Do Firms Mitigate or Magnify Capital Misallocation? Evidence from Plant-Level Data*,**

Matthias Kehrig† and Nicolas Vincent‡

December 12, 2016

Abstract

Almost two thirds of the cross-plant dispersion in marginal revenue products of capital occurs across plants within the same firm rather than between firms. Even though firms allocate investment very differently across their plants, they do not equalize marginal revenue products across their plants. We reconcile these findings in a model of multi-plant firms, physical adjustment costs and credit constraints. Credit constrained multi-plant firms can utilize internal capital markets by concentrating internal funds on investment projects in only a few of their plants in a given period and rotating funds to another set of plants in the future. The resulting increase in within-firm dispersion of marginal revenue products of capital is hence not a symptom of misallocation within the firm, but rather actions taken by the firm to mitigate external credit constraints and adjustment costs of capital. Economies with multi-plant firms produce more aggregate output despite *higher* dispersion in marginal revenue products of capital compared to economies with single-plant firms. Because emerging economies are predominantly populated by single-plant firms, the gains from reducing their distortions to the level of developed are larger than previously thought.

Keywords: Misallocation, Productivity Dispersion, Multi-Plants Firms, Internal Capital Markets.

Jel Codes: E2, G3

*This paper grew out of our earlier paper "Financial Frictions and Investment Dynamics in Multi-Plant Firms." We thank seminar participants at CREI, Birmingham, Wharton, Duke, Oxford, Penn State, Iowa, UW Madison, Arizona State, Ente Einaudi, Boston College, UT Austin, the Federal Reserve, the FRB of St. Louis, the Dutch Central Bank, the Barcelona Summer Forum, the NBER Summer Institute, the Ifo Conference on Macroeconomics and Survey Data, the Midwest Macro Meetings, the Empirical Macro Workshop and the Annual Meeting of German Economists Abroad as well as Andrew Abel, Klaus Adam, Christian Bayer, David Berger, John Cochrane, Russ Cooper, John Haltiwanger, Hugo Hopenhayn, Pete Klenow, Ben Moll, Gordon Phillips and Michèle Tertilt for helpful discussions.

**Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The latest version of this paper can be downloaded at http://papers.ssrn.com/abstract=2731594

†Duke University and University of Mannheim, matthias.kehrig@gmail.com.

‡Institute of Applied Economics, HEC Montréal, 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 in turn to a revival of interest in the causes and consequences of resource misallocation. For instance, Hsieh and Klenow (2009) have found that lowering this dispersion through better allocation of resources across manufacturing plants could yield substantial aggregate output gains. Alternatively interpreting dispersion as exogenous increases in uncertainty, Bloom (2009) has shown how such uncertainty shocks may cause recessions.

The implicit assumption in the literatures growing out of these seminal papers is that a high level of productivity dispersion is a sign of resource misallocation in an economy and therefore reduces welfare. But does higher productivity dispersion always depress output and welfare? The present paper questions this line of thought: we show that *removing* constraints from the optimization problem of firms may in fact *increase* productivity dispersion across production plants, and we provide novel empirical evidence supporting this argument. Our quantitative exercises reveal that a sizable share of the empirically observed productivity dispersion can be attributed to optimal firm behavior rather than misallocation. As a consequence, these theoretical, empirical and quantitative findings suggest that previous work may have substantially miscalculated the gains from lowering productivity dispersion.

While the central point of our paper could potentially apply in many settings, our analysis focuses on the role of multi-plant firms in generating productivity dispersion.The reasons are two-fold: First, even if the literature has historically made little distinction between them, firms and plants are fundamentally different institutions. While, in markets, firms compete for resources, 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. 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. The within-firm margin is also important over the cycle: most of the increase in overall dispersion during recessions occurs within rather than between firms. These empirical facts are quantitatively significant: a typical firm may increase output by about 25% if it shifted an optimal amount of resources towards its high-productivity plants. In strong downturns such as the 1980/82 recession, this potential output gain is even larger.

This set of novel empirical findings leads us to believe that any study of productivity dispersion across manufacturing plants should arguably pay close attention to the role played by the allocative decisions of firms. They also pose a theoretical challenge, however: Why do firms forego such 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 than markets may therefore shed new light on the causes of resource (mis)allocation and how the economy copes with aggregate disturbances such as credit or uncertainty shocks.

To answer these questions, we build a model of capital allocation in a multi-plant firm facing various constraints and study its properties. In shaping our framework, we are mindful of the many frictions and imperfections that have been suggested in models of investment (see Caballero (1999) for an overview). These comprise "technological" frictions such as a fixed cost or capital irreversibilities (see among others Abel and Eberly (1994), Caballero et al. (1995), Doms and Dunne (1998), Cooper and Haltiwanger (2006), Gourio and Kashyap (2007)) or financing frictions (see among others Fazzari et al. (1988), Gilchrist and Himmelberg (1995)). Most of these papers have considered frictions in isolation.[1] Moreover, frictions do not all matter at the same level of aggregation. While some, mostly technological, are most relevant at the level of the plant, others such as financial or organizational frictions affect the firm as a whole. In our model, the firm organizes internal and costly external financing across plants that face both fixed and convex adjustment costs of adjustment.

We then turn our attention to studying the properties of the model, and in particular the interaction of the technological and credit frictions. When firms are credit constrained, 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*. This optimal investment policy leads to a rise in the dispersion of both investment rates and marginal revenue products of capital within the firm. In addition, credit-constrained firms will exhibit less correlated investment rates across their plants than their unconstrained counterparts. Using micro data from the U.S. Annual Survey of Manufactures, we find empirical support for these predictions at the micro and macro levels.

Next, we use the framework to conduct various counterfactual exercises aimed at quantifying the social value of firms and highlighting how the presence of multi-plant firms alters the response of an economy to aggregate disturbances. To do so, we calibrate the model

[1] Gomes (2001) and Eisfeldt and Muir (forthcoming) who have combined real and financial frictions in a unified model of a firm operating one plant are notable exceptions.

based on moments of the U.S. Economic Census and standard parameter values in the literature. As a first step, we show that introducing firm-level financial frictions in the form of a collateral constraint increases the dispersion of marginal revenue products of capital within the firm by more than 45% relative to the level of dispersion stemming solely from technological frictions at the plant level as in Asker et al. (2014).

Our model is able to explain a significant portion of the within-firm dispersion of marginal revenue products, but this finding *per se* is not a confirmation that higher productivity dispersion may come hand in hand with a more efficient allocation of productive resources. This is what we turn our attention to next. Using our calibrated model, we are now in a position to quantitatively address the key question of our study: can dispersion be beneficial? In our economy, the answer is unambiguously 'yes'. To illustrate this point, we compare the dispersion of marginal products of capital as well as aggregate quantities in two economies: one in which multi-plant firms can pool resources and allocate them freely across their production plants, and one in which these internal capital markets are shut down. We find that overall dispersion is 40% higher in the multi-plant-firm economy, *yet* aggregate investment and output are larger by 4.2% and 2.4%, respectively. This is because with fixed investment adjustment costs, only large investment projects are profitable. Since they can leverage internal capital markets, multi-plant firms are more likely to invest in new capital, and boost aggregate output. In other words, we find that firms optimally raise dispersion by exploiting internal capital markets to overcome the financial and technological frictions they face.

Our quantitative exercises suggest two things. First, the interaction of plant- and firm-level frictions is quite powerful in generating substantial within-firm heterogeneity. Second, dispersion of marginal revenue products of capital is not necessarily an indicator of misallocation or inefficient investment policies. It could be the outcome of constrained-efficient decisions that are, in fact, welfare improving. In our setting, they are the by-product of the firm's effort to mitigate external financial frictions. Hence, we view our results as a cautionary tale to interpreting higher productivity dispersion as a sign of resource misallocation.

Our findings have potentially important implications for the literature on resource allocation. Starting with Hsieh and Klenow (2009), a large literature has studied productivity dispersion across plants without considering the positive aspects of dispersion in marginal revenue products. Bayer et al. (2015) distinguish the long-term from the temporary components of dispersion, while Buera et al. (2011), Khan and Thomas (2013), Moll (2014), Midrigan and Xu (2014) have studied the impact of financial frictions on misallocation, usually ignoring their effects within firms that operate several plants as we do. In an additional

quantitative exercise, we show that Hsieh and Klenow (2009) may in fact have *underestimated* the gains from reallocation in emerging economies such as India. When we make the U.S. economy more comparable to that of India by populating it with single- instead of multi-plant firms, dispersion in the U.S. is even smaller relative to that of India. This, in turn, implies that output losses from a distorted resource allocation in India are arguably higher than initially estimated.

Ultimately, we see our project as a first step into modeling 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. Buera et al. (2011), Khan and Thomas (2013), Moll (2014), Buera and Moll (2015), Midrigan and Xu (2014) have studied the impact of financial frictions on misallocation, but in the process have ignored how such firm-level frictions can in turn shape plant-level dynamics within the firm. In related work, some theoretical research has been done on the efficiency of internal versus external capital markets: Stein (1997), Malenko (2016) study mostly principal-agent problems between a firm's owner and manager in a single-plant setup. Whether internal capital markets are more or less efficient is theoretically ambiguous: Gertner et al. (1994) show that division managers may exploit imperfect monitoring by firm headquarters to build up "inefficient empires." On the other hand, Scharfstein and Stein (2000) say that firms are likely to monitor investment spending better than outside financiers, thus improving capital allocation. 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 Müller (2015), empirical research on within-firm dynamics is scarce and often limited to studying major business divisions of conglomerates. While we take the organizational structure of the firm as given, Ševčík (2015) considers the endogenous formation of multi-plant firms (business groups).

Our paper is organized as follows. In Section 2, we describe the data and show evidence on the importance of the within-firm dimension for the dispersion of marginal revenue products of capital and investment. Section 3 describes our multi-plant model of the firm and analyses its predictions when an external financing constraint is introduced. Section 4 conducts various quantitative exercises geared towards understand the nature of productivity dispersion and carry out counterfactual exercises. In Section 5, we investigate whether the model predictions are borne out in the micro-level data. Section 6 concludes.

2 Dispersion between and within firms

2.1 Variables of interest and data sources

Our main objects of interest are the dispersions of real marginal revenue products of capital and that of real investment rates across plants. They reflect differences in productivity of capital as well as divergence in capital allocation activity, respectively. In the standard economic models generally used in the literature on misallocation, when marginal revenue products of capital ($MRPK$) differ across plants in the economy, reallocating capital from low-$MRPK$ to high-$MRPK$ plants may increase aggregate output. In a standard frictionless economy, agents should choose investment to equate the expected $MRPK$.[2] Following the literature, we assume a Cobb-Douglas production technology, approximate the marginal revenue product of plant n in time t with its real value added per unit of capital and study the variance of its logarithm: $Var_t\left(mrpk_{nt}\right) \equiv Var_t\left(\log MRPK_{nt}\right) = Var_t\left(\log(y_{nt}/k_{nt})\right)$.

Our data source is the Annual Survey of Manufactures (ASM), an annual Census dataset covering manufacturing plants and described in detail in Appendix A.2. Census 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 refer to establishments as "plants." Each plant also carries information about its parent firm which is defined by Census as the collection of plants under common ownership or control.

2.2 Dispersion of revenue products and investment within firms

The main objective of this empirical section is to document how much of economy-wide dispersion is accounted for by heterogeneity across plants within the same firm in contrast to heterogeneity across firms. By shedding light on the sources of misallocation, the answer has potentially important implications for future research efforts in this literature, as well as the structure and calibration of micro-founded macroeconomic models. Throughout the section, we study not only dispersion in $mrpk$ but also i/k, an object that has received relatively little attention so far in the misallocation literature. This is surprising since investment activity is the direct channel through which firms and plants can overcome the possible misallocation

[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 towards the units with the highest $TFPR$ if they already operate a large capital stock. In fact, investment should flow to units with the highest expected capital return.

underlying *mrpk* dispersion. We also present evidence on the cyclical properties of dispersion in *mrpk* and i/k. In doing so, we expand on Eisfeldt and Rampini (2006), who document the cyclical capital reallocation between firms, by studying the within-firm patterns of *mrpk* dispersion and capital reallocation over the business cycle.

Our emphasis on contrasting the dispersion across firms to that occuring within the firm is only relevant if a significant portion of aggregate economic activity is accounted for by multi-plant firms. Table 1 provides evidence that it is. While single-plant firms dominate in numbers, it is the multi-plant firms that 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 half of all the capital stock in U.S. manufacturing.

Table 1: 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.

We now investigate to what extent marginal revenue products of capital and investment rates differ across plants within multi-plant firms and how it compares to differences across firms. This can be done by decomposing the total variance of marginal revenue products of capital, denoted by V_t, into the variance between firms, denoted by V_t^B, and the average variance between plants within firms, denoted by V_t^W. Since marginal revenue products are expressed in units of industry good per unit of capital, comparing dispersion of *mrpk* across industries is not very meaningful. Any arbitrary definition of what constitutes a typical good in an industry will automatically impact the measured *mrpk* dispersion in the economy. We therefore perform the decomposition within sufficiently homogeneous industries and report results on *mrpk* dispersion averaged across industries. We define an industry at the 4-

digit NAICS level in order to make goods within such industries sufficiently substitutable to compare their marginal revenue products while leaving enough observations in an industry-year cell to reliably study heterogeneity. In a given 4-digit NAICS industry our variance decomposition is then:

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

where n indicates the plant, j the firm and t the year. 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 take into account "economic relevance." In the former case, we have $\omega_{njt} = 1/N_t$ (where N_t is the number of observations), in the latter $\omega_{njt} = k_{njt}/k_t$ and accordingly for ω_{jt} and ω_{nt}^j. We aggregate the industry-specific between-firm and within-firm decompositions across the 86 industries at the 4-digit NAICS level as detailed in Appendix A.4. The basic results of that decomposition are displayed in Panel A. of Table 2; Panel B. of that same table displays general cross-sectional moments of $mrpk$ and i/k.

Two results from our simple accounting exercise stand out: First, overall dispersion in marginal revenue products is large. In the average industry and year, the standard deviation of its logarithm is 0.915. This means that a plant that is one standard deviation above the mean produces $e^{0.889} \approx 2.4$ times the value added as the average plant with the same capital stock; the difference between the plant at the 90th percentile compared to that at the 10th percentile even implies an $e^{2.134} \approx 8.5$-fold value-added difference. Reallocating capital to high-$mrpk$ plants in the same industry could hence result in a considerable boost in aggregate output. If one interprets $mrpk$ dispersion as misallocation, this means that the economy potentially foregoes a considerable amount of income. To illustrate this loss, imagine there are two production units in the economy. Calibrating their Cobb-Douglas technology and productivity to the empirically estimated parameters in Section 4, we find that cutting the observed dispersion in marginal revenue products just in half would increase aggregate output by 19.3%.

Table 2: Cross sectional moments of capital and investment

Variable	*A. Share of Variance*		*B. General moments across plants*					
	between firms	within firms	Mean	StD	IDR	Skew-ness	Kelley Skewn.	Excess Kurtosis
$mrpk$	0.399	0.601	-/-	0.889	2.134	0.424	0.099	1.361
				(0.002)	(0.008)	(0.006)		(0.102)
i/k	0.321	0.679	0.096	0.230	0.175	5.234	0.479	38.945
			(0.003)	(0.002)	(0.008)	(0.006)		(0.012)

Note: 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 all Census years 1982-2007 are computed for each industry and years first before being aggregated by industry and the averaged across years. For details see Appendix A.3.

The second takeaway from our accounting exercise is that the majority of dispersion occurs inside firms. More precisely, over 60% of marginal revenue products of capital dispersion in the typical industry occurs within firms, with the remainder accounted for by between-firm variations. The implication is that most of the benefits from reallocation would take place by shifting of capital across the production plants within a given firm, instead of between firms. Analogously to the above exercise, we document that a plant that is one standard deviation (0.709) above the firm's average produces twice the value added with the same capital stock as the firm on average.

Interestingly, the cross-sectional distribution of $mrpk$ is positively skewed, which is consistent with the presence of credit constraints: some productive plants may have too little capital, which thus drives up their marginal product and skews the $mrpk$ distribution to the right. Following Kelley (1947), p. 250, we define the Kelley skewness as $\gamma^{Kelley} = \frac{mrpk^{90}+mrpk^{10}-2mrpk^{50}}{mrpk^{90}-mrpk^{10}}$. This quantile-based measure of skewness is 0.099 on average which means that the top half of the inter-decile range, $mrpk^{90} - mrpk^{50}$ is about $1.099/0.901 = 22\%$ more spread out as the bottom half of the distribution, $mrpk^{50} - mrpk^{10}$.

Investment rates also differ substantially across production plants, which implies that allocative activity of capital differs greatly across units within a typical year. Note that since investment rates are measured in percentage points, they are not subject to the same industry-specific measurement problems present with marginal revenue products. The cross-sectional standard deviation of 23% is large given that the average plant in the economy

has an investment rate of 9.6%, an indication of the lumpy nature of investment. Like with *mrpk* dispersion, the majority of investment rate dispersion originates within firms rather than between firms: within the typical industry, the within-firm share of investment rate variance amounts to almost 70%.

2.3 Robustness

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 to our main empirical result as well as identify possible causes behind the importance of within-firm dispersion. Details about these exercises can be found in Appendix A.5.1. Table 3 provides an overview of our robustness checks.

Table 3: Dispersion of *mrpk* within and between firms

	Sample	$Var(mrpk)$		$Var(i/k)$	
		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)	5-year averages	0.554	0.446	0.656	0.344
(6)	5-plant firms	0.668	0.332	0.786	0.214
(7)	Homog. Industries	0.559	0.441	0.689	0.311
(8)	y is physical output	0.625	0.376		
(9)	K-weighted	0.540	0.460	0.680	0.320
(10)	Equipment	0.591	0.409	0.631	0.369
(11)	Private firms	0.713	0.287	0.795	0.205

The first row in Table 3 reiterates the baseline result of within-firm versus between-firm dispersion of both marginal revenue products of capital (navy blue on the left) and investment rates (light blue on the right): 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 if 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 or 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 if 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 that 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 considered a five-year distance to entry and exit. If we turn to a strongly balanced panel (Row (4)), the share of within-firm dispersion is even larger.

Ruling out measurement error In our next robustness check, we want to address the possibility that the high within-firm share primarily reflects measurement error at the plant level. Indeed, transitory plant-level noise would "wash out" at the firm level and thus artificially increase the within-firm share of dispersion. To rule out that this effect drives our result, we construct rolling 5-year windows of averaged $mrpk$ and i/k for each plant. This time aggregation should filter out most of high-frequency noise at the plant level. Performing the between-firm/within-firm decomposition on this averaged data shows that the importance of the within-firm share of dispersion persists. As shown in Row (5) of Table 3, the within-firm share of overall dispersion in an industry are now 55% and 66%, respectively. This suggests that the presence of plant-level measurement error is unlikely to be driving our findings. Further investigations of measurement error can be found in Appendix A.5.1.

Ruling out mechanical aggregation 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 does not probably 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.[3] Row (6) 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.

Ruling out multi-product-firm bias Even though our benchmark definition of an industry is fairly fine, products within 4-digit NAICS industries are still heterogeneous. This could potentially lead to spurious differences of marginal revenue products of capital arising from differences in product composition within 4-digit NAICS industries. Whether such within-industry product differences are most likely to occur between or within firms is ambiguous. For robustness purposes, we repeat the decomposition but this time focusing on plants that produce only one homogeneous standardized good. We follow Foster et al. (2008) and consider industries that produce almost perfectly homogeneous goods such as cement, sugar, coffee beans etc.[4] Naturally, we expect the within-firm share of dispersion to be smaller then in our benchmark since many firms within a given 4-digit NAICS industry are spread out across several of these narrowly-defined product codes. But even in these homogeneous industries, the within-firm share of dispersion in marginal revenue products of capital and investment rates displayed in Row (7) amounts to 56% and 69% respectively.

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 differ-

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

[4] More specifically, these "industries" are defined as 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).

ential markups or transfer prices – rather than physical productivity differences. Row (8) 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 stake 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 (9) shows that while the within-firm share of capital-weighted dispersion is slightly lower, it remains still dominant at 54% and 68%, respectively.

Examining different capital types In Row (10), 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 structure capital which would lower dispersion. Again, the results of the unweighted between-firm/within-firm decomposition are almost unchanged at 59% and 63%, respectively.

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 was even higher, suggesting that our baseline results may in fact represent lower bounds on the actual within-firm share of dispersion.

2.4 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 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), TFP levels by Kehrig (2015), and TFP innovations by Bloom et al. (2012), while Bachmann and Bayer (2014) have shown that the

dispersion in investment rates is procyclical. These findings have important implications for the literatures on Schumpeterian creative destruction, misallocation, development or uncertainty-driven business cycles. Yet, to our knowledge, no one has investigated seperately the cyclicality of dispersion within firms. We close that 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[5] which we then use for time-series analysis. Table 4 displays properties of both the long-run averages, autocorrelations and time series standard deviations for each measure.

Table 4: Dynamic properties of within-firm and between-firm dispersion

	$V^B(mrpk)$	$V^W(mrpk)$	$V^B(i/k)$	$V^W(i/k)$
A. Time-series moments				
Average	0.341	0.516	0.009	0.021
Autocorrelation	0.732	0.657	0.667	0.679
Volatility	0.042	0.084	0.006	0.016
B. Cyclicality				
$Corr(\Delta Y^{mfg}_{t+1}, ...)$	-0.252	-0.147	0.306	0.330
$Corr(\Delta Y^{mfg}_{t}, ...)$	-0.582	-0.304	0.389	0.394
$Corr(\Delta Y^{mfg}_{t-1}, ...)$	-0.413	-0.229	0.222	0.241

Note: The table reports time-series moments of the between-firm and the 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 the the cyclicality is with respect to the growth rate of aggregate manufacturing value added, denoted ΔY^{mfg}_t.

Consistent with the evidence from Section 2.2, 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 conmpares coefficients of variations instead of variances, and similar patterns are observed in the between-firm and within-firm dispersion of investment rates.

As can be seen from Panel B. of Table 4, 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 Bloom et al. (2012) and Christiano et al. (2014), then our within-firm result suggests that cycles manifest themselves within as well

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

as between firms. This means that looking at the granular level of the firm, as is done in Gabaix (2011) or Eisfeldt and Rampini (2006), may underestimate the role of dispersion for aggregate fluctuations. Since fluctuations in heterogeneity at the sub-granular level of plants within firms are larger and as countercyclical, this suggests one should not discard the dynamics inside multi-unit firms when studying business cycles.

2.5 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 impeding the reallocation of productive capital. Arguably, some frictions such as technological ones as in Bloom (2009), Bloom et al. (2012) 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 we 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 less plants than their publicly-traded counterparts, which, as we discussed early, 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.

3 A model of the multi-plant firm

In this section, we describe, solve, simulate and analyze a simple model of a firm comprised of several plants. We study how various plant- and firm-level frictions interact with the optimal allocation of capital by the firm across its plants. The interdependency of a plant's capital allocation varies with constraints at the level of the firm. At one extreme, the firm is a collection of disconnected plants: decisions are made on a plant-by-plant basis, without any interactions between them. In that case, it does not matter for a plant's investment and production decisions to what firm it belongs to, how many plants that firm operates and what other common factors may characterize the firm. We show that in the presence of frictions at the firm level, however, the firm alters the size and timing of plant-level investment plans which may result in dispersed capital returns across the firm's plants. These constraints at the firm level may prove so strong that, at the other extreme, the firm will treat several perfectly identical plants differently. After describing and solving the problem of the multi-plant firm, we will examine how much dispersion in marginal revenue products of capital a calibrated version of our model can generate and how much marginal revenue products of capital dispersion remains unexplained.

3.1 The problem of the firm

We focus on the basic problem of a firm that operates two plants n, $n = A, B$. 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 the technology level of each one of its plants, without adding insight into the underlying fundamental economic mechanisms. We start by describing the technology and constraints at the level of the individual plant before analyzing the problem of the firm. In what follows, lower case letters refer to plant variables, upper case letters to firm variables and bold upper case letters refer to vectors of a firm's plant variables.

3.1.1 Technology and frictions at the plant level

Each plant n is characterized by total factor productivity which may include a firm-specific component that is common to both plants. The plant operates a Cobb-Douglas production function which 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.[6] This means we can substitute out any static first-order condition for variable inputs and write plant revenues net of variable factor costs as

$$y_{nt} = e^{z_{nt}} k_{nt}^{\alpha}. \tag{2}$$

z_{nt} contains plant (log) total factor productivity and prices of other statically chosen production factors and α is the scaled production elasticity of capital. The productivity level of plant n which belongs to firm j consists of a firm component common to both plants in the firm and an idiosyncratic plant component; both evolve as follows:

$$z_{njt} = \rho^p z_{njt-1} + \eta_{njt}$$
$$z_{jt} = \rho^f z_{jt-1} + \eta_{jt}.$$

with $\mathbb{E}\eta_{njt} = \mathbb{E}\eta_{jt} = 0$, $V(\eta_{njt}) = (V^f)^2$ and $V(\eta_{jt}) = (V^f)^2$.

The capital stock of plant n depreciates every period at rate δ and grows with investment i_{nt}, so it evolves over time according to the conventional expression

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

As documented in a number of studies (see Cooper and Haltiwanger (1993), Cooper et al. (1999), Doms and Dunne (1998), 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."[7] 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, ψk_{nt}, if investment is greater than zero.

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

[7] Investment spikes are usually defined as investment rates exceeding 15% or 20%.

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 regulates how much revenue is foregone when the plants 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 in Eberly et al. (2012).

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.[8] The parameter γ below captures the importance of this margin.

To summarize, frictions at the plant level will be expressed as:

$$\theta\left(i_{nt}, k_{nt}\right) = \left[\psi \mathbb{I}\left\{\frac{i_{nt}}{k_{nt}} > \vartheta\right\} + \gamma\left(\frac{i_{nt}}{k_{nt}}\right)^2\right] k_{nt} \tag{3}$$

where $\mathbb{I}$ is an indicator function equal to 1 if the plant investment rate is above ϑ; ψ is a parameter regulating the foregone sales if the plant undergoes an investment, and γ regulates the impact of the quadratic adjustment cost. Everything is scaled by the plant's capital stock k_{nt} in order to eliminate size differences.

Combining equations (2) and (3) above, plant cash flow is

$$\pi_{nt} = z_{nt} k_{nt}^{\alpha} - \theta(i_{nt}, k_{nt}). \tag{4}$$

3.1.2 Technology and frictions at the firm level

What sets plants in a multi-unit firm apart from their identical counterparts in single-unit firms? What are the economic benefits the firm provides to its own plants? This section discusses the allocation of funds within the firm.

A firm collects the cash flow from all plants and decides on how to allocate funds across its plants for investment projects. This means firm cash flow is

$$\Pi_t = \pi_{At} + \pi_{Bt} - \Phi. \tag{5}$$

Managing firm-wide operations requires overhead costs such as firm management which every firm needs to pay in order to be functional. This overhead fixed cost, denoted by Φ,

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

reflects items such as management costs, overall firm infrastructure, expenditures for R&D, marketing etc. Hence, compared to single-unit firms, multi-unit firms benefit from operating two plants with the same corporate overhead.

Second, while all production and investment activities take place at the level of the individual plant, only the firm is capable of organizing external finance. This assumption is realistic and sensible: while large and complex firms like General Electric operate hundreds of plants, only the firm issues bonds, borrows from banks or raises equity. Typically, the firms then allocate these funds in an internal capital market to individual plants. Consistent with this empirical pattern, we assume that it is the firm that co-ordinates investment plans across all its plants, organizes financing of investment through either internal cash flow or external borrowing and allocates funds to plants where investment is put in place. Only if desired firm-wide investment exceeds firm cash flow, the firm attempts to borrow amount B_{t+1} at net interest rate R_t so that all investment gets financed:[9]

$$i_{At} + i_{Bt} = I_t \leq \Pi_t + B_{t+1}. \tag{6}$$

Organizing external financing, however, is an imperfect process. Following the literature on financial frictions, we assume there are two types of frictions: First, there is a financial participation cost, denoted by ζK_t, if the firm wants to borrow at all. This cost reflects the effort to establish a relationship with a lender, expenditures for various information and disclosure requirements and other administrative expenses independent of the loan amount. Like the real investment adjustment cost above, this cost is scaled by the capital stock. Second, we assume that the firm can divert a fraction $1/\eta$ of the loan amount B_{t+1} to its private benefit. This diversion of resources cannot be observed or prevented by the lender. As a consequence, the lender will require collateral which it can seize in case it discovers ex post that the firm did divert funds. We assume that firms pledge a fraction ξ of the value of their capital stock, $K_t = (k_{At} + k_{Bt})$, as collateral. Given the posted collateral, lenders will limit the loan such that the diverted loan amount never exceeds the collateral in order to avoid ex post moral hazard unprofitable for the borrower:

$$\frac{B_{t+1}}{\eta} \leq \xi K_t. \tag{7}$$

Note that one can think of $\lambda \equiv \eta\xi$ as the maximum leverage the lender is willing to tolerate.

[9]For computational reasons, we do not allow for savings by the firm, except through the accumulation of capital.

In our quantitative exercise, we will calibrate λ directly, since it is neither possible not necessary to separately identify η and ξ. We summarize the cost of external finance in the following function

$$\Theta(B_{t+1}, K_t) = \begin{cases} 0 & \text{if } B_{t+1} = 0 \\ \zeta K_t & \text{if } 0 < B_{t+1} \leq \lambda K_t \\ \infty & \text{if } B_{t+1} > \lambda K_t. \end{cases} \tag{8}$$

Note that $\Theta(\cdot)$ captures the net present value of the borrowing cost associated with B_{t+1}. The interest rate payments next periods have a net present value of $\beta R_t B_{t+1}$. Since there is no risk in our model, the firm's borrowing rate R_t is equal to the risk-free interest rate which in turn equals the inverse of the discount rate, so the net present value of the linear part of borrowing costs exactly equals B_{t+1}. Of course, this would change once we introduce any risk of bankruptcy along the lines of Townsend (1979) which would drive up R_t above the risk-free rate and make the middle portion of the borrowing cost in equation (8) monotonically increasing in the borrowing amount B_{t+1}.

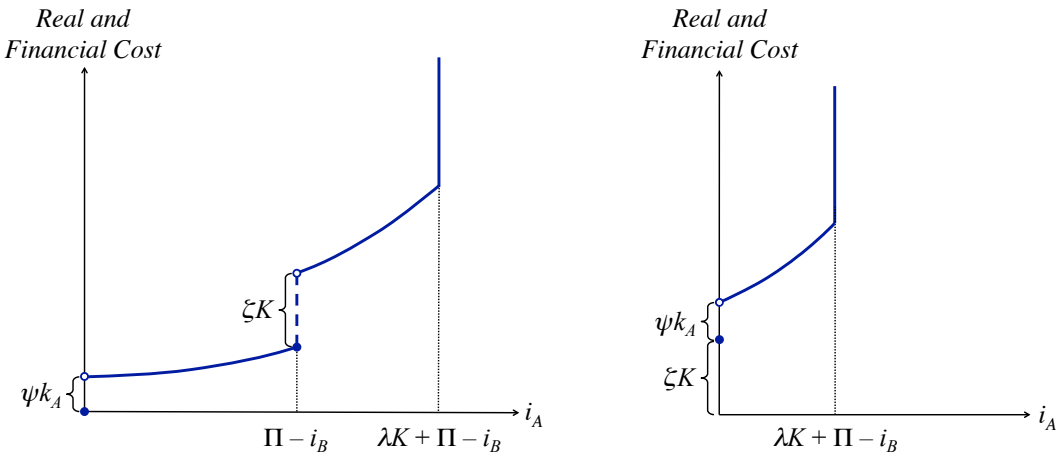
Total cost of investment in a given plant depends on the investment amount in that plant, the combined investment in the rest of the firm, whether or not the firm needs to borrow and if it possibly runs into the collateral constraint. The total cost of investment in plant A then consists of fixed and quadratic adjustment costs (real costs $\theta(i_{At}/k_{At})$ in equation (3)) as well as fixed borrowing costs (financial costs $\Theta(B_{t+1}, K_t)$ in equation (8)). The latter part depends on how much the other plant in the firm, plant B, invests as it dictates how fast and how much the firm needs to borrow. Thus, investment in one plant imposes an externality on investment in the rest of the firm because it depletes internal funds and imposes a borrowing cost that is shared by the entire firm.

We plot the total cost of investment in Figure 1 to illustrate the multiple non-convexities and how the interaction of investment across plants shape the cost of investment for the firm. Note that in these plots we assume that ψ, the parameter that regulates the fixed costs of investing in a plant, is "small" in the sense that the minimum investment in one plant can be financed using internal funds of the firm. If they were excessive, even the minimum investment to justify the fixed investment adjustment costs would require borrowing. In that case, the effective fixed cost of investing in any plant would be $(\psi + \zeta)k_{nt}$.

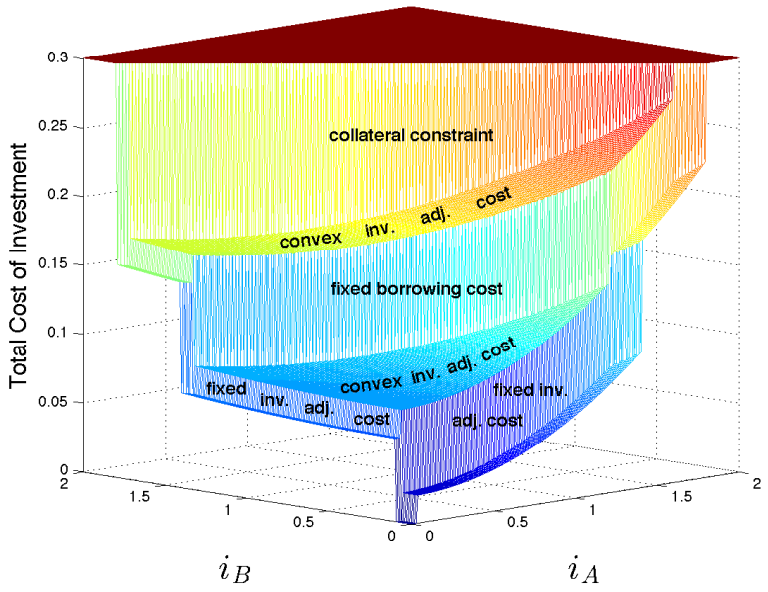
Figure 1: Total cost of investment

(a) Total cost of investing in plant A when i_B can be fully financed internally

(b) Total cost of investing in plant A when i_B cannot be financed internally



(c) Total investment costs for the firm



Note: Panel (a) on the left displays total cost of investing in plant A when investment in the rest of the firm does not exceed internal funds $0 \leq i_{Bt} \leq \Pi_t$. Then, small amounts of i_{At} can be financed with left over internal funds $(\Pi_t - i_{Bt})$ without incurring borrowing costs. Any investment exceeding that amount makes the cost level jump due to the fixed borrowing cost ζK_t. Panel (b) on the right displays the case when the firm already needs to borrow to finance investment in the rest of the firm. Even zero investment in plant A means fixed and linear borrowing costs $\zeta K_t + R(i_{Bt} - \Pi_t)$. Investment in either case is always limited by the collateral constraint: $B_{t+1} \leq \lambda K \Leftrightarrow i_{At} \leq \lambda K + \Pi_t - i_{Bt}$. Panel (c) shows the the total cost jointly for $k_{At} = 1$, $k_{Bt} = 3$, $\zeta = \psi = 0.02$, $\gamma = 0.04$, $\vartheta = 0.03$, $\lambda = 0.1$ and $\Pi_t = 1.4$.

3.1.3 Firm value and firm policy

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 will consist 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 either plant A and B in order to maximize firm value which consists of the net present value of discounted future gross profits net of investment and borrowing costs. When choosing the investment levels in either plant, the firm takes into account the various adjustment costs and whether or not borrowing is required to finance the desired level of investment. The firm's problem can be written in recursive form as:

$$
\begin{aligned}
V(\mathbf{Z_t}, \mathbf{K_t}) = \max_{i_{At}, i_{Bt}, B_{t+1}} & \left\{ \Pi_t - I_t - \Phi - \Theta(B_{t+1}, K_t) + \beta \mathbb{E} V(\mathbf{Z_{t+1}}, \mathbf{K_{t+1}}) \right\} \\
s.t. \quad & k'_{nt} = (1-\delta) k_{nt} + i_{nt} \quad \forall n = A, B \\
& B_{t+1} \leq \lambda K_t. \\
& I_t \leq \Pi_t + B_{t+1}.
\end{aligned}
$$

This value function embeds three non-convexities: the fixed investment adjustment cost, ψ, the fixed borrowing cost, ζ, and the collateral constraint λ.[10] All three non-convexities give rise to "inaction regions" in the firm's state space where the firm may choose to not change investment and/or borrowing even though the underlying productivity shocks z do. In those regions, the realized marginal revenue product of capital will also differ from that of a frictionless investment model. The three non-convex investment/borrowing costs make the solution of the firm problem fairly complex and there are consequentially at most seven different[11] cases of investment-borrowing decisions each of which reflect investment and borrowing (in)activity in the various parts of the firm.

We illustrate the investment policy for plant A and the firm's borrowing policy in Figure 2 for a given amount of investment in plant B. In principle, investment in either plant is mono-

[10] We refer to these frictions as non-convexities even though the last one is strictly speaking only a non-differentiability; but it has similar results for the dispersion of marginal revenue products of capital within the firm as the non-convexities.

[11] With three binary decisions, there are eight cases, but when neither plant invests the firm would never borrow thus making one case redundant. We restrict the investment fixed adjustment cost ψ to be small. If it were extremely high, even the minimum investment which justifies paying the fixed cost would require firm borrowing and there would only be four cases. Then, there is no role for an internal capital market to finance investment because it never suffices even for the minimum investment in one of the two plants.

tone in its TFP. The fixed investment adjustment costs prevent small amounts of investment that would not justify paying the fixed cost ψk_{At}. So investment will not change unless z exceeds a threshold $z^0(k_{At})$. At $z^0(k_{At})$ the expected benefit from investing, $\mathbb{E}\frac{\partial V'}{\partial k'_{At}} i_{At}$, equals the fixed costs, ψk_{At}. Because the latter are modeled proportional to k_{At} and because of decreasing returns to scale, the first threshold will depend on k_{At}. Note that z^0 does not depend on firm variables. Investment jumps and then increases monotonically in productivity. The quadratic investment adjustment cost regulated by γ increase the marginal cost of investing but do not break the monotonicity. Plant investment increases until the firm's internal funds, Π_t, are exhausted at $z^1_{At}(k_{At}; z_{Bt}, k_{Bt})$. Notice that this second threshold of productivity for plant A depends on investment in the rest of the firm, i.e. i_{Bt}, and vice versa. This is because investment expenditures in the other plants entail a negative externality for plant A as it depletes internal funds of the firm.

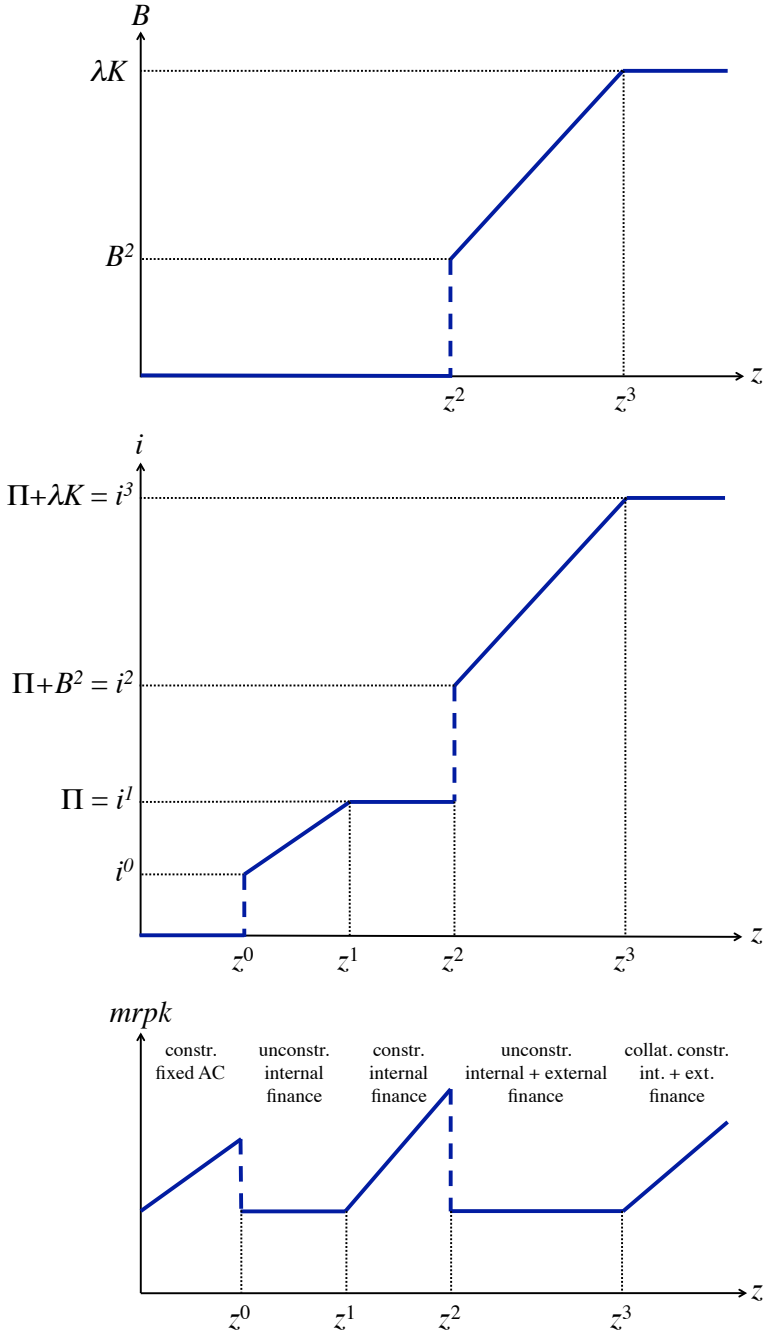
As with the fixed adjustment cost at the plant level, the fixed borrowing costs, ζK_t, prevent the firm from borrowing small external funds. This means there is another region of productivities $[z^1_{At}(\cdot), z^2_{At}(\cdot)]$ where investment does not respond to productivity. Only at productivities above $z^2_{At}(z_{At}, k_{At}; z_{Bt}, k_{Bt})$ is it worthwhile for the firm to pay the fixed borrowing cost. This third threshold depends on investment needs in the entire firm and hence depends on the levels and the distribution of capital stocks and the productivity shocks. Investment again increases monotonically in productivity above $z^2_{At}(\cdot)$ until productivity threshold $z^3_{At}(z_{At}, k_{At}; z_{Bt}, k_{Bt})$ at which point the firm's borrowing demands are limited by the collateral constraint.

3.2 The effects of external financial constraints in multi-plant firms

In the previous section, we described how investment at the individual plant is affected by firm-level frictions. Understanding these micro-level dynamics is helpful to gain intuition about the *joint dynamics* of marginal revenue products of capital and capital reallocation within the firm, which is central to the model. Next, we study the effect of each firm-level financial constraint on the dispersion of returns and investment rates as well as additional moments that can be empirically checked in the data. We qualitatively explain how interactions between plants shape the firm's investment and borrowing policies, leaving our quantitative assessment to Section 4.

Our benchmark for the analysis is a version of the model in which the only frictions are the quadratic and fixed investment adjustment costs. This makes our comparisons more

Figure 2: Firm borrowing and plant investment



Note: Schematic illustration of firm borrowing, plant investment policy functions and realized marginal revenue products of capital. This illustration assumes that the other plant does not invest at all and that physical adjustment costs ψ are smaller than fixed borrowing costs ζ, so that firms start investing without having to borrow; otherwise z^0 and z^2 would coincide.

meaningful, since in a completely frictionless world the dispersion of the marginal revenue products of capital would be zero. In addition, the non-convexity from the fixed cost is in itself a known source of *mrpk* dispersion (see Asker et al. (2014)).

The framework we described in Section 3 includes two types of financial frictions. The first is a fixed borrowing cost that must be paid every time the firm engages in external borrowing. The impact of this constraint are intuitive. First, our model predicts that fixed borrowing costs lead to investment spikes that are synchronized across plants within the firm. Because accessing capital markets is onerous, the firm minimizes these costs by borrowing less often but larger amounts. It then uses the funds obtained to invest in both plants at once. In those periods of firm-wide investment, the correlation between investment and marginal revenue products of capital is high, while the opposite is true during the long periods of internally-financed small investment projects. The latter tend to dominate, leading to lower overall correlation between investment and marginal revenue products of capital. Finally, because the fixed borrowing costs make investment at the plant level even lumpier, the autocorrelation of investment becomes even more negative, at both the plant and firm level.

While it is enlightening to study its effects, the fixed cost of accessing external capital markets plays little role in our quantitative analysis. Instead, our focus is on a type of financial friction that is more standard in the literature: a collateral constraint that limits the amount of borrowing that can be undertaken by the firm. This constraint is most likely to bind when both plants within the firm are hit by positive technology shocks: the firm would like to invest heavily in both units, yet faces limited internal funds. As a consequence, the firm is forced to focus on one plant at a time, leveraging both internal and external financial markets in the process. In other words, it rotates investment projects across its plants. Seen over time, investment across the individual plants within a firm looks staggered.[12]

What does this staggered investment policy imply for the dispersion of marginal revenue products of capital across plants? By rotating investment across its plants, the firm is in effect creating higher within-dispersion in *mrpk*. More interesting is the fact that this increased dispersion is the optimal response of the firm to the constraints it faces rather than a consequence of inefficient distortions. In fact, were the firm unable to leverage internal capital markets by pooling plant revenues, it may not be in a position to invest at all because of its limited access to external finance. The result would be a lower dispersion of *mrpk* within the firm. In addition, the staggering of investment activity naturally leads to a rise in the

[12] Besley et al. (1993) study the "Rotating Savings and Credit Associations" (Roscas), a financial institution mostly found in developing economies which relies on a similar mechanism.

within-firm dispersion of investment rates and a drop in their correlation across plants.

The two crucial elements in this beneficial interpretation of *mrpk* dispersion are the fixed investment adjustment costs ψ and the collateral constraint λ. This leads to investment spikes that the firm can stagger over time thus resulting in higher *mrpk* dispersion across the plants within the firm. If one the two crucial components of our model mechanism would be absent, there would not be any role for "beneficial *mrpk* dispersion." For example, without a collateral constraint, the firm would have unlimited access to financial resources and invest to equalize *mrpk* across its plants. In other words, absent distortions à la Hsieh and Klenow (2009), any dispersion in *mrpk* would be detrimental but merely reflect the presence of the technological fixed investment adjustment costs ψ as in Asker et al. (2014).[13] Conversely, without the presence of fixed adjustment costs, a credit-constrained firm would invest suboptimally low amounts, but would still allocate investment to equalize marginal revenue products across its plants.

Our discussion so far has been purely qualitative. In the next sections, we explore the quantitative relevance of the channels highlighted above and discuss the link between *mrpk* dispersion and aggregate welfare in a multi-firm economy. Finally, in Section 5.1 we present supporting evidence for the mechanisms at play in our model.

4 Quantitative analysis

4.1 Calibration

Table 5 summarizes the parameter values used to calibrate our model for the quantitative analysis. Most values are based on moments from the ASM dataset and are in line with calibrated parameters generally used in the investment literature.

In order to learn about the production function elasticity α, we extend the structural framework of Cooper and Haltiwanger (2006) to accommodate multi-plant firms. Our GMM estimate puts α at 0.627 which is fairly close to their value. The parameters governing persistence, ρ^p and ρ^f, and volatility, σ^p and σ^f, of the plant and firm shock processes are chosen to match the persistence and volatility of *mrpk* at the plant and firm levels in two-plant firms in the ASM.

[13] Similar results could be obtained in the case of a 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.

Table 5: Model Calibration

Parameter	Meaning	Value	Target/Source
β	Discount rate	0.95	Long-run real interest rate
α	Production elasticity	0.627	ASM data
ρ^p	TFP persistence plant	0.60	serial correlation of $mrpk_p$: 0.25
ρ^f	TFP persistence firm	0.85	serial correlation of $mrpk_f$: 0.31
σ^p	TFP shock plant	0.25	volatility of $mrpk_p$: 0.33
σ^f	TFP shock firm	0.24	volatility of $mrpk_f$: 0.26
δ	Depreciation rate	0.067	Mean investment rate in ASM data
ψ	Fixed inv. adj. cost	0.039	Cooper and Haltiwanger (2006)
γ	Quadratic inv. adj. cost	0.049	Cooper and Haltiwanger (2006)
ζ	Fixed borrowing cost	10^{-4}	See text
λ	Collateral constraint	0.30	Li et al. (2016)
Φ	Firm fixed cost	0	See text

Of note, the fixed investment adjustment costs are lower than those estimated in Cooper and Haltiwanger (2006), Asker et al. (2014). This is because the fixed borrowing costs ζ reinforce the effects of fixed investment adjustment costs ψ. As a result, estimating a model that does not take into account firm structure will confound the two and bias the inferred value for ψ upwards. In addition, when we use the value from Cooper and Haltiwanger (2006), we find that the minimum investment projects needed to justify the fixed adjustment costs are so large that they usually require borrowing. While this case may be theoretically possible, it seems not very plausible that even minimal investment cannot be financed out of the combined firm's internal funds.

Given the lack of hard evidence to support the calibration of the borrowing fixed cost ζ, we remain conservative and pick a small value equal to only 0.01% of the firm capital stock. Also, we start with no firm fixed production cost Φ, which mostly acts as a scale factor anyway. Regarding the collateral constraint, our benchmark case relies on $\lambda = 0.3$, while we will be experimenting with other values as well to illustrate the key driver of within-firm capital allocation. This value is in line with the evidence from Li et al. (2016) who estimate a dynamic model of the firm with a similar borrowing constraint. They find values of the collateral constraint parameter between 0.22 and 0.32 across manufacturing industries.[14]

[14]In the case of Li et al. (2016), it is the non-depreciated value of the capital stock that is collateralizable. Also, while Rampini and Viswanathan (2013), D'Acunto et al. (2016) also provide information on leverage, their approach is not model-based which makes it difficult to transpose their findings into our framework.

4.2 Financial constraints, investment activity and dispersion

Next, we turn our attention to a quantitative analysis of frictions and dispersion through model simulations. The problem of a representative firm is solved and simulated using a value function iteration procedure which is described in detail in Appendix B. Then, we simulate an economy composed of 1,000 two-plant firms for 1,000 periods. 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 2). While it allows for more meaningful dispersion comparisons across various scenarios, no other moments are affected by this adjustment.

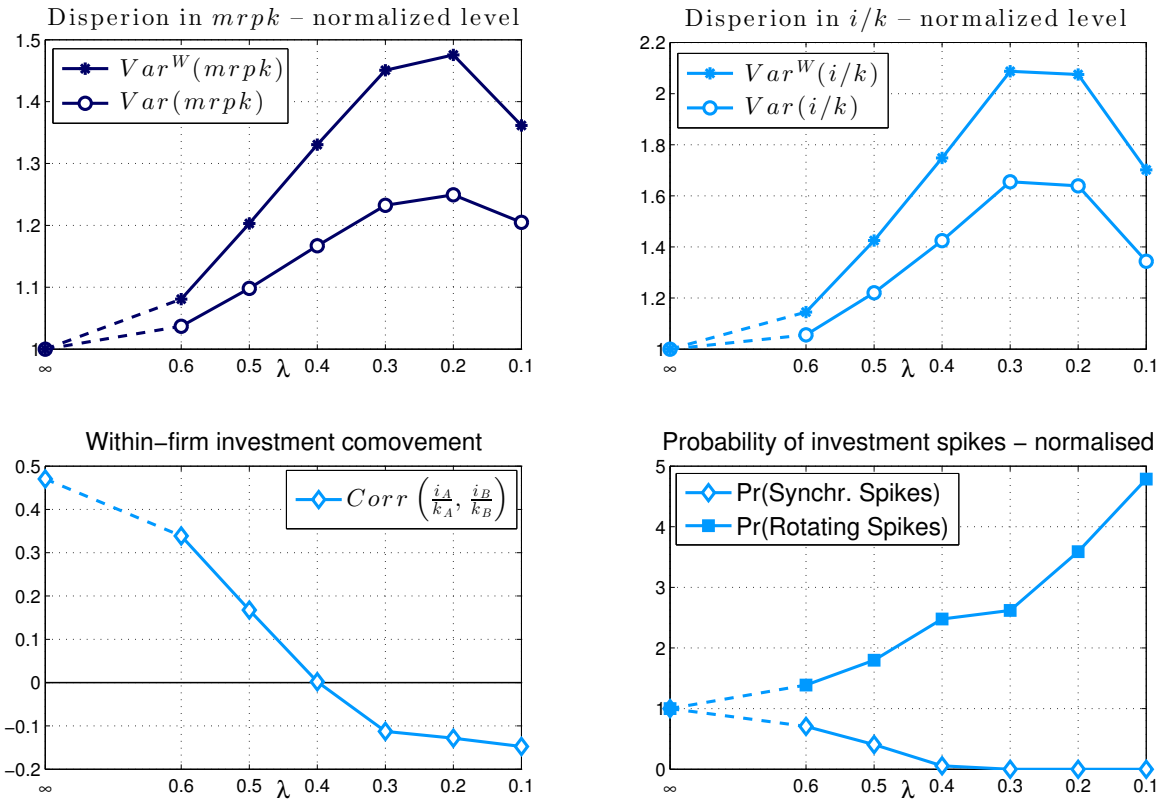
In what follows, we study how the firm's capital allocation responds to a tightening of their collateral constraint in their capital allocation. As we will see, the firm will leverage its internal capital markets, shaping the dispersion of $mrpk$ of its plants in the process. In the top panel of Figure 3, we plot the relationship between the tightness of the collateral constraint, determined by the parameter λ, and the degree of dispersion of logged expected marginal revenue products of capital $mrpk_{nt}$ and investment i_{nt}/k_{nt}, both within the typical firm and across all plants. Both moments are normalized to 1 in the scenario where the collateral constraint is not binding ($\lambda = \infty$). As the constraint is tightened (λ gets smaller), dispersion within the typical firm rises sharply, more than doubling for investment and increasing by more than 45% for marginal revenue products of capital at $\lambda = 0.3$ (the firm cannot borrow more than 30% of its collateralizable capital stock) which is the value that corresponds to the U.S. economy. Note that even if the constraint is external to the firm, it is dispersion of $mrpk$ within the firm that increases. Within-firm dispersion reaches a maximum at around $\lambda = 0.2$, before eventually falling as the constraint becomes so binding that the firm rarely invests at all. This can be also be witnessed in the rapid drop of the within-firm dispersion for $\lambda > 0.3$.[15] This rise in within-firm dispersion in turn translates in higher overall dispersion across plants in our simulated economy: it rises by up to 25% for $mrpk$ and more than 60% for i/k.

The bottom-left panel of Figure 3 hints at the mechanism behind the increase in dispersion: as the collateral constraint limits the ability of the firm to borrow, the correlation in investment activity across plants within the firm drops dramatically from its benchmark

[15] We find a very similar pattern if we instead compute the dispersion of 5-year moving averages of logged expected $MRPK$ thus being consistent with Row (5) in Table 3.

value of +0.48 in the unconstrained case. In fact, for low-enough values of λ, the cross-plant investment correlation turns *negative*, reaching a trough of −0.12 with our parameterization.

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



Note: This figure plots various dispersion and capital allocation moments within firms as a function of the external collateral constraint λ. Top panels display the resulting dispersion we are interested in explaining, bottom panels highlight the mechanism behind that: more constrained firms use the internal capital market to switch towards "rotating investment spikes" within the firm.

The second panel at the bottom illustrates the mechanics behind the drop in correlation. Here, we plot two additional moments. The first one represents the probability of observing synchronized investment spikes, i.e. investment rates above 15% in both plants at the same time. As the collateral constraint is tightened and access to internal finance is limited, the firm cannot allocate large amounts of capital to both plants at the same time, even following a positive firm-level shock that increases internal funds. Conceivably, an option for the firm would be to invest smaller amounts in both firms to restrain how much it needs to borrow on capital markets. But because of the presence of investment fixed costs, the firm would rather

instead opt for desynchronization of investment activity across its plants. That is, in the event of a large positive firm-level shock, the firm allocates capital first to the plant with the highest (expected) return on capital, then does the same for the second plant in the following period. This is why the frequency of "rotating investment spikes," i.e. an investment spike in one plant followed by a spike in the other plant next period, rises as external financing becomes more limited. This is the key mechanism driving the rise in within-firm dispersion. The likelihood of "rotating investment spikes" in the constrained economy ($\lambda = 0.3$) is 2.5 times as high as it is in the unconstrained economy ($\lambda = \infty$). Conversely, "synchronized investment spikes" virtually disappear in the constrained economy as the firm finds itself unable to gather the internal and external funds to finance investment projects in both plants at the same time.

4.3 Dispersion and aggregates: the social value of multi-plant firms

In the previous section, we studied how a multi-unit firm optimally allocates capital across its plants when it is faced with limited access to external financing in addition to the traditional frictions found in the literature. We found that a credit constrained firm tends to rotate investment activity across its plants to limit borrowing activity. As a result, dispersion of investment but also marginal revenue products of capital within the firm rises.

In what follows, we ask ourselves a different but related question: could the rise in dispersion observed within the firm be socially optimal, i.e. associated with higher aggregate output? After all, the staggering of investment projects helps the multi-plant firm overcome external financial frictions in a way that single-plant firms cannot. This is because when plants are arranged into firms, it becomes possible for them to pool financial resources, to leverage the internal capital market that the firm creates. On their own, single plants can only count on their own revenues. As a result, could it be that multi-plant firms generate both higher *mrpk* dispersion *and* higher output?

To investigate the social value of the firm in the presence of external financing contraints, we simulate two almost identical economies, but for the presence of internal capital markets. The first economy is similar to the one we have simulated so far. The second panel also consists of 1,000 two-plant firms subjected to the same shocks. This time, however, each plant within the firm is considered to be a standalone unit: it must borrow on its own by posting its plant-specific capital stock as collateral. Otherwise, it faces the exact same

constraints and frictions as the plants in multi-unit firms. In other words, we create a "wall" between the plants that bars them from pooling financial resources when access to external financial markets is too expensive. This exercise allows us to specifically isolate the role of the firm in creating internal capital markets which enables new investment projects across its plants.

We then investigate two questions. First, do two-plants firms with internal capital markets produce more dispersion in marginal revenue products than firms without any intenal ability to pool resources? Second, how much more aggregate output can the economy with internal capital markets produce? Naturally, two-plant firms with internal capital markets cannot do worse since they convexify the choice set and can always reproduce the allocation of the single-plant economy. Our objective is to determine how large these gains can be.

Table 6: Comparing multi-plant (MUF) and single-plant (SUF) firms

Moment	No frictions; $\lambda = \infty$ (MUF = SUF)	$\lambda = 0.3$ MUF	$\lambda = 0.3$ SUF
A. Plant-level moments			
Share of within-firm dispersion of $E[\log(y/k)]$	0.510	0.600	0.467
Share of within-firm dispersion of i/k	0.789	0.923	0.795
$Corr(i/k_{At}, i/k_{Bt})$	0.437	-0.113	0.317
B. Aggregate Moments			
Aggr I	0.559	0.556	0.534
Aggr K	0.835	0.831	0.797
Aggr Y	0.172	0.171	0.167
Aggr V	344.2	342.6	332.8

Table 6 and Figure 4 compare a number of moments across the two economies. To identify the role of credit frictions, we contrast allocations under $\lambda = \infty$ (collateral constraint not binding) to $\lambda = 0.3$ (collateral constraint binding as estimated for U.S. manufacturing). In the former scenario, the single-plant and multi-plant firm economies are by definition identical: When external finance is the only value of the firm and credit is abundant, then firms cannot improve upon the allocation of resources. Only when credit becomes scarce does the firm's internal capital market gain value and the firm with several plants may do better. The economy with internal capital markets may raise aggregate investment, capital or output or a combination of these. Our exercise is thus similar in spirit to Buera and Moll (2015) who also examine the incidence of a "credit crunch."

In line with our previous results, we find that the presence of financial frictions raises dispersion in a world of multi-unit firms due to the optimal decision of rotating investment spikes across plants. With standalone plants, however, dispersion of i/k is virtually unchanged once the collateral constraint binds, while the dispersion of $mrpk$ declines slightly. Also, while the correlation of investment across plants turns negative within multi-unit firms, it is only midly lower when the plants cannot pool resources and need to borrow on their own.

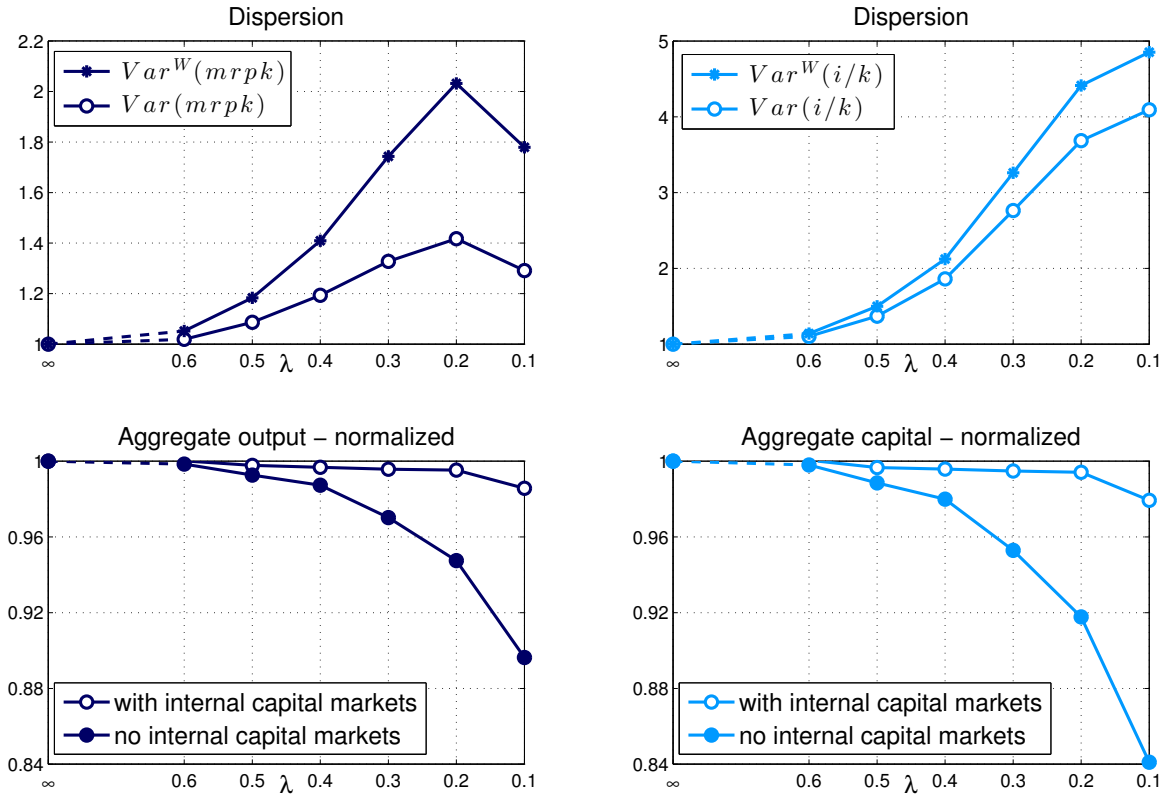
The bottom panel of the table pertains to the aggregates across our economies: we sum up output, investment and capital across all firms for each period, and take the time-series average. One can notice a striking result: the multi-plant firm is capable of undoing most if not all the effets of financial frictions for a wide range of values of λ, with aggregate variables remaining almost unchanged relative to the unconstrained benchmark. This is true even if the presence of collateral constraints *raises* dispersion of marginal revenue products of capital. This indicates that a multi-plant firm's internal capital market is quite a powerful tool to overcome frictions in external markets. On the other hand, in an economy populated by plants that cannot pool resources, tightening the collateral constraint (lowering λ) leads to sharply lower aggregate output *and* lower dispersion. Our results indicate that a higher dispersion of marginal revenue products of capital need not be associated with lower aggregate output.

We investigate this finding further by simulating the impact of a tightening of the collateral constraint in our two economies. The left panel of Figure 4 shows the relative dispersion of $mrpk$ as λ decreases, i.e. borrowing becomes more limited. In line with our previous findings, we can see that as the constraint tightens, the multi-plant firm optimally decides to generate 40% *more* dispersion in $mrpk$, even if it could technically replicate the standalone-plants case by shutting down internal capital markets. This optimal strategy at the firm level translates into 4.2% more aggregate capital and 2.4% more aggregate output relative to an identical economy in which plants cannot pool internal resources to fund targeted investment projects and access external financing.

To summarize, our model illustrates that a higher dispersion of marginal revenue products of capital can be an outcome of constrained efficient behavior by multi-plant firms. This higher dispersion is not a sign of resource misallocation in the economy, but instead of an optimal decision by the firm to stagger investment activity in order to circumvent financial constraints.

Our findings therefore cast doubt on the notion that dispersion in marginal revenue prod-

Figure 4: Quantitative effects in the multi-plant-firm vs. single-plant-firm economy



Note: Each value corresponds to the ratio of a specific moment in the multi-plant-firms economy (MUF) relative to the same moment for the economy composed of firms that cannot pool capital across plants (SUF). Ratios of average within-firm marginal revenue products of capital dispersion (left panel) and aggregate investment, capital and output (right panel) are plotted for various values of the collateral constraint parameter λ.

ucts of capital necessarily indicates inefficiencies. In our calibrated model, up to 31%=0.45/1.45 of within-firm dispersion and 19%=0.23/1.23 of overall dispersion of *mrpk* in an economy composed of two-plant firms can be traced back to the optimal response to external financial frictions.

4.4 Rethinking the gains from eliminating misallocation

Next, we investigate the potential implications of our findings for assessing aggregate losses from misallocation. In their seminal work, Hsieh and Klenow (2009) estimate that lowering distorsions in the Indian manufacturing sector to levels found in the U.S. would lead to TFP gains of 47% on average.[16] These is a sizable number, yet our findings suggest that it may in fact be a lower bound on potential gains. The reason stems from differences in the size of manufacturing firms between India and the U.S. [ACTUAL NUMBERS HERE]

We showed in Section 4.3 that moving from a single-plant-firm economy to one populated by two-plant firms leads to an increase in both dispersion *and* aggregate output. In conjunction with the fact that the single-plant firms are more prevalent in India than the U.S., this implies that the true difference in marginal revenue products dispersion between the two countries should in fact be higher than the metric Hsieh and Klenow (2009) use for their analysis. In other words, the potential gains from reallocation are even higher once differences in firm structure are taken into account.

[OPTION 1: we blow up India's dispersion] To illustrate how quantitatively important this margin could be, we make use again of the two aggregate economies we simulated in the previous section. Our question: how much more dispersion should we expect to observe across Indian manufacturing plants if they faced the same technological and financial constraints as their U.S. counterparts and were organized into similar firm structures? To answer this question, we simulate a panel of single-plant firms using the same parameter values as those found in Table 5, except for those pertaining to the shock process, which should be specific to India. Instead, we adjust the volatility parameters () in proportion to the numbers found in Table XX of Asker et al. (2014).The between-firm dispersions are once again scaled down by a factor that allows to match the within-firm share of overall dispersion for the U.S. As we mentioned earlier, this is to take into account the fact that the focus of our model is not on factors that may make firms behave in a more synchronised fashion, such as aggregate or sectoral shocks. This scaling only impacts between-firm and overall dispersion moments,

[16] 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.

not aggregates.

Table XX summarizes and compares the relevant moments for our two economies. We find that for India, making its manufacturing sector move from a single- to a multi-plant firm economy would lift overall dispersion of marginal revenue products of capital by about 50% (from 0.299 to 0.455).

[OPTION 2: we reduce US' dispersion] To illustrate how quantitatively important this margin could be, we make use again of the two aggregate economies we simulated in the previous section. Our question: what is the level of marginal revenue product of capital dispersion for the U.S. that should be compared to that of India? Given our earlier findings, we know that one needs to control for differences in firm structure before trying to assess the extent of misallocation due to frictions. As can be seen in 4, at our benchmark level of $\lambda = 0.3$, our model predicts that populating the U.S. manufacturing sector of single-plant firms instead of two-plant firms, a structure closer to that of India [references], would lower overall dispersion of $mrpk$ by close to 25% (from 1.327 to 1).

Next, we use this dispersion adjustment to compute new gains from reallocation in the spirit of Hsieh and Klenow (2009).....

[IDEALLY, we would know the proportion of single-plant and two-plant firms in both India and the U.S. Then we would simulate both types with country-specific shock processes. This would give us four types of firms. Next, we would create two panels of plants (India and US), each one matching the relative proportion of SUF and MUF in the respective countries. Finally, we would compute overall dispersion across plants in both panels, and compare them. Right now everything is really approximate, we are reducing dispersion based on going from 2-plant firms to single-plant firms, knowing full well (we show it in the text) that larger firms are where the action really is.]

that because manufacturing firms in India count on average a smaller number of plants than in the U.S.,

- Previous sections showed that measured dispersion reflects both inefficient distortions and constrained efficient behavior of multi-plant firms.
- Developing countries have almost no multi-plant firms while in the U.S. multi-plant firms account for he majority of economic activity.
- Thus, dispersion in developing economies reflects solely distortions while in the U.S. almost a quarter of dispersion can be attributed to constrained-efficient firm behavior (see quantitative analysis in Section 4.3 above).

- Fostering output in developing countries can be done in two ways:
  First, give them multi-plant firms, so these firms can reallocate internally; this would increase dispersion, output and welfare. Difference in returns dispersion now reflects true distortions only, not differences in firm complexity; this difference is larger than previously thought.
  Second, reduce inefficient distortions. Because they were previously underestimated, bringing down distortions to U.S. levels comes with larger output and welfare gains.

- Conventional welfare gains exercises only consider the second avenue and underestimated the true dimension of distortions.

- So, there are two additional output and welfare gains here:
  First, gain from giving developing economies a more complex firm structure which facilitates within-firm reallocation and increases output (previously not considered).
  Second, gain which results from bringing down distortions to U.S. levels (previously underestimated).

- We intend to quantify both gains.

4.5 Discussion

4.5.1 Multi-plant firms in general equilibrium

The previous section explained how the staggered investment policy of multi-plant firms depends on fixed investment adjustment costs. While a plethora of research has documented the lumpiness of investment at the level of the individual plant (see e.g. Doms and Dunne (1998), Cooper and Haltiwanger (2006), Gourio and Kashyap (2007)), work by Thomas (2002), Khan and Thomas (2008) has shown that general equilibrium forces make the micro-level investment spikes irrelevant at the aggregate level. Investment spikes that rotate across plants in the firm are at the core of the "beneficial dispersion of $mrpk$" in our partial equilibrium model. Hence it is a valid question to ask if general equilibrium forces would make our model mechanism irrelevant.

We argue that the organizational structure of multi-plant firms itself makes investment smoothing motives that play out in general equilibrium less pressing. How is that? The firm will allocate investment across its plants in a way to maximize the value of the firm. Doing so will involve allocating investment projects to different parts of the firm which smoothes

the return of the firm as a whole. The more plants a firm has, the more it can smooth its returns. But the firm's return would be more volatile – and thus disliked by the firm owner, the household – when the firm operates a single plant and this single plant undergoes an investment spike as in the above-mentioned papers. We hence argue that the staggered investment spike behavior of multi-plant firms in our model world would not change their investment behavior much if factor prices were endogenous in general equilibrium. This also means that a lot of the endogenous dispersion in capital revenue products and investment rates *within* multi-plant firms would not change much.

4.5.2 Multi-plant firms and investment adjustment cost

The organizational structure of multi-plant firms also has important consequences for the nature of adjustment costs. The landmark paper of Cooper and Haltiwanger (2006) has structurally estimated a variety of physical investment frictions; Cooper and Ejarque (2003) have expanded this work to include financial frictions in the investment model. Like the overwhelming majority of these papers assume that individual plants are independent decision makers. If plants belong to the same firm, however, our model shows that credit constrained firms smooth out investment over time by rotating investment spikes across their plants. Credit constraints thus filter through the orgainizational structure of the firm to smooth out investment *at the firm level*. This makes the firm look like as if it was subjected to investment flow adjustment costs of the form $\kappa(I_t/I_{t-1})^2$. Such costs have been found to fit the dynamics of investment at more aggregate levels much better than other types of adjustment costs (see Christiano et al. (2005)). It appears that our model of investment spikes within multi-plant firms are a novel micro foundation for such adjustment costs at the firm or more aggregate levels. With the exception of Lucca (2007), we are unaware of alternative micro-foundations. How structurally estimated investment adjustment costs such as flow adjustment costs or partial irrevesabilities would change once one takes into account the within-firm interdependencies of individual plants, is an interesting and important question for future research.

5 Empirical model support

The previous sections documented a set of new facts and proposed a new theoretical model whose quantitative analysis challenged the notion that high dispersion in marginal revenue products of capital necessarily reflects misallocation. In fact, we showed in the previous

section that economies with multi-plant firms may enable more aggregate investment and produce higher aggregate output while at the same time increasing dispersion in revenue products. With our calibrated model capable of explaining a considerable quantity of the empirically observed data facts, we now want to demonstrate its empirical relevance more directly. To do that, we verify the crucial features of our model mechanism in firms which operate exactly two plants as in our model. The crux of the model mechanism is a collateral constraint which limits external finance. The tighter this financing constraint, the more relevant a firm's internal capital market becomes which the firm uses to finance investment in at least a subset of its plants. This means that the more financially constrained a firm, the more dispersed are its marginal revenue products of capital as well as investment rates and the more prevalent are "rotating investment spikes" as opposed to "synchronized investment spikes." We check for these patterns in two-plant firms in the data. As an empirical proxy for the tightness of the collateral constraint, we choose the amount of the pledgeable collateral which we identify as proportional to the plant's capital stock.

5.1 Empirical evidence I: Staggering investment

We first study how a firm responds to a firm-wide technology shock, that is, all the firm's plants experience a boost in plant productivity. Controlling for a plant's capital stock such a firm-wide technology shock should lead to "synchronized investment spike" if the firm has a large debt capacity. But, when the firm is credit constrained, such a firm-wide technology shock should lead to "rotating investment spikes."

We first estimate the likelihood of synchronized investment spikes after a firm-wide technology shock using the following probit model:

$$Pr(Y_{jt}^{sync} = 1|X_{jt}) = \Phi\left(X_{jt}\beta\right) \tag{9}$$

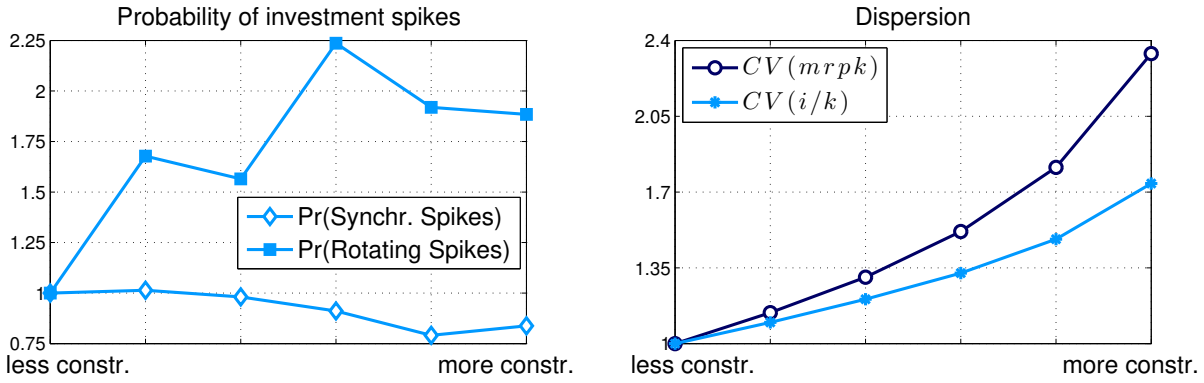
where X_{jt} is a vector of controls which includes is the firm's level of the productivity shock, η_{jt}, and the dummy variable Y_{jt}^{sync} is defined as follows

$$Y_{jt}^{sync} = \begin{cases} 1 & \text{if } \frac{i_{At}}{k_{At}} > 0.15 \text{ and } \frac{i_{Bt}}{k_{Bt}} > 0.15 \\ 0 & \text{otherwise.} \end{cases}$$

We focus on the firm's investment response to a firm shock because this provides the cleanest example of the changing investment patterns as the financial constraint tightens. We estimate

equation (9) separately for ten size deciles of two-plant firms where size is defined as the value of the pledgeable collateral. We then evaluate the marginal likelihood of a synchronized investment spike in the wake of a firm-wide technology shock, i.e. a shock experienced by both plants. To ease disclosure requirements we smooth the results of these deciles using a cross-sectional rolling-window average estiamte of the five adjacent deciles. Normalizing the probability of a "synchronized investment spike" to unity for the least financially constrained firms, we plot the normalized probabilities in the left panel of Figure 5 (line with hollow diamonds). It shows that the financially most constrained firms are only 80% as likely as the least constrained firms to respond to a firm-wide technology shock with an investment spike in both plants.

Figure 5: Dispersion and the nature of investment spikes after firm TFP shocks



Note: Left panel: normalized probabilities of "synchronized investment spikes" and "rotating investment spikes" after firm technology shocks by collateral size. Right panel: Dispersion of marginal revenue products of capital and investment rates by collateral size.

In a similar vein, we estimate probit models where we regress a dummy variable if the firm experiences a "rotating investment spike," that is, both plants experience an investment spike in the wake of a firm productivity shock, but just in subsequent periods. We estimate

$$Pr(Y_{jt}^{rotate} = 1|X_{jt}) = \Phi\left(X_{jt}\beta\right) \tag{10}$$

where X_{jt} is a vector of controls which includes is the firm's level of the productivity shock,

η_{jt}, and the dummy variable Y_{jt}^{rotate} is defined as follows

$$Y_{jt}^{rotate} = \begin{cases} 1 & \text{if } \frac{i_{At}}{k_{At}} > 0.15 \text{ and } \frac{i_{Bt}}{k_{Bt}} < 0.15 \text{ and } \frac{i_{At+1}}{k_{At+1}} < 0.15 \text{ and } \frac{i_{Bt+1}}{k_{Bt+1}} > 0.15 \\ 1 & \text{if } \frac{i_{At}}{k_{At}} < 0.15 \text{ and } \frac{i_{Bt}}{k_{Bt}} > 0.15 \text{ and } \frac{i_{At+1}}{k_{At+1}} > 0.15 \text{ and } \frac{i_{Bt+1}}{k_{Bt+1}} < 0.15 \\ 0 & \text{otherwise.} \end{cases}$$

Again, we evaluate the marginal probabilities, smooth them across deciles and normalize the probability in the least constrained group to unity. The results are plotted in the left panel of Figure 5 (line with solid squares). In line with the model, the financially most constrained firms are more likely to respond to firm shocks with a rotating investment spike. That is, the firms contemporaneously invests in only one plant *and* invests in the other plant in the subsequent period. The empirical difference of the likelihood of a rotating investment spike increases a lot with our proxy measure of financial constraint: The most constrained firms are twice as likely to experience a rotating investment spike as financially unconstrained ones. All in all, we view this evidence as strong support for the presence of the key model mechanism.

5.2 Empirical evidence II: Within-firm dispersion

Lastly, we consider the within-firm dispersion in marginal revenue products of capital and capital reallocation. According to our model, firms which are more credit constrained should experience more dispersion while firms with a lot of pledgeable collateral should have an easier time managing to equate returns. We compute the two dispersion measures by decile of pledgeable collateral and divide by the mean level in each decile to account for level differences. Unlike the standard deviation, the resulting coefficient of variation is dimensionless and can be easily compared across deciles. Again, we smooth out the results across deciles as in the previous subsection and plot the results in the right panel of Figure 5.

Indeed, the within-firm dispersion measures show that investment dispersion monotonically increases in our proxy for credit constraints: the most constrained firms have investment dispersion within firms that is about 1.7 times as large as that for the least constrained firms. The difference in marginal revenue products of capital is even stronger: the most constrained firms are about 2.5 times as dispersed as the least constrained firms.

6 Conclusion

This paper showed that dispersion in marginal revenue products of capital need not indicate distortions. Motivated by the evidence that dispersion mostly occurred within firms rather than across firms, we built a model of a firm operating several plants. Such firms dispose over an internal capital market that helps easing external financial constraints and supports aggregate investment, capital and 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.

References

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

John Asker, Allan Collard-Wexler, and Jan De Loecker. Dynamic inputs and resource (mis)allocation. *Journal of Political Economy*, 122(5), October 2014.

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

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.

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

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

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

Nicholas Bloom, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry. Really uncertain business cycles. *NBER Working Paper No. 18245*, July 2012.

Francisco J. Buera and Benjamin Moll. Aggregate implications of a credit crunch: The importance of heterogeneity. *American Economic Journal: Macroeconomics*, 7(3):1–42, July 2015.

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.

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.

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.

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.

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.

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

Russell Cooper and João Ejarque. Financial frictions and investment: Requiem in Q. *Review of Economic Dynamics*, 6(4):710–728, October 2003.

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

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

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.

Francesco D'Acunto, Ryan Liu, Carolin Pflueger, and Michael Weber. Flexible prices and leverage. *Working Paper*, 2016.

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

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

Andrea L. Eisfeldt and Tyler Muir. Aggregate issuance and savings waves. *Journal of Monetary Economics*, forthcoming.

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

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

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.

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.

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

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

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.

Xavier Giroud. Proximity and investment: Evidence from plant-level data. *Quarterly Journal of Economics*, 128(2):861–915, May 2013.

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

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

Gita Gopinath, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. Capital allocation and productivity in South Europe. *Working Paper*, 2015.

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

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4), November 2009.

Matthias Kehrig. The cyclical nature of the productivity distribution. *Working Paper*, 2015.

Truman L. Kelley. *Fundamentals of Statistics*. Harvard University Press, Cambridge, Massachusetts, 1947.

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

Aubhik Khan and Julia K. Thomas. Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436, March 2008.

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.

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

Shaojin Li, Toni M. Whited, and Yufeng Wu. Collateral, taxes, and leverage. *Review of Financial Studies*, 29(6):1453–1500, June 2016.

David O. Lucca. Resuscitating time-to-build. *Working Paper*, 2007.

Andrey Malenko. Optimal design of internal capital markets. *Working Paper*, February 2016.

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.

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

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

Ezra Oberfield and Devesh Raval. Micro data and macro technology. *NBER Working Paper No. 20452*, 2014.

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

Adriano A. Rampini and S. Viswanathan. Collateral and capital structure. *Journal of Financial Economics*, 109(2):466–492, August 2013.

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

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.

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

Jae Song, David J. Price, Fatih Guvenen, and Nicholas Bloom. Firming up inequality. *Working Paper*, 2015.

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

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

Robert Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265–293, October 1979.

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

Appendix

A Additional empirical evidence

A.1 Investment and equalizing $MRPK$

The assumption that investment should ideally be undertaken to equalize marginal revenue products of capital rests on some assumptions which 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 make the realized and the expected marginal revenue product of capital different. All 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: $Var_t\left(mrpk_{nt}\right) = Var_t\left(\log(y_{nt}/k_{nt})\right)$.

A.2 Data

We mainly use confidential data on manufacturing establishments collected by the U.S. Census Bureau which 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, Census defines an "establishment" as a business location where the principal activity is production; we hence think of an "establishment" as a production plant. The Census data also contain 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 Census of Manufactures (CMF) and the Annual Survey of Manufactures (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 dynamic implications of our model of investment in multi-plant firms the highest possible frequency. Many aspects of out 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 55k plants which total up 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 the manufacturing data are merely 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[17], 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 BLS's price deflators for capital

[17]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.

goods at the 3-digit NAICS industry level.[18] 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 the 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`[19] 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 industry operations as separate firms. Our within-firm dispersion measures are hence an understatement because we ignore the between-industry component if within-firm dispersion.

A.3 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 the 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 along-run industry-specific moments, we first aggregate over years in order to take out any industry-specific trends. As Kehrig (2015), Gopinath et al. (2015) we note an upward trend in dispersion and – to a lesser extent – in skewness and a downward trend in kurtosis. Behold that the cross-plant standard deviation increases about 10 log points per decade; both between-firm and within-firm dispersion increase evenly, so that 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 a bottom tail) in 2007. We compute the typical cross-sectional moment in a given NAICS-4 industry in 1990 which 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.

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

[19] Song et al. (2015) 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.

A.4 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 which underly Table 2, Panel A. and the robustness exercises in Section 2.3.

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.

$$
\begin{aligned}
V_t &= \sum_n \omega_{njit} \left(mrpk_{njit} - mrpk_t\right)^2 \\
&= \underbrace{\sum_i \omega_{it} \left(mrpk_{it} - mrpk_t\right)^2}_{V_t^{Ind} \text{ between-industry}} + \underbrace{\sum_i \omega_{it} \underbrace{\sum_{j \in i} \omega_{jt}^i \left(mrpk_{jit} - mrpk_{it}\right)^2}_{V_{it}^B \text{ between firms within ind. } i}}_{V_t^B \text{ average between-firm}} + \underbrace{\sum_i \omega_{it} \sum_{j \in i} \omega_{jt}^i \underbrace{\sum_{n \in j,i}^{N_j} \omega_{nt}^{ji} \left(mrpk_{njit} - mrpk_{jit}\right)^2}_{V_{jit}^W \text{ within firm } j \text{ and industry } i}}_{V_t^W \text{ average within-firm}}
\end{aligned} \tag{11}
$$

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 which we ignore for its lack of economic meaning. In our empirical analysis in Section 2, we focus on V_i^B and V_i^W of plants with at least two plants only as 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 if we are looking at unweighted or capital weighted dispersion.

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

A.5 Robustness

A.5.1 Accounting for measurement error

In Section 2.3 we already dealt with some measurement error. If plant-level variables are measured with noise, then the firm-level averages will be measured more precisely thus artificially inflating the within-firm variance. Time aggregation should filter out that 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 which 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 dominance of within-firm share is true, then this should show up in all of these measures. Song et al. (2015) follow a similar procedure.

Table 7: 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: 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 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% 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 we are comparing the moments for the same underlying sample where we have both our benchmark measure as well as the alternative. It turns out, 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 the benchmark differs 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.5.2 Marginal versus average revenue products

Our empirical work of Section 2 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; like 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 where 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) documented that overheads in production are quite powerful in explaining differences between micro production units and they showed significant aggregate consequences. Most overhead inputs are likely at the headquarter level of a firm rather than the plant level. Though this cast some doubt that all our results could be driven by constant inputs, we don't want to dismiss this possibility.

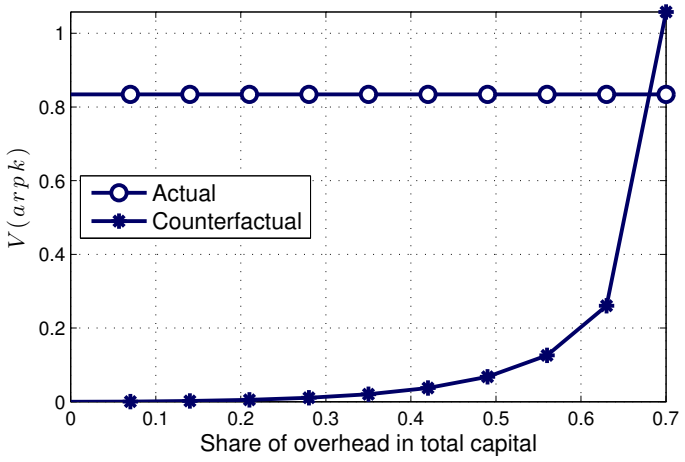
Overheads Any constant input requirements at the plant level are hard to identify empirically. We therefore carry out a quantitative exercise to examine how large overheads would have ot be in order to explain all or a portion of the empirically observed dispersion. Suppose the true technology is $y_n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

$$V(arpk_n) = V\left(\log\left(1 - \frac{\overline{k}}{k_n}\right)\right). \tag{12}$$

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 (12) is for different levels of $\overline{k}$.

Figure 6 plots the RHS of equation (12) as a function of $\mathbb{E}\left[\overline{k}/k_n\right]$. Naturally, when $\overline{k} = 0$, then this variance will be zero and the observed variance of average revenue products must

Figure 6: How much overheads are necessary to explain the observed $V(arpk)$?



Note: Simulation of the right hand side of equation (12) against the share of overhead ion total inputs.

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 overheads, less than a tenth of the empirically observed variance in average revenue products can be explained by overheads. Clearly, this much 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 overheads. We conclude that overheads 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:

$$V(arpk_n) = \sigma^2 V(mrpk_n). \tag{13}$$

The true dispersion of marginal revenue products could then be much lower is the elasticity of substitution is larger than unity. However, Oberfield and Raval (2014) that estimates that elasticity of substitution at the plant level, puts 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.

B Numerical solution of model

To solve the model, we discretize the plant-level capital stock using a N_k-point grid, where $N_k = 100$ to produce the results in this paper. This implies that the two-plant capital grid is comprised of 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.[20] 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 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.

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