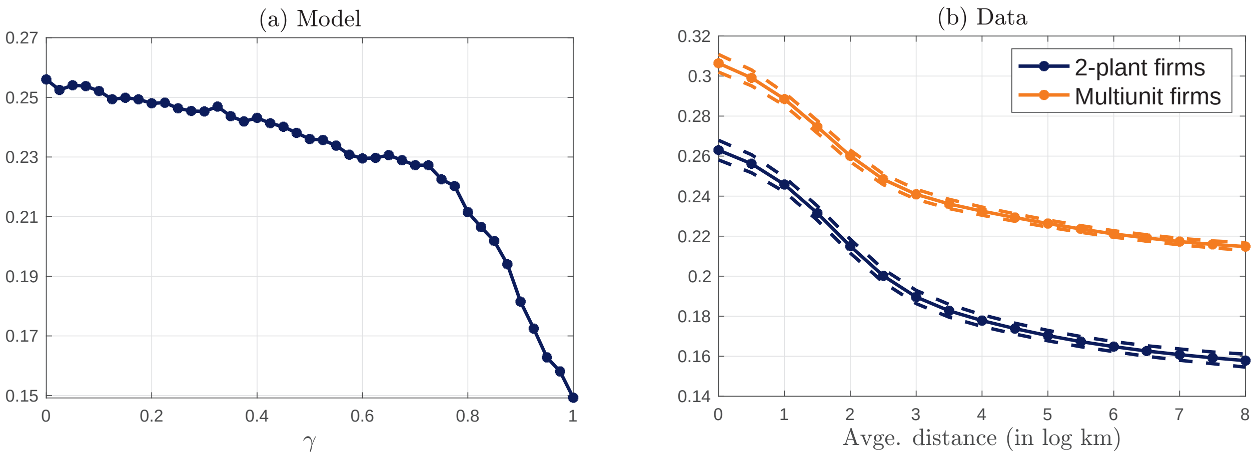
However, since shocks are completely uncorrelated across firms in our panel, dispersion *across* firms is mechanically higher than it would be if we allowed for aggregate disturbances. Because our focus is not on aggregate time-series properties, we follow a different route and instead adjust the between-firm dispersion measures in order to match the within-firm share of *mrpk* dispersion found in the data (equal to 0.6; see Table 2). While this allows for more meaningful dispersion comparisons across various scenarios, no other moments are affected by this adjustment.

C Additional empirical evidence of the model mechanism

In this appendix, we reproduce Figures 4-9, which compare the model implied moments to the analogous moments in the data for the sample of all multi-plant firms.

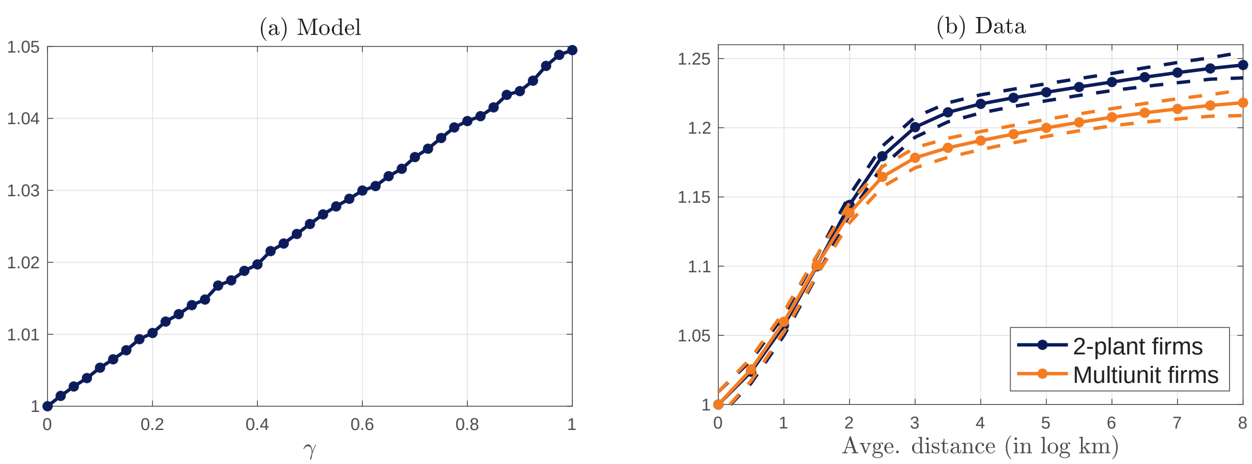
[38] Because we report the value of firms under various economic environments, we cannot solely rely on the convergence of the policy function.

Figure C1: Dispersion of *mrpk* within firms



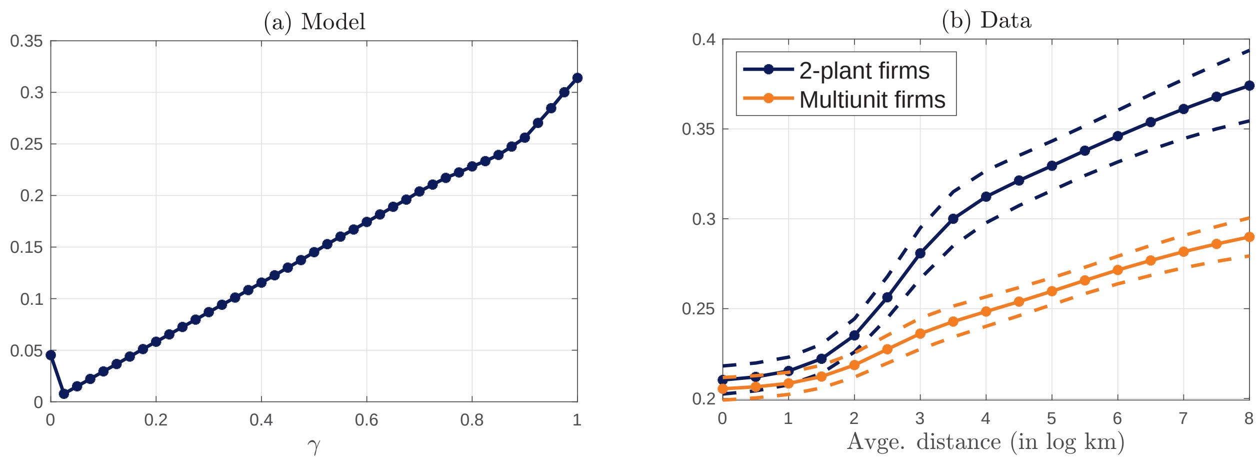
Note: Left panel: Dispersion of expected *mrpk* within firms as in Figure 2. The right panel plots its best empirical analogue for both 2-plant firms and multi-plant firms, the variance of average revenue products of capital across plants within multi-unit firms vis-a-vis geographic distance of the plants from the firm's center point. Underlying sample comprises all multi-unit firms in the ASM.

Figure C2: Firm-level *mrpk*



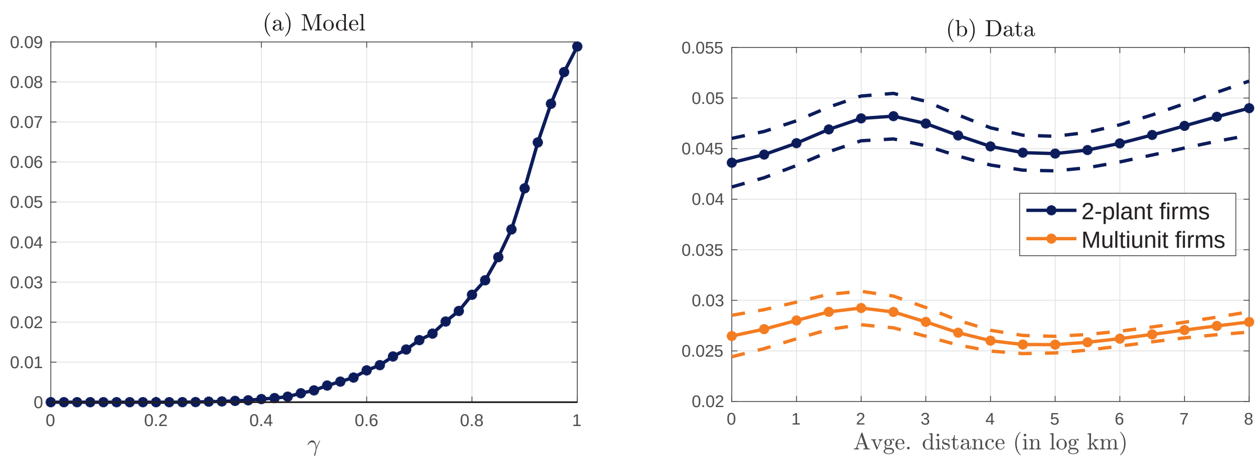
Note: See notes to Figure C1.

Figure C3: Volume of external financing conditional on external finance



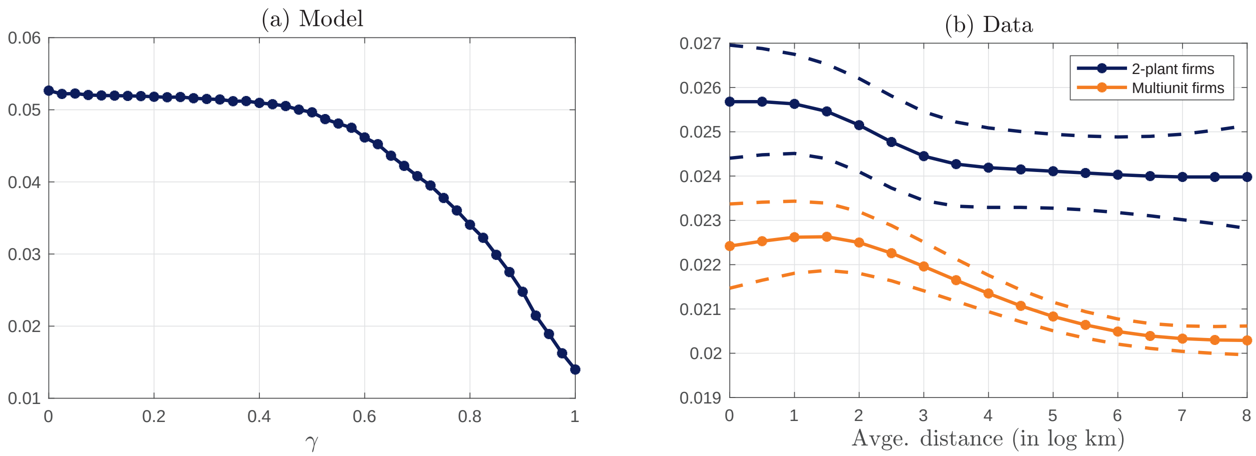
Note: See notes to Figure C1.

Figure C4: Probability of synchronized investment spikes



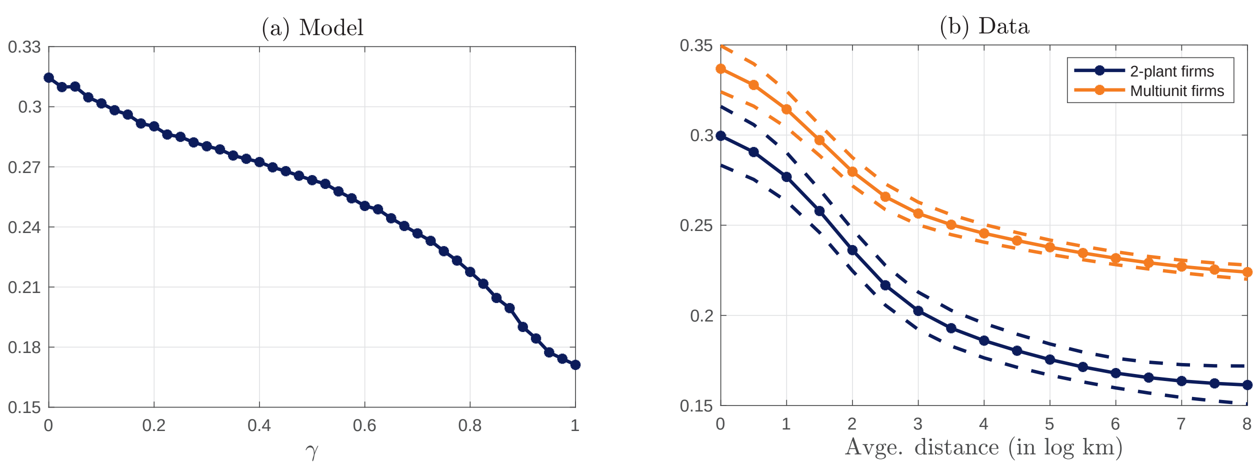
Note: See notes to Figure C1.

Figure C5: Within-firm co-movement of investment $Var(\Delta i/k)$



Note: See notes to Figure C1.

Figure C6: Within-firm co-movement of $mrpk$: $Var(\Delta mrpk)$



Note: See notes to Figure C1.