# Competition, Financial Constraints, and Misallocation: Plant-Level Evidence from Indian Manufacturing

Simon Galle *

BI Norwegian Business School

March 2019

**Abstract**

This paper demonstrates a dual impact of increased competition on misallocation in a setting with both oligopolistic competition and financial constraints. I develop a novel model where, in case there are no financial constraints, more competition unambiguously increases aggregate output by reducing markup levels and markup dispersion. However, with financial constraints, increased competition reduces the profitability of constrained firms, and thereby slows down their rate of self-financed investment and convergence to their optimal capital levels. To test the model's predictions, I leverage the pro-competitive impact of India's dereservation reform on incumbent plants exposed to the reform. In line with the theory, this reform leads to reduced markup levels and markup dispersion, and to slower capital convergence. To examine the external validity of the result on capital convergence beyond the sample of incumbent plants, I present further evidence for this prediction on the full panel of plants. My results help understand why misallocation in India is strongly persistent, despite multiple liberalization reforms.

*simon.galle@bi.no For invaluable advice, I am grateful to Ben Faber, Yuriy Gorodnichenko, Edward Miguel, and especially Andrés Rodríguez-Clare, and I am deeply thankful to Ishani Tewari and Hunt Allcott for sharing data. I also thank Juan-Pablo Atal, Pierre Bachas, Dorian Carloni, Fenella Carpena, Cécile Gaubert, Daniel Haanwinckel, Ben Handel, Alfonso Irarrázabal, Jeremy Magruder, Yusuf Mercan, Plamen Nenov, Louis Raes, Marco Schwarz, Yury Yatsynovich and Moises Yi for many helpful discussions and suggestions; the International Growth Centre and the PEDL Initiative by CEPR and DFID for financial support to purchase the ASI data; and seminar participants at BI, Bocconi, CREI, ECARES, ECORES Summer School, Edinburgh, Erasmus Rotterdam, Gothenburg, HEC Lausanne, HEC Paris, IHEID Geneva, Johns Hopkins SAIS, KULeuven, NHH Bergen, PEDL-CEPR, Surrey, Tilburg, UC Berkeley, UCLA Anderson and University of Oslo for excellent comment. All errors are my own.

# 1 Introduction

Aggregate productivity is central to understanding why some countries are rich while others are poor. Since plant-level marginal productivities tend to be much more misaligned in poorer countries, resource misallocation has become a prominent candidate for explaining differences in countries' aggregate productivity.[1] While the potential factors contributing to misallocation are varied, the predominant view in the literature is that competition would be a beneficial force in reducing misallocation. After all, it is highly intuitive that competition will help shift resources from low-performing to high-performing plants, for instance by reducing markup dispersion (Peters, 2016; Asturias, García-Santana, and Ramos, 2018), or by enhancing selection of high-productivity firms.

While the mechanisms driving competition's beneficial impact on aggregate productivity are undeniable, these beneficial mechanisms do not seem to cover the full story. Since limited access to finance is pervasive in developing countries (Levine, 2005), financially constrained firms often need to rely on retained earnings to finance their investments. Hence, profit levels determine how fast firms are able to save themselves out of their financially constrained position. Since competition reduces firms' profitability, it then also slows down investment for these firms. This, in turn, has negative implications for aggregate output.

This downside of competition may be especially salient in India, a large economy with strongly persistent levels of misallocation (Hsieh and Klenow, 2009). Strikingly, most of that country's liberalization reforms, including an extensive licensing reform and a trade liberalization, had a null-effect on the degree of allocative efficiency in its manufacturing sector (Bollard, Klenow, and Sharma, 2013). From the predominant perspective in the misallocation literature, this finding is puzzling. However, since even large Indian firms tend to be credit constrained (Banerjee and Duflo, 2014), it is important to take the interplay between competition and financial constraints into account in our understanding of the impact of liberalization on misallocation.

I develop a novel model to formally examine this interplay of competition and financial constraints. To allow for variation in competition, the market structure is oligopolistically competitive, which generates "markup misallocation" as in Atkeson and Burstein (2008). In the absence of financial constraints, intensified competition reduces markup misallocation by lowering markups toward their lower bound and thereby depressing markup dispersion. While this beneficial impact of competition on markup misallocation remains central in my framework, the introduction of financial constraints crucially leads to a second, harmful impact of competition on misallocation.

These financial constraints become binding because firms experience random shocks to their idiosyncratic productivity, and after a positive productivity shock, they optimally choose to grow their capital stock. Critically though, their limited access to finance hampers their ability to do so, which leads to "capital misallocation" as in Midrigan and Xu (2014). Since financially constrained firms then rely on retained earnings to finance their investment, their rate of self-financed capital growth becomes a function of their optimal markup. Increased competition, by reducing firms' markups, negatively affects their speed of capital convergence in response to a positive produc-

[1] The misallocation literature started with the seminal contribution by Restuccia and Rogerson (2008). For the case of India, which will be the country of interest for this paper, Bils, Klenow, and Ruane (2017) argue that misallocation of resources, after controlling for potential measurement error, accounts for 30 to 40% of the difference in aggregate manufacturing output per capita between the United States and India.

tivity shock. This way, competition amplifies the difference between a constrained firm's optimal and actual level of capital, i.e. the "capital wedge," and worsens capital misallocation. Interestingly, I derive these analytical results on the dual impact of competition - it reduces markup misallocation but amplifies capital misallocation - in a setting where there is no closed-form solution for the distribution of markups and capital.

In the second part of the paper, I test and confirm the predictions of the model in the context of the Indian manufacturing sector. To this end, I use a natural experiment arising from the staggered implementation of an industrial policy reform: the dereservation episode. Starting in 1997, this reform removed the investment ceilings imposed for the production of certain product categories, which led to the entry of new, larger firms in the production of the now dereserved product categories. Hence, the reform exposed incumbent plants to stiffer competition. Empirically, I examine the impact of the reform on incumbents' markups and their capital convergence. First I demonstrate in an event study that the dereservation reform led to lower markups for incumbent plants, which confirms the pro-competitive impact of the reform. Moreover, markups for plants with an initially higher markup fell more than for plants with a lower initial markup. Hence, the reform reduced markup dispersion, in line with the theory's predictions.

Next, I turn to testing the novel prediction of my model, namely the negative impact of competition on capital convergence. Since a plant's optimal level of capital is unobserved, I focus on convergence in marginal revenue product of capital (MRPK), inspired by Asker, Collard-Wexler, and De Loecker (2014). I proxy for a plant's optimal MRPK with a flexible function, including a plant fixed-effect to allow for maximal cross-plant heterogeneity, and find that plants exhibit strong and robust convergence to this measure of optimal MRPK. This enables me to use the speed at which a plant converges back to its optimal MRPK as an empirical counterpart for the model's speed of convergence to optimal capital levels. I then find that MRPK convergence is indeed slower after a plant's products have been dereserved.

To strengthen the external validity of the empirical analysis, I also study MRPK convergence on the full panel of plants, since I can only examine the pro-competitive impact of the dereservation reform on the subset of incumbent plants described above. For the full panel, the measure of competition is the median markup across plants observed in the same state, sector and year. This median value is plausibly exogenous from the perspective of the individual plant. I document that a higher median markup, indicating less competition, is associated with faster plant-level MRPK convergence. To further corroborate the theoretical mechanism, I explore how the impact of competition varies by a plant's degree of financial dependence. Using the standard Rajan and Zingales (1998) measures, I find that plants in sectors with higher degrees of financial dependence exhibit a stronger sensitivity to the degree of competition.

The measurement in this first battery of empirical tests relies on standard Cobb-Douglas assumptions for the MRPK values, and on an autoregression framework for the speed of convergence. The advantage of this empirical approach is that it links closely to the model with productivity volatility, but a disadvantage is that this framework may be less transparent than a more reduced-form approach. For this reason, I implement all the above tests also on capital growth for young plants, the benefit being that capital growth is a reduced-form object in the data. This empirical strategy is motivated by a version of the model where newborn firms are undercapitalized, and therefore competition will slow down their capital growth. Here, the assumption that newborn firms are undercapitalized is in line with the stylized facts in the data. For all three tests, namely the impact of dereservation on incumbent plants, the impact of the median markup on the full panel of plants, and for the heterogeneity along financial dependence, the evidence once more lines up with the model's predictions.

A closely related paper to mine is Itskhoki and Moll (2019), which also analyzes capital misallocation and examines how policy can optimally influence firm profitability. However, their main focus is on tax policy, in a setting with perfect competition. In contrast, the key contribution of this paper is to examine the impact of competition on capital misallocation in an oligopolistic setting. Interestingly, this oligopolistic setting implies that the closed-form results from Moll (2014) no longer apply here. Where a standard approach would then rely on simulations, I am still able to derive analytical results on the interplay between the distribution of capital and the distribution of markups. More generally, my paper also relates to the macro-development literature on financial frictions, surveyed by Buera, Kaboski, and Shin (2015).

My choice to model oligopolistic competition as in Atkeson and Burstein (2008) is an increasingly common theoretical strategy to examine variations in competition at the macroeconomic level, see for instance Edmond, Midrigan, and Xu (2015); Brooks, Kaboski, and Li (2016) and Hottman, Redding, and Weinstein (2016).[2] Importantly, none of these papers feature financial frictions.

Empirically, this paper focuses on testing the novel prediction of competition's negative impact on capital convergence, and I document robust support for this prediction across a series of plant-level tests. As indicated earlier, this evidence can help us understand why misallocation has been persistent in India, despite several liberalization reforms. From that perspective, the paper is complementary to the existing studies on allocative efficiency in Indian manufacturing, going from the analysis of markup misallocation (De Loecker, Goldberg, Khandelwal, and Pavcnik, 2016; Asturias et al., 2018), over the role of financial constraints (Banerjee, Cole, and Duflo, 2005; Banerjee and Duflo, 2014), to the impact of structural reforms (Aghion, Burgess, Redding, and Zilibotti, 2008; Sivadasan, 2009; Chari, 2011; Bollard et al., 2013; Alfaro and Chari, 2014), and the role of formal and informal institutions (Akcigit, Alp, and Peters, 2016; Boehm and Oberfield, 2018). Here, my paper is most closely related to the studies of the dereservation reform (García-Santana and Pijoan-Mas, 2014; Martin, Nataraj, and Harrison, 2017; Tewari and Wilde, 2017; Balasundharam, 2018). These studies document the beneficial impact of this dereservation reform, while my analysis leverages the pro-competitive impact of the reform to test my model's predictions.

Taken together, the contribution of this paper is to develop a more nuanced understanding of the positive, as well as the underexamined negative impact of competition on misallocation. Outside the misallocation literature, other areas of economics already have a more nuanced understanding of the ambiguous welfare impact of competition.[3] For instance, shielding an infant industry from competition may be beneficial if that industry has a latent comparative advantage. Importantly though, my model is more widely applicable than the infant industry argument, since financial constraints and market power are a general and robust feature of the data (Levine,

[2] Peters (2016), Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2017), Mrázová, Neary, and Parenti (2017) and Dhingra and Morrow (2019) employ alternative frameworks for analyzing the markup distribution.

[3] In industrial organization, it is well-established that increasing competition can have both positive and negative effects on aggregate output or welfare. Negative effects can arise from business stealing (Mankiw and Whinston, 1986; Dhingra and Morrow, 2019), or by decreasing incentives to innovate (see e.g. Gilbert (2006); Aghion, Akcigit, and Howitt (2014)). In international trade, Foellmi and Oechslin (2016) show that increased competition due to trade can hamper credit access and thereby firm productivity, while Epifani and Gancia (2011) show that it can amplify cross-sectoral markup misallocation. In a more recent contribution, Jungherr and Strauss (2017) argue that higher market power is associated with higher growth in the Korean manufacturing sector.

2005; De Loecker and Eeckhout, 2018), whereas the evidence on industries having a latent comparative advantage is mixed at best (Harrison and Rodríguez-Clare, 2010).

The next section presents the theory. Then, Section 3 discusses the data and Section 4 examines the impact of the industrial policy reform. Section 5 analyzes the impact of competition for the full panel, Section 6 studies capital convergence for young plants and Section 7 concludes.

# 2 Theory

## 2.1 Setup of the economy

Following Hottman et al. (2016), I assume that the economy has a continuum of sectors and within each sector there is a finite number of firms that produce differentiated goods. The final good $Q_t^F$ is produced in a competitive market according to the following Cobb-Douglas production function:

$$\ln Q_t^F = \int_{s \in S} \phi_s \ln Q_{st} ds, \qquad \text{with} \int_{s \in S} \phi_s ds = 1, \tag{1}$$

where $S$ is the set of sectors and $Q_{st}$ is a sector-level composite good for sector $s$ in period $t$. Time is discrete. The standard price index $P_t^F$ for the final good is $\ln P_t^F = \int_{s \in S} \phi_s \ln(P_{st}/\phi_s) ds$, where $P_{st}$ is the price index for sector $s$. A direct implication of this setup is that the optimal expenditure shares on goods for sector $s$ are constant at $\phi_s = \frac{P_{st} Q_{st}}{P_t^F Q_t^F}$.

The sector-level composite good for each sector, $Q_{st}$, is given by

$$Q_{st} = M_s^{\frac{1}{1-\sigma}} \left[ \sum_{i=1}^{M_s} q_{ist}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where $q_{ist}$ is consumption of the variety from firm $i$ in sector $s$ at time $t$, $\sigma$ is the elasticity of substitution, with $\sigma > 1$, and $M_s$ is the number of firms in sector $s$. The fact that a sector's CES aggregate consists of a finite number of firms, as in Atkeson and Burstein (2008), implies that the intensity of competition is a function of that number of firms. The term $M_s^{\frac{1}{1-\sigma}}$ eliminates love of variety, as in Blanchard and Kiyotaki (1987). In the analysis below, this elimination of love of variety allows me to isolate the pro-competitive effects of changes in $M_s$. The inverse demand function and associated revenue function $v_{ist}$ for variety $i$ are then given by:

$$p_{ist}(q_{ist}, P_{st}) = q_{ist}^{-1/\sigma} P_{st}^{\frac{\sigma-1}{\sigma}} \left( \frac{\phi_s P_t^F Q_t^F}{M_s} \right)^{1/\sigma},$$

$$v_{ist}(q_{ist}, P_{st}) = (q_{ist} P_{st})^{\frac{\sigma-1}{\sigma}} \left( \frac{\phi_s P_t^F Q_t^F}{M_s} \right)^{1/\sigma}, \tag{3}$$

where the sectoral price index is

$$P_{st} = M_s^{\frac{1}{\sigma-1}} \left( \sum_{i=1}^{M_s} p_{ist}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{4}$$

In the oligopolistic setting under consideration, a firm's demand will become more inelastic as its market share $m_{ist}$ increases:[4]

$$\varepsilon_{ist}(m_{ist}) \equiv -\frac{\partial q_{ist}}{\partial p_{ist}}\frac{p_{ist}}{q_{ist}} \quad \text{with} \quad \frac{\partial \varepsilon_{ist}(m_{ist})}{\partial m_{ist}} < 0, \quad \text{and} \quad 1 \leq \varepsilon_{ist}(m_{ist}) < \sigma, \tag{5}$$

where the market share is defined as

$$m_{ist} \equiv \frac{v_{ist}}{\sum_{j=1}^{M_s} v_{jst}} = \frac{q_{ist}^{\frac{\sigma-1}{\sigma}}}{\sum_{j=1}^{M_s} q_{jst}^{\frac{\sigma-1}{\sigma}}}. \tag{6}$$

Below we will see how variation in the demand elasticity, given productivity differences between firms, leads to variation in markups across firms, and thereby to "markup misallocation." I derive this result for the benchmark case of Cournot competition, but the analytical results hold for Bertrand competition as well.

The economy has two types of infinitely lived agents: workers and firm-owners. A measure $L$ of workers supplies labor inelastically, and each worker is hired at a wage $w_t$. Next, in each sector there is an exogenous number $M_s$ of firm-owners. Both workers and firm-owners consume the final good, and have the following intertemporal preferences over their consumption:

$$U_{jt} = \sum_{t=r}^{\infty} \beta^{t-r} c_{jt}, \tag{7}$$

where $\beta$ is the discount factor and $j$ denotes either a specific worker $l$, or a certain firm-owner $i$ in sector $s$. As I explain below, the assumption of linear preferences will ensure that when firm owners are financially constrained, they optimally choose to set their consumption to zero and use all retained earnings for capital growth. This corner solution for consumption allows me to handle the comparative statics on the distribution of firms' capital growth analytically.

Each firm produces $y_{ist}$, the output for its variety, using capital $k_{ist}$ and labor $l_{ist}$ according to a Cobb-Douglas production function:[5]

$$y_{ist}(l_{ist}, k_{ist}) = z_{ist} k_{ist}^{\alpha} l_{ist}^{1-\alpha}, \tag{8}$$

where each firm's productivity $z_{ist}$ follows a Markov process over the state space $\{z_{sL}, z_{sH}\}$, with $z_{sL} < z_{sH}$, and switching probabilities between the two states are strictly positive.[6] The assumptions on this Markov process are such that, even though there are a finite number of firms in each sector and therefore the law of large numbers does not hold, the number of firms of a given "type" is constant over time (see Appendix B). Here, the definition of a firm's type will become clear after Lemma 2. Together with the financial frictions that I introduce below, the volatility in productivity can lead to some firms being financially constrained, even in steady state, as in Midrigan and Xu (2014) or Moll (2014).

[4]While the precise value of the demand elasticity will depend on the details of the oligopolistic market structure, the qualitative relation between market share and demand elasticity in equation (5) holds under both Bertrand and Cournot competition. In case firms engage in Cournot competition, their demand elasticity is $\varepsilon_{ist}(q_{ist}) = \left[\frac{1}{\sigma}(1-m_{ist}) + m_{ist}\right]^{-1}$, and in case of Bertrand competition, it is $\varepsilon_{ist}(q_{ist}) = \sigma(1-m_{ist}) + m_{ist}$ (see Atkeson and Burstein (2008), and Amiti, Itskhoki, and Konings (2016) for derivations.)

[5]In this production function, $\alpha$ is allowed to be sector specific, but I drop the subscript $s$ to ease the notational burden.

[6]Increasing the dimensionality of the state space would add complexity to the analytical solution of the comparative statics, without yielding additional economic insight.

A firm accumulates capital according to the standard equation of motion:

$$k_{ist+1} = x_{ist} + (1-\delta)k_{ist},$$

where $\delta$ is the depreciation rate and where investment $x_{ist}$ is financed at the end of period $t$. Investment can be funded using retained earnings or debt, since firms can borrow from workers using a one-period risk-free security at an interest rate $r_t^d$. I assume that a firm's debt $d_{ist}$ is constrained to be weakly positive ($d_{ist} \geq 0$). If instead the firm-owner would be able to save, given her linear preferences, the firm-owner could choose to accumulate sufficient savings such as never to be financially constrained in steady state.[7] On the debt market, firms borrow from workers, who thereby accumulate assets $b_{lt}$. The equilibrium in the debt market holds when

$$\int_{l\in L} b_{lt} dl = \int_{s\in S} \left( \sum_{i=1}^{M_s} d_{ist} \right) ds.$$

Importantly, the firm's borrowing is subject to a collateral constraint as in Moll (2014), which puts a limit on the firm's leverage ratio:

$$\frac{d_{ist}}{k_{ist}} \leq \lambda, \qquad \lambda \geq 0. \tag{9}$$

In contrast to investment, payments to labor $l_{ist}$ are only made after revenue in period $t$ is realized. Before the end of period $t$, i.e. after revenue is realized, debt is repaid and capital has depreciated, but before decisions about borrowing, consumption, investment and labor hiring are made, a firm owner's net real wealth is then

$$a_{ist}(l_{ist}, k_{ist}, \mathbf{y_{-ist}}) \equiv \frac{\pi_{ist}(l_{ist}, k_{ist}, \mathbf{y_{-ist}})}{P_t^F} + (1-\delta)k_{ist} - (1+r_t^d)d_{ist},$$

where $\pi_{ist}(l_{ist}, k_{ist}, \mathbf{y_{-ist}}) \equiv v_{ist}(y_{ist}, P_{st}(y_{ist}, \mathbf{y_{-ist}})) - w_t l_{ist}$, i.e. revenue net of payments to labor, $y_{ist}$ is a function of $l_{ist}$ and $k_{ist}$, and $\mathbf{y_{-ist}}$ is the vector of output choices of firm $i$'s competitors. Therefore, the owner faces the following period-by-period budget constraint:[8]

$$k_{ist+1} + c_{ist} \leq a_{ist}(l_{ist}, k_{ist}, \mathbf{y_{-ist}}) + d_{ist+1}, \tag{10}$$

Since there is a non-negativity constraint on firm-owner consumption ($c_{ist} \geq 0$), given the collateral constraint, next period's capital level is bounded above as follows:

$$k_{ist+1} \leq \frac{a_{ist}(l_{ist}, k_{ist}, \mathbf{y_{-ist}})}{1-\lambda}.$$

Hence, for $\lambda \geq 1$, firms are not constrained in terms of their capital choice.

In terms of timing of the decision process, firms' productivities for $t+1$ are revealed at the end of period $t$, and at that point firms make decisions, subject to their budget constraint, about consumption in period $t$ and labor, capital and debt for period $t+1$. Given this set-up, the indi-

[7] An alternative assumption would be to allow firms to save, but introduce a finite instead of an infinite intertemporal elasticity of substitution for firm-owners' consumption. In fact, these are the assumptions in Midrigan and Xu (2014). I strongly suspect that my results would continue to hold in that setup, but the methodology would have to rely on simulations instead of the derivation of analytical results.

[8] In the subsequent analysis, I only consider cases where any firm owner's initial wealth is strictly positive, i.e. $a_{is0} > 0$, such that the optimization implies that any firm's wealth is positive in any period. In combination with the Inada conditions implied by the revenue function, this ensures that firms always set strictly positive levels of capital and labor.

vidual firm's relevant state variables are future productivity $z_{ist+1}$ and wealth $a_{ist}$. The state of a firm's competitors can be summarized by $D^s_{-i}(z_{jst+1}, a_{jst})$, the joint distribution of productivity and wealth for all firms in sector $s$ excluding firm $i$. In terms of notation, $D^s(z_{ist+1}, a_{ist})$ denotes the joint distribution of productivity and wealth for *all* firms in industry $s$.

**Workers' problem** Workers optimize their intertemporal utility function from equation (7) subject to their period by period budget constraint

$$c_{lt} + b_{lt+1} \leq \frac{w_t}{P_t^F} + (1 + r_t^d) b_{lt}. \tag{11}$$

Their linear utility implies the following optimal choices for saving and consumption

$$\begin{aligned} \left(r_{t+1}^d > \frac{1}{\beta} - 1\right) &\implies \left(b_{lt+1}^* > 0, c_{lt+1}^* = 0\right) \\ \left(r_{t+1}^d < \frac{1}{\beta} - 1\right) &\implies \left(b_{lt+1}^* = 0, c_{lt+1}^* > 0\right) \\ \left(r_{t+1}^d = \frac{1}{\beta} - 1\right) &\implies \left(b_{lt+1}^* \geq 0, c_{lt+1}^* \geq 0\right) \end{aligned} \tag{12}$$

**Market structure and the firm's problem** The firms play an infinitely repeated Cournot-type game where they decide each period about investment and labor hiring. In this setting, a strategy $\psi_{ist}$ of a firm consists of a set of decision rules for capital, labor and debt, valid for all current and future periods, that are conditional on the firm's own state, the state of its competitors, the state of the macroeconomy, and the history of the game $h_{st}$. The state of the macroeconomy is summarized in $\mathbf{F_t} \equiv \{w_t, P_t^F, Q_t^F, r_t^d\}$. For notational convenience, I write these decision rules as a function of $D^s(z_{ist+1}, a_{ist})$, which includes both the firm's own state and the state of its competitors.[9] In each sector, the firms then choose the strategy that maximizes their present value of consumption, conditional on $\psi_{-ist}$, the strategies of the firm's competitors:

$$\max_{\psi_{ist}} E_t[U_{ist}(\psi_{ist}, \psi_{-ist})], \tag{13}$$

subject to the budget constraint from equation (10) and the collateral constraint from equation (9). This optimization will imply that the budget constraint is satisfied with equality, and therefore decisions for capital and debt immediately imply a decision for consumption as well. In addition, decisions on labor and capital for next period imply a decision for output and associated revenue.

**Definition.** ***An industry equilibrium*** *consists of a set of strategies $\psi_{ist}$ for all firms $i$ in sector $s$ that constitute a Nash equilibrium given the optimization problem in* (13)*, conditional on a specific path for the macroeconomy $\{\mathbf{F_t}\}$.*

In the next sections, I will discuss the firms' optimization process and the resulting industry equilibrium in greater detail for extreme cases of the collateral constraint $\lambda$. For now, I note that the firms' equilibrium decision rules for capital, labor and debt are a function of the state of the sector, the history of the game, and the state of the macroeconomy:

[9] While a firm's decision depends on the macroeconomic prices $w_{t+1}, P_{t+1}^F, r_{t+1}^d$, each sector is atomistic, and therefore these factor prices are exogenous to the individual sector. For convenience I will therefore, omit these prices from the notation of the decision rules, except in the definition of general equilibrium below, where the aggregate factor prices play a central role.

$$
\begin{aligned}
&k^*_{ist+1}(D^s(z_{ist+1}, a_{ist}), h_{st}, \mathbf{F_{t+1}}) \\
&l^*_{ist+1}(D^s(z_{ist+1}, a_{ist}), h_{st}, \mathbf{F_{t+1}}) \\
&d^*_{ist+1}(D^s(z_{ist+1}, a_{ist}), h_{st}, \mathbf{F_{t+1}})
\end{aligned} \tag{14}
$$

Within each sector, the decision rules on labor and capital, given the state of the sector, result in the joint distribution of productivity, capital and labor, denoted by $H^s(a_{ist}, k_{ist}, l_{ist})$.

**Definition.** ***A steady state equilibrium*** *consists of, first, stable industry equilibria for all sectors, where within each sector the joint distribution of productivity and wealth is stable:*

$$D^s(z_{ist+1}, a_{ist}) = D^s(z', a),$$

*and so is the joint distribution of productivity, capital and labor:*

$$H^s(z_{ist}, k_{ist}, l_{ist}) = H^s(z, k, l).$$

*Second, workers' decision rules for saving and consumption described in* (12) *that satisfy the budget constraint in equation* (11) *with equality. Third, a stable macroeconomic state* $\mathbf{F}$*: a wage* $w$*, a final good price* $P^F$ *an interest rate* $r^d$*, such that the labor market clears in every period:*

$$L = \int_{s\in S} \sum_{i=1}^{M_s} l^*_{ist+1}(D^s(z', a), h_{st}, \mathbf{F})ds, \tag{15}$$

*and the debt market clears given decisions about investment and consumption:*

$$\int_{l\in L} b^*_{lt+1}(\mathbf{F})dl = \int_{s\in S} \sum_{i=1}^{M_s} d^*_{ist+1}(D^s(z', a), h_{st}, \mathbf{F})ds, \tag{16}$$

The interest rate on debt in a steady state equilibrium will be $r^d = \frac{1}{\beta} - 1$. If the interest rate would deviate from this level, then in each period, workers would either only accumulate savings and have zero consumption, or aim to consume only in the present period, both of which cannot be a steady state equilibrium. Also note that since all firms and all workers satisfy their budget constraints with equality, when the labor and debt market for period $t + 1$ clear, then the goods market for period $t$ also clears given the real wage $\frac{w}{P^F}$ determined in the previous period:

$$Q^F - \frac{w}{P^F}L = \int_{s\in S} \sum_{i=1}^{M_s} \left(x^*_{ist}(D^s(a', z), h_{st}, \mathbf{F}) + d^*_{ist}(D^s(a', z), h_{st}, \mathbf{F})\right) ds.$$

In the next section, I examine two types of steady state equilibria. First I look into the benchmark case where firms can finance any desired level of capital, and afterwards I consider the limit case where firms have no access to external finance.

## 2.2 No capital constraints

In this section, I set $\lambda = 1$ to examine the benchmark case where firms face no limit on their investment in capital. The collateral constraint then only constrains the firm-owner's consumption level, but imposes no restrictions on her investment in capital. It is convenient to solve the firm's problem via the dual approach of first minimizing costs for any output and then determining optimal output levels. The marginal cost, in terms of utility from consumption in period $t+1$, of the inputs are the real wage $w_{t+1}/P_{t+1}$ for labor, and the rental rate of capital $r^k \equiv \frac{1}{\beta} + \delta - 1$. Given these factor costs, standard cost minimization for Cobb-Douglas production functions implies that firms have the following factor demands:

$$k^*(y_{ist+1}) = \frac{w_{t+1}}{P^F_{t+1}} \frac{1}{r^k} \frac{\alpha}{1-\alpha} y_{ist} \tag{17}$$

$$l^*(y_{ist+1}) = \frac{P^F_{t+1}}{w_{t+1}} r^k \frac{1-\alpha}{\alpha} y_{ist} \tag{18}$$

which result in the following constant marginal cost:

$$MC^u_{st}(z_{ist}) = \frac{1}{z_{ist}} \frac{(r^k)^\alpha \left(\frac{w_{t+1}}{P^F_{t+1}}\right)^{1-\alpha}}{\alpha^\alpha (1-\alpha)^{1-\alpha}}, \tag{19}$$

where the superscript $u$ stands for unconstrained. Given this marginal cost, firms then maximize profits as a function only of output produced, where we can rewrite the profit maximization problem as:

$$\max_{\{y_{ist+1}\}} \sum_{t=v}^{\infty} \beta^{t-v} E_v[v_{ist}(y_{ist}, P_{st}(y_{ist}, \mathbf{y_{-ist}})) - MC^u_{st}(z_{ist}) y_{ist}].$$

This problem is closely related to a standard one-shot Cournot game, except that here the strategic interaction among firms is also dynamic. To understand the potential Nash equilibria, it is useful to first consider the one-period version of this game, where each firm's objective function is to maximize the following profit function: $v_{is}(y_{is}, P_s(y_{is}, \mathbf{y_{-is}})) - MC^u_{st}(z_{ist}) y_{is}$. This optimization implies that each firm sets the optimal markup $\mu_{is}(q_{is}) = \frac{\varepsilon_{is}(q_{is})}{\varepsilon_{is}(q_{is})-1}$. Given the demand function, the firm's best response, or reaction function is then implicitly characterized by:

$$\frac{\varepsilon_{is}(m_{is}(q_{is}))}{\varepsilon_{is}(m_{is}(q_{is}))-1} = \frac{p_{is}(q_{is})}{MC^u_{st}(z_{ist})}. \tag{20}$$

Collecting the reaction functions for all firms in sector $s$, we then have $M_s$ equations for $M_s$ unknown values for $q_{is}$, and the solution to this system of equations is the unique Nash equilibrium for this one-shot Cournot oligopoly game (see also Atkeson and Burstein (2008)).[10]

Returning to the dynamic setting, it is clear that the dynamic game is an infinite repetition of the one-period game, except that firms' productivities are stochastic. Analogous to the argument in Friedman (1971), an infinite repetition of the equilibrium in the one-period game is a subgame perfect equilibrium in the dynamic game. The reason is that when all other firms play the best response from the stage game in every period, a deviation by one firm from this best response in one period lowers the firm's profit in that period, and leaves unaltered its profits in

[10] Abstracting from factor prices that are exogenous to the individual industry, in this Cournot game a firm's payoff depends on two variables, namely his own output and the output of his competitors as summarizes by the $P_s(\mathbf{y_{ist}})$. It is therefore straightforward to show that this game falls under the class of Aggregative Games, as defined by Corchon (1994) or Acemoglu and Jensen (2013). For this class of games, Corchon (1994) explains the uniqueness of the equilibrium, before examining comparative statics on the number of players.

the other periods. In this equilibrium, each firm chooses the strategy where they play the stage game's best response to other firm's output choices in each period, regardless of the history of the game. Since this equilibrium consists of the infinite repetition of the equilibrium strategies from the one-period game, there is necessarily a Nash equilibrium in each subgame of the game. Moreover, since the strategies and actions are indepedent of the history of the game, this subgame perfect equilibrium is also Markov perfect. Of course, as implied by the Folk Theorems, the above equilibrium is not the unique subgame perfect equilibrium, but as a repetition of the stage game equilibrium, it is a natural choice as benchmark equilibrium. I also return to the focus on this specific equilibrium in the next section.[11]

## 2.3 Markup misallocation

In equilibrium, high-productivity firms have a higher market share than the low-productivity firms,[12] and therefore also higher markups : $\mu_{sH} > \mu_{sL}$, where subscripts $H, L$ henceforth refer to the high- and low-productivity firms' optimal choices. Since this markup dispersion is associated with unequal marginal products across high and low productivity firms, there is within-sector resource misallocation.

A formal way to diagnose this "markup misallocation" is by comparing the optimal to the actual input factor ratios across firms. It is straightforward to show that, for any given $M_s$, optimal within-sector resource allocation would imply equalized marginal products across firms and therefore the following factor ratios:[13]

$$\frac{\tilde{l}_{sH}}{\tilde{l}_{sL}} = \frac{\tilde{k}_{sH}}{\tilde{k}_{sL}} = \left(\frac{z_{sH}}{z_{sL}}\right)^{\sigma-1},$$

with the tilde referring to the socially optimal input choices. In contrast, given the factor demand functions from equations (17) and (18), firms' best responses in equation (20) result in the following equilibrium factor ratios:

$$\frac{l_{sH}}{l_{sL}} = \frac{k_{sH}}{k_{sL}} = \left(\frac{z_{sH}}{z_{sL}}\right)^{\sigma-1} \left(\frac{\mu_{sL}}{\mu_{sH}}\right)^{\sigma}.$$

Since high-productivity firms set higher markups, it follows that $\frac{\mu_{sL}}{\mu_{sH}} < 1$ as long as $M_s$ is finite. Hence, in the decentralized equilibrium with finite number of firms, the factor ratios are too low compared to the socially optimal allocation. In other words, high-productivity firms are too small relative to the low-productivity firms. However, when competition goes to its upper limit, i.e. when $M_s \to \infty$, firms' market shares become atomistic and each firm's demand elasticity $\varepsilon_{ist}(m_{ist}(q_{ist})) \to \sigma$. Hence, in this limit case we arrive at monopolostic competition, where all firms set an identical, constant markup equal to $\frac{\sigma}{\sigma-1}$. Hence, markup dispersion, and associated

[11] Note that the expression for the reaction functions in equation (20) is robust to the choice of market structure, as they nest the reaction functions from both the Cournot and the Bertrand oligopoly game. The same will hold when I analyze the equilibrium with financial constraints and the related comparative statics.

[12] Consider the ratio of their reaction functions:

$$\frac{y_{sH}}{y_{sL}} = \left(\frac{z_{sH}}{z_{sL}}\frac{\mu_{sL}(m_{sL})}{\mu_{sH}(m_{sH})}\right)^{\sigma}$$

This relation implies that $y_{sH} > y_{sL}$. To see why, suppose to the contrary that $y_{sH} < y_{sL}$. This would require $\mu_{sH} < \mu_{sL}$. Given that $\frac{\partial \varepsilon_{ist}(m_{ist})}{\partial m_{ist}} < 0$, this in turn would imply that $m_{sH} < m_{sL}$, and this results in a contradiction with $y_{sH} < y_{sL}$. Hence, the supposition is false and its opposite must be true.

[13] To obtain this well-known result, one can maximize total industry output $Q_{st}$, for any given amount of capital and labor available for sector $s$.

markup misallocation, disappears when $M_s \to \infty$.

**Lemma 1.** *When firms have no capital constraints, taking competition to its upper limit, i.e.* $M_s \to \infty$*, ensures that marginal products are equalized across firms within sector* $s$*.*

This analysis of optimal resource allocation does not take into account the issue of selection. Naturally, comparing across equilibria with only atomistic firms, a sector's productivity is highest when only high-productivity firms are producing. The current model is not set up to analyze this type of productivity-enhancing selection. The analysis of markup misallocation here, requires the existence of productivity differences, and such differences typically persist even after the lowest productivity plants disappear through selection. In the next section, the key mechanism is that individual firms can become financially constrained due to productivity volatility, and then start investing in capital using retained earnings. This mechanism can also occur in a setting with selection. In fact, Midrigan and Xu (2014) provide a model of financial constraints and productivity volatility, where there is a selection mechanism operating on the firms in the "modern" sector.

## 2.4 No external finance

In the previous section, we examined the case with no capital constraints. Now, we consider the opposite extreme, where firms have no access to external finance, i.e. $\lambda = 0$.[14] Without access to external financing, each firm's budget constraint simplifies to:

$$k_{ist+1} + c_{ist} \leq a_{ist}(l_{ist}, k_{ist}, \mathbf{y_{-ist}}),$$

where $c_{ist} \geq 0$.

I again start by first performing cost-minimization for any specific output level, denoted by $\bar{y}_{ist}$. For this output level, the firm can either be unconstrained or constrained. Whether a firm is constrained, depends on its maximum capital level, defined as:

$$k^c_{ist+1} \equiv (1-\delta)k_{ist} + \frac{\pi_{ist}}{P^F_t}.$$

If for a given output level $\bar{y}_{ist}$, a firm's capital demand from equation (17) is below $k^c_{ist}$, then its demand functions for capital and labor are exactly as in the case with no financial constraints, and this gives rise to the standard marginal cost function for a Cobb-Douglas, as in equation (19). In contrast, the firm is constrained for a specific $\bar{y}_{ist}$ when the capital demand from equation (17) is larger than its maximum capital amount. Next, define the return on the constrained capital level, denoted $r^{kc}_t$, such that the following equality holds:

$$k^c_{ist+1} = \frac{w_{t+1}}{P^F_{t+1}} \frac{1}{r^{kc}_{t+1}} \frac{\alpha}{1-\alpha} \bar{y}_{ist+1},$$

which naturally implies that $r^{kc}_{t+1} > r^k$. In turn, the fact that the actual return on capital is higher than $r^k$, which is the required return on capital to ensure the intertemporal optimality of investment, entails that it will be optimal for the constrained firm to invest all its wealth in capital by setting it at $k^c_{ist+1}$. If the firm would invest less, it would have a strictly lower net present value

[14] In models with heterogeneous agents and financial constraints, it is becoming increasingly common to analyze the limit case of no external financing to obtain analytical results. See e.g. Krusell, Mukoyama, and Smith Jr (2011); Werning (2015) and Ravn and Sterk (2016) in the context of income risk for consumers. This paper is one of the first to employ this strategy in the context of firm heterogeneity and resource misallocation.

of expected utility. Note that these input demand functions for capital implicitly define the consumption path for firms. When firms are constrained, they invest all their revenue in order to attain $k^c_{ist+1}$. When they are not constrained, they invest until they reach their optimal capital choice and consume the remaining revenue.

Since the capital level of the constrained firm is at its maximum, at the margin the firm can only adjust its labor input. For any quantity $\bar{y}_{ist}$, its total variable costs are $\frac{w}{P^F} l(\bar{y}_{ist})$, which implies the following marginal cost function:

$$MC^c_{st}(z_{ist}, \bar{y}_{ist}, k^c_{ist}) = \frac{w_t}{(1-\alpha)P^F_t} \left( \frac{\bar{y}^{\alpha}_{ist}}{z_{ist}(k^c_{ist})^{\alpha}} \right)^{\frac{1}{1-\alpha}} \tag{21}$$

Combining both the unconstrained and the constrained case, and noting that $k^c_{ist}$ is a function of $a_{ist-1}$, a firm's marginal cost is a function of its productivity, output level, and wealth in the previous period: $MC_{ist}(z_{ist}, y_{ist}, a_{ist-1})$.

The fact that firms can be financially constrained, implies that their actions depend not only on their productivity, but also on their wealth. This implies that firms may have an incentive to influence the future wealth distribution, thereby affecting other firms' future actions. For instance, firms may expand their production and drive down prices in order to slow down the growth path of its financially constrained competitors. Importantly, when this happens off the equilibrium path, namely constrained firms' capital growth being slowed down, then the joint distribution of capital and productivity becomes non-stationary. This non-stationarity renders the analytical derivation of subgame perfect equilibria highly challenging in the case of financial constraints.

To sidestep this analytical challenge, I focus on the subset of Nash equilibria where firms have the reaction functions from the static game in each period:

**Assumption 1.** *On the equilibrium path, firms' reaction functions take the following form:*

$$\frac{\varepsilon(m_{ist}(y_{ist}))}{\varepsilon(m_{ist}(y_{ist}))-1} = \frac{p_{ist}(y_{ist})}{MC_{ist}(z_{ist}, y_{ist}, a_{ist-1})}, \tag{22}$$

Consider a strategy where firms choose the above reaction function as long as their competitors also opt for this reaction function. When a firm deviates from this reaction function however, for instance to slow down the growth rate of constrained firms, the other firms increase output and thereby punish the deviator. It is straightforward to show that such an equilibrium is a Nash equilibrium, given a condition on the discount factor $\beta$. What is analytically challenging however, is the application of equilibrium refinements, because $D^s(z_{ist}, k_{ist}, a_{ist})$ becomes non-stationary in the subgame off the equilibrium path.[15] For this reason, I focus on the equilibria containing the reaction functions in Assumption 1. An important advantage of this approach, is that it allows for an analytical solution of the comparative statics across steady states with different $M_s$. Additionally, the analysis remains robust to alternative assumptions on market structure, as the results will hold for both Cournot and Bertrand competition.

In each period, there is a unique $y_{ist}$ that solves the reaction function in (22), conditional on a firm's productivity and other firms' output choices, since the left-hand side and the right-hand side of equation (22) are monotonically increasing and decreasing, respectively. For identical

[15] In settings with firm heterogeneity, subgame or Markov perfect equilibria in oligopolistics industries are often solved numerically instead of analytically. See e.g. Doraszelski and Pakes (2007) for an overview.

mathematical reasons as why the equilibrium under no financial constraints was unique, the equilibrium outcomes for any given period continue to be unique in this setting, given the above reaction functions. The intuition for the unique equilibrium is that each firm's output choice is a monotonic function of the sectoral price index $P_{st}(\mathbf{y_{ist}})$, which in turn is a concave function of each firm's output, as explained in Corchon (1994).

## 2.5 Steady state distribution of capital

What will the distribution of input levels look like in steady state equilibrium? First, a firm's capital depends on whether it is unconstrained or constrained. When it is constrained, it grows its capital by setting it at $k^c_{ist}$. For unconstrained firms, as derived in Section 2.3, high-productivity firms have a higher capital level:

$$k_{sH} > k_{sL},$$

where the subscripts $H, L$ refer to optimal choices in steady state by the unconstrained high- and low-productivity firms respectively.[16]

In the steady state equilibrium with infinitely lived firms, these capital choices will imply that either firms are at their unconstrained capital level, or that they are growing their capital from $k_{sL}$ to $k_{sH}$, in the manner described by the following lemma:

**Lemma 2.** *In steady state, the joint distribution of capital and productivity within a sector is as follows:*

- *all low-productivity firms are unconstrained: if $z_{ist} = z_{sL}$, then $k_{ist} = k_{sL}$*
- *high-productivity firms can be constrained or unconstrained, depending on the number of periods $\tau$ since their most recent productivity shock, and constrained firms invest all their wealth into capital growth: if $z_{ist} = z_{sH}$, then $\forall i$ with $\tau = t - v$, where $v \equiv \max r$ s.t. $z_{isr+1} = z_{sH} \& z_{isr} = z_{sL}$:*
  - *if constrained, then $k_{ist} = G_{s\tau} k_{sL}$, with $G_{s\tau} \equiv \Pi_{r=s}^{s+\tau} \left( \frac{\pi_{sv}}{P^F k_{sv}} + 1 - \delta \right)$*
  - *if unconstrained, then $k_{ist} = k_{sH}$*

To see why this lemma holds, start by supposing that a firm is currently unconstrained at $k_{sL}$, and receives a positive productivity shock and knows that in the next period, it will be at high productivity $z_{sH}$. There are then two possibilities. One is that the firm is immediately unconstrained and able to set $k_{sH}$. The other possibility is that the firm is constrained, and in that case it will start growing its capital level by setting it at $k^c_{ist+1}$. Moreover, for any future period $\tau$ after its most recent positive productivity shock where the firm is still constrained, the firm will set its capital at $k^c_{ist}$. Then, consider firms that are at low productivity $z_{sL}$ for the current period, which should be unconstrained at $k_{sL}$ according to the Lemma. Suppose they were not at $k_{sL}$, then these firms would all be growing their capital to the next period, even if they remain at $z_{sL}$, and this will lead to a contradiction with $H^s(z, k, l)$ being stable over time.[17]

[16] The fact that the minimum markup $\sigma/(\sigma - 1)$ is above unity implies that firms can afford an investment rate at least as high as the depreciation rate, since it implies that a return at least as high as $r^k$ is available for each unit of capital.

[17] The stability of $H^s(z, k, l)$ requires that when *all* the $z_{sL}$ firms are growing their capital, there should be other firms, which were $z_{sH}$ in the previous period, that take their spot in the $H^s(z, k, l)$ distribution, with a capital level below $k_{sL}$. This is inconsistent with being in steady state. Firms have a wealth level either above or below $k_{sL}$. If it is above, then when they experience a negative productivity shock, they will choose exactly $k_{sL}$ in the next period. If their wealth level is below $k_{sL}$, then we can again ask who will take their spot in the $H^s(z, k, l)$ distribution. We have an inductive step then here, with a requirement for more deeply constrained firms at each step. Since firms with higher wealth will never return to wealth below $k_{sL}$ and given that firms are infinitely lived, this inductive step leads to a contradiction with $H^s(z, k, l)$ being stable.

## 2.6 Comparative statics on the degree of competition

To start the analysis of comparative statics across industry equilibria, I consider industry equilibria under two different values for the number of firms: $M_s \neq M'_s$. The equilibrium values under $M'_s$ are denoted with a prime. Comparing these equilibria, I aim to examine how capital growth for constrained firms in each "bin" $\tau$, as well as markup levels for both constrained and unconstrained firms behave as a function of the number of firms in the industry. Initially, I will be agnostic about whether $M_s > M'_s$ or not, and start instead by supposing, without loss of generality, that the low-productivity firm's market share is higher in the former equilibrium ($m_{sL} \geq m'_{sL}$), and examining the logical implications of that supposition on other firms' market shares and capital growth rates. Those logical implications, summarized in Lemma 3, will in turn imply that $M_s < M'_s$, which will allow me to conduct comparative statics across equilibria with different numbers of firms. Note that from now on, only constrained firms are denoted with the subscript $\tau$.

**Lemma 3.** *If the market share of the unconstrained low-productivity firm is higher in one industry equilibrium compared to another, then the market share of the high-productivity unconstrained firms is also higher:*

$$(m_{sL} > m'_{sL}) \implies (m_{sH} > m'_{sH}). \tag{23}$$

*Moreover, the capital growth rate and the market share for all constrained firms in any bin $\tau$ will be higher in the former equilibrium as well:*

$$(m_{sL} > m'_{sL}) \implies \forall \tau > 0 : (G_{s\tau} > G'_{s\tau}) \wedge (m_{s\tau} > m'_{s\tau}) \tag{24}$$

To see why equation (23) holds, consider the implication of the factor demand functions (17) and (18) on the output ratio of the high- and low-productivity firm:

$$\frac{y_{sH}}{y_{sL}} = \left(\frac{z_{sH}}{z_{sL}}\right)^{\sigma} \left(\frac{\mu_{sL}}{\mu_{sH}}\right)^{\sigma} \tag{25}$$

This implies that the market shares of firm types $H, L$ are linked, and that when $m_{sL}$ increases, $m_{sH}$ can only decrease if the relative markup $\frac{\mu_{sL}}{\mu_{sH}}$ increases. The proof in Appendix Section A.1 formally demonstrates that this entails a contradiction and that therefore Equation (23) holds.

Next, I demonstrate that equation (24) holds by employing a proof by induction, available in Appendix Section A.2. The proof starts by observing that if the low-productivity firms have a higher market share, their higher markup implies that they have higher capital growth immediately after a positive productivity shock: $G_{s1} > G'_{s1}$. This is because their capital growth depends on their revenue net of labor payments, per unit of capital, which can be written as the sum of profits and payments to capital:

$$\frac{\pi_{sL}}{P^F k_{sL}} = (\mu_{sL} - 1) MC_{sL} \frac{y_{sL}}{k_{sL}} + r^k.$$

Hence, since the rental rate of capital $r^k$ is constant, variation in the capital growth depends on profits per unit of capital: $(\mu_{sL} - 1) MC_{sL} \frac{y_{sL}}{k_{sL}}$. Since marginal cost and the capital-output ratio are constant for unconstrained firms, the higher markup associated with a higher market share will entail faster capital growth.

After establishing that capital growth increases for firms in the period when they first learn about their positive productivity shock, the proof shows that capital growth continues to be elevated in subsequent periods. It does so by demonstrating the following inductive step:

$$((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (G_{s\tau+1} > G'_{s\tau+1}).$$

Proving this inductive step starts from noting that capital growth increases in revenue net of labor payments per unit of capital, since $G_{s\tau} \equiv \Pi_{r=s}^{s+\tau} \left( \frac{\pi_{sv}}{P^F k_{sv}} + 1 - \delta \right)$. It then proceeds by showing that revenue net of labor payments per unit of capital increases when the premise of the inductive step holds, i.e. when $((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}))$. The intuition for this argument is as follows. When the low-productivity firm has a higher market share and the firm in bin $\tau$ has a higher capital growth rate $G_{s\tau}$, the firm in bin $\tau$ has the option of setting its marginal cost at the same level as the firm in bin $\tau$ under $M'_s$. In that case, this firm would see its markup increase, $\mu_{s\tau} > \mu'_{s\tau}$, as it faces higher demand than the firm under $M'_s$. A higher markup under constant marginal cost implies that $\frac{\pi_{s\tau}}{P^F k_{sL}}$ increases, and thereby capital growth increases as well. Of course, the firm can choose a different level of production at a different marginal cost. Critically though, the firm will not be worse off in terms of revenue net of labor cost if its optimal choice implies a different marginal cost, and therefore its capital growth will also be increasing in that case. In addition, the proof demonstrates that:

$$((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (m_{s\tau} > m'_{s\tau}).$$

Here, the intuition is that a higher $G_{s\tau}$ reduces a firm's marginal cost, which translates into a higher market share through the reaction function in Equation (22).

Equations (23) and (24) from Lemma 3 together imply that the market shares of all types of firms, i.e. low-productivity firms $L$, unconstrained high-productivity firms $H$, and constrained firms in any bin $\tau$, jointly increase or decrease across industry equilibria.[18] When we have $(m_{sL} > m'_{sL})$, market shares for all these types of firms increase. When market shares for all types of firms increase, it implies that there are fewer firms in the industry, and therefore:

$$(m_{sL} > m'_{sL}) \implies (M_s < M'_s).$$

The converse also holds, since $(M_s < M'_s)$ implies that the market share of at least one type of firm needs to strictly decrease. Since all market shares increase and decrease together (see footnote 18), it follows that:

$$(M_s < M'_s) \implies (m_{sL} > m'_{sL}).$$

In combination with Lemma 3, this directly implies that when the number of firms falls, the market share of all types of firms increases:

$$(M_s < M'_s) \implies ((m_{sL} > m'_{sL}) \wedge (m_{sH} > m'_{sH}) \wedge (\forall \tau : m_{s\tau} > m'_{s\tau})) \tag{26}$$

This result, combined with the monotonically increasing relationship between market shares and markups, implied by equations (5) and (22), then directly implies that markup levels for all types of firms fall as the number of firms increases. Moreover, together with equation (24) from

[18] Naturally, the statements $(m_{sL} > m'_{sL}) \wedge (m_{sH} \leq m'_{sH})$, and $(m_{sL} > m'_{sL}) \wedge (m_{s\tau} \leq m'_{s\tau})$, for any $\tau > 0$, are in contradiction with Equation (23) and Equation (24) respectively.

Lemma 3, it implies that capital growth rates of constrained firms fall as well when the number of firms increases. Finally, Appendix Section A.3 demonstrates that markup dispersion also falls with the number of firms. This is intuitive, since as $M_s \to \infty$, all markups converge to $\frac{\sigma}{\sigma-1}$, the markup under monopolistic competition. Taking all the above together, this demonstrates the following Proposition:

**Proposition 1.** *For any $M_s' > M_s$, and for unconstrained firm-types $L, H$, and for constrained firms in bin $\tau > 0$:*

- *Markup levels fall with $M_s$:*

$$\mu_{sL}' < \mu_{sL} \, ; \, \mu_{sH}' < \mu_{sH} \, ; \, \mu_{s\tau}' < \mu_{s\tau}$$

- *Markup dispersion falls with $M_s$:*

$$\frac{\mu_{sH}'}{\mu_{sL}'} < \frac{\mu_{sH}}{\mu_{sL}} \, ; \, \frac{\mu_{s\tau}'}{\mu_{sL}'} \leq \frac{\mu_{s\tau}}{\mu_{sL}}$$

- *Capital growth rates for all financially constrained firms fall with $M_s$:*

$$G_{s\tau}' < G_{s\tau}.$$

The results in Proposition 1 are highly intuitive, but not obvious. After all, since the capital growth rate $G_{s\tau}$ falls with $M_s$, it could have been possible for the market shares of the unconstrained firms to increase with $M_s$. The above analysis verifies that this is not the case, and that both capital growth rates and markup levels fall monotonically with the number of firms.

These results also have important welfare implications, in particular for understanding the gains from taking competition to its upper limit. Recall from Lemma 1, that when firms face no capital constraints, setting $M_s \to \infty$ equalizes marginal products across firms. In contrast, Proposition 1 entails that $\frac{k_{s\tau}}{k_{sL}}$ falls with $M_s$, such that the wedge between the socially optimal capital ratio $\frac{\tilde{k}_{s\tau}}{\tilde{k}_{sL}} = \left(\frac{z_{sH}}{z_{sL}}\right)^{\sigma-1}$ and the actual capital ratio actually deepens. While previous research argues for unambiguously positive productivity gains from increasing competition (e.g. Peters (2016); Edmond, Midrigan, and Xu (2018)), Proposition 1 demonstrates that the productivity effects are ambiguous instead in the presence of financial constraints.

## 3 Panel data on Indian plants

I employ data on manufacturing establishments, or plants, from the Indian Annual Survey of Industries (ASI), for the period 1990-2011. The ASI provides data by accounting year, which starts on April 1st, and I refer to each accounting year by the calendar year on which it starts. The ASI sampling scheme consists of two components. The first component is a census of all manufacturing establishments located in a small number of specific geographic areas, or above a certain size threshold. For the large majority of years in my dataset, the size threshold for the census component is an employment level of 100 workers. Only for the years 1997 till 1999, this threshold is 200 workers instead. The second component of the sampling scheme includes, with a certain probability, each formally registered establishment that is not part of the census component. All establishments with more than 20 workers, or 10 workers if the establishment uses electricity, are required to be formally registered. For my analysis, I restrict the sample to plants in manufacturing sectors that are operational, and have non-missing positive values for the logarithm of three critical variables in the analysis, namely revenue, capital and labor cost. I also ensure that definitions of geographical units, i.e. states or union territories, are consistent over time. Appendix C gives a full overview of the data cleaning procedure.

Central to my analysis are the establishment identifiers, used to construct the panel for the entire 1990-2011 period. To construct the panel, I use the establishment identifiers provided by the Indian Statistical Office for all years from 1998 onward. For the years prior to 1998, I use the establishment identifiers employed by Allcott, Collard-Wexler, and O'Connell (2016), who gained access to these identifiers while working in India. I thank Hunt Allcott for generously making these panel identifiers available.

# 4 Testing the theory with a natural experiment

In this section, I use a natural experiment to test the theory's predictions on the impact of competition at the plant level. In particular, I exploit the natural variation in competition arising from India's dereservation reform. In line with the theory, among incumbent plants exposed to this pro-competitive reform, markup levels and markup dispersion fall after the reform, while their capital convergence slows down.

## 4.1 Background on the dereservation reform

The dereservation reform consists of the staggered removal of the small-scale industry (SSI) reservation policy. This reservation policy mandated that only industrial undertakings below a certain ceiling for accumulated investment - 10 million Rupees at historical cost in 1999 (Martin et al., 2017) - were allowed to produce products within certain product categories.[19] As such, this policy was one of the most important aspects of India's economic agenda of promoting small-scale industries (Mohan, 2002). This agenda was launched in the 1950s in an attempt to promote social equity and to boost employment growth by stimulating intensive use of labor in manufacturing. The reservation policy itself was introduced in the Third Five Year Plan (1961-1966). In 1996, before the start of dereservation, more than 1000 product categories were reserved for SSI. These product categories crossed many different industries, such as chemicals, car parts, electronics, food and textiles. In total, reserved products constituted around 12% of Indian manufacturing output (Tewari and Wilde, 2017).

Dereservation started in 1997, several years after a first wave of liberalization policies in the early 1990s (Bollard et al., 2013). According to Martin et al. (2017), stiffer foreign competition following the trade liberalization and increased complexity of industrial production convinced policymakers to gradually start abandoning the reservation policy. The decision to then actually dereserve a particular product is only taken after a series of meetings between relevant stakeholders, review up a chain of bureaucrats, and final approval by the central government minister (Tewari and Wilde, 2017). Appendix Figure D.1 provides an overview of the timing of dereservation. The process of dereservation clearly peaks between 2002 and 2008. By 2015, no products are

[19] At the time of reservation, an exception was made for large industrial undertakings already producing the product. These undertakings were allowed to continue production, but with output capped at existing levels.

reserved anymore.

## 4.2 Data for dereservation analysis

Data on the timing of dereservation for each dereserved product is available on the website of the Indian Ministry of Micro, Small and Medium Enterprises.[20] Importantly, the ministry sets the time of dereservation by SSI product code. To match these SSI product codes to the product classification in the ASI data, I use the concordance available from Martin et al. (2017). Appendix Section C.2 provides more detail on the construction of this concordance.

Starting 2010, the ASI changes its product classification from ASICC (A Standard Industrial Commodity Classification) to NPCMS (National Product Classification for Manufacturing Sector). Since the concordance between the ASICC and NPCMS classifications is not one-to-one, I only consider products that were dereserved prior to 2010. This covers 97.9% of reserved products. An additional complication is that the ASI data has incomplete product coverage for the years 1998 and 1999. A potential concern is that the covered products in those years are a non-randomly selected sample, which could introduce bias into the estimation results. To address this issue, I implement robustness checks where I restrict the sample to incumbent plants with dereservation years after 1999 (see Appendix D).

I define a plant as an "incumbent" if it produces at least one reserved product prior to dereservation. For incumbent plants, I define their year of dereservation as the accounting year in which a first product of theirs is dereserved.[21] Importantly, as described by Martin et al. (2017), the large majority of plants who produce reserved products - 89.5% of them in my sample - only produce one such product. Moreover, 93.7% of the plants make products that are dereserved in the same accounting year. Appendix Table D.1 presents summary statistics for the panel of incumbent plants used in the estimation of capital convergence below. The average incumbent in my sample is 20 years old and has 327 workers.

## 4.3 Semi-random nature of the timing of dereservation

Since the decision to dereserve a product is made by policymakers in consultation with stakeholders, one may be worried that the endogeneity of this political process undermines causal identification of the reform's effects. A first observation is that all reserved products are eventually dereserved, so on the policy side there is no selection of some products being dereserved and others not. Nevertheless, the exact timing of dereservation may be non-random. For instance, a product may get dereserved when the economic circumstances are more favorable for that specific product category. However, there appear to be no systematic patterns in the data regarding the timing of dereservation. To start, Tewari and Wilde (2017) demonstrate that there is considerable variation in the timing of dereservation for strongly related product categories, e.g. different types of vegetable oils. As products within these narrow product categories arguably share similar demand and supply characteristics, this limits the scope for a structural explanation of the timing of dereservation. More importantly, I find that there are no pre-trends associated with the timing of dereservation for plant-level revenue, capital, labor or labor cost (see Appendix Figure

[20] This list is available at `http://dcmsme.gov.in/publications/circulars/newcir.htm#RESERVED`, as retrieved on February 15, 2019.

[21] Since the accounting years in the ASI data start on April 1st, the dereservation year for products dereserved between January 1st and March 31st, is the accounting year starting in the previous calendar year. For instance, products dereserved on February 3d 1999, have the 1998-1999 accounting year as their year of dereservation.

D.2). This evidence is in line with an older result from Martin, Nataraj, and Harrison (2014) on the absence of pre-trends associated with dereservation timing for plant-level employment. Taking everything together, this lends credible support to the timing of dereservation being semi-random and considering this reform a natural experiment.

## 4.4 Dereservation as a pro-competitive shock

Which are the main effects of deservation? First, dereservation implies the removal of a size restriction, and removing this distortion has direct positive effects on allocative efficiency (Guner, Ventura, and Xu, 2008; García-Santana and Pijoan-Mas, 2014). Second, dereservation also entails an increase in the degree of competition, since incumbent plants are allowed to grow their capital stock, and particularly because larger firms can now enter the previously reserved product categories. The large magnitude of this pro-competitive shock is already indicated by the fact that over the entire sample period, there are only 38,592 unique incumbent plants in my dataset, compared to 113,967 unique plants entering the previously reserved product categories after dereservation. Martin et al. (2017) provide a detailed analysis of the pro-competitive impact of dereservation, demonstrating that for entrant plants, employment, output and capital grow substantially after dereservation, whereas the average incumbent plant shrinks on these dimensions.

In this paper, I leverage the pro-competitive impact of the dereservation reform to test the theoretical predictions of my model. Critically then, I am not performing a complete analysis of the impact of the reform on allocative efficiency in the manufacturing sector. After all, this would require bringing the additional complication of size restrictions into the model. Since the steady state of the model is already described by a system of non-linear equations, it is unclear if analytical results would still be available in that case. Instead of performing a welfare analysis of the dereservation reform, I am interested in examining if the analytical model predictions of my model are borne out by the empirical reality. To that end I exploit the natural variation in competition for incumbent plants arising from the dereservation reform.

I start by investigating how dereservation affects the markups of incumbent plants. I measure $\mu_{it}$, the markup for plant $i$, in year $t$, as in De Loecker and Warzynski (2012) application of Hall (1986)'s insight:

$$\mu_{it} = \alpha_i^L \frac{S_{it}}{w_{it} L_{it}} \tag{27}$$

where $\alpha_i^L$ is the elasticity of value added with respect to labor, and $\frac{w_{it}L_{it}}{S_{it}}$ is labor's share of revenue $S_{it}$.[22] Here, it is intuitive that for a constant output elasticity of labor, a higher labor share implies a lower markup. This expression for the markup rests on the assumptions of a Cobb-Douglas production function, cost minimization, and having labor as a variable input.[23] The latter assumption appears plausible for my data, since after dereservation the number of employees immediately starts falling significantly and substantially for incumbent plants (see Panel c of Appendix Figure D.2.)

[22] In my data, the distribution of the inverse of the labor share of revenue appears roughly lognormal (see Appendix Figure D.3).

[23] Given a Cobb-Douglas production function, the Lagrangian of the cost-minimization problem with variable labor input and a predetermined capital level is $\min_{l_{it}} \mathcal{L}_{it} = w_{it} l_{it} + \lambda_{it}(Y_{it} - a_{it} k_{it}^{\alpha_i^K} l_{it}^{\alpha_i^L})$. Optimization sets: $w_{it} = \lambda_{it} \alpha_i^L \frac{a_{it} k_{it}^{\alpha_i^K} l_{it}^{\alpha_i^L}}{l_{it}}$, and therefore $\frac{p_{it}}{\lambda_{it}} = \alpha_i^L \frac{p_{it} a_{it} k_{it}^{\alpha_i^K} l_{it}^{\alpha_i^L}}{w_{it} l_{it}}$. Since $\lambda_{it}$ is the marginal cost of output, $\frac{p_{it}}{\lambda_{it}} = \mu_{it} = \alpha_i^L \frac{p_{it} y_{it}}{w_{it} l_{it}}$.

Empirically, I allow for maximal cross-plant heterogeneity in $\alpha_i^L$ by absorbing it in a plant fixed-effect. Using this markup measure, I run the following event-study on dereservation:

$$\ln \mu_{it} = \gamma_i + \nu_t + \sum_{\tau=-5}^{4} \beta_\tau 1[t = e_{it} + \tau] + \varepsilon_{it} \tag{28}$$

Here, $\gamma_i$ is a plant fixed-effect and $\nu_t$ is a year fixed effect. The year in which a plant's first product is dereserved is denoted by $e_{it}$, and I bin up the end points and normalize $\beta_{-1} = 0$. For the purpose of this event study, I restrict attention to a balanced sample of incumbent plants which are observed at least three years before and after they were dereserved. An incumbent plant is any plant that produced a reserved product prior to dereservation. Since dereservation status is determined at the product level, I cluster standard errors at that level.

I find that on average, markups indeed fall due to the dereservation reform (see Figure 1, Panel a). The initial decline is modest, but eventually the average markup declines by 0.08 log points (p=0.086). This impact of dereservation is in line with the theoretical prediction that markup levels fall when competition increases, and the economic magnitude of the impact of the reform is substantial. Note also that there is no pre-trend for markups prior to dereservation.[24]

In addition to leading to a decrease in the average markup, dereservation also reduces markup dispersion. Recall that the model predicts that as the degree of competition increases, all markups converge to a lower bound. To test this prediction, I split the set of incumbent plants into two subsets, depending on whether before dereservation a plant's markup is above or below the median initial markup.[25] I find that plants who have higher markups in the periods before dereservation, exhibit a stronger average decline in their markup after dereservation, as predicted by the theory. More specifically, for the subset of plants with below-median initial markup, dereservation appears to have no effect on markup levels (see Figure 1, Panel (c)). In contrast, plants with above-median initial markups, experience a strong decline in their markup. After three years, their average markup is 0.14 log points lower (see Figure 1, Panel (b)).[26]

### 4.5 Dereservation and capital misallocation

I now examine if the increase in competition associated with dereservation also leads to a slowdown in plants' capital convergence. In the model, firms optimally choose to grow their capital stock in response to positive productivity shocks until they reach their optimal, unconstrained level of capital. The empirical challenge is that the optimal level of capital is unobserved. Interestingly, while a positive productivity shock leads to a first-order increase in the optimal level of capital, the change in the optimal marginal revenue product of capital (MRPK) is second order. As a result, it is feasible to find valid proxies for optimal MRPK. For this reason, and inspired by Asker et al. (2014), I focus on convergence in MRPK in my empirical analysis of capital con-

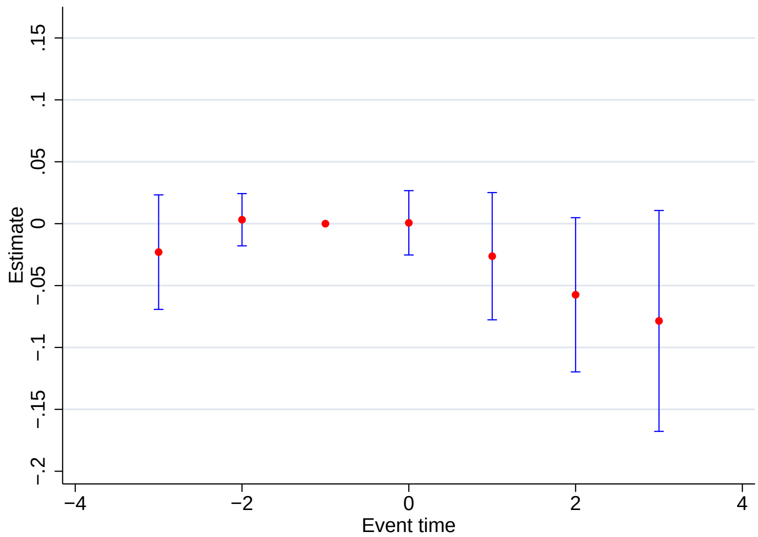
[24]To further corroborate the absence of a pre-trend, I re-estimate the event study over a longer time horizon in Appendix Figure D.4. The results are qualitatively very similar for the plants below or above the median initial markup. The longer time horizon implies that the number of observations included in the balanced panel shrinks from 20,937 for the analysis in Figure 1 to 13,522 in Appendix Figure D.4. Together with the heterogeneity across plants with high or low initial markup, this reduced sample size renders the estimates in Panel (a) less statistically significant.

[25]The initial markup is averaged over event times $\tau = -3$ and $\tau = -2$, and the median initial markup is then set after taking out sector and year fixed effects.

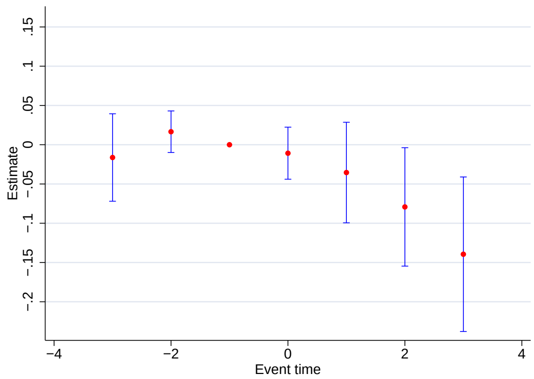
[26]What is driving this drop in markups for this latter group of plants? Interestingly, while these plants shed labor, if anything their labor cost seems to increase (see Appendix Figure D.5 Panels a and b). Combined with the downward pressure on revenue (Panel c), this leads to lower markups.

Figure 1: Event study for the impact of dereservation on markups

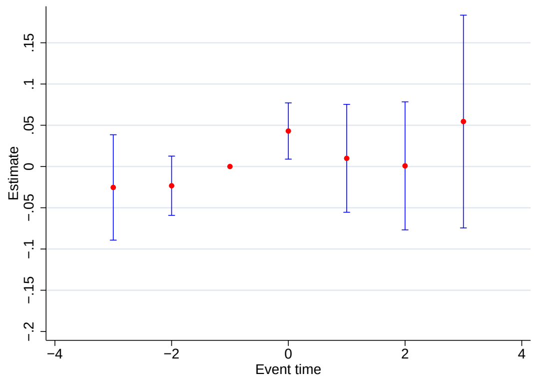
(a) Full sample of incumbents




(b) Incumbents with above-median initial markup




(c) Incumbents with below-median initial markup




The figure displays the coefficients and 95% confidence intervals for the $\beta_\tau$ coefficients from the following event-study regression: $\ln \mu_{it} = \gamma_i + \nu_t + \sum_{\tau=-3}^{3} \beta_\tau 1[t = e_{it} + \tau] + \varepsilon_{it}$, where $\gamma_i$ is a plant fixed effect and $\nu_t$ is a year fixed effect. I define the time at which a plant's first product of plant $i$ is dereserved as $e_{it}$, and I impose the normalization that $\beta_{-1} = 0$. Since dereservation status is defined at the product-level, I also cluster standard errors at that level. I restrict the sample to a balanced sample of incumbent plants that are observed at least three years before and after they are dereserved. Panel (a) shows results for the full sample of incumbent plants. Panel (b) displays results for plants with initial markups weakly above the median initial markup, and Panel (c) for the other plants. The initial markup is an average over event times $\tau = -3$ and $\tau = -2$, and the median initial markup is then set after taking out sector and year fixed effects.

vergence. When a firm is financially constrained, its actual $MRPK_{it}$ will be above its optimal, unconstrained MRPK, denoted by $MRPK^*_{it}$. Since $MRPK_{it}$ is a strictly monotonic function of a firm's capital level, when capital convergence in the model slows down, convergence in terms of MRPK also slows down.

As in Asker et al. (2014), I measure MRPK in logs after assuming Cobb Douglas production functions:

$$MRPK_{it} = \ln \alpha_i^K + s_{it} - k_{it}, \tag{29}$$

where $\alpha_i^K$ is the output elasticity of capital, $s_{it}$ is log revenue, and $k_{it}$ is log capital. To allow for maximal cross-plant heterogeneity in $\alpha_i^K$, I absorb this output elasticity in a plant fixed effect in the regression analysis. Hence, within-plant variation in MRPK will be driven by the log ratio of revenue to capital. To examine the sensitivity of the results to the precise measurement of MRPK, I also measure it using log value-added instead of log revenue. Since the measurement in logarithms requires value added to be positive, this leading robustness check is also performed on a more restrictive selection of "well-performing" plants.

To examine if MRPK convergence slows down due to dereservation, I use the following autoregression framework:

$$\begin{aligned} MRPK_{it} = \gamma_i + \nu_t + \beta_1 Deres_{it-1} + \rho_0 MRPK_{it-1} \\ + \rho_1 MRPK_{it-1} * Deres_{it-1} + \beta_2 \ln age_{it} + \varepsilon_{irt} \end{aligned} \tag{30}$$

where $\gamma_i$ and $\nu_t$ are plant and year fixed effects respectively, and $Deres_{it-1}$ indicates if a first product of plant $i$ has been dereserved in period $t-1$ or earlier. I estimate equation (30) both on a sample with only incumbents, as well as on the full sample of plants. In the latter setup, I can control for economic shocks at the region-sector-year level, which is impossible when restricting the sample to incumbents, due to collinearity issues.[27]

The main coefficient of interest is $\rho_1$, which estimates how the speed of convergence to $MRPK^*_{it}$ changes after dereservation. To understand the estimation strategy, consider the case when $\rho_0 = \rho_1 = 0$. In that case, plants exhibit immediate convergence to $MRPK^*_{it} \equiv E[MRPK_{it}|(\rho_0 = \rho_1 = 0)]$, regardless of $MRPK_{it-1}$. In practice however, the average plant experiences a delayed ad-

[27]Collinearity issues arise from small numbers of incumbent plants in many region-sector-year observations. For the estimation on the full sample, I distinguish between three types of plants. A first type is the incumbent plant, defined above. A second type is the "entrant" plant, which after dereservation starts producing a previously reserved product. The third type of plant - labeled as "outsider" - includes all remaining plants. For this full sample of plants, I employ the following estimation specification:

$$\begin{aligned} MRPK_{irst} = \gamma_{irs} + \nu_{rst} + \beta_1 Deres_{irst-1} + \beta_2 Deres_{irst-1} * entrant_{irs} \\ + \rho_0 MRPK_{irst-1} + \rho_2 MRPK_{irst-1} * entrant_{irs} + \rho_1 MRPK_{irst-1} * outsider_{irs} \\ + \rho_3 MRPK_{irst-1} * Deres_{irst-1} + \rho_4 MRPK_{irst-1} * Deres_{irst-1} * entrant_{irs} \\ + \beta_3 X_{irst} + \varepsilon_{irst} \end{aligned} \tag{31}$$

Here, $\gamma_{irs}$ is a plant fixed-effect, $entrant_{irs}$ and $outsider_{irs}$ are indicators for plant $i$ being entrants or outsiders. Next, $\nu_{rst}$ is a region-sector-year fixed effect that absorbs local economic shocks. While lengthy, the above specification is still intuitive. The top row is a standard difference-in-difference framework, where I allow for different $MRPK^*_{it}$ levels post dereservation for incumbents and entrants. The middle row estimates convergence speeds prior to dereservation, allowing for different speeds of convergence for incumbents, entrants and outsiders. The third row then estimates how speeds of convergence change after dereservation, where $\rho_3$ - the coefficient of interest - estimates how speed of converges changes for incumbent firms.

Table 1: Dereservation and MRPK convergence

| | $MRPK_{it}$ - Gross Revenue (GR) | | $MRPK_{it}$ - Value Added (VA) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $Deres_{it-1}$ | -0.118** | -0.061* | -0.165** | -0.029 |
| | (0.023) | (0.027) | (0.018) | (0.021) |
| $MRPK_{it-1}(GR)$ | 0.439** | 0.369** | | |
| | (0.016) | (0.020) | | |
| $MRPK_{it-1}(GR) * Deres_{it-1}$ | 0.031** | 0.034* | | |
| | (0.011) | (0.014) | | |
| $MRPK_{it-1}(VA)$ | | | 0.306** | 0.244** |
| | | | (0.012) | (0.020) |
| $MRPK_{it-1}(VA) * Deres_{it-1}$ | | | 0.032** | 0.032+ |
| | | | (0.011) | (0.017) |
| Plant Fixed Effects | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | – | Yes | – |
| State-sector-year Fixed Effects | No | Yes | No | Yes |
| Observations | 62482 | 186688 | 56482 | 168314 |

In specifications 1 and 2, MRPK is measured based on gross revenue, and based on value added in specifications 3 and 4. Due to data limitations described in section 4.2, specifications 1 and 3 estimate equation (30) on a sample restricted to a sample of incumbent plants dereserved before 2010, while specifications 2 and 4 estimate equation (31) on the full sample for all sample years before 2010. All specifications control for the logarithm of a plant's age. Standard errors, in parentheses, are clustered at the product level, which is the level at which dereservation status is defined. $^{+}p < 0.1$;$^{*}\ p < 0.05$;$^{**}\ p < 0.01$.

justment to $MRPK^*_{it}$, with $\rho_0 > 0$. The closer $\rho_0$ is to unity, the slower convergence to $MRPK^*_{it}$. Crucially, when $\rho_0 > 0$, then $\rho_1 > 0$ indicates that the speed of MRPK convergence slows down after dereservation. In equation (30), I proxy for $MRPK^*_{it}$ with $\gamma_i + \nu_t + \beta_1 Deres_{it-1} + \beta_2 \ln age_{it}$, though the findings on $\rho_1$ are robust to the exact choice of proxy. Still, it is important to include the indicator variable for dereservation status. After all, dereservation implies the removal of the size restriction on capital, which may have implications for the optimal capital share. Including the $Deres_{it-1}$ indicator, entails that $MRPK^*_{it}$ can adjust accordingly after the dereservation reform.

The empirical results are in line with the theoretical predictions of the model (see Table 1). First, I find that there is indeed convergence to $MRPK^*_{it}$, since $\rho_0$ is significantly below 1, but this convergence is not immediate as $\rho_0$ is also significantly above 0. This is consistent with the speed of convergence being limited by the presence of financial constraints. The point estimates for $\rho_0$ are below 0.5, which implies that convergence to $MRPK^*_{it}$ is relatively fast. Hence, the proxy for $MRPK^*_{it}$ appears to be empirically valid. Most importantly, all coefficients on the interaction of dereservation with $MRPK_{it-1}$ are positive - as predicted by the theory - and statistically significant. The estimated magnitude of the effect of dereservation is modest but economically meaningful. For specifications 1 and 2 specifically, dereservation increases the half-life of the autoregressive process by 9% and 9.7% respectively.[28] [29]

One concern with my estimation procedure arises from the downward bias on autoregression

[28]The following formula, which is derived from the AR(1) convergence process, computes the percentage increase in the half-life: $\frac{\log(0.5)/\log(\rho_0+\rho_1)}{\log(0.5)/\log(\rho_0)}$.

[29]Since product coverage in the ASI data is incomplete for the years 1998 and 1999, I perform a robustness check in Appendix Table D.2, where I only focus on incumbents dereserved after 1999. The results continue to be in line with the theoretical predictions, but the point estimates are slightly lower.

coefficients described by Nickell (1981). While in principle it may be possible to obtain consistent estimates using GMM methods, these econometric approaches come with their own pitfalls (Roodman, 2009). Note however that my primary objective here is to provide qualitative support for the model's predictions. First then, observe that my panel spans the period 1990-2011, so the downward bias, which is of order $1/T$, becomes small. Second and most importantly, the downward bias on $\rho_1$ works against finding evidence for dereservation slowing down MRPK convergence. Still, I obtain strongly significant support for my theoretical predictions. Third, in Section 6, I also present evidence for my theoretical predictions based on capital growth from young plants, which does not require the use of an autoregression framework.

As mentioned, my autoregression framework is inspired by the analysis on capital convergence in Asker et al. (2014). They employ this specification to show that heterogeneity in MRPK can be driven by delayed adjustment to productivity shocks. Hence, the purpose of their empirical analysis is very closely related to the one in this paper. The main difference between their setup and mine, is that they assume that the delayed adjustment to productivity shocks is driven by capital adjustment costs. However, as implied by my model, such delayed adjustment can also be caused by financial constraints. In my empirical setting, financial constraints are a much more likely driver of the results than adjustment costs. After all, I have provided evidence that plants' markups indeed fall after dereservation, which directly affects their retained earnings. Moreover, MRPK convergence slows down after dereservation. From the perspective of adjustment costs, this would require increases in adjustment costs to coincide with the timing of dereservation. Finally, in the next section I extend the analysis of MRPK convergence to the full panel of plants employing a different measure for competition. It is again unlikely that this measure for competition is correlated with the severity of adjustment costs.

# 5 Competition and MRPK convergence in the full panel

In this section, I strengthen the external validity of my results by extending the analysis to the full panel of plants, instead of focusing only on incumbent plants whose products become dereserved.[30] In this analysis, I set aside the model's predictions on markup misallocation and focus on MRPK convergence for two reasons. First, the prediction on capital convergence is my model's most novel one, while the predictions on markup levels and dispersion have been examined in previous research (e.g. Peters (2016); Schaumans and Verboven (2015)). Second, I will use the median markup in a market as my inverse measure of competition, and this choice does not allow me to examine markup misallocation. My competition measure is consistent with the model however, since the model features a monotonic relationship between the number of firms, which governs the degree of competition, and the first moments of the markup distribution. From the point of view of the model, an alternative measure of competition could have been the number of firms. Note however that what matters for the degree of competition is not only the number of firms, but also market size, which depends on sectoral expenditure shares and income per capita, among others. The median markup incorporates these factors directly.

As my inverse measure of competition, I use the median markup at the region-sector-year level, $Median_{rst}[\ln \mu_{irst}]$, which is arguably exogenous from the plant's point of view. The indices $r$ and $s$ stand for a region and a 3-digit sector respectively. The markup is measured as

[30] Appendix Table D.1 shows that these two groups of plants are relatively similar, except for the revenue to capital ratio, which is more than twice as high in the full panel.

$$\mu_{irst} = \alpha_s^L \frac{S_{irst}}{w_{irst} L_{irst}}$$

This markup measure is identical to equation (27), except that the elasticity $\alpha_s^L$ is now measured as a cost share at the sector level. Because the median markup will only enter in interactions terms in the specification below, I need to demean $Median_{rst}[\ln \mu_{irst}]$. To avoid results being driven by the measurement of $\alpha_s^L$, demeaning happens within sectors. Hence, I am leveraging within-sector variation in the median markup, which is insensitive to $\alpha_s^L$. Next, to ensure that the median markup is plausibly exogenous to the individual plant, I restrict the sample to cases where at least 7 plants are observed in a given region-sector-year. Finally, I normalize the median markup to standard deviation units.

To implement the empirical test on MRPK convergence, I update the autoregression framework from specification (30) in the following way:

$$\begin{aligned} MRPK_{irst} = \alpha_{irs} + \gamma_{rst} + \beta \ln age_{irst} + \rho_0 MRPK_{irst-1} \\ + \rho_1 MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}] + \varepsilon_{irst} \end{aligned} \tag{32}$$

As before, the main coefficient of interest is $\rho_1$. This coefficient estimates how the speed of convergence changes as a function of $Median_{rst-1}[\ln \mu_{irst-1}]$. Recall that when $\rho_0 = \rho_1 = 0$, plants exhibit immediate convergence to the empirical proxy for $MRPK^*_{irst}$, regardless of $MRPK_{irst-1}$. In practice, plants experience delayed adjustment to $MRPK^*_{irst}$. The theoretical prediction is then that $\rho_1 < 0$, as this implies that the speed of MRPK convergence increases with $Median_{rst-1}[\ln \mu_{irst-1}]$. I cluster standard errors at the sector level in the estimation.

## 5.1 Heterogeneity along financial dependence

The above tests on MRPK convergence have all implicitly assumed that the average plant in the sample is financially constrained. However, there is empirical heterogeneity in the degree to which plants are financially constrained, which I can leverage to further corroborate the mechanism driving the link between competition and MRPK convergence. The motivating idea is that in sectors with higher levels of financial dependence, measured as $Fin\ Dep_s$, changes in the level of sector-level competition have a stronger impact on the rate of MRPK convergence.

I employ the standard Rajan and Zingales (1998) measure of sectoral financial dependence:

$$Fin\ Dep_s = \frac{Capital\ Expenditures_s - Cash\ Flow_s}{Capital\ Expenditures_s},$$

based on data for US sectors over the 1980's.[31] Here, $Fin\ Dep_s$ captures the share of external finance in a firm's investments in a setting with highly developed financial markets, namely the US. The central idea in Rajan and Zingales (1998) is then that in economies with less developed financial markets, such as India, financial constraints become especially binding in sectors with high levels of $Fin\ Dep_s$.

[31] I use the original Rajan and Zingales (1998) measures of financial dependence for ISIC Rev.2 sector definitions, except that I trim the financial dependence measure such that $Fin\ Dep_s \geq 0$. This ensures a clean identification of the effect of competition in the triple interaction term in specification (33). The ISIC Rev.2 sector definitions match closely with India's NIC 1987 sector definitions. The concordance between ISIC Rev.2 and NIC 1987 is provided by the Indian Statistical Office.

To examine the role of financial dependence in the setting of MRPK convergence, I augment the earlier specification to allow for heterogeneous effects along financial dependence:

$$
\begin{aligned}
MRPK_{irst} &= \alpha_{irs} + \gamma_{rst} + \beta \ln age_{irst} + \rho_0 MRPK_{irst-1} \\
&\quad + \rho_1 MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}] + \rho_2 MRPK_{irst-1} * Fin\ Dep_s \\
&\quad + \rho_3 MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}] * Fin\ Dep_s + \varepsilon_{irst}
\end{aligned} \tag{33}
$$

For this specification, the expectation is that $\rho_3 < 0$, as a decrease in competition would speed up convergence more for plants in sectors with higher levels of financial dependence.

## 5.2 Estimation results

The results for MRPK convergence in the full panel (see Table 2) confirm the results from the analysis of dereservation. First, across all specifications, MRPK converges strongly to the empirical proxy for $MRPK^*_{irst}$ ($\rho < 1$), but this convergence is not immediate ($\rho > 0$). Second, for baseline specification (32), the speed of convergence always increases with the median markup. Specifically, the coefficient on $\rho_1$ is always negative and strongly statistically significant in three of the four specifications (see columns 2, 5, and 6). This confirms the qualitative prediction of the model that the speed of convergence slows down with competition. The magnitude of this effect is modest but economically meaningful, as in the case of dereservation. As an example, in specifications 2 and 6, an increase in the median markup by two standard deviations, decreases the half-life of MRPK convergence by 4.8% and 13.4% respectively.[32]

The results for heterogeneity along financial dependence are also in line with expectations (see columns 3,4,7,8). First note that the coefficient on $MRPK_{irst-1} * Fin\ Dep_s$ is always positive, which is consistent with MRPK convergence being slower in more financially dependent sectors. More importantly, the coefficient $\rho_3$, estimated on the triple interaction term, is always significantly negative. These results imply that the median markup speeds up MRPK convergence more in sectors with higher financial dependence. Consider for instance the industry producing electric machinery, which has a relatively high level of financial dependence at $Fin\ Dep_s = 0.77$. For this sector, an increase in the median markup by two standard deviations, decreases the half-life of MRPK convergence by 12.5% or 21.8%, according to specifications 4 and 8 respectively.

# 6 Additional evidence from undercapitalized young plants

So far, I focused on MRPK convergence to test the effect of competition on capital convergence. The advantage of examining MRPK convergence is that any plant optimally converges to $MRPK^*_{it}$. Hence, the tests on MRPK convergence are valid in general. Still, the empirical measurement of MRPK convergence is based on the assumption of Cobb-Douglas production functions, and requires the use of an autoregression framework. I now complement the evidence from MRPK convergence with more transparent reduced-form evidence on competition's effect on capital growth for young plants. The data suggests that these young plants are undercapitalized, as a

[32]Given the presence of the Nickell bias in this autocorrelation framework, it is also noteworthy that the predicted sign for $\rho_1$ switches from the dereservation setting to the full panel. Despite the potential downward bias on this coefficient, the estimation results pick up this sign switch.

Table 2: Competition and Speed of MRPK Convergence

| | $MRPK_{irst}$ (Gross Revenue (GR)) | | | | $MRPK_{irst}$ (Value added (VA)) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $MRPK_{irst-1}(GR)$ | 0.392** | 0.372** | 0.394** | 0.367** | | | | |
| | (0.013) | (0.014) | (0.015) | (0.019) | | | | |
| $MRPK_{irst-1}$ (GR) * $Median_{rst-1}[\ln \mu_{irst-1}]$ | -0.000 | -0.009* | 0.014** | -0.000 | | | | |
| | (0.003) | (0.004) | (0.005) | (0.005) | | | | |
| $MRPK_{irst-1}$ (GR) * $Fin\ Dep_s$ | | | 0.005 | 0.012 | | | | |
| | | | (0.013) | (0.017) | | | | |
| $MRPK_{irst-1}$ (GR)* $Median_{rst}[\ln(\mu_{irst}] * Fin\ Dep_s$ | | | -0.028** | -0.032** | | | | |
| | | | (0.008) | (0.009) | | | | |
| $MRPK_{irst-1}(VA)$ | | | | | 0.248** | 0.219** | 0.225** | 0.185** |
| | | | | | (0.016) | (0.016) | (0.023) | (0.023) |
| $MRPK_{irst-1}$ (VA) * $Median_{rst-1}[\ln \mu_{irst-1}]$ | | | | | -0.015** | -0.023** | 0.014+ | -0.006 |
| | | | | | (0.005) | (0.007) | (0.008) | (0.010) |
| $MRPK_{irst-1}$ (VA) * $Fin\ Dep_s$ | | | | | | | 0.049* | 0.067** |
| | | | | | | | (0.023) | (0.023) |
| $MRPK_{irst-1}$ (VA) * $Median_{rst}[\ln(\mu_{irst})] * Fin\ Dep_s$ | | | | | | | -0.054** | -0.043* |
| | | | | | | | (0.018) | (0.018) |
| Plant Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year Fixed Effects | Yes | – | Yes | – | Yes | – | Yes | – |
| State-sector-year Fixed Effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 217942 | 183052 | 177192 | 144807 | 188261 | 155925 | 153012 | 122963 |

In specifications 1 - 4, MRPK is measured based on gross revenue, and based on value added in specifications 5-8. All specifications control for the logarithm of a plant's age. The variable $Median_{rst-1}[\ln(\mu_{irst-1}]$ is demeaned within 3-digit sectors and measured in standard deviation units. To ensure that the median markup is plausibly exogenous to the individual plant, the sample is restricted to region-sector-years consisting of at least 7 plants. Standard errors, in parentheses, are clustered at the level of 3-digit sectors. $^{+}p < 0.1;^{*} p < 0.05;^{**} p < 0.01$.

stylized fact in the existing literature is that young plants exhibit higher capital growth rates than older plants (see e.g. Evans (1987); Geurts and Van Biesebroeck (2016); Haltiwanger, Jarmin, and Miranda (2013)). As I show below, this stylized fact is also present in Indian manufacturing.

To motivate the empirical analysis of capital growth theoretically, I present a model in Appendix E where new, undercapitalized firms are born each period. Intuitively, when firms are born with suboptimally low levels of capital, then firms' optimizing behavior implies that firms grow their capital to its optimal level while they are young and financially constrained. In this setting, increased competition also reduces optimal markups and thereby slows down internally financed capital growth for young firms. This is the prediction I take to the data.

I measure capital growth as:

$$g(k_{irst}) = k_{irst+1} - k_{irst}.$$

Here, capital is measured as the book value of assets, which is observed both at the start of year $t$, and at the end. The latter value is used as measure for $k_{irst+1}$. The estimation will therefore not require observing plants in the previous years, which increases the sample size compared to the MRPK convergence analysis. I deflate the book value of capital using the capital deflator from the Indian Handbook of Industrial Statistics.

**Dereservation reform** To examine the effect of dereservation on capital growth for young plants, I run the following difference-in-difference specification on a sample of incumbent plants:

$$\begin{aligned} g(k_{irst}) = \alpha_{irs} + \gamma_{rst} + \beta_1 young_{irst} + \beta_2 Deres_{irst-1} \\ + \beta_3 Deres_{irst-1} * young_{irst} + \beta_4 \ln age_{irst} + \varepsilon_{irst}, \end{aligned} \tag{34}$$

where $\alpha_{irs}$ is a plant fixed effect, and $\gamma_{rst}$ a state-sector-year fixed effect. I consider two different measures for $young_{irst}$, namely $[-\ln(age_{irst})]$ and the indicator variable $1(age_{irst} \leq 5)$. The prediction is that the increase in competition due to dereservation leads to slower capital growth for young plants, namely $\beta_3 < 0$. I estimate specification (34) both with and without the plant-level fixed effects, as it is ambiguous what the optimal approach is. Plant fixed effects are the best way to control for unobserved plant-level characteristics. However, the theory predicts that capital growth ends once a plant reaches its optimal capital level. Hence, capital growth should not have a stable trend for a plant.[33]

Table 3 demonstrates that dereservation has a negative impact on the capital growth for young plants, in particular for plants younger than five years old. For instance, specification 3 estimates that for these plants, dereservation leads to a reduction in the capital growth rate of 0.021 log points (p=0.009).[34]

[33]Given that capital growth is observed within a given year, this specification allows estimation on a larger sample than the MRPK analysis for the dereservation reform. Collinearity issues will prove not to be an issue, which is why I present results only for the sample of incumbents. Results for the full sample are comparable to those for the sample of incumbents.

[34]To address the incomplete product coverage for the years 1998 and 1999, Appendix Table D.3 provides a robustness check for incumbent plants dereserved after 1999. Results are very similar, but statistical significance drops slightly for specifications 3 and 4.

Table 3: Dereservation and capital growth for young plants

| | Capital growth $g(k)_{it}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $Deres_{it-1}$ | 0.005 | -0.011 | 0.005 | 0.001 |
| | (0.010) | (0.019) | (0.006) | (0.008) |
| $Deres_{it} * [-\ln(age_{it})]$ | -0.001 | -0.003 | | |
| | (0.003) | (0.006) | | |
| $Deres_{it-1} * 1(age_{it} \leq 5)$ | | | -0.021$^{**}$ | -0.040$^{+}$ |
| | | | (0.008) | (0.023) |
| $[-\ln(age_{it})]$ | -0.000 | 0.008 | | |
| | (0.005) | (0.006) | | |
| $1(age_{it} \leq 5)$ | | | 0.018 | 0.012 |
| | | | (0.011) | (0.010) |
| State-sector-year Fixed Effects | Yes | Yes | Yes | Yes |
| Plant Fixed Effects | No | Yes | No | Yes |
| Observations | 135894 | 119817 | 137853 | 121731 |

Due to data limitations described in section 4.2, I restrict the sample to incumbent plants who were dereserved before 2010. Standard errors, in parentheses, are clustered at the product level, which is the level at which dereservation status is defined. $^{+}p < 0.1;^{*} p < 0.05;^{**} p < 0.01$.

**Full panel** I also examine the impact of the median markup on young plants' capital growth in the full panel. To this end, I update the regression analysis to the following specification, and predict that $\beta_2 > 0$.

$$g(k_{irst}) = \alpha_{irs} + \gamma_{rst} + \beta_1 young_{irst} + \beta_2 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} + \varepsilon_{irst} \tag{35}$$

Finally, I also examine the heterogeneous impact of competition across sectors with different levels of financial dependence. The prediction is that the impact of competition on capital growth for young plants is increasing with the degree of financial dependence ($\beta_3 > 0$).

$$\begin{aligned} g(k_{irst}) =& \alpha_{rst} + \beta_1 young_{irst} + \beta_2 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} \\ &+ \beta_3 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} * Fin\ Dep_s + \varepsilon_{irst} \end{aligned} \tag{36}$$

The estimation results are again generally in line with the theoretical predictions (see Table 4). Capital growth for young plants typically increases with the median markup, and this effect is stronger in sectors with higher levels of financial dependence. These results are stronger for specifications that omit the plant fixed effect (columns 1,3,5,7). As mentioned above, it is ambiguous if including such a fixed effect is the optimal estimation strategy, since capital growth is not constant across a plant's lifetime.

Table 4: Competition and capital growth of young plants

| | Capital growth $g(k)_{irst}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Median_{rst}[\ln(\mu_{irst})] * [-\ln(age_{irst})]$ | 0.003** | 0.002 | | | 0.001 | 0.001 | | |
| | (0.001) | (0.002) | | | (0.002) | (0.002) | | |
| $Median_{rst}[\ln(\mu_{irst-1})] * 1(age_{irst} \leq 5)$ | | | 0.009** | -0.002 | | | 0.006 | -0.000 |
| | | | (0.003) | (0.003) | | | (0.004) | (0.005) |
| $Median_{rst}[\ln(\mu_{irst})] * [-\ln(age_{irst})] * Fin\ Dep_s$ | | | | | 0.009* | 0.006 | | |
| | | | | | (0.004) | (0.005) | | |
| $Median_{rst}[\ln(\mu_{irst})] * 1(age_{irst} \leq 5) * Fin\ Dep_s$ | | | | | | | 0.017* | -0.000 |
| | | | | | | | (0.008) | (0.008) |
| $-\ln(age_{irst})$ | 0.002 | 0.012** | | | -0.003 | 0.008+ | | |
| | (0.002) | (0.003) | | | (0.004) | (0.004) | | |
| $1(age_{irst} \leq 5)$ | | | 0.029** | 0.012** | | | 0.019* | 0.005 |
| | | | (0.005) | (0.003) | | | (0.008) | (0.004) |
| $-\ln(age_{irst}) * Fin\ Dep_s$ | | | | | 0.016* | 0.013* | | |
| | | | | | (0.007) | (0.005) | | |
| $1(age_{irst} \leq 5) * Fin\ Dep_s$ | | | | | | | 0.031** | 0.015* |
| | | | | | | | (0.012) | (0.006) |
| State-sector-year Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Plant Fixed Effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 607273 | 485880 | 620387 | 496878 | 515139 | 405881 | 526675 | 415494 |

The variable $Median_{rst-1}[\ln(\mu_{irst-1}]$ is demeaned by sector, and then measured in standard deviation units. Columns 1-4 display estimation results for specification (35), while columns 5-8 show results for specification (36). To ensure that the median markup is plausibly exogenous to the individual plant, the sample is restricted to region-sector-years consisting of at least 7 plants. Standard errors, in parentheses, are clustered at the level of 3-digit sectors. $^{+}p < 0.1;^{*} p < 0.05;^{**} p < 0.01$.

# 7 Conclusion

Misallocation of resources is a pervasive challenge throughout the developing world. While the exact contribution of misallocation to cross-country differences in aggregate productivity is debatable, it seems uncontroversial that shifting resources to firms with higher marginal products would be particularly beneficial for developing economies. When quantifying the benefits of reducing misallocation, a typical approach is to focus on one particular friction, for instance market power, financial constraints, or a particular government intervention, and ask how the removal of this friction could contribute to aggregate productivity. This paper shows that this type of approach can lead to misleading conclusions.

I have examined how the interplay of competition with financial constraints affects misallocation, and found that increased competition reduces markup misallocation, but amplifies capital misallocation. These countervailing forces make the impact of intensified competition on misallocation ambiguous. I have also documented a range of empirical results for India's manufacturing sector, all indicating that increased competition indeed slows down firms' convergence to their optimal capital level. Hence, the empirical relevance of competition's negative impact on capital convergence is robustly demonstrated.

Deriving precise policy implications is outside the scope of this paper. Nevertheless, the analysis suggests that reaping the full gains of pro-competitive reforms depends very much on the precise implementation of these reforms. More precisely, it is essential that reforms do not stifle the growth of high-potential firms. In thinking about sequencing of reforms, the take-away could be to first optimize financial access for firms, before enhancing competition in the real economy.

# References


Acemoglu, D. and M. K. Jensen (2013). Aggregate comparative statics. *Games and Economic Behavior 81*, 27–49.

Aghion, P., U. Akcigit, and P. Howitt (2014). What do we learn from Schumpeterian growth theory? In *Handbook of Economic Growth*, Volume 2, pp. 515–563. Elsevier.

Aghion, P., R. Burgess, S. J. Redding, and F. Zilibotti (2008). The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India. *American Economic Review 98*(4), 1397–1412.

Akcigit, U., H. Alp, and M. Peters (2016). Lack of selection and limits to delegation: Firm dynamics in developing countries. *NBER working paper*.

Alfaro, L. and A. Chari (2014). Deregulation, Misallocation, and Size: Evidence from India. *The Journal of Law and Economics 57*(4), 897–936.

Allcott, H., A. Collard-Wexler, and S. D. O'Connell (2016). How do electricity shortages affect industry? Evidence from India. *American Economic Review 106*(3), 587–624.

Amiti, M., O. Itskhoki, and J. Konings (2016). International shocks and domestic prices: how large are strategic complementarities? *NBER working paper*.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodríguez-Clare (2017). The elusive pro-competitive effects of trade. *The Review of Economic Studies (forthcoming)*.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014). Dynamic inputs and resource (mis) allocation. *Journal of Political Economy 122*(5), 1013–1063.

Asturias, J., M. García-Santana, and R. Ramos (2018). Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India. *Journal of the European Economic Association (forthcoming)*.

Atkeson, A. and A. Burstein (2008). Pricing-to-market, trade costs, and international relative prices. *The American Economic Review 98*(5), 1998–2031.

Balasundharam, V. (2018). Bottlenecks versus ripple effects: The role of linkages in india's product market liberalization. *working paper*.

Banerjee, A., S. Cole, and E. Duflo (2005). Bank Financing in India. In *India's and China's recent experience with reform and growth*, pp. 138–157. Springer.

Banerjee, A. V. and E. Duflo (2014). Do firms want to borrow more? Testing credit constraints using a directed lending program. *The Review of Economic Studies 81*(2), 572–607.

Bils, M., P. J. Klenow, and C. Ruane (2017). Misallocation or mismeasurement? *Working Paper*.

Blanchard, O. J. and N. Kiyotaki (1987). Monopolistic competition and the effects of aggregate demand. *The American Economic Review*, 647–666.

Boehm, J. and E. Oberfield (2018). Misallocation in the market for inputs: Enforcement and the organization of production. *NBER working paper*.

Bollard, A., P. J. Klenow, and G. Sharma (2013). India's Mysterious Manufacturing Miracle. *Review of Economic Dynamics 16*(1), 59–85.

Brooks, W. J., J. P. Kaboski, and Y. A. Li (2016). Growth policy, agglomeration, and (the lack of) competition. *NBER working paper*.

Buera, F. J., J. P. Kaboski, and Y. Shin (2015). Entrepreneurship and financial frictions: A macrodevelopment perspective. *Annual Review of Economics 7*(1), 409–436.

Chari, A. (2011). Identifying the aggregate productivity effects of entry and size restrictions: An empirical analysis of license reform in India. *American Economic Journal: Economic Policy 3*(2), 66–96.

Corchon, L. C. (1994). Comparative statics for aggregative games the strong concavity case. *Mathematical Social Sciences 28*(3), 151–165.

De Loecker, J. and J. Eeckhout (2018). Global market power. Technical report, National Bureau of Economic Research.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, markups, and trade reform. *Econometrica 84*(2), 445–510.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *The American Economic Review 102*(6), 2437–2471.

Dhingra, S. and J. Morrow (2019). Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy 127*(1), 000–000.

Doraszelski, U. and A. Pakes (2007). A framework for applied dynamic analysis in io. *Handbook of industrial organization 3*, 1887–1966.

Edmond, C., V. Midrigan, and D. Y. Xu (2015). Competition, markups, and the gains from international trade. *American Economic Review 105*(10), 3183–3221.

Edmond, C., V. Midrigan, and D. Y. Xu (2018). How costly are markups? Technical report, National Bureau of Economic Research.

Epifani, P. and G. Gancia (2011). Trade, markup heterogeneity and misallocations. *Journal of International Economics 83*(1), 1–13.

Evans, D. S. (1987). Tests of alternative theories of firm growth. *The Journal of Political Economy*, 657–674.

Foellmi, R. and M. Oechslin (2016). Harmful pro-competitive effects of trade in presence of credit market frictions. *manuscript, University of St. Gallen*.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *The Review of Economic Studies 38*(1), 1–12.

García-Santana, M. and J. Pijoan-Mas (2014). The reservation laws in India and the misallocation of production factors. *Journal of Monetary Economics 66*, 193–209.

Geurts, K. and J. Van Biesebroeck (2016). Firm creation and post-entry dynamics of de novo entrants. *International Journal of Industrial Organization 49*, 59–104.

Gilbert, R. (2006). Looking for mr. schumpeter: Where are we in the competition-innovation debate? In *Innovation Policy and the Economy, Volume 6*, pp. 159–215. The MIT Press.

Guner, N., G. Ventura, and Y. Xu (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics 11*(4), 721–744.

Hall, R. (1986). Market structure and macroeconomic fluctuations. *Brookings papers on economic activity 1986*(2), 285–338.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics 95*(2), 347–361.

Harrison, A. and A. Rodríguez-Clare (2010). Trade, foreign investment, and industrial policy for developing countries. In *Handbook of development economics*, Volume 5, pp. 4039–4214. Elsevier.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics 131*(3), 1291–1364.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics 124*(4).

Itskhoki, O. and B. Moll (2019). Optimal development policies with financial frictions. *Econometrica 87*(1), 139–173.

Jungherr, J. and D. Strauss (2017). A Blessing in Disguise? Market Power and Growth with Financial Frictions. *working paper*.

Krusell, P., T. Mukoyama, and A. A. Smith Jr (2011). Asset prices in a huggett economy. *Journal of Economic Theory 146*(3), 812–844.

Levine, R. (2005). Finance and growth: theory and evidence. *Handbook of economic growth 1*, 865–934.

Mankiw, N. G. and M. D. Whinston (1986). Free entry and social inefficiency. *The RAND Journal of Economics*, 48–58.

Martin, L. A., S. Nataraj, and A. E. Harrison (2014). In with the Big, Out with the Small: Removing Small-Scale Reservations in India. *NBER Working Paper Series*.

Martin, L. A., S. Nataraj, and A. E. Harrison (2017). In with the big, out with the small: Removing small-scale reservations in India. *American Economic Review 107*(2), 354–86.

Midrigan, V. and D. Y. Xu (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review 104*(2), 422–458.

Mohan, R. (2002). Small-scale industry policy in india. *Economic policy reforms and the Indian economy 213*.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *The American Economic Review 104*(10), 3186–3221.

Mrázová, M., J. P. Neary, and M. Parenti (2017). Sales and markup dispersion: Theory and empirics. *working paper*.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 1417–1426.

Peters, M. (2016). Heterogeneous Mark-ups, Growth and Endogenous Misallocation. *mimeo, Yale University*.

Rajan, R. and L. Zingales (1998). Financial dependence and growth. *The American Economic Review 88*(3), 559–586.

Ravn, M. O. and V. Sterk (2016). Macroeconomic Fluctuations with HANK & SAM: An Analytical Approach. *working paper*.

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics 11*(4), 707–720.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics 71*(1), 135–158.

Schaumans, C. and F. Verboven (2015). Entry and competition in differentiated products markets. *Review of Economics and Statistics 97*(1), 195–209.

Sivadasan, J. (2009). Barriers to competition and productivity: Evidence from India. *The BE Journal of Economic Analysis & Policy 9*(1).

Tewari, I. and J. Wilde (2017). Product Scope and Productivity: Evidence from India's Product Reservation Policy. *working paper, University of South Florida*.

Werning, I. (2015). Incomplete markets and aggregate demand. *NBER Working Paper*.

# Appendix A Proof for Lemma 3

## A.1 Proof for Equation (23)

In this proof by contradiction, I will demonstrate that Equation (23) holds, namely $(m_{sL} > m'_{sL}) \implies (m_{sH} > m'_{sH})$. Suppose to the contrary[35] that $(m_{sL} > m'_{sL}) \wedge (m_{sH} \leq m'_{sH})$. First defining the output ratio:

$$\mathcal{G}_{sH} \equiv \left(\frac{y_{sH}}{y_{sL}}\right)^{\frac{\sigma-1}{\sigma}},$$

which from Equation (6) implies that $m_{sH} = \mathcal{G}_{sH} m_{sL}$. Hence the supposition implies that the output ratio shrinks:

$$(m_{sL} \geq m'_{sL}) \wedge (\mathcal{G}_{sH} < \mathcal{G}'_{sH}).$$

Equation (25) then entails that $(\mathcal{G}_{sH} < \mathcal{G}'_{sH}) \implies \left(\frac{\mu'_{sL}}{\mu'_{sH}} > \frac{\mu_{sL}}{\mu_{sH}}\right)$. Or alternatively, since all markups are positive:

$$(\mathcal{G}_{sH} < \mathcal{G}'_{sH}) \implies \left(\frac{\mu_{sH}}{\mu'_{sH}} > \frac{\mu_{sL}}{\mu'_{sL}}\right).$$

However, note that $(m_{sL} > m'_{sL}) \implies \left(\frac{\mu_{sL}}{\mu'_{sL}} > 1\right)$. After also observing that $\left(\frac{\mu_{sH}}{\mu'_{sH}} > 1\right) \iff (m_{sH} > m'_{sH})$, the combination of the previous results yields:

$$((m_{sL} > m'_{sL}) \wedge (\mathcal{G}_{sH} < \mathcal{G}'_{sH})) \implies (m_{sH} > m'_{sH}).$$

This entails a contradiction with the supposition. Hence, its opposite must be true, which proves the statement in Equation (23).

## A.2 Proof for Equation (24)

I start by demonstrating the first component of the implication in Equation (24), namely:

$$(m_{sL} > m'_{sL}) \implies \forall \tau > 0 : (G_{s\tau} > G'_{s\tau}) \tag{37}$$

To show this, I start from the expression for capital growth, $k_{s\tau+1} = (1-\delta)k_{s\tau} + \frac{\pi_{s\tau}}{PF}$, which implies:

[35] Recall that $\neg(p \implies q) \iff (p \wedge \neg q)$.

$$G_{s\tau+1} \equiv \frac{k_{s\tau+1}}{k_{sL}} = (1-\delta)G_{s\tau} + \frac{\pi_{s\tau}}{P^F k_{sL}} \tag{38}$$

I then demonstrate Equation (37) by induction. In a first step, I show that $(m_{sL} > m'_{sL}) \implies (G_{s1} > G'_{s1})$, and afterwards I demonstrate the inductive step that $(m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}) \implies (G_{s\tau+1} > G'_{s\tau+1})$. The first step and the inductive step together imply that Equation (37) holds.

**Step 1** First, notice that revenue net of labor costs of the unconstrained low-productivity firm is $\pi_{sL}/P^F = (\mu_{sL} - 1)y_{sL}MC_{sL} + r^k k_{sL}$, where the first term are profits, and the second term is "payments to capital." Hence

$$\frac{\pi_{sL}}{p^F k_{sL}} = (\mu_{sL} - 1)MC_{sL}\frac{y_{sL}}{k_{sL}} + r^k.$$

Recall that $(m_{sL} > m'_{sL}) \implies (\mu_{sL} > \mu'_{sL})$, while $MC_{sL} = MC'_{sL}$ (see Equation (19)). In addition, Equations (17) and (18) imply that the output to capital ratio is constant at $\frac{y_{sL}}{k_{sL}} = \frac{P^F}{w} r^k \frac{(1-\alpha)}{\alpha}$, which implies $\frac{y_{sL}}{k_{sL}} = \frac{y'_{sL}}{k'_{sL}}$. Taken together, $(m_{sL} > m'_{sL}) \implies \left(\frac{\pi_{sL}}{P^F y_{sL}} > \frac{\pi'_{sL}}{P^F y'_{sL}}\right)$. Since $G_{s1} \equiv \frac{k_{s1}}{k_{sL}} = (1-\delta) + \frac{\pi_{sL}}{P^F y_{sL}}$, I obtain that

$$(m_{sL} > m'_{sL}) \implies (G_{s1} > G'_{s1}).$$

**Inductive Step** I now demonstrate that the inductive step holds, i.e.:

$$(m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}) \implies (G_{s\tau+1} > G'_{s\tau+1}).$$

Rewriting Equation (38), note that $G_{s\tau+1} = (1-\delta)G_{s\tau} + \frac{\pi_{s\tau}}{P^F k_{sL}}$. If $G_{s\tau} > G'_{s\tau}$, then a sufficient condition to have $(G_{s\tau+1} > G'_{s\tau+1})$ is therefore that

$$\frac{\pi_{s\tau}}{P^F k_{sL}} \geq \frac{\pi'_{s\tau}}{P^F k'_{sL}}.$$

It is relatively straightforward to show that:[36]

$$\frac{\pi_{s\tau}}{P^F k_{sL}} = (\mu_{s\tau} - (1-\alpha))\frac{l_{s\tau}}{k_{s\tau}} G_{s\tau} \frac{w}{P^F(1-\alpha)}, \tag{39}$$

and in what follows I examine how $\frac{\pi_{s\tau}}{P^F k_{sL}}$ is related to $m_{sL}$ and $G_{s\tau}$.

In order to demonstrate the inductive step, I first show that:

[36] Revenue net of labor cost is $\frac{\pi_{s\tau}}{P^F} = (p_{s\tau} - ALC_{s\tau})\,y_{s\tau}$, where $ALC_{s\tau}$ is the average cost of labor input. It is useful to rewrite this as:

$$\frac{\pi_{s\tau}}{P^F} = \left(\mu_{s\tau} - \frac{ALC_{s\tau}}{MC_{s\tau}}\right) y_{s\tau} MC_{s\tau},$$

Given the Cobb-Douglas production function, we have that total labor costs for any given quantity $\bar{y}_{s\tau}$ are $TLC(\bar{y}_{s\tau}) = \frac{w}{P^F} l(\bar{y}_{s\tau})$. For constrained firms, setting $\bar{y}_{s\tau}$ directly implies setting the amount of labor in the following function: $l(\bar{y}_{s\tau}) = \left(\frac{\bar{y}_{s\tau}}{z_{sH} k^{\alpha}_{s\tau}}\right)^{\frac{1}{1-\alpha}}$, such that $TLC(\bar{y}_{s\tau}) = \frac{w}{P^F}\left(\frac{\bar{y}_{s\tau}}{z_{sH} k^{\alpha}_{s\tau}}\right)^{\frac{1}{1-\alpha}}$. Hence, $ALC_{s\tau}(\bar{y}_{s\tau}) = \frac{w}{P^F}\left(\frac{\bar{y}^{\alpha}_{s\tau}}{z_{sH} k^{\alpha}_{s\tau}}\right)^{\frac{1}{1-\alpha}}$, and given these firms' marginal cost is as in Equation (21), $\frac{ALC_{s\tau}}{MC_{s\tau}} = (1-\alpha)$. Together with $\frac{\bar{y}_{s\tau}}{z_{sH} k^{\alpha}_{s\tau}} = l^{1-\alpha}_{s\tau}$, this implies that:

$$\frac{\pi_{s\tau}}{P^F} = (\mu_{s\tau} - (1-\alpha))\, l_{s\tau} \frac{w}{P^F(1-\alpha)}.$$

$$((m_{sL} > m'_{sL}) \wedge (G_{s\tau} = G'_{s\tau})) \implies (G_{s\tau+1} > G'_{s\tau+1}),$$

and then afterwards demonstrate that $((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (G_{s\tau+1} > G'_{s\tau+1})$ holds a fortiori. Recall that $(m_{sL} > m'_{sL}) \implies (\mu_{sL} > \mu_{sL})$. I then analyze the implications of $(m_{sL} > m'_{sL}) \wedge (G_{s\tau} = G'_{s\tau})$ under two exhaustive cases; first $\frac{l_{s\tau}}{k_{s\tau}} = \frac{l'_{s\tau}}{k'_{s\tau}}$, and second $\frac{l_{s\tau}}{k_{s\tau}} \neq \frac{l'_{s\tau}}{k'_{s\tau}}$.

- Case (i): suppose $\frac{l_{s\tau}}{k_{s\tau}} = \frac{l'_{s\tau}}{k'_{s\tau}}$. First, recall that $MC_{sL} = MC'_{sL}$. Combined with $\mu_{sL} > \mu'_{sL}$ this implies that $p_{sL} > p'_{sL}$. Next, the inverse demand function in Equation (2.1) implies that $\frac{p_{s\tau}}{p_{sL}} = \left(\frac{q_{s\tau}}{q_{sL}}\right)^{-1/\sigma} = \left(\frac{z_{sH} l_{s\tau}}{a_L l_{sL}} G_{s\tau}\right)^{-1/\sigma}$, which I rewrite as

$$\frac{p_{s\tau}}{p_{sL}} = \left(\frac{z_{sH}}{a_L} G_{s\tau} \frac{l_{s\tau}}{k_{s\tau}} \frac{k_{sL}}{l_{sL}} G_{s\tau}\right)^{-1/\sigma}.$$

  This implies $\frac{p_{s\tau}}{p_{sL}} = \frac{p'_{s\tau}}{p'_{sL}}$, since by assumption $G_{s\tau} = G'_{s\tau}$ and $\frac{l_{s\tau}}{k_{s\tau}} = \frac{l'_{s\tau}}{k'_{s\tau}}$, while $\frac{l_{sL}}{k_{sL}} = \frac{l'_{sL}}{k'_{sL}}$ from factor demand Equations (17) and (18). Since $p_{sL} > p'_{sL}$ and $\frac{p_{s\tau}}{p_{sL}} = \frac{p'_{s\tau}}{p'_{sL}}$, I find that

$$p_{s\tau} > p'_{s\tau}.$$

  Note then finally from Equation (21) that $MC_{s\tau} = \frac{w}{z_{sH}(1-\alpha)P^F}\left(\frac{l_{s\tau}}{k_{s\tau}}\right)^{\alpha}$. So $\frac{l_{s\tau}}{k_{s\tau}} = \frac{l'_{s\tau}}{k'_{s\tau}}$ implies a constant marginal cost at $MC_{s\tau} = MC'_{s\tau}$. Taken together then, $MC_{s\tau} = MC'_{s\tau}$ and $p_{s\tau} > p'_{s\tau}$ imply that

$$\mu_{s\tau} > \mu'_{s\tau}.$$

  Therefore, using Equation (39), I find that

$$\left((m_{sL} > m'_{sL}) \wedge (G_{s\tau} = G'_{s\tau}) \wedge \left(\frac{l_{s\tau}}{k_{s\tau}} = \frac{l'_{s\tau}}{k'_{s\tau}}\right)\right) \implies \left(\frac{\pi_{s\tau}}{P^F k_{sL}} > \frac{\pi'_{s\tau}}{P^F k'_{sL}}\right). \quad (40)$$

- Case (ii): suppose the firm, following the reaction function in Equation (22), chooses optimally to have $\frac{l_{s\tau}}{k_{s\tau}} \neq \frac{l'_{s\tau}}{k'_{s\tau}}$. Since the reaction function in Equation (22) maximizes the revenue net of labor costs of the firm within period $\tau$, and since $\frac{l_{s\tau}}{k_{s\tau}} = \frac{l'_{s\tau}}{k'_{s\tau}}$ is within the firm's choice set, setting $\frac{l_{s\tau}}{k_{s\tau}} \neq \frac{l'_{s\tau}}{k'_{s\tau}}$ implies that its revenue net of labor costs is weakly higher than under case (i), so Equation (40) continues to hold.

So far, I have assumed that $G_{s\tau} = G'_{s\tau}$. Now, consider instead $G_{s\tau} > G'_{s\tau}$. In this case, the firm with $G_{s\tau}$ can produce an identical output level as when $G_{s\tau} = G'_{s\tau}$, but with less labor input, which would lead to identical revenue at a lower labor cost. Hence, $\pi_{s\tau}$, revenue net of labor cost, is increasing. Moreover, $k_{sL}$ is fixed, and hence, the firm's capital growth rate, $G_{s\tau+1} = (1-\delta)G_{s\tau} + \frac{\pi_{s\tau}}{P^F k_{sL}}$, will increase, and Equation (40) continues to hold. This completes the demonstration of the inductive step.

**Impact on market shares** Having shown that $((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (G_{s\tau+1} > G'_{s\tau+1})$, what remains to be shown is that $((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (m_{s\tau} > m'_{s\tau})$. In order to do so, start by examining $\mathcal{G}_{s\tau} \equiv \left(\frac{y_{s\tau}}{y_{sL}}\right)^{\frac{\sigma-1}{\sigma}}$, which based on the reaction function in Equation (22), is equal to:

$$\mathcal{G}_{s\tau} \equiv \left(\frac{y_{s\tau}}{y_{sL}}\right)^{\frac{\sigma-1}{\sigma}} = \left(\frac{MC_{sL}}{MC_{s\tau}}\frac{\mu_{sL}}{\mu_{s\tau}}\right)^{\sigma-1}$$

Since $m_{s\tau} = \mathcal{G}_{s\tau} m_{sL}$, when $m_{sL} > m'_{sL}$, a necessary condition to have $m_{s\tau} \leq m'_{s\tau}$ is that:

$$\left(\frac{MC_{sL}}{MC_{s\tau}} < \frac{MC'_{sL}}{MC'_{s\tau}}\right) \vee \left(\frac{\mu_{sL}}{\mu_{s\tau}} < \frac{\mu'_{sL}}{\mu'_{s\tau}}\right).$$

I now demonstrate that given $(m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})$, $\left(\left(\frac{MC_{sL}}{MC_{s\tau}} < \frac{MC'_{sL}}{MC'_{s\tau}}\right) \vee \left(\frac{\mu_{sL}}{\mu_{s\tau}} < \frac{\mu'_{sL}}{\mu'_{s\tau}}\right)\right) \implies (m_{s\tau} > m'_{s\tau})$. Hence, $(m_{s\tau} \leq m'_{s\tau})$ cannot hold, and therefore I will have shown what I set out to show, namely $((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (m_{s\tau} > m'_{s\tau})$.

- Here I show that $\left((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}) \wedge \left(\frac{\mu_{sL}}{\mu_{s\tau}} < \frac{\mu'_{sL}}{\mu'_{s\tau}}\right)\right) \implies (m_{s\tau} > m'_{s\tau})$. Recall that $(m_{sL} > m'_{sL}) \implies (\frac{\mu_{sL}}{\mu'_{sL}} > 1)$, and note that given that all markups are positive: $\left(\frac{\mu_{sL}}{\mu_{s\tau}} < \frac{\mu'_{sL}}{\mu'_{s\tau}}\right) \implies \left(\frac{\mu_{sL}}{\mu'_{sL}} < \frac{\mu_{s\tau}}{\mu'_{s\tau}}\right)$. Hence, $\left((m_{sL} > m'_{sL}) \wedge \left(\frac{\mu_{sL}}{\mu_{s\tau}} < \frac{\mu'_{sL}}{\mu'_{s\tau}}\right)\right) \implies (1 < \frac{\mu_{s\tau}}{\mu'_{s\tau}})$. Since $(1 < \frac{\mu_{s\tau}}{\mu'_{s\tau}}) \iff (m_{s\tau} > m'_{s\tau})$, and since $(G_{s\tau} > G'_{s\tau})$ is consistent with $(m_{s\tau} > m'_{s\tau})$, I obtain:

$$\left((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}) \wedge \left(\frac{\mu_{sL}}{\mu_{s\tau}} < \frac{\mu'_{sL}}{\mu'_{s\tau}}\right)\right) \implies (m_{s\tau} > m'_{s\tau})$$

- Here I show that $\left((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}) \wedge \left(\frac{MC_{sL}}{MC_{s\tau}} < \frac{MC'_{sL}}{MC'_{s\tau}}\right)\right) \implies (m_{s\tau} > m'_{s\tau})$. Note that

$$MC_{s\tau} = \frac{w}{(1-\alpha)P^F}\frac{1}{z_{sH}}\left(\frac{l_{s\tau}}{k_{s\tau}}\right)^{\alpha}$$

Since $MC_{sL} = MC'_{sL}$, $\left(\frac{MC_{sL}}{MC_{s\tau}} < \frac{MC'_{sL}}{MC'_{s\tau}}\right) \implies (\frac{l_{s\tau}}{k_{s\tau}} > \frac{l'_{s\tau}}{k'_{s\tau}})$, with $\frac{l_{s\tau}}{k_{s\tau}} = \frac{l_{s\tau}}{G_{s\tau}k_{sL}}$. Hence

$$\left((G_{s\tau} > G'_{s\tau}) \wedge \left(\frac{MC_{sL}}{MC_{s\tau}} < \frac{MC'_{sL}}{MC'_{s\tau}}\right)\right) \implies \left(\frac{l_{s\tau}}{k_{sL}} > \frac{l'_{s\tau}}{k'_{sL}}\right).$$

Now, note that $\mathcal{G}_{s\tau} = \frac{z_{sH}}{a_L}G^{\alpha}_{s\tau}\left(\frac{l_{s\tau}}{l_{sL}}\right)^{\alpha}$. From cost minimization, it follows that $l_{sL} = \left(\frac{r^k}{\alpha}\right)^{\frac{1}{1-\alpha}} k_{sL}$. Hence, $\mathcal{G}_{s\tau} = \frac{z_{sH}}{a_L}G^{\alpha}_{s\tau}\left(\frac{l_{s\tau}}{\left(\frac{r^k}{\alpha}\right)^{\frac{1}{1-\alpha}}k_{sL}}\right)^{\alpha}$. Therefore, $\left((G_{s\tau} > G'_{s\tau}) \wedge (\frac{l_{s\tau}}{k_{sL}} > \frac{l'_{s\tau}}{k'_{sL}})\right) \implies (\mathcal{G}_{s\tau} > \mathcal{G}'_{s\tau})$. Hence:

$$\left((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau}) \wedge \left(\frac{MC_{sL}}{MC_{s\tau}} < \frac{MC'_{sL}}{MC'_{s\tau}}\right)\right) \implies (m_{s\tau} > m'_{s\tau})$$

Combining the previous two bullet points, given $(m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})$, the necessary condition to have $(m_{s\tau} \leq m'_{s\tau})$ itself implies $(m_{s\tau} > m'_{s\tau})$, and therefore:

$$((m_{sL} > m'_{sL}) \wedge (G_{s\tau} > G'_{s\tau})) \implies (m_{s\tau} > m'_{s\tau})$$

### A.3 Proof on Markup Dispersion

First, I examine how the markups of constrained firms behave relative to those of unconstrained firms:

$$\frac{\partial \frac{\mu_{s\mathcal{T}}}{\mu_{sL}}}{\partial M_s} = \frac{\partial \frac{\varepsilon(m_{s\mathcal{T}})\varepsilon(m_{sL})-\varepsilon(m_{s\mathcal{T}})}{\varepsilon(m_{s\mathcal{T}})\varepsilon(m_{sL})-\varepsilon(m_{sL})}}{\partial M_s}$$

Working out the derivative and simplifying:

$$\frac{\partial \frac{\mu_{s\mathcal{T}}}{\mu_{sL}}}{\partial M_s} = \frac{\varepsilon(m_{s\mathcal{T}})(\varepsilon(m_{s\mathcal{T}})-1)\frac{\partial\varepsilon(m_{sL})}{\partial m_{sL}}\frac{\partial m_{sL}}{\partial M_s} - \varepsilon(m_{sL})(\varepsilon(m_{sL})-1)\frac{\partial\varepsilon(m_{s\mathcal{T}})}{\partial m_{s\mathcal{T}}}\frac{\partial m_{s\mathcal{T}}}{\partial M_s}}{(\varepsilon(m_{s\mathcal{T}})\varepsilon(m_{sL})-\varepsilon(m_{sL}))^2}$$

Plugging in the values for $\frac{\partial\varepsilon(m_{ist})}{\partial m_{ist}}$ for the Cournot demand elasticity:[37]

$$\frac{\partial \frac{\mu_{s\mathcal{T}}}{\mu_{sL}}}{\partial M_s} = \frac{(1-\frac{1}{\sigma})\frac{\varepsilon(m_{sL})(\varepsilon(m_{sL})-1)}{\varepsilon(m_{s\mathcal{T}})^2}\frac{\partial m_{s\mathcal{T}}}{\partial M_s} - (1-\frac{1}{\sigma})\frac{\varepsilon(m_{s\mathcal{T}})(\varepsilon(m_{s\mathcal{T}})-1)}{\varepsilon(m_{sL})^2}\frac{\partial m_{sL}}{\partial M_s}}{(\varepsilon(m_{s\mathcal{T}})\varepsilon(m_{sL})-\varepsilon(m_{sL}))^2} < 0 \tag{41}$$

Here we have that $\frac{\varepsilon(m_{s\mathcal{T}})(\varepsilon(m_{s\mathcal{T}})-1)}{\varepsilon(m_{sL})^2} < \frac{\varepsilon(m_{sL})(\varepsilon(m_{sL})-1)}{\varepsilon(m_{s\mathcal{T}})^2}$. I then consider two cases, with first $\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} > 0$, and then its opposite.

- Case (i), suppose $\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} > 0$. Note that $\frac{y_{s\mathcal{T}}}{y_{sL}} = \left(\frac{MC_{sL}}{MC_{s\mathcal{T}}}\frac{\mu_{sL}}{\mu_{s\mathcal{T}}}\right)^\sigma$ and $\frac{\partial G_{s\mathcal{T}}}{\partial M_s} < 0$. There are then two subcases:
  - $\frac{\partial l_{s\mathcal{T}}/k_{sL}}{\partial M_s} < 0$, in which case $\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} < 0$, which would entail a contradiction with the supposition, so it cannot hold.
  - $\frac{\partial l_{s\mathcal{T}}/k_{sL}}{\partial M_s} \geq 0$, which entails $\frac{\partial MC_{s\mathcal{T}}}{\partial M_s} < 0$. In that case, having $\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} > 0$ requires that

$$\frac{\partial \frac{\mu_{s\mathcal{T}}}{\mu_{sL}}}{\partial M_s} < 0.$$

- In case (ii), suppose $\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} \leq 0$. It is then straightforward that $\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} \leq 0 \iff \frac{\partial \frac{m_{s\mathcal{T}}}{m_{sL}}}{\partial M_s} \leq 0$, where $\frac{\partial \frac{m_{s\mathcal{T}}}{m_{sL}}}{\partial M_s} \leq 0 \implies \left(\frac{\partial m_{s\mathcal{T}}}{\partial M_s} m_{sL} \leq \frac{\partial m_{sL}}{\partial M_s} m_{s\mathcal{T}}\right)$. Since $m_{sL} < m_{s\mathcal{T}}$ and $\frac{\partial m_{sL}}{\partial M_s}, \frac{\partial m_{s\mathcal{T}}}{\partial M_s} < 0$ from Lemma 3, in this second case it holds that:

$$\frac{\partial \frac{y_{s\mathcal{T}}}{y_{sL}}}{\partial M_s} \implies \frac{\partial m_{s\mathcal{T}}}{\partial M_s} \leq \frac{\partial m_{sL}}{\partial M_s} < 0.$$

  Given Equation (41), this has the same implication as in the first case, namely that:

$$\frac{\partial \frac{\mu_{s\mathcal{T}}}{\mu_{sL}}}{\partial M_s} < 0.$$

The remaining question is how the relative markups of unconstrained firms behave. Here, note that $m_{sH} = \mathcal{G}_{sH} m_{sL}$, such that $\frac{\partial m_{sH}}{\partial M_s} = \mathcal{G}_{sH}\frac{\partial m_{sL}}{\partial M_s} + m_{sL}\frac{\partial \mathcal{G}_{sH}}{\partial M_s}$. Now, suppose $\frac{\partial \frac{\mu_{sH}}{\mu_{sL}}}{\partial M_s} \geq 0$. Then the factor ratios $\frac{k_{sH}}{k_{sL}} = \frac{l_{sH}}{l_{sL}} = \left(\frac{z_{sH}}{z_{sL}}\right)^{\sigma-1}\left(\frac{\mu_{sL}}{\mu_H}\right)^\sigma$ imply that $\frac{\partial \mathcal{G}_{sH}}{\partial M_s} < 0$. In that case, however, $\frac{\partial m_{sH}}{\partial M_s} < \frac{\partial m_{sL}}{\partial M_s}$, which from equation (41) implies that $\frac{\partial \frac{\mu_{sH}}{\mu_{sL}}}{\partial M_s} < 0$. Hence, the supposition that $\frac{\partial \frac{\mu_{sH}}{\mu_{sL}}}{\partial M_s} \geq 0$ entails a contradiction, and therefore its opposite must be true:

[37] It is straightforward to verify that the following result also holds for the Bertrand demand elasticity.

$$\frac{\partial \frac{\mu_{sH}}{\mu_{sL}}}{\partial M_s} < 0.$$

# Appendix B Assumptions on the productivity volatility process

By definition, in steady state, $D^s(a', k, z)$ is stable over time for each sector, and this distribution is described by Lemma 2. Given that $M_s$ is finite, the law of large numbers does not hold, and for a stochastic process with iid transition probabilities, $D^s(a', k, z)$ will not be exactly stable. To sidestep this issue, I do not assume that transition probabilities are iid. Instead, I make the following assumption:

**Assumption.** *Productivity shocks are such that, if a certain transition probability $Pr_{xy}$ to go from state x to state y applies to a set of firms of size $N_x$, then exactly $Pr_{xy}N_x$ firms will transition from state $x$ to state $y$.* [38]

Here, the different states $x$ and $y$ are defined in relation to Lemma 2. Before giving an exact definition of these states, we first need to keep track of following implications of Proposition 1.[39] Since these implications hold for any sector, we drop the subscript $s$.

- Implication 1: convergence to the optimal level of capital is reached in a finite number of periods and therefore the maximal $\tau$ is finite. This is because $\frac{k^*_H}{k^*_L}$ is finite and $g_\tau$ is always strictly positive.
- Implication 2: The number of periods it takes for a high productivity firm to become unconstrained is increasing with $M_s$.[40] Let therefore $T^M$ denote the number of periods it takes for a high productivity firm to grow out of its financial constraint in a steady state with $M$ firms.
- Implication 3: there are then in total $(T^M + 2)$ types of firms: $L, T^M, H$

Consider a sufficiently high $M$, $\bar{M}$, such that $\bar{M}$ will be the highest value of $M$ considered in the comparative statics on $M$. Importantly, given implications 1 and 2, we have that $\forall M < \bar{M}$ : $T^M \leq T^{\bar{M}}$.

Based on implication 3, the productivity volatility process will then be defined by transition probabilities across the low-productivity bin $L$ and the high-productivity bins $T^{\bar{M}}, H^{\bar{M}}$, where $H^{\bar{M}}$ denotes the bin with all the firms that are still high-productivity beyond bin $T^{\bar{M}}$. Note that this implies that for $M < \bar{M}$, firms might be in e.g. bin $T^{\bar{M}}$ for the definition of their transition probabilities, although they are already unconstrained.

**Definition.** ***The transition probabilities*** *across bins are:*

- *Probability to transition from $a_L$ to $z_{sH}$, i.e. probability to transition from L to $\tau = 1$: $P_{LH}$. Then, $(1 - P_{LH})$ is probability to remain within L.*

[38]One could think of a divine entity setting up a lottery such that exactly $Pr_{xy}N_x$ firms are selected to transition from $x$ to $y$.

[39]This proposition is demonstrated conditional on the steady state existing for different values of $M_s$, as well as the productivity volatility process being identical across $M$. Hence, if the characteristics of the productivity volatility process are such that the steady state exists, and that the process is identical across $M_s$, then Proposition 1 holds.

[40]This is because $G_\tau$ is weakly decreasing in $M$ and $\frac{k^*_H}{k^*_L}$ is increasing with $M$.

- *For firms with $z_{sH}$, the transition probabilities are dependent on $\tau$. Conditional on having $z_{sH}$ in period $\tau$, the probability of continuing having $z_{sH}$ is $P_{HH\tau}$.*
  - *Therefore, conditional on having $z_{sH}$ in $\tau = 1$, the probability of still having $z_{sH}$ in $\tau > 1$, is $\prod_{r=1}^{\tau-1} P_{HHr}$.*
- *Finally, the transition probability of moving from bin $H^{\bar{M}}$ to bin L, is $P_{HL}$.*

By specifying these bins and defining transition probabilities across them, I have assured that the number of firms in each bin is stable across periods.[41] To see this, denote the number of firms in L by $M_L$. Then, the number of firms in bin $\tau$: $M_L P_{LH} \prod_{r=1}^{\tau-1} P_{HHr}$. Next, the number of firms in $H^U_{\bar{M}}$ can be found by setting the number of exiters from $H^{\bar{M}}$ equal to the number of entrants in $H^{\bar{M}}$: $M_L P_{LH} \prod_{r=1}^{T^{\bar{M}}} P_{HHr} = P_{HL} M_H$. Hence

$$M_H = \frac{M_L}{P_{HL}} P_{LH} \prod_{r=1}^{T^{\bar{M}}} P_{HHr}$$

One can then also observe that the number of entrants in $L$ equals the number of exiters from $L$:

$$P_{LH} M_L = P_{LH} M_L \left[ (1 - P_{HH1}) + \sum_{\tau=2}^{T^{\bar{M}}} (1 - P_{HH\tau}) \prod_{r=1}^{\tau-1} P_{HHr} \right] + \frac{M_L}{P_{HL}} P_{LH} \prod_{r=1}^{T^{\bar{M}}} P_{HHr} P_{HL}$$

$$1 = (1 - P_{HH1}) + \sum_{\tau=2}^{T^{\bar{M}}} \left[ (1 - P_{HH\tau}) \prod_{r=1}^{\tau-1} P_{HHr} \right] + \prod_{r=1}^{T^{\bar{M}}} P_{HHr}$$

Now, note first that $\prod_{r=1}^{T^{\bar{M}}} P_{HHr}$ is the probability, conditional on a firm moving from $L$ to $H$ productivity, that after $T^{\bar{M}}$ periods it still has $H$ productivity. Then note that $\sum_{\tau=1}^{T^{\bar{M}}} (1 - P_{HH\tau}) \prod_{r=1}^{\tau-1} P_{HHr}$ is the probability that a firm moves back to low productivity at some point before $T^{\bar{M}}$. Hence we always have that $1 - \prod_{r=1}^{T^{\bar{M}}} P_{HHr} = (1 - P_{HH1}) + \sum_{\tau=2}^{T^{\bar{M}}} (1 - P_{HH\tau}) \prod_{r=1}^{\tau-1} P_{HHr}$. In other words, the condition for stability of the share of number of firms is always satisfied.[42]

While the described productivity volatility process will ensure that the number of firms in each bin is stable over time, it does not necessarily imply that the number of firms in a bin is an integer. Hence, one has to impose additional restrictions on the values of $M$ under consideration, or on the transition probabilities. One such possible restriction is to set $\forall \tau P_{HH\tau} = P_{LH} = P_{LH} = \frac{1}{n}$ and $M_L = n^{(T^{\bar{M}} + x)}$ with $x \geq 2$ and $n \in \mathbb{N}$. This implies that the number of firms in any bin $\tau$ is $n^{-\tau} M_L = n^{(T^{\bar{M}} + x - \tau)}$ and in bin $H$ it is $n^x$.

[41] Note that transition probabilities $P_{HH\tau}$ are allowed to be constant across $\tau$.

[42] Note that $\sum_{\tau=2}^{T^{\bar{M}}} (1 - P_{HH\tau}) \prod_{r=1}^{\tau-1} P_{HHr} = 1 - P_{HH1} + (1 - P_{HH2}) P_{HH1} + (1 - P_{HH3}) P_{HH1} P_{HH2} + .... + (1 - P_{HHT^{\bar{M}}}) \prod_{r=1}^{T^{\bar{M}}-1} P_{HHr}$, which confirms the equality.

# Appendix C Data description

## C.1 Annual Survey of Industries

In the empirical analysis, I employ plant-level data from the Annual Survey of Industries (ASI), obtained from India's Ministry of Statistics and Programme Implementation (MOSPI). I use data from each survey year for the accounting years 1990-91 till 2011-2012. However, in the analysis of the dereservation reform, I drop the final two years from the dataset due to changes in the product classification (see Section 4.2). As explained in Section 3, the ASI sampling design has been updated from time to time. Most importantly, for most years the census scheme includes all plants with 100 or more employees. The exceptions are the years 1997-98 till 1999-00, when this threshold was 200 workers instead. Further details on the exact sampling design for each year can be found on the MOSPI website.[43]

For the cleaning of the data, I follow a procedure that is highly similar to the data cleaning in Allcott et al. (2016). First, I correct observations in the 1993-94 to 1997-98 survey years, whose values have been provided in "pre-multiplied" form by MOSPI.[44] Second, I drop duplicate observations - a small but non-trivial issue in the early years of the sample - and observations missing state identifiers. Third, I restrict the sample to all plants (i) in the manufacturing sector, (ii) who are listed as open, and (iii) who have non-missing and positive values for three critical variables, namely the logarithms of revenue, capital and labor cost. Fourth, I drop observations whose labor cost or material cost exceeds 250 percent of their revenue.

Two of the main plant-level variables of interest in the analysis are defined as ratios: the labor cost share, and the ratio of revenue to capital. Since these are ratios, price inflation is an irrelevant concern in their computation. When computing these ratios, I therefore measure both the numerator and denominator in nominal terms. However, a third variable of interest - the capital growth rate - is sensitive to inflation in its measurement and hence I calculate it after deflating the book value of capital with the capital deflator from the Indian Handbook of Industrial Statistics, setting 2004 as the base year. To mitigate the influence of outliers, I winsorize all three variables - the labor cost share, the ratio of revenue to capital, and the capital growth rate - at the 1st and 99th percentile.

In the analysis, I define sectors at the 3-digit level, using India's 1987 National Industrial Classification (NIC). There are 191 separate 3-digit sectors in my data. I also ensure that plants are assigned to a consistent geographic region (state or union territory) over all years. To make the definitions of regions consistent over time, I employ the concordance provided by the Indian Statistical Office. In addition, I take into account the creation of Jharkhand, Chhattisgarh and Uttaranchal in 2001 from Bihar, Madhya Pradesh and Uttar Pradesh, respectively. I assign these newly created states the state code from the state they separated from. There are 29 separate geographic regions in my data.

[43] The sampling designs for all years prior to 1998-99 is described on http://mospi.nic.in/salient-features-sampling-designs-asi-1973-74-asi-1998-99, and the designs from 1998-99 onward on http://www.csoisw.gov.in/cms/cms/files/554.pdf. Both links were retrieved on February 22, 2019.

[44] For further detail on this pre-multiplication, see for instance the documentation on MOSPI's microdata catalogue http://mail.mospi.gov.in/index.php/catalog/75#page=sampling&tab=study-desc (retrieved on February 22, 2019.)

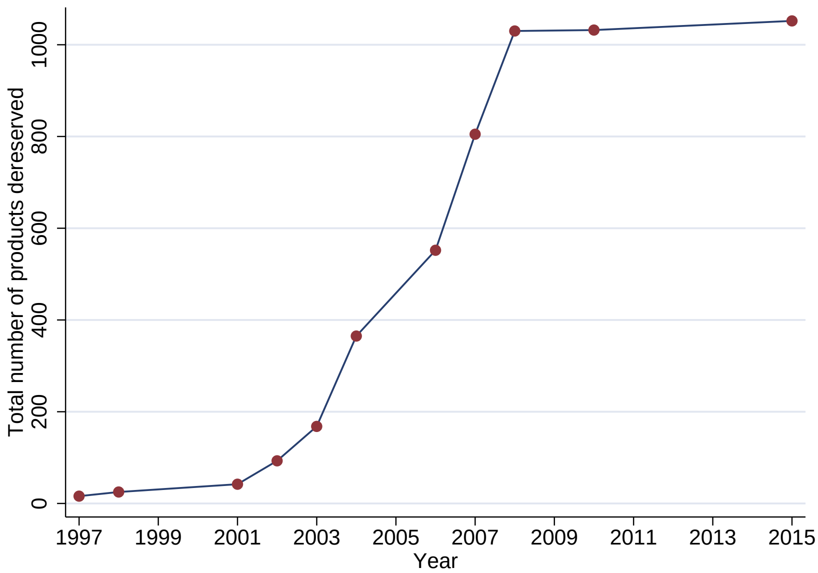
## C.2 Data on dereservation

As described in Section 3, I employ the concordance between SSI product codes, used to determine dereservation status, and ASICC product codes, used in the ASI data, available from Martin et al. (2017). A substantial subset of the ASICC codes is relatively broad. For this reason, Martin et al. (2017) construct a first set of matches between the SSI product codes and pairs of an ASICC code and a 5-digit NIC code. Remaining product codes are simple matches between the SSI and ASICC code.

I cross-checked the Martin et al. (2017) dereservation list with the list from Tewari and Wilde (2017), which Ishani Tewari generously made available to me. The two lists correspond very closely, except for seven products that are included in Tewari and Wilde (2017), but not in Martin et al. (2017). These seven products are also included in the dereservation notifications on the website of the Indian Ministry for Micro, Small and Medium Enterprises (referenced in footnote 20). I therefore add these omitted products to the product-level concordance.

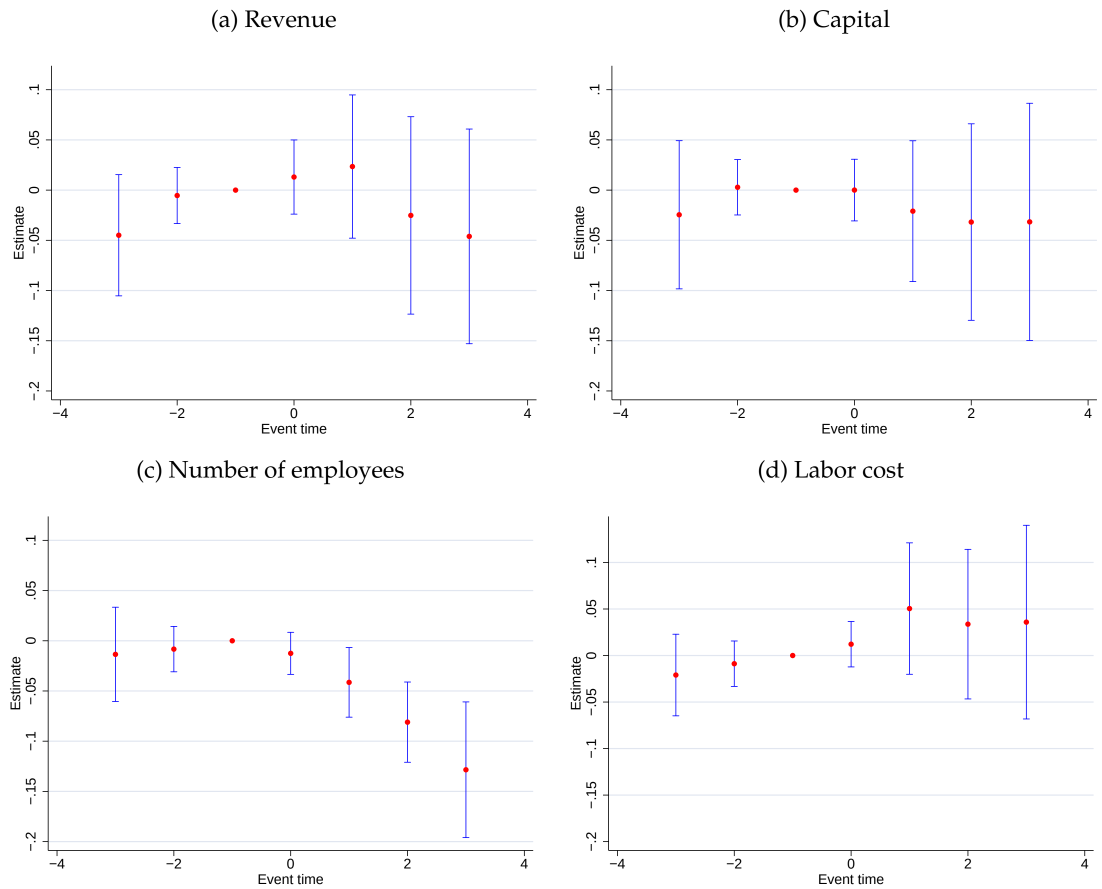
# Appendix D Supplementary tables and figures

Figure D.1: Timing of dereservation



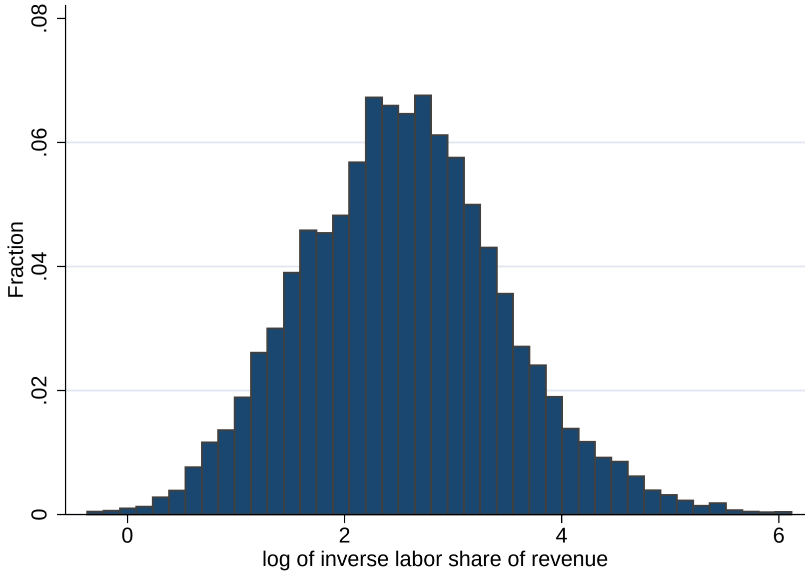

The figure displays the cumulative number of products dereserved in a given year, from the start of the dereservation reform in 1997 till the completion of dereservation in 2015. Each dot indicates a year when products were dereserved.

Figure D.2: Absence of pre-trends for dereserved plants




The figure displays the coefficients and 95% confidence intervals for the $\beta_\tau$ coefficients from the following event-study regressions: $\ln y_{it} = \gamma_i + \nu_t + \sum_{\tau=-3}^{3} \beta_\tau 1[t = e_{it} + \tau] + \varepsilon_{it}$, where $\gamma_i$ is a plant fixed-effect and $\nu_t$ is a year fixed effect. The dependent variable $y_{it}$ is revenue, capital, number of employees, or total labor cost, for panels a,b,c and d respectively. The regression is estimated on a balanced sample of between 20913 to 20937 incumbent plants, depending on the outcome variable. I define the time at which the first product of plant $i$ is dereserved as $e_{it}$, and impose the normalization that $\beta_{-1} = 0$. Since dereservation status is defined at the product level, I also cluster standard errors at that level. Revenue, capital and labor cost are deflated using the deflators of the Indian Handbook of Industrial Statistics, with 2004 as the base year.

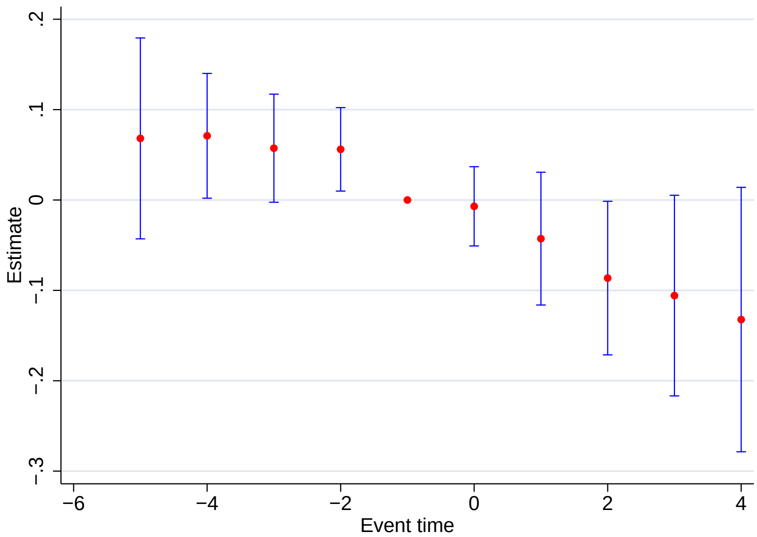
Figure D.3: Inverse labor share among plants in the dereservation event study



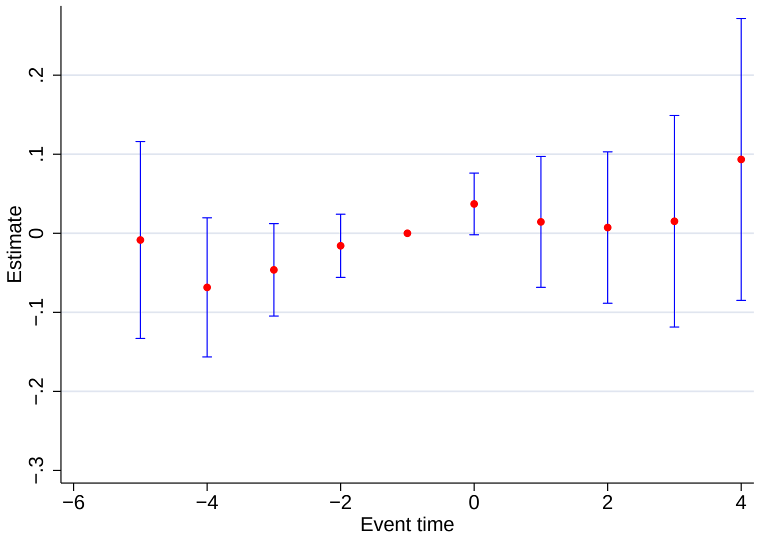

The figure displays the histogram for $\ln(S_{it}/w_t L_{it})$ among plants in the dereservation event study. This includes all plants who were incumbents producing reserved products, and who were observed each of the three periods before and three periods after their first product was dereserved. The variable $\ln(S_{it}/w_t L_{it})$ is driving the within-plant variation in the markups in the event study in Figure 1, since the output elasticity is absorbed in a plant fixed-effect.

Figure D.4: Event study of dereservation over a longer time horizon

(a) Incumbents with above-median initial markup



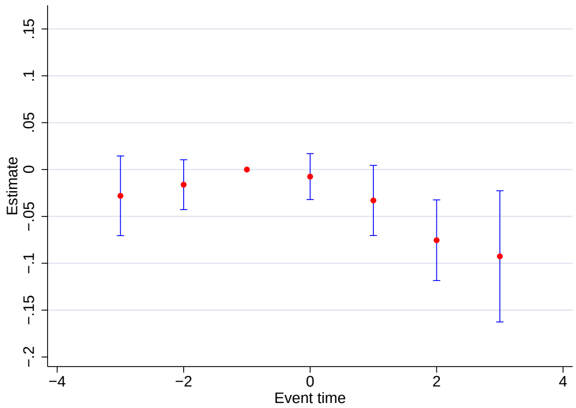

(b) Incumbents with below-median initial markup



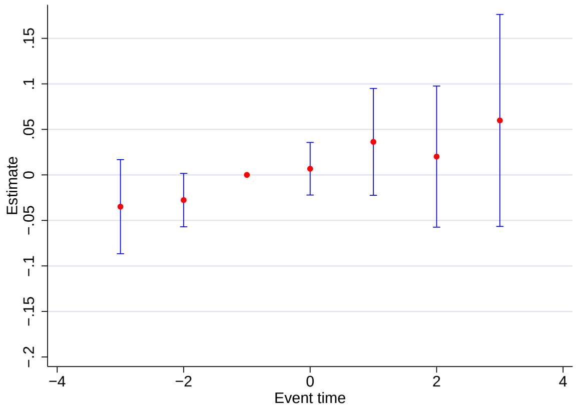

The figure displays the coefficients and 95% confidence intervals for the $\beta_\tau$ coefficients from the following event-study regression: $\ln \mu_{it} = \gamma_i + \nu_t + \sum_{\tau=-5}^{4} \beta_\tau 1[t = e_{it} + \tau] + \varepsilon_{it}$, where $\gamma_i$ is a plant fixed-effect and $\nu_t$ is a year fixed effect. I define the time at which a plant's first product of plant $i$ is dereserved as $e_{it}$, and impose the normalization that $\beta_{-1} = 0$. Since dereservation status is defined at the product level, I also cluster standard errors at that level. I restrict the sample to a balanced sample of incumbent plants that are observed at least five years before and four years after they are dereserved. Panel (a) shows results for plants with initial markups weakly above the median initial markup, and Panel (b) for the other plants. The initial markup is an average over event times $\tau = -4$ till $\tau = -2$, and the median initial markup is then set after taking out sector and year fixed effects.

Figure D.5: Event study of dereservation for incumbents with high initial markups

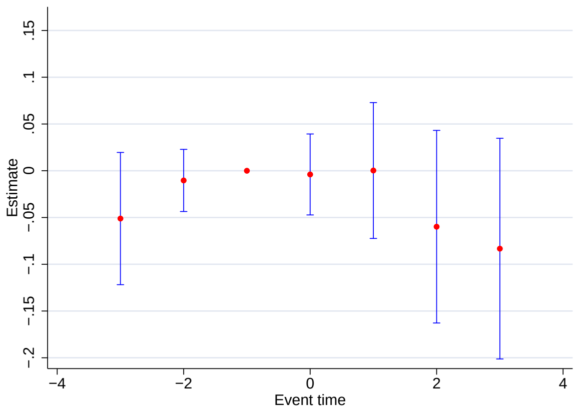
(a) Number of Employees




(b) Labor Cost




(c) Revenue




The figure displays the coefficients and 95% confidence intervals for the $\beta_\tau$ coefficients from the following event-study regressions: $\ln y_{it} = \gamma_i + \nu_t + \sum_{\tau=-3}^{3} \beta_\tau 1[t = e_{it} + \tau] + \varepsilon_{it}$, where $\gamma_i$ is a plant fixed-effect and $\nu_t$ is a year fixed effect. The dependent variable $y_{it}$ is number of employees, total labor cost, or revenue for panels a, b and c respectively. The regression is estimated on a balanced sample of plants above the median initial markup, as defined in Figure 1. I define the time at which the first product of plant $i$ is dereserved as $e_{it}$, and impose the normalization that $\beta_{-1} = 0$. Since dereservation status is defined at the product level, I also cluster standard errors at that level. Revenue and labor cost are deflated using the deflators of the Indian Handbook of Industrial Statistics, with 2004 as the base year.

Table D.1: Summary statistics

| | Full Panel | Dereservation Incumbents |
|---|---|---|
| Number of workers | 331<br>(1074) | 327<br>(1084) |
| Age of Plant | 20<br>(20) | 20<br>(16) |
| Revenue to capital ratio | 593<br>(38893) | 269<br>(26626) |
| Labor cost share of gross revenue | .13<br>(.15) | .13<br>(.15) |
| Observations | 217942 | 62482 |

The table shows average values, with standard deviations in parentheses, for the variables in each column. The full panel covers all plants with non-missing lagged MRPK, included in the estimation of column 1 of Table 2. The sample of dereservation incumbents covers all plants that are incumbents in the dereservation analysis with non-missing lagged MRPK, included in the estimation of column 1 of Table 1. I include the revenue to capital ratio and the labor cost share in this table because these two variables are central to the measurement of MRPK and markups, respectively.

Table D.2: MRPK convergence for dereservation years 2000-2009

| | $MRPK_{it}$ - Gross Revenue (GR) | $MRPK_{it}$ - Value Added (VA) |
|---|---|---|
| | (1) | (2) |
| $Deres_{it-1}$ | -0.104$^{**}$ | -0.176$^{**}$ |
| | (0.021) | (0.021) |
| $MRPK_{it-1}(GR)$ | 0.447$^{**}$ | |
| | (0.016) | |
| $MRPK_{it-1}(GR) * Deres_{it-1}$ | 0.020$^{+}$ | |
| | (0.011) | |
| $MRPK_{it-1}(VA)$ | | 0.313$^{**}$ |
| | | (0.011) |
| $MRPK_{it-1}(VA) * Deres_{it-1}$ | | 0.023$^{*}$ |
| | | (0.010) |
| Plant Fixed Effects | Yes | Yes |
| Year Fixed Effects | Yes | Yes |
| Observations | 57756 | 52461 |

Robustness analysis for the columns 1 and 3 in Table 1. Since product coverage in the ASI is incomplete for the years 1998 and 1999, the estimation in this table is restricted to the incumbent plants that are dereserved during the years 2000-2009. Since this restriction is based on the year of dereservation, there is no natural counterpart to this restriction for columns 2 and 4 in Table 1, which are estimated on the full panel, including entrants and outsiders. In specification 1, MRPK is measured based on gross revenue, and based on value added in specification 2. Both specifications control for the logarithm of a plant's age. Since dereservation status is defined at the product level, I also cluster standard errors at that level. Standard errors in parentheses. $^{+}p < 0.1;^{*} p < 0.05;^{**} p < 0.01$.

Table D.3: Young plants' capital growth for dereservation years 2000-2009

| | Capital growth $g(k)_{it}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $Deres_{it-1}$ | 0.006 | -0.011 | 0.007 | 0.002 |
| | (0.010) | (0.021) | (0.007) | (0.009) |
| $Deres_{it} * [-\ln(age_{it})]$ | -0.001 | -0.003 | | |
| | (0.004) | (0.007) | | |
| $Deres_{it-1} * 1(age_{it} \leq 5)$ | | | -0.020$^{*}$ | -0.038 |
| | | | (0.008) | (0.024) |
| $[-\ln(age_{it})]$ | 0.000 | 0.008 | | |
| | (0.005) | (0.006) | | |
| $1(age_{it} \leq 5)$ | | | 0.019 | 0.011 |
| | | | (0.012) | (0.010) |
| State-sector-year Fixed Effects | Yes | Yes | Yes | Yes |
| Plant Fixed Effects | No | Yes | No | Yes |
| Observations | 122989 | 107873 | 124832 | 109672 |

Robustness analysis for Table 3. Since product coverage in the ASI is incomplete for the years 1998 and 1999, the estimation in this table is restricted to the incumbent plants that are dereserved during the years 2000-2009. Since dereservation status is defined at the product level, I also cluster standard errors at that level. Standard errors in parentheses. $^{+}p < 0.1;^{*} p < 0.05;^{**} p < 0.01$.

# Appendix E A model with undercapitalized newborn firms

This model is analogous to the baseline model, except for the following modifications. First, there is no productivity volatility, and all firms have the same productivity. Second, each period in each sector, $qM_s$ firms die and are replaced by newborn firms. The ex-ante probability that any firm dies is constant at $q$. However, related to the assumptions on transition probabilities in the baseline model, the death probability is not independent across firms. Specifically, I assume that each period, the same number of firms of each "type" die. Here, the definition of firm types is identical to the types of high-productivity firms in Appendix B, and the death probabilities are identical to the transition probabilities from high-productivity to low-productivity in that appendix.

The dead firms together transfer a fraction of their last period earnings to the newborn firms, such that these are born with capital levels $k_{s0} \equiv \zeta \frac{\phi_s P^F Q^F}{M_s}$, where $K_s$ is aggregate capital in sector $s$ and $0 < \zeta < 1$. Finally, firm-owner $is$ has the following intertemporal preferences at time $t$:

$$U_{ist} = \sum_{v=t}^{\infty} (q\beta)^{v-t} c_{isv}$$

Where $\beta$ is the discount factor, $q$ is the ex-ante probability a firm dies in any given period and $c_{ist}$ is firm-owner consumption. Otherwise, the model is exactly as the baseline model. The optimization problem and the solution to the steady state equilibrium are therefore highly similar, except that there are no productivity differences anymore. Now, the only variation in marginal costs is coming from differences in capital levels. Otherwise, reaction functions are still as in equation (22). The solution to the steady state equilibrium then implies:

**Lemma 4.** *In steady state, the joint distribution of capital and productivity within a sector is as follows:*

- *When firms are unconstrained $\tau$ periods after their birth , then $k_{is\tau} = k_s^*$*
- *When firms are constrained $\tau$ periods after their birth, then $k_{is\tau} = G_{s\tau} k_0$, with $G_{s\tau} = \Pi_{s=0}^{\tau-1} (\frac{\pi_{sv}}{P^F k_{sv}} + 1 - \delta)$*

The initial capital growth equation is then:

$$\frac{\pi_{s\tau}}{P^F G_{s\tau} k_{s0}} = (\mu_{s\tau} - (1-\alpha)) \frac{w}{P^F(1-\alpha)} \frac{l_{s\tau}}{G_{s\tau} k_{s0}},$$

from which one can derive analogous comparative statics results as before. In particular, as $M_s$ increases, $G_{s\tau}$ will decline, which is the prediction that I take to the data.