

Heterogeneous Mark-Ups and Endogenous Misallocation

Michael Peters*

January 20, 2011

First Version: September 2009

Abstract

Why are resources misallocated across firms? I study an economy where misallocation stems from firms charging heterogeneous mark-ups. The distribution of mark-ups affects both aggregate TFP as well as equilibrium factor prices. While TFP is solely affected by the cross-sectional dispersion of mark-ups, factor prices depend only on the average mark-up. Mark-ups however are the result of firms' pricing decision and hence endogenous. In my model, they are determined by the process of entry and exit. If entry is more intense, aggregate TFP and factor prices are high, as product market competition reduces both the level and the dispersion of mark-ups. The static degree of misallocation is therefore closely linked to the dynamic evolution of the economy. I test the model's prediction using regional variation in the extent of entry into the manufacturing sector in Indonesia. I show that regional differences in the entry rate are negatively correlated with the average mark-up and the mark-up dispersion as the model predicts. In terms of welfare, the theory implies that the observed differences in mark-ups between high and low entry regions in Indonesia translate into TFP differences of 2.5%, differences in the steady state level of capital of 13% and differences in aggregate consumption of 6%.

Keywords: Endogenous mark-ups, Entry, Firm-level distortions, Measuring misallocation, TFP differences

1 Introduction

Misallocation of resources across firms reduces economic efficiency and welfare. Misallocation occurs, whenever relative prices are distorted in that they do not reflect the scarcity of the respective commodity. Hence, any theory of misallocation is a theory of mispricing. Analyzing misallocation as an equilibrium phenomenon therefore requires a theory for why prices are distorted, what equilibrium prices look like and how they depend on the exogenous characteristics of the environment.

In this paper, I study a general equilibrium economy with heterogeneous firms, where prices are distorted because monopolistic price setting endogenously induces heterogeneous mark-ups.¹ This

*Massachusetts Institute of Technology, Department of Economics, 50 Memorial Drive, E52-391, Cambridge, MA 02142, eMail: mipeters@mit.edu.

¹Recently, Epifaniy and Ganica (forthcoming) also studied an economy with heterogeneous mark-ups. While I will come back to their contribution in more detail below when I discuss the related literature, let me briefly note that the main difference is that Epifaniy and Ganica (forthcoming) only consider the case of exogenous mark-ups, while these will be endogenous in my paper.

cross-sectional heterogeneity in mark-ups adversely affects economic efficiency, because the economy's resources are not allocated towards the most productive units but rather to the ones whose monopolistic power is relatively low. While this basic intuition is of course not new, the general equilibrium effects of firms charging heterogeneous mark-ups are less well known. In particular, it is not obvious whether mark-ups manifest themselves only in lower factor prices or if they also reduce total factor productivity (TFP) compared to the competitive benchmark. While both factor prices and TFP in general depend on the entire distribution of mark-ups, some robust results emerge. As also argued in Epifani and Ganica (forthcoming), mark-up *heterogeneity* mostly affects aggregate TFP and a higher *level* of mark-ups is entirely absorbed by a reduction in factor prices. Intuitively, TFP is determined by firms' relative prices. If all mark-ups increase, prices still signal the appropriate relative scarcity of resources across producers so that there is no misallocation across firms and the economy's TFP is equal to the first-best benchmark. By setting higher prices however, firms produce less and hence reduce their factor demand which in turn lowers equilibrium factor prices. If on the other hand the average mark-up is low but mark-ups are very dispersed, the misallocation of resources across producers will cause TFP to decline but leaves equilibrium factor prices unaffected. Hence, the *distribution* of mark-ups determines allocative efficiency, has distributive implications and affects dynamic accumulation incentives.

So how does the equilibrium distribution of mark-ups look like? Equilibrium mark-ups depend on firms' relative market power, which in my model is fully determined by the *relative* productivity across firms. If the productivity distribution across firms is very dispersed, the most productive firm enjoys substantial monopolistic power as the competitive threat is low. If on the other hand, productivity is highly concentrated across producers, the scope of monopolistic pricing is reduced. The *static* degree of misallocation across firms is therefore inextricably linked to the *dynamic* properties of the evolution of productivity of heterogeneous firms. In particular, an environment, which is conducive to a large productivity dispersion across firms, reduces allocative efficiency by fostering monopolistically induced misallocation. Examples include slow technology diffusion among producers, limits to the learning of best-practice or frontier technologies or barriers to the entry of new firms.

In this paper I will focus on the last aspect, i.e. the importance of entry. In particular, I study an environment where productivity growth can either be generated by current producers or stem from new firms entering the market. I show that in my economy, the entire endogenous distribution of mark-ups can be characterized in closed form and is fully determined by a single sufficient statistic. This statistic is the economy's entry intensity, which is given by the equilibrium entry rate relative to the innovation rate of existing firms. If an economy's productivity growth is largely accounted for by new entrants, the entry intensity is high and the distorting effects of mark-ups are low because entering firms keep the productivity distribution compressed and monopoly power limited. In such an environment, aggregate TFP and factor prices are high. If on the other hand most of productivity growth is generated by current producers, the efficiency costs of mark-up distortions are high as most firms will have a large degree of monopoly power as their competitors are relatively unproductive. This causes TFP and factor prices to be low.

To study these issues formally, I propose a model which despite the non-trivial pricing rules generating heterogeneous mark-ups, is highly tractable. In particular, the aggregate economy looks exactly like the single-sector neoclassical growth model with Cobb-Douglas technology and an aggregate TFP term. This aggregate TFP term is the product of two components, where the first one depends only on the distribution of firm-level productivity and the second one is entirely determined by the distribution of mark-ups. It is this second term (and only this second term) which

is not present in the efficient allocation and is hence a sufficient statistic for the TFP losses of non-constant mark-ups. Similarly, there is a simple sufficient statistic for the effect on equilibrium factor prices. These two statistics completely summarize the degree of misallocation induced by monopolistic mark-ups.

The existence of such sufficient statistics is especially convenient to transparently characterize the link between the entry intensity of the economy and the degree of misallocation. In general, the former determines the equilibrium distribution of prices, which in turn affects the economic quantities of interest. In my model, there is a sufficiency results for each of these aspects. The entry intensity is a sufficient statistic for the entire mark-up distribution, which in turn affects the aggregate economy also only via two sufficient statistics itself. To study which characteristics of the environment reduce TFP and factor prices by fostering monopolistically induced misallocation, we only need a theory how the entry intensity is determined as an equilibrium phenomenon. In this paper I consider a particular microfoundation, namely an endogenous growth model in the Schumpeterian tradition (Aghion and Howitt, 1992; Grossman and Helpman, 1991; Klette and Kortum, 2004). In such an environment, the equilibrium entry intensity is determined by the entry costs, the total supply of innovation resources and the efficiency of the innovation and entry technology. In particular, the model implies that the entry intensity is decreasing in the entry costs and increasing in the total supply of innovative resources. Higher entry costs and scarce innovative resources therefore reduce the allocative efficiency of the economy by increasing the equilibrium distribution of mark-ups.²

The model makes tight predictions about the distribution of mark-ups and their correlation with the entry intensity of the economy. To test the model empirically, I analyze a comprehensive data set of manufacturing firms in Indonesia. The model predicts that a higher degree of entry into the manufacturing sector improves the allocation of resources by reducing mark-ups. Using regional variation in the entry intensity, I find evidence in favor of this prediction. In particular, I show that mark-ups in high-entry regions are both lower and less dispersed than in their low-entry counterparts and that the sufficient statistics suggested by the model are positively correlated with the entry intensity. These findings are robust to controlling for observable regional characteristics like population density, the degree of financial development as measured by the density of the banking network and the regional quality of the transportation infrastructure.

While these findings are consistent with the theory, I also consider two competing explanations namely regional differences in credit market imperfections and policy distortions. First of all I show how these competing microfoundations of misallocation can be represented as theories of mispricing which fit in my framework and whose efficiency consequences can therefore be expressed via the same two sufficient statistics. Then I test the implications of those competing theories and show that they cannot account for my results. In particular, the regional entry intensity remains an important determinant of the distribution of mark-ups. This is of course not to say that financial imperfections or differences in regional policy are unimportant. In fact, the entry rate is likely to be affected by both of them. What my results do show however, is that the entry rate has an independent effect on top of the direct consequences of say financial frictions. The theoretical part of the paper suggests that this effect might be causal by making product markets more competitive.

Finally I conduct a welfare comparison to analyze if regional differences in the entry intensity

²The former prediction is for example in line with Djankov, Porta, Lopez-De-Silvaes, and Shleifer (2002) and Porta and Shleifer (2009), who claim that entry of new firms is particularly inefficient in underdeveloped economies, and Barseghyan (2008) and Barseghyan and DiCecio (2009), who provide empirical evidence for the negative relation between aggregate TFP and entry costs.

are quantitatively important. Specifically, the observed differences in the distribution of mark-ups between high and low-entry regions and imply that the former have a higher total factor productivity by 2.5% and a higher steady state level of capital by 13%. In terms of welfare, steady state consumption in the high-entry regions is predicted to be about 6% higher. To put this into perspective, I also consider an exercise similar in spirit to Hsieh and Klenow (2009). By reducing the degree of pricing distortions from the most distorted region in Indonesia to the least distorted one, TFP would increase by 5.5%, the steady state capital stock by 45% and aggregate consumption by 17%. Hence, by focusing on the entry rate alone, I can predict roughly one third of the cross-regional efficiency differences in Indonesia.

Related Literature This paper is mostly related to the literature on long-run productivity differences across countries (Banerjee and Duflo, 2005; Caselli, 2005; Klenow and Rodriguez-Clare, 1997). While a large literature aimed to explain these differences in the context of an aggregate production function,³ recently various contributions focused explicitly on the behavior of individual firms and their interaction. On the one hand, Hsieh and Klenow (2009) and Restuccia and Rogerson (2008) argue that misallocation of resources across individual firms can have sizable negative effects on aggregate TFP but are agnostic about the reasons for this misallocation. On the other hand, there are papers taking a more structural perspective in that they explicitly model the microstructure of the economy. Acemoglu, Antras, and Helpman (2007) for example argue that contractual incompleteness reduces the incentives to invest into technology and Lagos (2006) considers a labor market with search frictions and shows that aggregate TFP depends on the primitives of the underlying search technology. Among others Buera, Kaboski, and Shin (2009), Midrigan and Xu (2010), Jeong and Townsend (2007) and Moll (2010) consider dynamic models of entrepreneurship and study how capital market imperfections affect aggregate productivity by distorting the allocation of talent across occupations and the allocation of capital across firms.

This paper tries to build a bridge between these two approaches. While my model is structural in that firm-level distortions are generated endogenously, the reduced form of my model looks exactly like the economy considered in Hsieh and Klenow (2009) or Restuccia and Rogerson (2008). The endogenous distribution of mark-ups therefore provides one microfoundation for the distribution of “firm-specific taxes” employed in the literature. Furthermore, the empirical analysis also identifies firm-specific mark-ups as a “wedge” (Chari, Kehoe, and McGrattan, 2007) but then uses the structural model to test for specific properties of the distribution of wedges, which are implied by the theory.

One very interesting recent paper, which is particularly relevant, is Epifaniy and Ganica (forthcoming). They also consider an economy with heterogeneous mark-ups, show that only particular moments of the mark-up distribution affect welfare and then study the welfare effect of trade liberalization in such a framework. In particular, they find that if trade opening reduces mark-ups in sectors, whose mark-up was already below average to begin with, opening up to trade might actually decrease welfare as the higher dispersion in mark-ups reduces TFP. In particular, the result that it is not the level but only the dispersion in mark-ups, which is relevant for aggregate productivity is also contained in their paper. While they also discuss the link between the distribution of mark-ups and firm-level distortions, their paper is different in many respects. Most importantly, the distri-

³To name a few, Hall and Jones (1999) and Acemoglu, Johnson, and Robinson (2001) argue in favor for the importance of social capital or institutions, Lucas (1988, 1990) stresses the importance of human capital externalities, Parente and Prescott (1994) focus on barriers of technology adoption and Ganica and Zilibotti (2009) identify an important role for the direction of technological progress.

bution of mark-ups is entirely exogenous, while my paper is mostly about a microfoundation for the endogenous stationary distribution of mark-ups. This is not only informative from a theoretical point of view, but it is also this equilibrium relationship between the entry intensity and the degree of misallocation which motivates my empirical analysis using firm-level data.⁴ Finally, they only consider a static economy, while the model in this paper is dynamic to also study the effect of heterogeneous mark-ups on the incentives to accumulate capital.

The structure of the paper is as follows. In the next section I present the basic model and show that the impact of mark-ups on the aggregate economy is summarized by two sufficient statistics. The first one is exactly the difference between equilibrium TFP and TFP in the first best allocation, the second one is the difference between each production factors' marginal product and the equilibrium factor price. I then turn to the dynamic evolution of equilibrium mark-ups. As mark-ups depend only on the cross-sectional distribution of productivity, this requires a model how relative technologies evolve. I consider a model with both entry and productivity improvements by incumbent firms, derive the stationary distribution of mark-ups in closed form and show how this distribution is fully determined by the equilibrium entry intensity. To endogenize this entry intensity as an equilibrium object, I then provide a simple microfoundation in the context of a Schumpeterian growth model. Section three is devoted to the empirical analysis, which proceeds in three steps. I first show that the entry intensity affects the two sufficient statistics as the theory predicts. Then I test four additional implications of the theory. Finally I consider competing explanations for the relationship between the entry rate and the degree of misallocation. After quantifying my empirical findings in terms a brief welfare calculation in Section four, I offer some concluding remarks in Section five. An appendix gathers the mathematical details and contains various robustness checks for the empirical findings.

2 The Model

My model has two main aspects to study the effects of monopolistic market power on aggregate statistics like TFP and factor prices. The first one concerns the *static* allocation, i.e. how do factor prices and TFP depend on the distribution of mark-ups across producers. The second one concerns the *dynamic* allocation, i.e. how do mark-ups evolve and how does the stationary distribution of mark-ups look like. I will first describe the static equilibrium and for notational simplicity I will drop time subscripts, whenever this does not cause any confusion.

2.1 Static Allocations: Heterogeneous Mark-ups and Misallocation

I consider a economy which produces a unique final good using a continuum of intermediary inputs according to

$$Y = \exp \left(\int_0^1 \ln \left(\sum_{j \in S_\nu} y_j(\nu) \right) d\nu \right), \quad (1)$$

where $y_j(\nu)$ is the quantity of the intermediary variety ν bought from producer j . Hence S_ν denotes the number of firms active in sector ν . It is clear from (1) that varieties of different intermediaries ν

⁴While Epifaniy and Ganica (forthcoming) provide evidence that mark-ups differ across sectors, they do not empirically analyze the effect of trade on the distribution of mark-ups and on social welfare. Using the empirical approach laid out in this paper, this analysis could be done if firm-level data and data on the “import penetration” across sectors was available.

and ν' are imperfect substitutes, whereas there is perfect substitutability between different brands within a variety. As will be seen below, this formulation generates non-constant equilibrium mark-ups in a very tractable way and is for example also used in Bernard, Eaton, Jensen, and Kortum (2003) and Acemoglu and Akcigit (2009). However, nothing substantial hinges on the Cobb-Douglas specification. I could consider a general CES production function but it would make the analysis more tedious (see also the discussion in footnote 5). The production of intermediaries is conducted by heterogeneous firms and requires capital and labor. The only source of heterogeneity across firms is their factor-neutral productivity. In particular, a firm in sector ν with current productivity q produces output according to a standard Cobb-Douglas production function

$$f(k, l; q) = qk^\alpha l^{1-\alpha}. \quad (2)$$

The market for intermediary goods is monopolistically competitive so that firms take the aggregate price index as given. Firms offering the same variety compete à la Bertrand. Let there be S_ν competing firms producing variety ν . Given that production takes place with a constant returns to scale technology and different brands of variety ν are perceived as perfect substitutes, in equilibrium only the most productive intermediary firm will be active. However, the presence of competing producers (even though they are less efficient) imposes a constraint on the quality leader's price setting. Factor markets are assumed to be competitive, i.e. all intermediary firms take wages w and the rental rate of capital R as given. From (2) it then follows immediately that the marginal cost of production of a firm with productivity q is given by

$$MC(q) = \left(\frac{w}{1-\alpha}\right)^{1-\alpha} \left(\frac{R}{\alpha}\right)^\alpha \frac{1}{q} \equiv \frac{\psi(R, w)}{q}. \quad (3)$$

Now consider variety ν . Given intermediary prices $\{p_j(\nu)\}_{j=1}^{S_\nu}$, the demand for intermediaries is given by

$$y(\nu) = \frac{Y}{p(\nu)}, \quad (4)$$

where $p(\nu) \equiv \min_{j \in S_\nu} \{p_j(\nu)\}$. As the demand function (4) has unitary elasticity, the most efficient firm has to resort to limit pricing where it will price its products equal to the marginal costs of the second most productive firm, which I will refer to as the *follower*. Hence,

$$p(\nu) = \frac{\psi(R, w)}{q_F(\nu)}, \quad (5)$$

where $q_F(\nu)$ is the follower's productivity.⁵ The equilibrium mark-up is therefore given by

$$\xi(\nu) = \frac{p(\nu)}{MC(\nu)} = \frac{q(\nu)}{q_F(\nu)}. \quad (6)$$

⁵It is at this point where the assumption of the aggregate production function being Cobb-Douglas simplifies the exposition because the producing firm will always be forced to use limit pricing. If the demand elasticity was to exceed unity, the firm might want to set the unconstrained monopoly price in case its productivity advantage over its closest competitor is big enough. To see this, suppose that $Y = \left(\int_0^1 y(\nu)^{\frac{\sigma-1}{\sigma}} d\nu\right)^{\frac{\sigma}{\sigma-1}}$. The unconstrained monopoly price is given by $p^M = \frac{\sigma}{\sigma-1} MC = \frac{\sigma}{\sigma-1} \frac{\psi}{q}$, so that the optimal price is given by $p = \min\left(\frac{\sigma}{\sigma-1} \frac{\psi}{q}, \frac{\psi}{q_F}\right) = \frac{\psi}{q_F} \min\left(\frac{\sigma}{\sigma-1} \frac{q_F}{q}, 1\right)$. Hence, for $q \gg q_F$, the firm will not have to use limit pricing. In the limit where $\sigma \rightarrow 1$, we get $\min\left(\frac{\sigma}{\sigma-1} \frac{q_F}{q}, 1\right) = 1$, which yields (5). See also the discussion in Bernard, Eaton, Jensen, and Kortum (2003, p. 1274).

Note in particular that the mark-up is not constant but depends on the productivity advantage of the producing firm relative to its closest competitor - a higher quality advantage shields the current producer from competition and allows him to post a higher mark-up. Equilibrium demand is therefore given by (see (4) and (5))

$$y(\nu) = \frac{Y}{\psi(R, w)} q_F(\nu) = \frac{Y}{\psi(R, w)} \frac{q(\nu)}{\xi(\nu)}, \quad (7)$$

so that the the producer's factor demands are

$$k(\nu) = \frac{1}{\xi(\nu)} \frac{\alpha Y}{R} \text{ and } l(\nu) = \frac{1}{\xi(\nu)} \frac{(1 - \alpha) Y}{w}. \quad (8)$$

Finally, the producing firm's profits are given by

$$\pi(\nu) = (p(\nu) - MC(\nu)) y(\nu) = \left(1 - \frac{1}{\xi(\nu)}\right) Y. \quad (9)$$

The expressions above show that the decision-relevant sector-specific state variables are the leaders' and follower's productivity $q(\nu)$ and $q_F(\nu)$ and that the cross-sectional variation in profits and factor demands can be entirely traced back to firm-specific mark-ups. Furthermore it is precisely these varying mark-ups which induce an inefficient allocation of resources across plants and hence have detrimental effects on aggregate TFP. This is most clearly seen from (8), which shows that producers' factor demands depend on their mark-up $\xi(\nu)$ and *not* on their productivity $q(\nu)$.

The General Equilibrium To characterize the general equilibrium of this economy, let $K(t)$ be the aggregate supply of capital and suppose that labor L is supplied inelastically. Aggregate output can then be written as (see Appendix)

$$Y(t) = Q(t) M(t) K(t)^\alpha L^{1-\alpha}, \quad (10)$$

where

$$Q(t) = \exp\left(\int_0^1 \ln(q(\nu, t)) d\nu\right),$$

is the CES aggregate of the sectoral productivity levels and

$$M(t) = \frac{\exp\left(\int_0^1 \ln(\xi(\nu, t)^{-1}) d\nu\right)}{\int_0^1 \xi(\nu, t)^{-1} d\nu} = \frac{\exp\left(E\left[\ln(\xi(\nu, t)^{-1})\right]\right)}{E\left[\xi(\nu, t)^{-1}\right]} \quad (11)$$

is the *TFP distortion index*. Hence, at the aggregate level, this economy looks exactly like the canonical neoclassical growth model where aggregate TFP is determined both by physical productivity $Q(t)$ and the monopolistic mark-ups summarized in the TFP distortion index $M(t)$. Furthermore, it can be easily verified, that given a productivity distribution $[q(\nu, t)]_{\nu=0}^1$ and level of capital $K(t)$, the efficient allocation of labor and capital across producers implies that first-best aggregate output is given by $Y^{FB}(t) = Q(t) K(t)^\alpha L^{1-\alpha}$, so that

$$Y(t) = Y^{FB}(t) M(t).$$

$M(t)$ therefore measures precisely the aggregate efficiency losses of monopolistic pricing. Monopolistic mark-ups also manifest themselves in equilibrium factor prices, as these are given by

$$R(t) = \frac{\alpha Y(t)}{K(t)} \left[\int_0^1 \xi(\nu, t)^{-1} d\nu \right] \equiv \frac{\alpha Y(t)}{K(t)} \Lambda(t) = R^{FB}(t) M(t) \Lambda(t), \quad (12)$$

$$w(t) = \frac{(1-\alpha) Y(t)}{L} \left[\int_0^1 \xi(\nu, t)^{-1} d\nu \right] \equiv \frac{(1-\alpha) Y(t)}{L} \Lambda(t) = w^{FB}(t) M(t) \Lambda(t), \quad (13)$$

where $w^{FB}(t)$ and $R^{FB}(t)$ denote the first-best equilibrium factor prices in the absence of monopolistic power and

$$\Lambda(t) = \int_0^1 \xi(\nu, t)^{-1} d\nu = E \left[\xi(\nu, t)^{-1} \right] \quad (14)$$

is the *factor price distortion index*. This term stresses that monopolistic pricing reduces equilibrium factor prices for two reasons.⁶ Not only is aggregate TFP (and hence each production factors' marginal product) lower compared to the competitive economy, but even conditional on aggregate TFP, equilibrium factor prices are below their social marginal products as (12) and (13) show that⁷

$$\frac{R(t)}{MPK(t)} = \frac{w(t)}{MPL(t)} = \int_0^1 \lambda^{-\Delta(\nu, t)} d\nu = \Lambda(t) < 1. \quad (15)$$

Hence, $M(t)$ and $\Lambda(t)$ are sufficient statistics for the TFP and factor price consequences of monopolistic power. I therefore refer to $M(t)$ and $\Lambda(t)$ as measuring the degree of misallocation. As both $M(t)$ and $\Lambda(t)$ are fully determined from the distribution of mark-ups, Proposition 1 contains some properties of how this distribution affects aggregate TFP and factor prices in this economy.

Proposition 1. *Let the distribution of productivity $[q(\nu, t)]_{\nu=0}^1$ and the level of capital $K(t)$ be given, let $Z(\xi, t)$ be a distribution of mark-ups at time t and let $M_Z(t) = \frac{\exp(E_Z[\ln(\xi^{-1})])}{E_Z[\xi^{-1}]}$ and $\Lambda_Z(t) = E_Z[\xi^{-1}]$ be the respective distortion indices if ξ is distributed according to Z . Then:*

1. *Any dispersion in mark-ups reduces aggregate TFP relative to the first best, i.e.*

$$M_Z(t) \leq 1 \text{ for all } Z, \quad (16)$$

and (16) holds with equality if and only if Z is degenerate.

2. *Increasing all mark-ups by a factor $\kappa > 1$ reduces factor prices by a factor κ but leaves TFP unchanged.*

⁶This discussion takes the level of capital $K(t)$ as given. Clearly, these distorted factor prices have a third, dynamic effect in that capital accumulation is not efficient. In particular, as the equilibrium return to capital is lower than its efficient counterpart, capital accumulation will be inefficiently low. I will come back to these dynamic aspects below, when I discuss the dynamic evolution of the economy (see Section 2.2).

⁷In particular, equilibrium interest rates will be depressed relative to the interest rate as imputed from the usual growth accounting calibration. More precisely: A researcher using the aggregate data generated by this economy would equate the interest rate with the marginal product of capital (net of depreciation), i.e. from (10) she takes the interest rate to be $r(t) = \frac{\partial F(K, L)}{\partial K} - \delta = \frac{\alpha Y(t)}{K(t)} - \delta$. However, the actual equilibrium interest rate is multiplied by the term $\Lambda(t)$ and hence lower. From an accounting point of view this is a desirable feature, because the mark-up distortions break the link between the average and marginal product of capital, which usually causes implied interest rates to be counterfactually high (see for example Banerjee and Duflo (2005); Lucas (1990))

3. *A mean preserving spread of the logarithm of mark-ups reduces TFP but leaves factor prices unchanged.*

Proof. The first part follows directly from (11). For the second part, note that $M(t) = \frac{\exp(E_Z[\ln(\xi^{-1})])}{E_Z[\xi^{-1}]} = \frac{\exp(E_Z[\ln((\kappa\xi)^{-1})])}{E_Z[(\kappa\xi)^{-1}]}$ and that factor prices are proportional to $\exp(E_Z[\ln(\xi^{-1})]) = \kappa \exp(E_Z[\ln((\kappa\xi)^{-1})])$. For the third part, note that $\exp(E_Z[\ln(\xi^{-1})]) = \exp(-E_Z[\ln(\xi)])$ is not affected by a mean-preserving spread. That $\Lambda_G(t) > \Lambda_F(t)$ if G is a mean-preserving spread of F follows from the definition of $\Lambda(t)$ and Jensen's Inequality. $\square$

Proposition 1 shows that the location and the dispersion of the mark-up distribution have very different impacts on the economy. In particular, factor prices are entirely insensitive with respect to a higher dispersion of the underlying mark-up distribution - they only depend on the mean. A shift in the level of mark-ups will therefore reduce factor prices and increase equilibrium profits. For the case of aggregate TFP, exactly the opposite is true: whereas a higher dispersion of mark-ups will show up as a lower TFP for the aggregate economy, TFP is not affected by level shifts. Intuitively, aggregate TFP is determined by the allocation of resources across firms. If all mark-ups increase by a constant proportion, relative prices will be unaffected and hence still provide the appropriate signals about firms' relative productivity. Note that the canonical case of constant mark-ups as generated by a CES demand system is contained in parts 1. and 2. of Proposition 1: TFP will be identical to its efficient competitive counterpart but monopolistic power drives down equilibrium factor prices. Proposition 1 can be seen as the dynamic version of Proposition 1 in Epifaniy and Ganica (forthcoming, p. 14). While in their static model, *welfare* is homogeneous of degree zero in the level of mark-ups, in my dynamic setting the same holds true for aggregate TFP. Welfare however is still affected by the level of mark-ups as dynamic trade-offs will be affected. This will be discussed below.

While Proposition 1 is useful in that it concisely shows where monopolistic distortions manifest themselves in the aggregate economy, the distribution of relative productivity (and hence mark-ups) is still entirely exogenous. This however is unsatisfactory because the distribution of productivity across firms is likely to be endogenous if firms have the opportunity to innovate or to adopt new technologies. The rest of this paper is therefore concerned with the endogenous evolution of firms' relative productivity. Before explicitly turning to the dynamic determination of productivity, let me briefly relate this model to the recent literature on misallocation and firm-level distortions.

Monopolistic mark-ups as firm-level distortions The TFP consequences of resource misallocation across heterogeneous firms have recently been analyzed in Hsieh and Klenow (2009). Rather than proposing a structural model of misallocation, Hsieh and Klenow (2009) follow Chari, Kehoe, and McGrattan (2007) and Restuccia and Rogerson (2008) by accounting for the degree of misallocation through exogenous firm-specific taxes. In particular, they identify firm-level distortions from the concept of revenue total factor productivity, which is defined as $TFPR = \frac{py}{K^\alpha L^{1-\alpha}}$. In their framework $TFPR$ is proportional to $\frac{(1+\tau_K(\nu,t))^\alpha}{1-\tau_Y(\nu,t)}$, where $\tau_K(\nu,t)$ and $\tau_Y(\nu,t)$ are the *exogenous* firm-specific taxes on capital and output. Hence, if there was no misallocation of resources, revenue total factor productivity should be equalized across firms. In my economy, $TFPR$ is given by $TFPR(\nu,t) = \frac{p(\nu,t)y(\nu,t)}{k(\nu,t)^\alpha l(\nu,t)^{1-\alpha}} \propto \xi(\nu,t)$. Hence, $TFPR$ is also not equalized across firms but is proportional to the mark-up. If we were to stop here, the only progress made was a relabeling of what the literature calls "distortions". However, by interpreting the distribution of mark-ups

$[\xi(\nu, t)]$ through the lens of a structural model, these are determined endogenously and we can therefore trace their evolution back to the underlying primitives of the economy. It is in that sense that this model generates firm-level distortions endogenously.

2.2 Dynamics: The Endogenous Evolution of Mark-Ups

Proposition 1 shows that the distribution of mark-ups affects aggregate TFP and equilibrium factor prices. Mark-ups in turn are given by $\xi(\nu, t) = \frac{q(\nu, t)}{q_F(\nu, t)}$ and are therefore fully determined by relative productivity across firms. The degree of misallocation across firms is therefore generated endogenously through the competition between producers of different productivity. Hence, *any* theory about the determination of firm-level productivity will also characterize the degree of misallocation reflected in the terms $M(t)$ and $\Lambda(t)$. To be more precise: The crucial aspect of the innovation environment is the *relative* speed with which leading firms improve their productivity relative to their most advanced potential competitors. If the leading firm innovates faster than its potential competitors, it will increase its productivity advantage, which in turn will allow for higher mark-ups. If on the other hand less productive firms increase their productivity faster, competition will tighten and aggregate TFP and factor price will tend to increase.⁸

To put more structure on these productivity differences across firms, suppose that firm productivity evolves on a quality-ladder (Aghion and Howitt, 1992; Grossman and Helpman, 1991). Formally, if a firm in sector ν has had $n(\nu, t)$ innovations in the interval $[0, t]$, its productivity is given by $q(\nu, t) = \lambda^{n(\nu, t)}$, where $\lambda > 1$. Hence, innovations are assumed to lead to proportional productivity improvements. This specification of the productivity process is convenient, because it implies that monopolistic mark-ups can be expressed as

$$\xi(\nu, t) = \frac{q(\nu, t)}{q_F(\nu, t)} = \frac{\lambda^{n_L(\nu, t)}}{\lambda^{n_F(\nu, t)}} = \lambda^{\Delta(\nu, t)}, \quad (17)$$

where $\Delta(\nu, t) \equiv n_L(\nu, t) - n_F(\nu, t) \geq 0$ is the quality gap between the leader and the follower. Monopolistic power can therefore be expressed by a single sufficient statistic - the productivity gap Δ . Intuitively: firms are able to set high prices if they managed to climb the quality ladder faster than their potential competitors. This formulation in particular implies that the two sufficient aggregate statistics $M(t)$ and $\Lambda(t)$ are also fully determined from the cross-sectional distribution of productivity gaps as (11) and (14) imply that

$$\begin{aligned} M(t) &= \frac{\exp\left(E\left[\ln\left(\xi(\nu, t)^{-1}\right)\right]\right)}{E\left[\xi(\nu, t)^{-1}\right]} = \frac{\lambda^{E[-\Delta]}}{E[\lambda^{-\Delta}]} \\ \Lambda(t) &= E[\lambda^{-\Delta}]. \end{aligned}$$

So how does firms' relative productivity evolve? In this paper, I focus on an admittedly special but nevertheless important case in that I assume that innovations can stem from two sources: either

⁸In fact, it is precisely the latter intuition, which Epifaniy and Ganica (forthcoming) seem to have in mind, when they argue that exposure to international trade reduces equilibrium mark-ups by making the market more competitive. In the framework of this paper, this intuition would be formalized by assuming that the productivity difference between the currently producing firm and the most efficient domestic competition exceeds the one between the current producer and exporting firms from overseas. Opening up to trade reduces the mean distortion but if the exposure to trade is heterogeneous, the distribution of mark-ups in the economy might become more dispersed. Proposition 1 then shows that equilibrium factor prices will increase, while the effect on TFP is ambiguous.

current producers (i.e. leading quality firms) can experience technological improvements, or new entrants can enter the market by coming up with new blueprints and thereby replace the current producer. I decided to focus on the process of entry for two reasons. The first reason is that entry is often associated with a competition-enhancing role (Aghion and Bolton, 1987; Blanchard and Giavazzi, 2003). The second reason is the empirical content of the model. In the model, the productivity gap depends on the productivity of the followers, i.e. firms who *do not* produce. Productivity improvements by these firms are clearly not observable. Entry on the other hand is observable and I can measure it in the empirical analysis in Section 3. However, I will point out below in how far the model can easily be thought of as a model of learning or technological diffusion.

Given the quality ladder structure of productivity, it is straight forward to model technological progress. If a producer with current productivity q experiences an innovation, he reaches the next step of the quality ladder so his new productivity is given by λq . For the case of entry, suppose that leading technologies are common knowledge for the process of innovation so that the entrant in sector ν enters the market with productivity $\lambda q(\nu, t)$, where $q(\nu, t)$ is the leading quality in sector ν . This formulation is appealing in the current context, because it focuses on the different allocational consequences of the two sources of productivity improvements. While both entrants and incumbents increase the frontier technology by the same amount (in both cases, the productivity of the most efficient producer increases by a factor of λ), the implications for equilibrium mark-ups and allocational efficiency are very different. In case the innovation stems from the current producer of variety ν with a quality advantage of $\Delta(\nu, t)$, the equilibrium mark-up for that variety *increases* by a factor λ . If on the other hand, productivity growth will be induced by entry, the equilibrium mark-up for variety ν *decreases* by the factor $\lambda^{\Delta(\nu, t)-1}$, as the new entrant only has a productivity advantage one.⁹ It is in this sense that this formulation focuses on the main allocational trade-off of productivity growth - the relative technological lead and hence the equilibrium mark-up depends crucially on the source of the technological advancement.

Suppose that incumbents generate an innovation with flow rate I , where I is constant over time and across incumbent firms. Similarly, suppose that each variety ν experiences entry at the constant rate z . These are of course equilibrium objects and I will provide a microfoundation in Section 2.3 below. However, it is useful to characterize the equilibrium allocations as a function of (I, z) to stress that whatever the particular microfoundation, it is these two equilibrium outcomes, which determine aggregate efficiency by shaping the distribution of productivity gaps and hence mark-ups. In particular, given (I, z) this distribution will be stationary and has a closed form representation. To see this, note that the productivity gap distribution is fully characterized by the collection $\{\mu(\Delta, t)\}_{\Delta=1}^{\infty}$, where $\mu(\Delta, t)$ denotes the measure of sectors with quality gap Δ at time t . These measures solve the flow equations

$$\dot{\mu}(\Delta, t) = \begin{cases} -(z + I)\mu(\Delta, t) + I\mu(\Delta - 1, t) & \text{if } \Delta \geq 2 \\ -I\mu(1, t) + z(1 - \mu(1, t)) & \text{if } \Delta = 1 \end{cases} \quad (18)$$

To understand these flow equations, consider first the upper branch. There are two ways to leave the state (Δ, t) : the current producer could have an innovation or there could be entry. The only way to get into this state is by being in state $\Delta - 1$ and then having an innovation. Similarly, all firms in state $(1, t)$ exit this state if they have an innovation and all sectors where entry occurs will enter the state $(1, t)$. This stochastic process has a unique stationary distribution $\{\mu(\Delta)\}_{\Delta=1}^{\infty}$,

⁹Note that the continuous time formulation of the model precludes the possibility that a variety experiences both entry and a productivity improvement by the current producer, which is of second order.

which is characterized by the condition that $\dot{\mu}(\Delta, t) = 0$ for all Δ . From (18) it then follows directly that this stationary distribution is given by

$$\mu(\Delta) = \left(\frac{I}{z + I} \right)^\Delta \frac{z}{I} \equiv \left(\frac{1}{1 + x} \right)^\Delta x, \quad (19)$$

where the second equality shows that the mark-up distribution is only dependent on the *entry intensity* defined by $x = \frac{z}{I}$. This is particular convenient, because both the cumulative distribution of productivity gaps F and the two distortion indices M and Λ will be constant and are entirely determined by the entry intensity x . As shown in the Appendix, these are given by

$$F(\Delta; x) = \sum_{i=1}^{\Delta} \mu(i) = 1 - \left(\frac{1}{x + 1} \right)^\Delta \quad (20)$$

$$\Lambda(x) = E[\lambda^{-\Delta}] = \frac{x}{(\lambda - 1) + \lambda x} \quad (21)$$

$$M(x) = \frac{\lambda^{-E[\Delta]}}{E[\lambda^{-\Delta}]} = \lambda^{-\frac{x+1}{x}} \frac{(\lambda - 1) + \lambda x}{x}. \quad (22)$$

Using (20)-(22) we can gather the main comparative static effects with respect to the entry intensity in the following proposition.

Proposition 2. *Consider the economy described above. For a given distribution of productivity $[q(\nu, t)]_\nu$ and level of capital $K(t)$, a higher entry intensity x*

1. *reduces equilibrium mark-ups in a first order dominance sense, as $F(\Delta; x)$ is increasing in x for all Δ*
2. *increases aggregate TFP and aggregate income*
3. *increases equilibrium factor prices*
4. *reduces the wedge between factor prices and their marginal product*

Proof. The proposition follows directly from (20)-(22), (10), (12), (13) and (15).¹⁰ □

Proposition 2 shows the importance of the allocation of innovative resources between entrants and incumbents and especially the efficiency-enhancing role of entry: In this economy both types of innovation bring about the same productivity gain but entry reduces factor misallocation by limiting monopolistic pricing. This formalizes the intuition that both TFP and equilibrium factor

¹⁰To see part 2., define $s = x^{-1}$ so that $M(s) = \lambda^{-(1+s)}((\lambda - 1)s + \lambda)$. $M(s)$ is decreasing in s if

$$\frac{\partial \ln(M^{BGP}(s))}{\partial s} \equiv h(\lambda, s) = \frac{(\lambda - 1)}{(\lambda - 1)s + \lambda} - \ln(\lambda) < 0.$$

As h is decreasing in both s and λ , it is sufficient to note that $h(\lambda, s) < h(1, 0) = 0$ and $\frac{\partial h(\lambda, 0)}{\partial \lambda} = \frac{1}{\lambda^2}(1 - \lambda) < 0$ for all $\lambda > 1$.

prices are especially low in economies which are characterized by an environment where the entry intensity x is low.¹¹

Furthermore, Proposition 2 also has dynamic implications for the incentives to accumulate capital and hence the long-run capital level. The equilibrium growth rate of the economy along a balanced growth path (BGP) is given by

$$g_Y = \frac{\dot{Y}(t)}{Y(t)} = \frac{\dot{Q}(t)}{Q(t)} + \alpha \frac{\dot{K}(t)}{K(t)} = g_Q + \alpha g_K = \frac{g_Q}{1 - \alpha},$$

where the last equality uses the fact that along a balanced growth path we have $g_Y = g_K$. Given the constant innovation and entry rates (I, z) , I show in the Appendix that productivity Q grows at a constant rate, which is given by

$$g_Q = \ln(\lambda)(I + z). \quad (23)$$

The equilibrium growth rate (23) simply formalizes the fact that entry and incumbents' innovation increase the productivity of the economy by the same factor λ and occur with a flow rate I and z respectively. To solve for the long-run capital stock of the economy, suppose that the economy admits the representation of a representative household who discounts the future at rate ρ and assume for simplicity that the per-period utility function is given by $u(c) = \ln(c)$. Standard arguments then show that the BGP level of capital in efficiency units is given by

$$\tilde{k} = \frac{K(t)}{LQ(t)^{1/(1-\alpha)}} = \left(\frac{\alpha \Lambda(x) M(x)}{g_Y + \delta + \rho} \right)^{\frac{1}{1-\alpha}} = \left(\frac{\alpha}{g_Y + \delta + \rho} \right)^{\frac{1}{1-\alpha}} \lambda^{-\frac{1}{1-\alpha} \frac{x+1}{x}}, \quad (24)$$

Hence: conditional on the growth rate g_Y (which of course also depends on the innovation and entry rates), the BGP level of capital is increasing in the entry intensity. Intuitively: if the entry component to generate a growth rate of g_Y is higher, there will be more competitive pressure. This in turn increases the demand for capital, hence also the equilibrium price of capital (the rental rate) and therefore the incentive to accumulate.

It is worth noting that the economic channel of this efficiency benefits of entry is very different from the ones stressed in e.g. Barseghyan and DiCecio (2009) or Buera, Kaboski, and Shin (2009). In these papers, inefficient entry is also at the heart of aggregate TFP losses but the mechanism is distinct. Barseghyan and DiCecio (2009) use a model a la Hopenhayn (1992) and show that higher entry costs lower aggregate TFP because it reduces the productivity of the marginal entrant as lower factor prices make production profitable even for relatively unproductive firms. Similarly, Buera, Kaboski, and Shin (2009) argue that credit market imperfections distort the allocation of talent across sectors if sectors are characterized by different entry costs. Hence, both of these papers link the costs of entry to the resulting equilibrium distribution of the productivity of active producers. This paper supplements these ideas by showing that lower entry intensities will have negative TFP consequences *conditional on the distribution of productivity and capital* simply by affecting the stationary distribution of mark-ups and with it allocational efficiency.

¹¹From this discussion it is easily seen how we can also think about learning or technological diffusion shaping the distribution of mark-ups. Suppose for simplicity that potential frontier technologies were growing at an exogenous rate and call $z + \theta$ the diffusion rate among current producers and z the diffusion rate applying to non-producing firms. If production and the adoption of new techniques are complements, we would expect that $\theta > 0$. If on the other hand considerations of say vintage capital are important, we would expect $\theta < 0$ as current producers have capital suited to the old technology in place. This model is of course identical to the one above with $z + \theta = I$.

2.3 A microfoundation for innovation and entry

Section 2.2 above traced the aggregate effect of mark-up distortions back to the equilibrium entry intensity. However, the equilibrium levels of innovation and entry were taken as given. I will now provide a microfoundation, which determines the entry intensity as an equilibrium outcome and therefore allows me to study its dependence on the economic fundamentals in particular the entry costs, the efficiency of the innovation technology and the amount of innovative resources the economy has access to. Specifically, I model the evolution of productivity as being governed by a Schumpeterian process of creative destruction. To do so I assume that both incumbents' productivity increases and entrants' new blueprints are brought about by scientists (or innovators), which are in fixed supply, say L_S .¹² The innovation technology is the following. Consider first the case of innovation by incumbents. As seen from (8) and (17), the quality gap Δ is the unique state variable for the static problem of the current producer. To also be able to characterize the dynamic problem of the firms using a single state variable only, I assume that if a firm with quality advantage Δ wants to achieve a flow rate of I , it has to hire $\Gamma(I, \Delta)$ innovators. I will be more specific about the functional form of Γ below. If the current wage rate of innovators is $w_I(t)$, it thus costs $w_I(t) \Gamma(I, \Delta)$ to achieve a flow rate of innovation of I . Now consider the process of entry. New blueprints are also brought about by innovators. In accordance with the formulation above I adopt the usual institutional structure of Schumpeterian growth models: whereas there is patent protection for the usage of technologies (i.e. no other firm than the "inventor" can use the technology q in production), the leading technologies are common knowledge for the process of innovation, i.e. innovators can try to improve upon them. More specifically: if successful, the innovator gets a blueprint of quality $\lambda q(\nu, t)$, where $q(\nu, t)$ is the leading quality in sector ν so that the entrant has a productivity advantage of $\Delta = 1$. The innovation technology of entrants is assumed to be linear, i.e. each innovator generates a new blue print with flow rate η . Conditional on success, the innovator can use the generated blueprint after paying an "entry costs" $(1 - \chi)V$, where V is the (endogenously determined) value of producing with the new blueprint. These entry costs can be either thought of various bureaucratic costs (Djankov, Porta, Lopez-De-Silvaes, and Shleifer, 2002) or as one-time start-up costs, for example through the necessity of building supplier relations (Banerjee and Duffo, 2000). Hence, the innovator earns χV in case the innovation is successful.

To solve for the equilibrium innovation and entry rates, we have to solve for the equilibrium value of a firm. Let $V(\Delta, t)$ denote the value of a firm with a quality gap Δ at time t . Consider first a quality leader, i.e. $\Delta > 0$. The value function solves the Hamilton-Jacobi-Bellman (HJB)

¹²To be consistent with the derivation of the BGP capital stock (24), suppose there is perfect risk sharing across the L production workers and the L_S innovators and they share the same preferences. Then we can represent the consumption side as a representative agent who maximizes lifetime utility $U = \int_0^\infty \exp(-\rho t) \ln(c(t)) dt$ subject to the intertemporal flow constraint $\dot{c}(t) + \dot{A}(t) = (1 + r(t))A(t) + w(t)L + w_I(t)L_S$ and the transversality condition. Here $w(t)$ is the equilibrium wage for production workers, $w_I(t)$ denotes the wage rate for innovators and $A(t)$ is the household's asset position which consists both of its capital holdings and the equity in the intermediary firms. It is useful to make production workers and innovators distinct production factors. If I had assumed that there is only one type of labor, the aggregate production function in (10) would be given by $Y(t) = Q(t)M(t)K(t)^\alpha L_P(t)^{1-\alpha}$, where $L_P(t)$ are the workers allocated towards the production sector. While the allocation of workers is constant along a BGP, it would have made the analysis more tedious. Assuming that production workers and innovators are distinct, allows me to characterize the product market without recourse to the exact microfoundation of the innovation sector.

equation

$$r(t)V(\Delta, t) - \dot{V}(\Delta, t) = \pi(\Delta, t) - z(\Delta, t)(V(\Delta, t) - V(-1, t)) + \max_I \{I(V(\Delta + 1, t) - V(\Delta, t)) - w_I(t)\Gamma(I, \Delta)\},$$

where $z(\Delta, t)$ is the (endogenous) entry rate for a sector with quality gap Δ , which current incumbents take as given.¹³ The return on the “asset”, $r(t)V(\Delta, t)$, consists of the per-period dividends $\pi(\Delta, t)$ and the asset’s appreciation $\dot{V}(\Delta, t)$. Additionally, with flow rate $z(\Delta, t)$ the current leader ceases to be in that position and instead is now one quality step behind the new entrant. Hence, with flow rate $z(\Delta, t)$ a value of $V(\Delta, t) - V(-1, t)$ is destroyed. Similarly, the possibility of investing in technological improvements represents an option value. By spending $w_I(t)\Gamma(I, \Delta)$, the firm generates a surplus of $V(\Delta + 1, t) - V(\Delta, t)$ with flow rate I . As usual, the occurrence of both a successful incumbent innovation and entry is of second order.

As lagging firms do not produce, the value of not being the leading-edge firms is simply given by option of entry, so that free entry implies that $V(-1, t) = 0$. Furthermore, I assume that

$$\Gamma(I, \Delta) = \lambda^{-\Delta} \frac{1}{\varphi} I^\gamma, \quad (25)$$

where $\gamma > 1$ and φ parametrizes the efficiency of the innovation technology.¹⁴ The functional form imposed in (25) implies that innovations are easier the bigger the productivity advantage Δ . This assumption is similar in spirit to the assumption of knowledge capital made in Klette and Kortum (2004) and formalizes the notion that firms build on the shoulders of those giants, which they themselves built. Note that the only way how a firm can have a productivity advantage of Δ is if it had Δ innovations *and* there has not been any entry in between. Hence, I assume that innovation gets easier if the firm itself had multiple innovations *in a row*, which basically assumes that there is an externality between successful research conducted yesterday and the research technology today. As in Klette and Kortum (2004), this externality makes the model consistent with Gibrat’s Law, i.e. with the fact the firms’ growth rates are independent of size (Sutton, 1997; Luttmer, 2010). Substituting (25) and the profit function (9) into the HJB equation yields

$$\begin{aligned} & (r(t) + z(\Delta, t))V(\Delta, t) - \dot{V}(\Delta, t) \\ &= Y(t)(1 - \lambda^{-\Delta}) + \max_I \left\{ I(V(\Delta + 1, t) - V(\Delta, t)) - w_I(t)\lambda^{-\Delta} \frac{1}{\varphi} I^\gamma \right\}. \end{aligned} \quad (26)$$

Along the BGP, interest rates are constant and aggregate output grows at a constant rate g_Y . So let us conjecture that both innovator wages $w_I(t)$ and the value functions $V(\Delta, t)$ also grow at rate g_Y and that innovation and entry rates are constant and equal across all sectors, i.e. $z(\Delta, t) = z$ and $I(\Delta, t) = I$. These conjectures will be verified below.

To characterize the BGP, I show in the appendix, that the value function defined in (26) is given by

$$V(\Delta, t) = \frac{(1 - \lambda^{-\Delta})Y(t) + (\gamma - 1)w_I(t)\frac{1}{\varphi}I^\gamma\lambda^{-\Delta}}{\rho + z}, \quad (27)$$

¹³In general this entry rate could depend on q . However, along the BGP this will not be the case so that I drop the dependence on q from the outset.

¹⁴The functional form φI^γ is for simplicity only. All the results go through with a more general specification of $\Gamma(\Delta, I) = \lambda^{-\Delta}h(I)$, with $h'(I)\frac{I}{h(I)} > 1$. The latter condition is needed for there to exist a positive interior solution to the incumbents’ innovation problem.

where ρ is the consumer's discount rate and I is the optimal innovation rate, which is implicitly defined by

$$\frac{1}{\varphi} \left((\rho + z) \gamma I^{\gamma-1} + \frac{\lambda-1}{\lambda} (\gamma-1) I^\gamma \right) = \frac{\lambda-1}{\lambda} \frac{Y(t)}{w_I(t)}, \quad (28)$$

and z is the equilibrium entry rate. As $\frac{Y(t)}{w_I(t)}$ is constant along the BGP, (28) shows that I is also constant. As the LHS is increasing in I , I is increasing in the innovation efficiency φ . Furthermore, I is increasing in $Y(t)$ as higher demand increases monopolistic profits, and is decreasing in $w_I(t)$ and z , where the latter reflects Arrow's replacement effect: more entry reduces the innovation incentives because it is less likely that the innovator can reap the benefits of his investment in the future.

To determine the equilibrium entry intensity x , we have to consider the allocation of researchers between incumbents and entrants. When working for an existing firm, researchers can earn a wage $w_I(t)$. If they instead try to develop a new blueprint to leapfrog the current quality leader, they are successful with flow rate η in which case they get paid the value of the firm net of entry costs, $\chi V(1, t)$. Along the BGP (where $I > 0$ and $z > 0$) innovators are indifferent where to work¹⁵ and we therefore have that

$$\frac{w_I(t)}{\eta} = \chi V(1, t) = \chi \left(\frac{(1 - \frac{1}{\lambda}) Y(t) + (\gamma-1) w_I(t) \frac{1}{\varphi} I^{\gamma-1}}{\rho + z} \right), \quad (29)$$

where the second equality substitutes the value function from (27) and uses that entrants always enter with a productivity gap of $\Delta = 1$. Substituting (28) into (29) not only shows that $\frac{Y(t)}{w_I(t)}$ is constant as required, but it delivers a first relationship between the innovation intensity I and the entry rate z . The second relationship comes from the market for innovators. Equilibrium market clearing requires that

$$L_S = \int_0^1 \lambda^{-\Delta(\nu)} \varphi I^\gamma d\nu + L_E = \varphi I^\gamma \Lambda\left(\frac{z}{I}\right) + L_E = \frac{1}{\varphi} I^\gamma \frac{z}{(\lambda-1)I + \lambda z} + \frac{z}{\eta}, \quad (30)$$

where L_E is the measure of entrants, $\int_0^1 \lambda^{-\Delta(\nu)} \varphi I^\gamma d\nu$ is the aggregate innovator demand by incumbents (see (25)) and the last equality uses the fact that $\eta L_E = z$ and substitutes the equilibrium value of $\Lambda\left(\frac{z}{I}\right)$ given in (21). Using (30) we can therefore determine the equilibrium levels of I and z . According to Proposition 2, however, the crucial endogenous variable is the entry intensity x . Hence, it is useful to characterize the BGP in the (I, x) -space, instead of the (I, z) -space. (28), (29) and (30) imply that the equilibrium innovation rate I and entry intensity x are determined from the two equations

$$\frac{1}{\eta} = \chi \frac{1}{\varphi} \left[\gamma I^{\gamma-1} + \frac{(\gamma-1)}{\rho + xI} I^\gamma \right] \quad (31)$$

$$L_S = \frac{1}{\varphi} I^\gamma \frac{x}{\lambda-1 + \lambda x} + \frac{x}{\eta} I. \quad (32)$$

¹⁵Note that for (29) to be the appropriate free entry condition, I assume that innovators are able to insure the idiosyncratic risk. This could be decentralized by a "mutual fund" hiring a continuum of innovators. If there is free entry to open mutual funds, each innovator will be paid the expected return $\eta \chi V$. Note that I anticipated this result already in the formulation of the households' budget constraint in footnote 12, where total labor income generated by innovators is given by $w_I(t) L_S$.

Hence, to show existence and uniqueness of the equilibrium, we simply have to establish that (31) and (32) have a unique solution. In the Appendix I show that the free entry condition in (31) describes a schedule $x = x^{FE}(I; \eta, \varphi, \chi)$, which is upward sloping. Intuitively: (31) is derived from respective innovation incentives, i.e. the incumbents first order condition and the entrants' free entry condition. The incentives to hire innovators is increasing in $\frac{Y(t)}{w_I(t)}$ for both incumbents and entrants. Hence, I and x are positively related. That the market clearing condition (32) on the other hand describes a decreasing schedule $x = x^{MC}(I; \eta, \varphi, L_S)$ is intuitive: For a given supply of innovative resources, there has to be less entry if incumbents increase their innovation efforts. This gives rise to the following

Proposition 3. *Consider the economy described above. There exists a unique BGP where the equilibrium innovation rate $I = I(\chi, \varphi, \eta, L_S)$ and the equilibrium entry intensity $x = x(\chi, \varphi, \eta, L_S)$ are constant. Furthermore, the equilibrium innovation rate and the entry intensity satisfy*

$$\begin{aligned} \frac{\partial I(\cdot)}{\partial \chi} < 0, \frac{\partial I(\cdot)}{\partial \varphi} > 0, \frac{\partial I(\cdot)}{\partial \eta} \leq 0, \frac{\partial I(\cdot)}{\partial L_S} > 0. \\ \frac{\partial x(\cdot)}{\partial \chi} > 0, \frac{\partial x(\cdot)}{\partial \varphi} \leq 0, \frac{\partial x(\cdot)}{\partial \eta} > 0, \frac{\partial x(\cdot)}{\partial L_S} > 0. \end{aligned}$$

Proof. That there exists a unique solution to (31) and (32) is shown in the Appendix. The comparative static results are illustrated in Figure 1. See also the discussion in the main body of the text. $\square$

Proposition 3 provides the main comparative static results of the crucial determinant of the degree of misallocation with respect to different characteristics of the underlying environment. That the entry intensity is decreasing in the entry costs (i.e. is increasing in χ) follows from the free entry condition and is illustrated in the first panel of Figure 1. Intuitively: a decrease in the entry costs increases the entry incentives so that the wage for innovators has to increase. This reduces the innovation rate by incumbent firms. As innovators are in fixed supply, more entry has to occur for the market to clear. Hence, the new entry intensity will be higher. That the comparative statics with respect to the improvements of the innovation technologies φ or η are ambiguous is expected. That an increase in the productivity of the incumbents' technology increases the innovation rate I (see Panel 2 of Figure 1) and that a better technology of creating new blueprints increases the entry intensity x (illustrated in Panel 3 of Figure 1), is intuitive. The respective effect on the other source of innovation depends on the strength of the respective innovator demand effects. Holding say the innovation rate constant, a increase in φ will increase entry, because it frees up innovative resources from current producers. However, I will not be constant but also increase. Depending on which effect dominates, equilibrium entry can either increase or decrease. Finally, both the entry intensity and the innovation rate is increasing in the economy's supply of innovative resources (Panel 4). Taken together, Propositions 2 and 3 illustrate, which economies suffer from a higher degree of misallocation through dispersed mark-ups.

Corollary 4. *For a given distribution of productivity $[q(\nu, t)]_\nu$ and level of capital $K(t)$, equilibrium TFP and equilibrium factor prices are increasing in the supply of innovative resources L_S and the efficiency of entry η and decreasing in the costs of entry $1 - \chi$. Furthermore, the wedge between equilibrium factor prices and their marginal product is increasing in the entry costs and decreasing in the supply of innovation resources and the efficiency of entry.*

Note that this Corollary only focuses on the interaction between the degree of misallocation and the underlying characteristics of the innovation environment. Obviously, the equilibrium growth rate will also be affected by say the supply of innovation resources. In this paper however, I focus entirely on the link between the different sources of productivity growth the implications for the degree of misallocation. Holding aggregate productivity $Q(t)$ and capital $K(t)$ fixed, a lower intensity of entry reduces the allocative efficiency of the economy, as product markets get less competitive, the equilibrium distribution of mark-ups increases (in a distributional sense) and becomes more dispersed.

3 Empirical Analysis

In this section I will take the model to the data to test its main implications. The major obstacle is that most micro-data sets do not contain information on firms' prices so that firm-specific mark-ups are not directly observable. My dataset, which I will describe below, is no exception. Hence, I will use the structural model to infer mark-ups from observable quantities. This of course raises the issue in how far competing explanations would also be able to rationalize the data. To address this question, I explicitly consider alternative explanations in section 3.4 below. The empirical analysis will have three parts. First of all, I will test the main implications of the theory contained in Proposition 2, namely that a higher entry intensity reduces the degree of misallocation. Then I will test four additional restriction, of the underlying theory. Finally I will turn to the competing explanations. In particular, I will focus on credit market imperfections and policy distortions, as these considerations have been given substantial attention recently (Hsieh and Klenow, 2009; Moll, 2010; Buera, Kaboski, and Shin, 2009). I will also discuss in how far measurement error might be responsible for my results.

For this empirical exercise, I exploit differences in the entry intensity of the manufacturing sector across regions of Indonesia and correlate these observed entry intensities with regional differences in the degree of misallocation.¹⁶ To partly address the obvious endogeneity problems in that other regional characteristics affect both the region's entry intensity and the degree of misallocation, I augment the firm level data with data on regional characteristics and control for those in the regression analysis. The obvious missing step between the theory and the empirical application is to test for the determinants of the equilibrium entry intensity. The comparative static results contained in Proposition 3 could in principle be tested. Doing so however, would require information which is not available in the data I have. In particular, my data sources do not contain convincing proxies for the entry costs, the supply of innovation resources or the efficiency of the entry technology. Hence, I am not able to present any evidence regarding the particular microfoundation of the determinants of the entry intensity.¹⁷

¹⁶While conceptually it is useful to think of each region as a replication of the model economy characterized above, I want to stress that most parts of this exercise do not require factors to be immobile. I will come back to this point below.

¹⁷Note also that even if the data was available, tests of Proposition 3 would require more stringent assumptions on the economic environment. In particular I would have to assume that the different regions in Indonesia were autarkic in the sense that the "innovation market" clears region by region. This assumption does not seem to be valid in the Indonesian context but might be applicable to the cross-country comparison.

3.1 The Data

The empirical analysis will be based on two data sources, which are described in more detail in the Appendix. My main data set is the Manufacturing Survey of Large and Medium-Sized Firms (Statistik Industri), which for example has also been used in Amiti and Konings (2007) and Blalock, Gertler, and Levine (2008). The Statistik Industri is an annual census of all manufacturing firms in Indonesia with 20 or more employees and I will use data from 1990 to 2001.¹⁸ It contains data on the firms' revenue, labor inputs, the wage bill and the capital stock and covers roughly 20,000 firms in each year.¹⁹ Most importantly, it contains information on the geographic region the firm is located in, which allows me to calculate regional entry intensities. To do so, I identify a firm as an entrant when it appears first in the data. The regional information is recorded at the level of a regency (*kabupaten*). There are roughly 240 regencies in Indonesia. As the manufacturing sector in Indonesia is fairly concentrated, many regencies contain too few establishments for a meaningful analysis. I therefore aggregate the data at the level of the 33 provinces.

To be able to control for regional characteristics, I augment this data with information from the Village Potential Statistics (PODES) dataset in 1996. The PODES dataset contains detailed information about characteristics for all of Indonesia's about 65,000 villages. Using the village level data I then aggregate this information to the province level and match these to the firm-level data. In particular, I exploit information about the financial environment, the state of the infrastructure, the sectoral composition and the density of small, informal firms, which are not in the Statistik Industri data but could of course still exert competitive pressure on the official manufacturing sector. Controlling for these factors should at least alleviate the most pressing concerns about omitted variables, but in the absence of a convincing instrument for the regional entry intensity I will not be able to rule out endogeneity bias entirely.

3.2 Main Implications of the Model

As stressed in Proposition 2, the primitive determinant of the degree of misallocation is the distribution of productivity gaps $\mu(\Delta)$, which is entirely determined by the (endogenous) entry intensity. In particular, according to the model, it is only the two moments

$$\Lambda = E[\lambda^{-\Delta}] \text{ and } \Psi = \lambda^{-E[\Delta]} \quad (33)$$

which affect aggregate allocations, as the two distortion indices $\Lambda = E[\lambda^{-\Delta}]$ and $M = \frac{\lambda^{-E[\Delta]}}{E[\lambda^{-\Delta}]}$ are sufficient statistics for the effect of the distribution of productivity gaps on the aggregate economy. The productivity gap Δ , which in the model varies in the cross-section of firms, is of course not directly observable. However, according to the model $\lambda^{-\Delta}$ is identified from the *firm-specific labor-shares*. In particular, (8) implies that

$$\frac{wl}{py} \frac{1}{1-\alpha} \equiv \rho(\Delta) = \lambda^{-\Delta}. \quad (34)$$

¹⁸As Indonesia experienced a substantial financial crisis in the late 90s, I exclude the data from 1998 (the main year of the crisis) from the main analysis, but I also report the results including this crisis year. The results are qualitatively similar.

¹⁹To be absolutely precise, the data is collected at the plant level. As more than 90% of the plants report to be single branch entities, I will for the following refer to each plant as a firm. In the context of the model, this distinction is important in that different plants within the same firm are unlikely to compete against each other on product markets.

Intuitively, (34) states that the marginal *revenue* product of labor (which in this setup is proportional to the average product) should be equalized across firms, if it was not for the monopolistic pricing rule as $\lambda^{-\Delta}$ is equal to the (inverse of) the monopolistic mark-up.²⁰ Importantly, $\rho(\Delta)$ is observable in the micro data up to scale. Setting $\alpha = 0.7$ as a normalization, Λ is directly identified from the data.²¹ If λ was known, Ψ could also be directly calculated from (34) as

$$\Psi = \lambda^{E[\ln(\rho(\Delta))]/\ln(\lambda)}.$$

As λ is of course unknown, I calibrate a value of $\lambda = 1.15$ in section 3.3 below.²² However, I also show in the Web Appendix that the results are insensitive to different values of λ . According to the model, there is a monotone relation between Λ and Ψ and the entry intensity of the economy as Proposition 2 showed that²³

$$\Lambda = \frac{x}{(\lambda - 1) + \lambda x} \text{ and } \Psi = \lambda^{-\frac{x+1}{x}}. \quad (35)$$

To test this prediction, I run regressions of the form

$$\begin{aligned} \Lambda_{r,t} &= \beta + \delta_t + \phi \text{Entry}_{r,t} + w'_{r,t} \gamma + u_{r,t} \\ \Psi_{r,t} &= \beta + \delta_t + \phi \text{Entry}_{r,t} + w'_{r,t} \gamma + u_{r,t}, \end{aligned}$$

where $\Lambda_{r,t}$ and $\Psi_{r,t}$ are the sample counterparts of (33) in region r at time t , δ_t is a year fixed effect, $\text{Entry}_{r,t}$ is the entry rate in region r at time t and $w_{r,t}$ contains additional regional characteristics, which do not vary over time. The model implies that $\phi > 0$.²⁴

The results are reported in Tables 1 and 2. Consider first the case of Λ , contained in Table 1. [TABLE 1 HERE] In the first column I report the simple correlation between the entry rate and $\Lambda_{r,t}$. The coefficient is positive and significant. That this relation is not purely driven by aggregate shocks over time is shown in column 2, where I include a full set of year fixed effect. The coefficient

²⁰Because I am exploiting variation across regions, I have to identify the the distribution of $\lambda^{-\Delta}$ from (34) and not from TFP as in for example Hsieh and Klenow (2009). (The inverse of) TFP in my economy is given by $\frac{k^\alpha l^{1-\alpha}}{py} \propto \lambda^{-\Delta} \frac{1}{w^{1-\alpha} R^\alpha}$, i.e. it would indeed identify the distribution of $\lambda^{-\Delta}$ up to scale if wages and interest rates were equalized across regions. However, with imperfect factor mobility, the model implies that both equilibrium wages and interest rates itself depend on the regional entry intensity.

²¹This normalization of course relies on firms having access to homogeneous technologies. I will consider the case of technological heterogeneity in section 3.4 below.

²²This value is also consistent with the implied growth rate. Recall that the model implies that $g_Q = \ln(\lambda)(I + z)$ see (23). The estimated rate of TFP growth in my data is about 8% (see footnote 36 for the construction). According to Bartelsman and Doms (2000), about 75% of productivity growth originates in innovation done by incumbent firms. Hence, $\lambda = \exp\left(\frac{0.08}{4} \frac{1}{z}\right)$ so that an entry rate of 12% (which is the mean entry rate in my sample) implies $\lambda = 1.18$.

²³Note that the derivation of (35) did not hinge on the economy being closed. Given any constant innovation and entry rates I and z , (35) simply followed from the quality ladder structure. Similarly, the measure $\rho(\Delta)$ does not rely on the way equilibrium wages are determined, but follows directly from the firms' labor demand taking wages as given. Hence, to study the relation between the entry intensity in region r and the moments $\Lambda = E[\lambda^{-\Delta}]$ and $\Psi = \lambda^{-E[\Delta]}$, I do not have to assume that factors are immobile across regions. I do require that there is no trade so that current producers in region r compete against potential entrants of region r and not other producers outside the region. It is only when I use these moments to infer factor prices and TFP in Section 4 that I we have to impose factor immobility.

²⁴Note that the model suggests that it is not the entry rate z but the entry intensity $\frac{z}{I}$ which matters. As the entry rate is directly observable in the data and conceptually simple, I take $z_{r,t}$ as my entry measure for the main specification. However, in the Web Appendix I show that the results are unchanged if I construct the entry intensity $\frac{z}{I}$ from the data. I will come back to this when I discuss the robustness of the results below.

hardly changes. The main worry is of course the potential endogeneity of the entry rate, i.e. are there unobserved regional characteristics, which increase both the entry rate and reduce the degree of misallocation? To partly address this issue, the next columns report the results where I control for different regional characteristics drawn from the PODES dataset. Column 3 simply controls for the regional population. This does not change any of the results. In column four I include four specific controls, which could affect the degree of misallocation and might be correlated with the entry rate. In particular I include a measure of infrastructure, namely the fraction of villages, which are accessible via asphalt roads. If better infrastructure (or a reduction in transport costs) makes it easier for firms to access other markets within their province, we would expect that a more integrated market improves the allocation of resources. In accordance with this hypothesis, the coefficient is positive and highly significant. In fact, it is noteworthy that this positive relation between allocative efficiency and this measure of transport infrastructure is very robust and is highly significant and remains so in all the robustness checks I conducted. This finding resonates well with the results of Donaldson (2010), who finds that infrastructural improvements increased trade and reduced price dispersion. Through the lens of this model and the work of Epifaniy and Ganica (forthcoming), this is consistent with the idea that market integration improves allocative efficiency by reducing monopolistic mark-ups. The effect of the entry intensity however does not change substantially. I also include a measure of the financial system, namely the average number of banks per village. As I will discuss below in how far a model with financial frictions would be consistent with my findings, let me here just note that Table 1 does not find evidence of any positive allocational effect of this particular measure of the financial environment. Finally, I include the average number of small and informal firms per village and the agricultural share of the population as controls. The former is motivated by the fact that informal firms (which are not part of my firm level data) might also put competitive pressure on the formal sector. I do not find evidence of this hypothesis, which is in line with the results of Porta and Shleifer (2009), who also find very little interaction between formal and informal firms. Column 5 finally uses the panel structure of the data, which allows me to control for time-invariant regional characteristics by including regional fixed effects. While the coefficient drops quite substantially it is still positive and significant. This could of course be the result of now properly controlling for regional characteristics. However, the large decrease in the coefficients is also suggestive of the importance of measurement error in my entry rate measure as stressed by Griliches and Hausman (1986). The analogous results for $\Psi_{r,t}$ are contained in Table 2. [TABLE 2 HERE] The results mirror the ones in Table 1. In particular the coefficient on the entry measure is always positive and significantly so. Overall, the results in Tables 1 and 2 are consistent with the predictions of the theory.

I conducted various robustness checks, which are reported in the Web Appendix. In particular, I redid the regressions using the usual robust standard errors instead of the clustered standard errors used in the main analysis, I estimated the regressions using weighted least squares, I included the crisis year 1998 and I used a measure for the entry intensity $x = \frac{z}{l}$ instead of the entry rate. The results are qualitatively unchanged in that all coefficients are positive and mostly significantly so. Furthermore I show there that the results are neither affected by my choice of $\lambda = 1.15$, nor the cutoff for the number of used to calculate the data on the province-year level. This Web Appendix also contains robustness checks for all the results presented in the next section. These are also qualitatively similar to the ones reported in the main text.

3.3 Other implications of the model

While Λ and Ψ are the two sufficient statistics for the effect of the mark-up distribution on the aggregate allocations, Proposition 2 puts much more structure on this distribution. In particular, I showed there that a higher entry intensity induces first order stochastic dominance shifts in this distribution. This restriction is of course testable given that $F_{\lambda-\Delta}$, the distribution of $\lambda^{-\Delta}$, is observed. Consider first Figure 2, where I display the estimated distribution function of $\lambda^{-\Delta}$ for “high” and “low” entry intensity regions. [FIGURE 2 HERE] Specifically, I calculate the mean entry rate for each region and sort the firm-level data according to this regional mean entry rate. Then I take the lower 25%-quantile of firms as the low-entry firms and the upper 25%-quantile as the high-entry firms and estimate the distribution function of $\lambda^{-\Delta}$ for each subsample. Figure 2 shows that the mark-up distribution of firms operating in high-entry regions is indeed dominated by the one of firms in a region, where little entry imposes less competitive discipline on the producing firm. Note also that the distribution function are precisely estimated in that the high-entry region distribution has (pointwise) confidence intervals around it. Hence, these functions do differ in a statistical sense.

While Figure only considered the raw mark-up comparison between high and low entry regions, the analysis above showed that other regional characteristics (e.g. the state of the infrastructure) are important. To focus on the partial effect of the regional entry rate on the distribution of mark-ups, consider Table 3, where I show that the effect of entry is present through the entire mark-up distribution even conditional on such regional characteristics. In particular, I estimate regressions of the form

$$q_{r,t}^{\tau} = \beta + \delta_t + \phi Entry_{r,t} + w_r' \gamma + u_{r,t}, \quad (36)$$

where $q_{r,t}^{\tau}$ is τ th quantile of the distribution $F_{\lambda-\Delta}$ in region r at time t , $Entry_{r,t}$ is the entry rate and w_r are the regional characteristics employed above. I run (36) for different values of τ separately. Proposition 2 implies that $\phi > 0$ for all τ . [TABLE 3 HERE] In the first panel of Table 3 I depict the simple correlation between the entry rate z and different quantiles. It is seen that all the coefficients are positive and - except for the highest quantile - significant. In the second panel I show that this remains the case when I control for year fixed effects and the regional controls from above.²⁵ Hence, even conditional on regional controls, the entry rate induces shifts in the distribution of mark-ups as the theory predicts.

Another implication of the model is that a higher entry intensity both reduces and compresses the distribution of mark-ups. In particular, from the distribution of productivity gaps $\mu(\Delta)$ it is easy to see that (see (19)) $E[\Delta] = \frac{1+x}{x}$ and $var(\Delta) = \frac{1+x}{x^2}$. This immediately implies that

$$var(\Delta) = (E[\Delta] - 1) E[\Delta]. \quad (37)$$

As $\ln(\xi) = \Delta \ln(\lambda)$, (37) states that the model predicts a positive relation between the dispersion of (log) mark-ups and the average (log) mark-up. This correlation is depicted in Figure 3, which shows a robust positive correlation as the model predicts. [FIGURE 3 HERE] While this is one implication of the model, the theory is of course stronger in that it also tells us where this positive correlation stems from: It is a low rate of entry, which increases both the mean and the dispersion of (log) mark-ups. To test this additional restriction, Panels 1 and 2 of Table 4 report the results of a regression of $E[\Delta]$ and $sd(\Delta)$ on the regional entry rate and controls. [TABLE 4 HERE] Panel 1 confirms that a higher entry rate indeed has a significant negative effect on the average (log) mark-up and

²⁵For brevity I only report the coefficient on the entry rate. The full results are available upon request.

the coefficients are pretty robust across the different specification. Panel 2 shows that this is only weakly borne out for the dispersion measure. While all the point estimates have the correct sign, none of them is significant. In Panel 3 of Table 4 I report another implication of the model, which is informative about the dispersion of mark-ups. As equilibrium factor demands are fully determined by firm-specific mark-ups (see (8)) it follows that $\text{var}(\ln(l)) = \text{var}(\ln(\xi)) = \ln(\lambda)^2 \text{var}(\Delta)$. Hence, the model implies that the within-region dispersion of employment is decreasing in the entry rate. That this prediction, which is based on a different moment as the analysis above, is strongly borne out by the data is further evidence about the specific channel of the model above.

With $\Lambda_{r,t}$ and $\Psi_{r,t}$ at hand, I can also study the effects of entry on the misallocation component of TFP. As $M = \frac{\Psi}{\Lambda}$ (see (22)), this term is the product of two counteracting effects. According to the theory, the increase in Ψ dominates the one of Λ , as M is increasing in x . From Tables 1 and 2 we can already expect that this horse race will also exist in the data. In Table 5 I show the results of the regression

$$M_{r,t} \equiv \Psi_{r,t}/\Lambda_{r,t} = \beta + \delta_t + \phi \text{Entry}_{r,t} + w'_r \gamma + u_{r,t}.$$

While the theory predicts that $\phi > 0$, it is seen that there is not much of an effect. [TABLE 5 HERE] In fact Proposition 1 is informative to explain this finding. There I showed that TFP will be unaffected by a uniform increase in the level of mark-ups. While the theory implies that a higher entry rate will both reduce the level and the dispersion of mark-ups, Table 4 above showed that the effect of entry on the mean mark-up was stronger than the one on the mark-up dispersion.

As a final test of the model, I take the model very seriously and consider the closed form expression for the distribution of mark-ups. In particular, in Figure 4 I depict the distribution of $\xi^{-1} = \lambda^{-\Delta}$ in the population of firms and the implied distribution from the model induced from the density $\mu(\Delta) = \left(\frac{1}{1+x}\right)^\Delta x$ (see (19)). [FIGURE 4 HERE] To construct Figure 4 I calibrate x such that $E[\lambda^{-\Delta}]$ coincides with its sample counterpart.²⁶ While the model does not match the distribution of observed mark-ups perfectly, it nevertheless captures some salient features of the observed distribution of firm-level “distortions”, in particular the concavity of the schedule.²⁷

3.4 Competing explanations

While the empirical results reported above are consistent with the model, there are of course other explanations, which have the potential to rationalize the data. In this section I will consider four specific ones, which seem to be the most relevant in the this context. While the first two concern different economic mechanisms, the last two address issues of mismeasurement and misspecification in general. In particular, I will first consider the case of credit market frictions and policy distortions as competing structural explanations. Then I will address the possibility of measurement error in the wagebill due to unpaid family members and technological heterogeneity across firms.

²⁶To do so, I need to calibrate a value for λ . As the model predicts that $\Delta \geq 1$ and $\max \rho(\Delta) = \rho(1) = \lambda^{-1} = 0.8763$ (see (34)), I take $\lambda = 1.15$, which implies that a successful innovation increases productivity by 15%. As $\hat{E}[\lambda^{-\Delta}] = N^{-1} \sum_i \frac{w l_i}{(1-\alpha) p y_i}$ is given by 0.261, the implied entry intensity is calibrated to be 5.25%. This implies that $\frac{z}{z+I} \approx 5\%$, i.e. 5% of productivity growth is driven by new entry. Empirical studies find that about 75% of productivity growth is due innovation by existing firms (Bartelsman and Doms, 2000; Acemoglu and Cao, 2010), hence this entry component seems to be a little too low.

²⁷In particular, the model overpredicts the heterogeneity of mark-ups in the economy as it puts more mass of firm on the tails of the distribution. Intuitively, to match the average distortion $E[\lambda^{-\Delta}] = \frac{x}{\lambda-1+\lambda x}$ using only the entry intensity x , the required entry intensity is calibrated to be very low. However, we know from Proposition 2 that a low entry intensity increases the mark-up heterogeneity in the economy.

To formalize the discussion, I want to think of a constant mark-up economy, which otherwise resembles the one described above. In particular, I consider the same economy as above but assume that technologies become common knowledge, once a firm has climbed a new step on the quality ladder. This directly implies that current producers always have a productivity advantage of λ and hence set a constant mark-up over marginal costs, i.e. $\xi(\nu) = \lambda$ (see (6)).²⁸

3.4.1 Credit market imperfections

That credit market imperfections can reduce aggregate TFP and factor prices by causing a misallocation of resources is of course well known. To see how credit market frictions could rationalize the results given above, suppose that firms face borrowing constraints of the form $wl + Rk \leq \theta Rev$, where θ is a firm-specific multiplier governing the pledgability of future revenue. Such a constraint naturally arises if factors have to be paid for before revenue is realized and θ specifies how much of the revenue can be pledged to pay those factors.²⁹ θ is firm-specific because firms might differ in their liquid wealth, collateralizable assets or “reputational capital”, which allows them to borrow more. It is then easy to verify that the mark-up of firm ν is given by $\xi(\nu, t) = \frac{p(\nu, t)}{MC(\nu, t)} = \max \left\{ \theta(\nu)^{-1}, \lambda \right\}$, i.e. unconstrained firms set the unconstrained markup λ and constraint firms (whose θ is low) set a mark-up which is exactly equal to (the inverse of) their borrowing limit $\theta(\nu)^{-1}$. Hence, in this world, heterogeneous mark-ups are direct consequences of some firms being credit constraint and it is those firms which set a high mark-up.³⁰ In fact, if we define $\tilde{\vartheta}(\nu)$ by $\theta(\nu) = \lambda^{-\tilde{\vartheta}(\nu)}$ and $\vartheta(\nu) \equiv \max \left\{ \tilde{\vartheta}(\nu), 1 \right\}$, then $\xi(\nu, t)^{-1} = \lambda^{-\vartheta(\nu, t)}$ so that the aggregate allocations in this economy are *as in the benchmark economy* once we substitute $\vartheta(\nu, t)$ for the productivity gap $\Delta(\nu, t)$. In particular, Proposition 1 is again immediately applicable: As long as ϑ is constant across firms, TFP in this economy will be equal to the first best level of TFP although all firms might be severely credit constraint. Capital market imperfections will only show up in the aggregate economy in that factor prices are reduced. Hence, it is only when *some firms are less constrained than others* that capital market imperfections reduce aggregate TFP.

If we think that θ is lower for “poor” firms or firms in financially underdeveloped regions, I would identify poor firms as having a lot monopoly power and if the entry intensity is positively correlated with the regional financial system, I would conclude that entry reduces monopoly power albeit it is just the case that high entry regions are such that less firms are constraint (or firms are less constraint). As a proxy for θ , I exploit information about both the ownership structure and the way firms raise funds to finance their investment. Using this information I construct 4 measures of

²⁸Similarly I could also have assumed that the production of the final good is given by $Y(t) = \left(\int y(\nu, t)^{\frac{\sigma-1}{\sigma}} d\nu \right)^{\frac{\sigma}{\sigma-1}}$ and that each firm is a monopolist in its variety ν . Then each firm would also have set a constant mark-up of $\frac{\sigma}{\sigma-1}$. However, by considering the same production technology the competing explanations can be more concisely mapped to the benchmark model.

²⁹Note in particular that labor inputs have to enter the borrowing constraint for credit market imperfections affecting the allocation of labor and hence the measured labor-shares ρ . Models where only the capital input is affected are unable to explain variation in labor shares because the marginal product of labor is equalized across firms (Moll, 2010; Buera, Kaboski, and Shin, 2009). A constraint similar to the one in the text is for example considered in Midrigan and Xu (2010).

³⁰My formulation of borrowing constraints is of course special and I adopted this formulation to make the point as clear as possible. In general, we can think of a borrowing constraint of the form $wl + Rk \leq H(s, f)$, where s is some firm-specific state variable and f parametrizes the financial system and varies for example across regions. Then it can be easily shown that $\xi(\nu, t) = \lambda(1 + \zeta(s, f))$, where $\zeta(s, f)$ is the Lagrange multiplier on the borrowing constraint so that constraint firms set higher mark-ups.

credit constraints. In particular, I will think of a firm as being relatively *unconstrained*, whenever it is (at least partly) owned by foreign investors or whenever its investment needs are (at least partly) financed via FDI, funds raised on the capital market or foreign loans. While these characteristics have intuitive appeal,³¹ in the Web Appendix I conduct two experiments to judge the validity of these firm-characteristics.³²

Given these proxies for credit constraints, I then investigate if the share of unconstrained firms is indeed positively correlated with the regional entry rate, and if the entry rate has an independent effect on the degree of misallocation above and beyond regional differences in the share of unconstrained firms. The results are shown in Table 6. [TABLE 6 HERE] In the first four columns I show that the entry rate is indeed positively correlated with the regional share of unconstrained firms. All four measures have a positive coefficient and two of them are statistically significant. In the last four columns, I show that it is not only regional differences in financial constraints active producers face, which explain the correlation between the entry rate and the two measures of misallocation. In columns (5) and (7) I report the results without regional controls. It is clearly seen that the entry rate still has a robust positive effect, which is statistically significant. Furthermore, the point estimates are almost identical to the ones reported in Tables 1 and 2 and controlling for regional characteristics in columns (6) and (8) does not change these results. Hence, the entry rate seems to have an independent effect on the degree of misallocation and is not only proxying for credit constraints. I want to stress that this does of course *not* imply that credit constraints are not important. First of all, I will show below in the welfare calculation that differences in the entry rate only explain roughly a third of the allocational differences across regions. Secondly, it might be precisely regional differences in the financial system which cause the entry rate to be different (and the first four columns in Table 6 provide evidence that this is the case). Table 6 only says that the entry rate is a significant predictor for allocational differences across regions even after differences in the credit constraints firms face are (at least crudely) controlled for.

3.4.2 Policy distortions

Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) model the degree of misallocation via “firm-specific taxes”, i.e. firms are envisioned to face a profit function of the form $\pi(\nu) = (1 - \tau_Y(\nu))p(\nu)y(\nu) - wl(\nu) - (1 + \tau_K(\nu))Rk(\nu)$. Here $\tau_Y(\nu)$ parametrizes a general, factor-neutral distortion and $\tau_K(\nu)$ denotes a distortion that affects the allocation of capital. As my identification relies only on the allocation of labor, let us suppose that $\tau_K = 0$. While τ_Y is of course only a modelling device to account for the misallocation of resources across firms, it is often argued that *actual* policies like regulation, taxes or bureaucratic red tape are important features. If it such actual policies, which I will think of in this section. Introducing such firm-specific policy distortions in my model is straight-forward. In particular, the firm-specific labor share was given by $\frac{wl(\nu)}{p(\nu)y(\nu)(1-\alpha)} = \lambda^{-1}(1 - \tau_Y(\nu)) \equiv \lambda^{-(1+\tilde{\tau}_Y(\nu))}$, where $\tilde{\tau}_Y = -\frac{\ln(1-\tau_Y)}{\ln(\lambda)}$. Again: This model with policy distortions is isomorphic to the benchmark model once we substitute $1 + \tilde{\tau}_Y(\nu)$ for the productivity gap $\Delta(\nu)$. Proposition 1 therefore directly applies. If all firms faced the same

³¹Foreign ownership is often taken as an indicator for relatively unconstrained firms. In fact, Blalock, Gertler, and Levine (2008) use precisely this legal status as an indicator for unconstrained firms in the same dataset, which I am using.

³²First of all, I show that the respective share of firms is positively correlated with the regional measure of financial development. Secondly, in the 1996 wave of my data, each firm was asked if they consider themselves credit constraint. Using this cross-section, I show that the above characteristics indeed reduce this (reported) probability of facing credit constraints.

“taxes”, TFP would again be equal to the first best and only factor prices were affected by the policy distortions. Only preferential treatment to some but not other firms reduces aggregate efficiency as measured by TFP.

To see how such policy distortions might drive my results, note that firms that I identify as having large monopoly power might in fact just face high policy distortions. If additionally regions with good policies (i.e. where the $\tilde{\tau}_Y(\nu)$ are low and relatively undispersed) are also characterized by more dynamic entry, policy distortions could generate the pattern observed in the data. To argue that this is unlikely to be the case, I follow a similar strategy as for the case of credit constraints. In particular, I observe if either the federal or the local government has a financial stake in the firm. So if we identify the taxes as being induced by policies, we expect that state firms appear less distorted, i.e. have higher labor shares. In the first two columns of Table 7 I show that firms owned by the state (column one) or firms where the central or local government has at least partial ownership (column two) indeed appear to be less distorted. [TABLE 7 HERE] Columns 3 and 4 however show that the regional share of firms owned by the state is uncorrelated with the entry rate. Hence it is not surprising that the remaining columns show that the correlation between the entry rate and the two misallocation measures Λ and Ψ remains strongly positive once I control for differences in the share of state firms. Like in the case of credit constraints, this of course does not mean that policy distortions do not induce a misallocation of resources across firms. First of all, my measure of policy distortions is of course very limited. Secondly (and more importantly), policy distortions might manifest themselves in differences in the entry intensity. In fact, the microfoundation given in Section 2.3 showed that higher entry costs $(1 - \chi)$ reduce the equilibrium entry intensity. So if there are regional differences in the regulation of entry, the model predicts that policy distortions determine the degree of misallocation: by reducing the extent of entry, they reduce the competitive pressure in the economy.

3.5 Measurement error

After considering two competing structural explanations for the results presented above, in this section I briefly discuss two types of measurement problems. In a developing economy like Indonesia, employment of family members is common. While this might be less of a problem for the formal manufacturing firms which I am studying, about 50% of firms do employ unpaid workers (many of which are probably family members) and these unpaid workers constitute on average around 5% of the labor force. To see how this might interfere with the identification of mark-ups, let $l^P(\nu)$ denote paid employees and $l^{HH}(\nu)$ the number of household members working in the firm so that $l(\nu) = l^P(\nu) + l^{HH}(\nu)$. While the economic costs of employing household members is of course the current wage rate, firms might not report this as their “wagebill”. Hence, while it is the case that firms hire labor on a spotless labor market so that $\frac{w(\nu)l(\nu)}{p(\nu)y(\nu)(1-\alpha)} = \lambda^{-1}$, in the data I observe $\frac{w(\nu)l^P(\nu)}{p(\nu)y(\nu)(1-\alpha)} = \lambda^{-1} \frac{l^P(\nu)}{l^P(\nu) + l^{HH}(\nu)}$. Hence, I will identify unpaid-worker intensive firms as firms with high monopoly power and if regions with a high entry rate are less reliant on “family firms”, I will find the positive relation established above. Table 14, reported in the Web Appendix, shows that this is not the case. While firms who employ unpaid workers and firms where such workers account for a higher share of the labor force do indeed have lower labor shares, the occurrence of such firms is uncorrelated with the entry rate and the entry rate remains significantly correlated with the misallocation measures once the regional shares of such firms are controlled for.

A second problem concerns technological heterogeneity, as my identification of firm-specific mark-ups relies strongly on the assumption of homogeneous technologies across firms. To see in

how far a failure of this assumption could generate my results, suppose that firms' production functions have constant returns, but that the factor share parameter in the production function α is variety-specific. Hence: some products are more capital-intensive than others. It is then easy to verify that $\frac{wl(\nu)}{p(\nu)y(\nu)(1-\alpha)} = \frac{\lambda^{-1}}{1-\alpha} \left(\frac{wl(\nu)}{Rk(\nu)+wl(\nu)} \right)$, which shows that firms producing different varieties will differ in their labor-shares. The apparent heterogeneity in mark-ups, which I identify as a lack of competition could be a solely a misspecification of the technology in that those firms that I identify as having lots of monopolistic power are just firms that have above-average capital-shares. If additionally the entry rate in region r is positively correlated with the labor-intensity in region r , I will find the same correlation as reported above. That this pattern is not present in the data is shown in Table 15, reported in the Web Appendix. There I show that the firm-level capital labor ratio is indeed negatively correlated with firm-specific labor shares. However, high entry regions are not regions where the mean (or median) capital-labor ratio is particularly low but there is no significant correlation between entry and firms' capital-intensity. The entry rate has still a positive and significant effect on the misallocation measures and the point estimates are identical to the one reported in Tables 1 and 2.³³ Hence, technological heterogeneity does not explain the particular correlation, which I interpret as being informative about the mechanism identified in this paper.

4 The welfare costs of low entry

Even if the results above are consistent with the economic mechanism stressed in this paper, are the observed differences in the entry rate across provinces in Indonesia quantitatively important? To get a sense of the implied economic magnitude, let us use the estimates from above for an admittedly crude welfare comparison. In particular we can ask by how much welfare is higher in the high entry regions, compared to their low entry counterpart simply because the resource allocation is improved through the competition enhancing effect of entry.³⁴ So suppose that all provinces are on their BGP, are characterized by the same level of aggregate productivity $Q(t)$, grow at the same rate and let us normalize the labor endowment to unity. Along the BGP all variables grow at the same rate, so let us denote the normalized variables by $\tilde{c}_r, \tilde{y}_r$ and $\tilde{k}_r$. Using the steady state capital stock $\tilde{k}_r$ given in (24), it is easy to verify that

$$\tilde{c}_r \propto M(x_r)^{\frac{1}{1-\alpha}} \Lambda(x_r)^{\frac{\alpha}{1-\alpha}} \left[1 - \alpha \frac{\delta + g_Y}{g_Y + \delta + \rho} \Lambda(x_r) \right]. \quad (38)$$

Hence, the level of consumption depends on the regional entry intensities via the respective distortion indices. When I measure $M(x_r)$ and $\Lambda(x_r)$ for the high and low entry regions in Indonesia

³³In the Web Appendix I also report the results of using the "residual" misallocation measures Λ^{Res} and Ψ^{Res} . Instead of constructing Λ and Ψ directly from the labor shares $\rho = \frac{wl}{(1-\alpha)py}$, I use the residual labor shares ρ^{Res} by $\rho^{Res} = \rho - \hat{\rho}$, where $\hat{\rho}$ are the predicted values from a regression of ρ on a full set of sector fixed effects and the capital-labor ratio. Hence, $\Lambda^{Res} = \hat{E}[\rho^{Res}]$ and $\Psi^{Res} = \lambda^{\hat{E}[\ln(\rho^{Res})] / \ln(\lambda)}$ are the two misallocation measures after technological heterogeneity is controlled for. There is still a strong positive correlation between these two measures and the entry rate.

³⁴I want to stress that there are two ways to interpret the following exercise. According to the mechanism identified in this model, the entry rate has a causal effect on the degree of misallocation. From that perspective, the reported differences in welfare could be eliminated by a reduction in monopolistic power in the low entry regions. In fact, Section 2.3 suggested that higher entry costs or a low productivity of the entry technology could be responsible for the effect. However, even we think that monopolistic power does not play an important role for the misallocation of resources, the welfare calculation is relevant. From that perspective, the calculation below implies that the entry rate has predictive power for regional differences in allocative efficiency, despite not being causal itself.

used in Figure 2, I find that

$$\frac{M^{High}}{M^{Low}} = 1.0264 \text{ and } \frac{\Lambda^{High}}{\Lambda^{Low}} = 1.0611.$$

Hence, high entry regions enjoy a TFP advantage of about 2.5% over their low entry counterparts. As the steady state capital stock is proportional to $(M(x)\Lambda(x))^{\frac{1}{1-\alpha}}$, both the high entry regions' capital stock and its equilibrium wage are about 13% higher. This high difference in the capital stock despite the modest TFP difference is due to the presence of the factor price distortion index $\Lambda(t)$, which reduces the return to capital above and beyond the reduction in TFP. Finally, (38) implies that the higher entry rate induces an increase in the steady state level of consumption by about 6%.³⁵

To put this number of 6% into perspective, I can also conduct the exercise of Hsieh and Klenow (2009) on my data. In particular, I can simply ask by how much TFP and welfare would be higher if we were to “move” the most distorted region to the degree of misallocation observed in the least distorted region in Indonesia. So if the ranking of regions according to the entry rate and their factor price distortion index Λ was identical, the TFP and welfare difference between the least and most distorted was also exactly equal to the 2.5% and 6% reported above. Doing this exercise however reveals that this is not the case. In contrast, reducing the degree of misallocation in the most distorted region to the one observed in the most efficient one, would increase TFP by 5.5%, the steady state capital stock by 45% and welfare by 17%. Hence, by focusing on the entry rate we can predict roughly a third of the cross-regional efficiency differences in Indonesia. The mechanism proposed in this paper provides one microfoundation why this predictive power of the regional entry rate might in fact be causal.

5 Conclusion

Misallocation of resources across firms reduces economic efficiency. In this paper I study an economy where misallocation is caused by monopolistic firms charging heterogeneous mark-ups. The heterogeneity of mark-ups is crucial. Specifically, I show that aggregate total factor productivity is only affected by the dispersion in mark-ups, while factor prices depend only on the mean of the mark-up distribution. Equilibrium mark-ups depend on the cross-sectional productivity distribution across firms as firms' market power is determined by their relative productivity advantage over their competitors. In particular, the scope of monopolistic pricing is reduced if the productivity distribution is highly concentrated. Hence, an environment, which is conducive to a large productivity dispersion across firms, reduces allocative efficiency. This insight provides the link between the static degree of misallocation and the dynamic properties of the evolution of productivity. In my model, productivity growth stems from either productivity improvements by current producers or entering firms. I show that the stationary distribution of mark-ups has a simple closed form expression and is parametrized by a single variable, namely the entry intensity of the economy. In particular, if the entry intensity is high (for example because entry costs are low or innovative resources are abundant), product markets become more competitive as the productivity dispersion between producing firms and their competitors is reduced. This lowers both the average mark-up and the dispersion of mark-ups in the cross-section of firm and via this channel shows up in higher

³⁵For this calculation I assume that $g_Y = 0.08$ as measured in the data and set $\delta = 0.1$, $\alpha = 0.3$ and $\rho = 0.05$.

TFP and higher factor prices. If on the other hand, entrants put little discipline on current producers, equilibrium mark-ups will be higher and more dispersed, allocative efficiency will suffer and both TFP and factor prices will be low.

I then test the model’s implications on Indonesian firm-level data. Using regional variation in the entry rate of the Indonesian manufacturing sector, I show that higher entry rates are robustly correlated with both lower average mark-ups and dispersion, that these results are not driven by regional differences in financial development or infrastructural quality and that the entry rate remains an important determinant of the distribution of mark-ups when policy distortions or imperfect credit markets are controlled for. Interpreted through the lens of my model, this independent effect of the entry rate is causal in that it improves the allocation of resources through competitive pressure. From a quantitative point of view, my results imply that differences in the allocational efficiency as explained by the regional variation in the entry intensity are consistent with TFP differences of 2.5%, differences in the capital stock of 13% and a difference in aggregate consumption of 6% across regions in Indonesia.

Acknowledgements: I especially thank Daron Acemoglu, Marios Angeletos, Ufuk Akcigit and Abhijit Banerjee for many helpful discussions. I also thank Thomas Chaney, Maya Eden, Chang-Tai Hsieh, Sam Kortum, Benjamin Moll, Ezra Oberfield and Rob Townsend as well as seminar participants at MIT and Chicago. A preliminary version of this paper was also presented at the 2010 SED meetings in Montreal and I thank participants for their comments. I am particularly grateful to Ida Fariana from the Badan Pusat Statistik (BPS-Statistics Indonesia) for her cooperation in obtaining the Indonesian micro data, to Monica Martinez-Bravo for providing me with me PODES data and for the financial assistance of the George and Obie Shultz fund to acquire the data necessary for this project.

References

- ACEMOGLU, D., AND U. AKCIGIT (2009): “State-Dependent Intellectual Property Rights Policy,” Working Paper.
- ACEMOGLU, D., P. ANTRAS, AND E. HELPMAN (2007): “Contracts and Technology Adoption,” *American Economic Review*, 97(3), 916–943.
- ACEMOGLU, D., AND D. CAO (2010): “Innovation by Entrants and Incumbents,” Working Paper.
- ACEMOGLU, D., S. JOHNSON, AND J. ROBINSON (2001): “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, 91(5), 1369–1401.
- AGHION, P., AND P. BOLTON (1987): “Contracts as a Barrier to Entry,” *American Economic Review*, 77(3), 388–401.
- AGHION, P., AND P. HOWITT (1992): “A Model of Growth through Creative Destruction,” *Econometrica*, 60(2), 323–351.
- AMITI, M., AND J. KONINGS (2007): “Trade Liberalization, Intermediate Inputs and Productivity,” *American Economic Review*, 97(5), 1611–1638.

- BANERJEE, A., AND E. DUFLO (2000): “Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry,” *The Quarterly Journal of Economics*, 115(3), 989–1017.
- (2005): “Growth Theory through the Lens of Development Economics,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1A, chap. 7, pp. 473–552. Elsevier B.V.
- BARSEGHYAN, L. (2008): “Entry Costs and Cross-Country Differences in Productivity and Output,” *Journal of Economic Growth*, 13(2), 145–167.
- BARSEGHYAN, L., AND R. DICECIO (2009): “Entry Costs, Misallocation and Cross-Country Income and TFP Differences,” Working paper, Feder Reserve Bank of St. Louis.
- BARTELSMAN, E. J., AND M. DOMS (2000): “Understanding Productivity: Lessons from Longitudinal Microdata,” *Journal of Economic Literature*, 38(3), 569–594.
- BERNARD, A., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): “Plants and Productivity in International Trade,” *American Economic Review*, 93(4), 1268–1290.
- BLALOCK, G., P. J. GERTLER, AND D. I. LEVINE (2008): “Financial Constraints on Investment in an Emerging Market Crisis,” *Journal of Monetary Economics*, 55(3), 568–591.
- BLANCHARD, O., AND F. GIAVAZZI (2003): “Macroeconomic Effects of Regulation and Deregulation in Goods and Labor Markets,” *Quarterly Journal of Economics*, 118(3), 879–907.
- BUERA, F., J. KABOSKI, AND Y. SHIN (2009): “Finance and Development: A Tale of Two Sectors,” Working Paper.
- CASELLI, F. (2005): “Accounting for Cross-Country Income Differences,” in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, vol. 1A, chap. 9, pp. 679–741. Elsevier B.V.
- CHARI, V., P. J. KEHOE, AND E. R. MCGRATTAN (2007): “Business Cycle Accounting,” *Econometrica*, 75(3), 781–836.
- DJANKOV, S., R. L. PORTA, F. LOPEZ-DE-SILVAES, AND A. SHLEIFER (2002): “The Regulation of Entry,” *Quarterly Journal of Economics*, 117(1), 1–37.
- DONALDSON, D. (2010): “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure,” NBER Working Paper 16487.
- EPIFANIY, P., AND G. GANICA (forthcoming): “Trade, Markup Heterogeneity and Misallocations,” *Journal of International Economics*.
- GANICA, G., AND F. ZILIBOTTI (2009): “Technological Change and the Wealth of Nations,” *Annual Review of Economics*, 1, 93–120.
- GOPINATH, G., AND B. NEIMAN (2010): “Trade Adjustments in Large Crises,” Working Paper.
- GRILICHES, Z., AND J. A. HAUSMAN (1986): “Errors in Variables in Panel Data,” *Journal of Econometrics*, 31, 93–118.
- GROSSMAN, G., AND E. HELPMAN (1991): “Quality Ladders in the Theory of Growth,” *Review of Economic Studies*, 58(1), 43–61.

- HALL, R., AND C. JONES (1999): “Why Do Some Countries Produce So Much More Output Per Worker Than Others?,” *Quarterly Journal of Economics*, 114(1), 83–116.
- HOPENHAYN, H. (1992): “Entry, Exit and Firm Dynamics in Long-Run Equilibrium,” *Econometrica*, 60, 1127–1150.
- HSIEH, C.-T., AND P. KLENOW (2009): “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124(2), 771–807.
- JEONG, H., AND R. TOWNSEND (2007): “Sources of TFP growth,” *Journal of Economic Theory*, 32(1), 179–221.
- KLENOW, P., AND A. RODRIGUEZ-CLARE (1997): “The Neoclassical Revival in Growth Economics: Has it Gone Too Far?,” in *NBER Macroeconomics Annual*, vol. 12, pp. 73–103.
- KLETTE, T. J., AND S. KORTUM (2004): “Innovating Firms and Aggregate Innovation,” *Journal of Political Economy*, 112(5), 986–1018.
- LAGOS, R. (2006): “A Model of TFP,” *Review of Economic Studies*, 73(4), 983–1007.
- LUCAS, R. E. (1988): “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 22, 3–42.
- (1990): “Why Doesn’t Capital Flow From Rich to Poor Countries,” *American Economic Review*, 80(2), 92–96.
- LUTTMER, E. G. J. (2010): “Models of Growth and Firm Heterogeneity,” *Annual Review of Economics*, 2, 547–576.
- MCKINSEY-GLOBAL-INSTITUTE (2001): “India: The Growth Imperative,” .
- MIDRIGAN, V., AND D. Y. XU (2010): “Finance and Misallocation: Evidence From Plant-Level Data,” NBER Working Paper 15647.
- MOLL, B. (2010): “Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation?,” Working Paper.
- PARENTE, S. L., AND E. C. PRESCOTT (1994): “Barriers to Technology Adoption and Development,” *Journal of Political Economy*, 102(2), 298–321.
- PORTA, R. L., AND A. SHLEIFER (2009): “The Unofficial Economy and Economic Development,” in *Brookings Papers on Economic Activity*, pp. 275–352.
- RESTUCCIA, D., AND R. ROGERSON (2008): “Policy distortions and aggregate productivity with heterogeneous establishments,” *Review of Economic Dynamics*, 11, 707–720.
- SANDLERIS, G., AND M. L. J. WRIGHT (2010): “The Costs of Financial Crises: Resource Misallocation, Productivity and Welfare in the 2001 Argentine Crisis,” Working Paper.
- SUTTON, J. (1997): “Gibrat’s Legacy,” *Journal of Economic Literature*, 35, 40–59.

6 Appendix

6.1 Data

As explained in the main text, the Statistik Industri dataset contains information on all manufacturing firms in Indonesia with a laborforce of more than 20 employees. For the main analysis of this paper, I only need information on firms' revenues, their wagebill and their entry behavior. For the robustness analysis I also utilize firms' capital level, their ownership structure and source of finance and their employment composition in terms of paid and unpaid workers. Both revenues and the total wagebill are directly taken from the data. As a measure of capital, I take firms' total assets, which consist of machines, working capital, buildings, vehicles and other forms of fixed capital, all measured at current market prices. The ownership structure is elicited from their capital structure. In particular, each firm reports if the federal or local government or foreign investors own stakes in the firm. The source of finance is also directly observable in the data as firms report the respective share of their investment expenditures, which are financed through foreign direct investment, retained earnings, domestic borrowing, government funds or assets raised on the capital market. Firms also report the share of paid and unpaid workers, which allows me to identify the employment composition. To identify entry, I rely on the panel nature of the data. My measure for the year-province specific entry rate, is the fraction of firms in province r at time t who appear in the data. This measure is not ideal for two reasons. From a conceptual point of view it is not clear why only a firm with at least 20 employees should put competitive pressure on active incumbents. From a measurement point of view, I might label a firm as an entrant if it laid off workers in some years and then started growing again. To address the latter concern, I experimented with other measures, which also condition on the reported age of the firm and other cutoffs and the results were similar. Furthermore, note that this problem only invalidates my empirical strategy if different regions are affected differentially. Regarding the former conceptual concern, Porta and Shleifer (2009) and McKinsey-Global-Institute (2001) argue that there is very little competition between small, informal firms and members of the formal industrial sector. While these claims mostly concern very small establishments, the 20 employee cutoff might not be such bad a measure of when a firm starts competing within the formal sector. At the end of the day however, this entry measure is simply the only one available in my data set. All nominal variables are deflated using the CPI Index from the World Development Indicators. To construct the final sample, some data cleaning steps were necessary. First of all I dropped all establishments, which did not report information on revenues, the wagebill or their asset position. Especially the latter requirement shrinks the sample considerably as roughly a third of plants have missing asset data. To eliminate some concerns about measurement error, I also dropped all firms, which report revenue or employment growth exceeding 200% per year. Finally, following Hsieh and Klenow (2009), I trim the 1% tails of the final laborshare distribution. As most regressions are done on the province-year level, I construct the respective dependent variables using the micro-data. As the manufacturing sector in Indonesia is relatively concentrated, many provinces contain only few firms. Using more firms to construct the averages obviously reduces the noise in the generated variables but also reduces the number of provinces I can use in each year. For the main analysis I drop all province-year cells, which contain less than 25 firms. However, I show below that the results are not dependent on this particular cutoff. In order to judge the validity of my credit constraints measures, I exploit the 1996 survey. The 1996 survey is special in that the Statistik Industri survey was done in conjunction with the economic census. Hence, the 1996 survey contains substantially more information. In particular, firms are asked if they are subject to a major constraint, which they could not overcome and if

this constraint refers to the scarcity of capital. They are also asked what efforts they undertook to overcome this capital constraint, e.g. if they applied for a bank loan, sold assets or issued shares.

The provincial characteristics are constructed using the PODES dataset for the year 1996. The unit of observation is one of the 65.000 villages. For each village, I record the total population, the number of banking branches, an indicator if the village is accessible by asphalted streets, an indicator if the village's main source of income is agriculture and the number of small industrial plants residing in the village. To generate the information on the level of the province, I consider the respective weighted averages of those variables, where the weights are the population sizes of the respective villages.

6.2 General Equilibrium Aggregation

To derive (10), note that (7) implies that

$$Y = \exp \left(\int \ln \left(\frac{Y}{\psi(R, w)} \frac{q(\nu)}{\xi(\nu)} \right) d\nu \right) = \frac{Y}{\psi(R, w)} \frac{\exp \left(\int \ln(q(\nu)) d\nu \right)}{\exp \left(\int \ln(\xi(\nu)) d\nu \right)},$$

so that $\psi(R, w) = \frac{Q}{\exp \left(\int \ln(\xi(\nu)) d\nu \right)}$. As factor markets have to clear, we get that $L = \frac{(1-\alpha)Y}{w} \int \xi^{-1}(\nu) d\nu$ and similarly $K = \frac{\alpha Y}{R} \int \xi^{-1}(\nu) d\nu$. Hence,

$$\psi(R, w) = \left(\frac{w}{1-\alpha} \right)^{1-\alpha} \left(\frac{R}{\alpha} \right)^{\alpha} = Y \left(\int \xi^{-1}(\nu) d\nu \right) L^{-(1-\alpha)} K^{-\alpha} = \frac{Q}{\exp \left(\int \ln(\xi(\nu)) d\nu \right)}.$$

Rearranging terms yields (10).

6.3 Proofs of Proposition 2

To derive the results of Proposition 2, note that

$$F(\Delta; x) = \sum_{i=1}^{\Delta} \mu(i) = x \sum_{i=1}^{\Delta} \left(\frac{1}{1+x} \right)^i = 1 - \left(\frac{1}{1+x} \right)^{\Delta+1},$$

that Λ is defined in (14) as $\Lambda = \left[\int_0^1 \lambda^{-\Delta(\nu)} d\nu \right]$ so that

$$\Lambda(x) = \sum_{i=1}^{\infty} \lambda^{-i} \mu(i) = \frac{x}{\lambda(x+1)} \sum_{i=0}^{\infty} \left(\frac{1}{\lambda(x+1)} \right)^i = \frac{x}{\lambda - 1 + \lambda x},$$

and that

$$\int_0^1 \Delta(\nu) d\nu = \sum_{i=1}^{\infty} i \mu(i) = x \sum_{i=1}^{\infty} i \left(\frac{1}{1+x} \right)^i = \sum_{i=0}^{\infty} \left(\frac{1}{1+x} \right)^i = \frac{1+x}{x}.$$

Therefore it follows from (11) that

$$M(x) = \frac{\lambda^{-\int_0^1 \Delta(\nu) d\nu}}{\int_0^1 \lambda^{-\Delta(\nu)} d\nu} = \frac{1}{\Lambda(x)} \lambda^{-\int_0^1 \Delta(\nu) d\nu} = \lambda^{-\frac{1+x}{x}} \frac{\lambda - 1 + \lambda x}{x}.$$

6.4 Characterization of the BGP

Along the BGP (26) reads

$$(r + z - g_Y) V(\Delta, t) = Y(t) (1 - \lambda^{-\Delta}) + \max_I \left\{ I (V(\Delta + 1, t) - V(\Delta, t)) - w_I(t) \lambda^{-\Delta} \frac{1}{\varphi} I^\gamma \right\}.$$

We now conjecture that the value function takes the form of

$$V(\Delta, t) = \kappa(t) - \phi(t) \lambda^{-\Delta}. \quad (39)$$

Using this conjecture, we get $V(\Delta + 1, t) - V(\Delta, t) = \phi(t) \frac{\lambda - 1}{\lambda} \lambda^{-\Delta}$. If V grows at rate g_Y , (39) implies that both $\kappa(t)$ and $\phi(t)$ grow at rate g_Y too. Hence let us normalize all variables by $\exp(g_Y t)$ so that they become stationary and t ceases to be a state variables. The Bellman equation therefore reads

$$(r + z - g_Y) V(\Delta) = Y(1 - \lambda^{-\Delta}) + \lambda^{-\Delta} \max_I \left\{ I \phi \frac{\lambda - 1}{\lambda} - w_I \frac{1}{\varphi} I^\gamma \right\}.$$

The optimal innovation flow rate I^* is implicitly defined by the necessary and sufficient FOC

$$\gamma \frac{1}{\varphi} I^{\gamma-1} = \frac{\phi}{w_I} \frac{\lambda - 1}{\lambda}, \quad (40)$$

where $\frac{\phi}{w_I}$ is constant along the BGP. Hence, I^* is independent of time and the same for all firms. Substituting the conjecture (39), we get that

$$\begin{aligned} (r + z - g_Y) \kappa - (r + z - g_Y) \phi \lambda^{-\Delta} &= Y(1 - \lambda^{-\Delta}) + \lambda^{-\Delta} \left[I \phi \frac{\lambda - 1}{\lambda} - w_I \frac{1}{\varphi} I^\gamma \right] \\ &= Y + \left[(\gamma - 1) w_I \frac{1}{\varphi} I^\gamma - Y \right] \lambda^{-\Delta}. \end{aligned}$$

As this has to hold for all Δ , it is clear that

$$-(r + z - g_Y) \phi = (\gamma - 1) w_I \frac{1}{\varphi} I^\gamma - Y,$$

which yields

$$\phi = \frac{Y - (\gamma - 1) w_I \frac{1}{\varphi} I^\gamma}{r + z - g_Y}. \quad (41)$$

Similarly we get $\kappa = \frac{Y}{r + z - g_Y}$, so that the stationary version of the value function is given by

$$V(\Delta) = \frac{(1 - \lambda^{-\Delta}) Y + (\gamma - 1) w_I \frac{1}{\varphi} I^\gamma \lambda^{-\Delta}}{r + z - g_Y}.$$

Hence,

$$V(\Delta, t) = \frac{(1 - \lambda^{-\Delta}) Y(t) + (\gamma - 1) w_I(t) \frac{1}{\varphi} I^\gamma \lambda^{-\Delta}}{\rho + z},$$

where the last equality uses that the Euler equation implies that $\frac{\dot{c}(t)}{c(t)} = g_Y = r - \rho$. For $V(\Delta, t)$ to grow at a constant rate, we need to show that $w_I(t)$ and output $Y(t)$ grow at a constant rate g_Y . This follows from 29 in the main text. That the growth rate along the BGP is constant is shown in Section 6.5.

6.5 BGP Growth Rate

To derive the BGP growth rate, note that

$$\begin{aligned} \ln(Q(t+\delta)) &= \int_0^1 \ln(q(\nu, t+\delta)) d\nu \\ &= \ln(Q(t)) + \ln(\lambda) [\zeta(I, \delta) + \zeta(E, \delta)] + \ln(\lambda^2) \zeta(B, \delta), \end{aligned}$$

where I, E and B denote the set of sectors which experience an innovation, entry or both in the interval of length δ and $\zeta(I, \delta), \zeta(E, \delta)$ and $\zeta(B, \delta)$ are the respective measures. With constant equilibrium flow rates I and z , the respective measures are given by $\zeta(I, \delta) = I\delta$, $\zeta(E, \delta) = z\delta$ and $\zeta(B, \delta) = zI\delta^2$. The BGP growth rate of aggregate productivity is therefore given by

$$g_Q = \lim_{\delta \rightarrow 0} \frac{\ln(Q(t+\delta)) - \ln(Q(t))}{\delta} = \ln(\lambda) (I + z).$$

6.6 Uniqueness and Existence of BGP

We have to establish that the two equations

$$\frac{1}{\eta} = \chi \frac{1}{\varphi} \left[\gamma I^{\gamma-1} + \frac{(\gamma-1)}{\rho + xI} I^\gamma \right] \quad (42)$$

$$L_S = \frac{1}{\varphi} I^\gamma \frac{x}{\lambda - 1 + \lambda x} + \frac{x}{\eta} I \quad (43)$$

have a unique solution. Consider first (43). Totally differentiating yields

$$0 = \left(\frac{1}{\varphi} I^\gamma \frac{\lambda - 1}{(\lambda - 1 + \lambda x)^2} + \frac{1}{\eta} I \right) dx + \left(\gamma \frac{1}{\varphi} I^{\gamma-1} \frac{x}{\lambda - 1 + \lambda x} + \frac{x}{\eta} \right) dI$$

so that

$$\frac{dx}{dI} = - \frac{\gamma \frac{1}{\varphi} I^{\gamma-1} \frac{x}{\lambda - 1 + \lambda x} + \frac{x}{\eta}}{\frac{1}{\varphi} I^\gamma \frac{\lambda - 1}{(\lambda - 1 + \lambda x)^2} + \frac{1}{\eta} I} < 0,$$

i.e. (43) defines a function $x(I; \varphi, \lambda, \eta)$, which is decreasing in I . This is intuitive: for a given supply of scientists, if more entry takes place, the incumbents' innovation rate has to decrease. Now consider (42). Totally differentiating implies that

$$\begin{aligned} 0 &= \left(\gamma(\gamma-1) I^{\gamma-2} + \frac{\gamma(\gamma-1) I^{\gamma-1} (\rho + xI) - (\gamma-1) I^\gamma x}{(\rho + xI)^2} \right) dI - \frac{(\gamma-1) I^{\gamma+1}}{(\rho + xI)^2} dx \\ &= (\gamma-1) \left[\left(\gamma I^{\gamma-2} + \frac{\gamma I^{\gamma-1} \rho + (\gamma-1) I^\gamma x}{(\rho + xI)^2} \right) dI - \frac{I^{\gamma+1}}{(\rho + xI)^2} dx \right], \end{aligned}$$

which gives that

$$\begin{aligned} \frac{dx}{dI} &= \frac{\gamma I^{\gamma-2} + \frac{\gamma I^{\gamma-1} \rho + (\gamma-1) I^\gamma x}{(\rho + xI)^2}}{\frac{I^{\gamma+1}}{(\rho + xI)^2}} = \frac{\gamma I^{\gamma-2} (\rho + xI)^2 + \gamma I^{\gamma-1} \rho + (\gamma-1) I^\gamma x}{I^{\gamma+1}} \\ &= \left(\frac{1}{I} \right)^3 \left(\gamma (\rho + xI)^2 + \gamma \rho I + (\gamma-1) x I^2 \right) > 0. \end{aligned}$$

Hence, (42) defines a function $x(I; \varphi, \chi, \eta)$, which is increasing in I . To conclude the proof note that (43) implies that $\lim_{I \rightarrow 0} x = \infty$ and that $\lim_{I \rightarrow \infty} x = 0$. Similarly, (42) implies that $\lim_{I \rightarrow \infty} x = \infty$. As both functions are continuous, there is a unique intersection.

Dependent Variable:	$\Lambda = E[\lambda^{-\Delta}]$				
	(1)	(2)	(3)	(4)	(5)
Entry rate z	0.0903** (0.0249)	0.102** (0.0364)	0.103** (0.0372)	0.115** (0.0372)	0.0679** (0.0221)
ln(population)			0.00239 (0.00676)	0.00190 (0.00529)	
share of asphalted streets				0.176* (0.0928)	
no of bank branches				-0.0186 (0.0258)	
ln(small firms)				-0.00713 (0.0110)	
agricultural share				-0.0560 (0.113)	
Year FE	No	Yes	Yes	Yes	Yes
Province FE	No	No	No	No	Yes
N	168	168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is the factor price distortion index $\Lambda_{r,t} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i$, where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t}^i l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i and α is normalized to 0.7. The entry rate $z_{r,t}$ is the share of firms who enter the manufacturing sector in region r at time t . “ln(population)” is the log of the total population in the region in 1996. “share of asphalted streets” is the share of villages in the region, which are accessible by asphalt streets in 1996. “no of bank branches” is the average number of banks per village in the region as of 1996. “ln(small firms)” is the log of the average number of informal firms per village in the region in 1996. “agricultural share” is the average share of the village population whose main income source is agricultural.

Table 1: Entry and the degree of misallocation: Λ

Dependent Variable:	$\Psi = \lambda^{-E[\Delta]}$				
	(1)	(2)	(3)	(4)	(5)
Entry rate z	0.0711** (0.0231)	0.0671* (0.0322)	0.0686** (0.0327)	0.0871** (0.0371)	0.0343* (0.0185)
ln(population)			0.00319 (0.00590)	0.00265 (0.00417)	
share of asphalted streets				0.162** (0.0756)	
no of bank branches				-0.0241 (0.0212)	
ln(small firms)				-0.000326 (0.0100)	
agricultural share				-0.0960 (0.0953)	
Year FE	No	Yes	Yes	Yes	Yes
Province FE	No	No	No	No	Yes
N	168	168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is $\Psi_{r,t} = \lambda^{\left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\rho_{r,t}^i)\right] / \ln(\lambda)}$, where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$,

$N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t} l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i , α is normalized to 0.7 and $\lambda = 1.15$. The entry rate $z_{r,t}$ is the share of firms who enter the manufacturing sector in region r at time t . “ln(population)” is the log of the total population in the region in 1996. “share of asphalted streets” is the share of villages in the region, which are accessible by asphalt streets in 1996. “no of bank branches” is the average number of banks per village in the region as of 1996. “ln(small firms)” is the log of the average number of informal firms per village in the region in 1996. “agricultural share” is the average share of the village population whose main income source is agricultural.

Table 2: Entry and the degree of misallocation: Ψ

	10%	25%	50%	75%	90%
No regional controls					
Entry rate z	0.0376** (0.00994)	0.0548** (0.0252)	0.0827** (0.0320)	0.0963** (0.0412)	0.104 (0.0679)
Regional controls and year fixed effect					
Entry rate z	0.0460** (0.0198)	0.0682* (0.0372)	0.106** (0.0416)	0.0982* (0.0528)	0.118 (0.0708)
N	168	168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is τ th quantile of the distribution of $\rho_{r,t}^i$, where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $w_{r,t} l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i , α is set to 0.7 and $\tau \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$. The entry rate $z_{r,t}$ is the share of firms who enter the manufacturing sector in region r at time t . The regional controls in the second panel are the total population, the share of villages, which are accessible by asphalt streets, the average number of banks per village, the average number of informal firms per village and the average share of the village population whose main income source is agricultural in each region in 1996.

Table 3: Entry and the distribution of mark-ups: Quantiles

	(1)	(2)	(3)	(4)
Dependent Variable: $E[\Delta]$				
Entry rate z	-0.366** (0.117)	-0.329* (0.164)	-0.340* (0.167)	-0.431** (0.184)
Dependent Variable: $sd(\Delta)$				
Entry rate z	-0.0664 (0.0687)	-0.00251 (0.101)	-0.00599 (0.102)	-0.0958 (0.114)
Dependent Variable: $sd(\ln(L))$				
Entry rate z	-0.355** (0.149)	-0.460** (0.182)	-0.469** (0.178)	-0.365* (0.183)
Year FE	No	Yes	Yes	Yes
$\ln(pop)$	No	No	Yes	Yes
Full regional controls	No	No	No	Yes
N	168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are: the average log mark-up $\ln(\rho)_{r,t} = \frac{1}{N_{r,t}} \sum_i \ln(\rho_{r,t}^i) \propto E[\Delta]$ (Panel A), the dispersion of log mark-ups $\hat{\sigma}_{r,t} = \left[\frac{1}{N_{r,t}} \sum_i \left(\ln(\rho_{r,t}^i) - \overline{\ln(\rho)}_{r,t} \right)^2 \right]^{1/2} \propto \sigma_{\Delta}$ (Panel B) and the dispersion of log employment $\left[\frac{1}{N_{r,t}} \sum_i \left(\ln(l_{r,t}^i) - \overline{\ln(l)}_{r,t} \right)^2 \right]^{1/2} \propto \sigma_{\Delta}$ (Panel C). Here, $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t} l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i , α is normalized to 0.7 and $l_{r,t}^i$ is total employment of firm i . “ $\ln(Pop)$ ” is the log of the total population in the region in 1996 and the other regional controls are the share of villages, which are accessible by asphalt streets, the average number of banks per village, the average number of informal firms per village and the average share of the village population whose main income source is agricultural in each region in 1996.

Table 4: Further implications of the theory: $E[\Delta]$, $var(\Delta)$ and $var(\ln L)$

Dependent Variable:	$M = \frac{\Psi}{\Lambda} = \frac{\lambda - E[\Delta]}{E[\lambda - \Delta]}$				
	(1)	(2)	(3)	(4)	(5)
Entry rate z	0.0121 (0.0414)	-0.0377 (0.0536)	-0.0351 (0.0546)	0.00282 (0.0587)	-0.0562 (0.0373)
ln(population)			0.00567 (0.00668)	0.00529 (0.00468)	
share of asphalted streets				0.0863 (0.0639)	
no of bank branches				-0.0401* (0.0220)	
ln(small firms)				0.0194* (0.00987)	
agricultural share				-0.218* (0.105)	
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	No	No	No	No	Yes
N	168	168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is $M_{r,t} = \frac{\Psi_{r,t}}{\Lambda_{r,t}} = \lambda \left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\rho_{r,t}^i) \right] / \ln(\lambda) / \left(\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i \right)$, where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t}^i l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i , α is normalized to 0.7 and $\lambda = 1.15$. The entry rate $z_{r,t}$ is the share of firms who enter the manufacturing sector in region r at time t . “ln(population)” is the log of the total population in the region in 1996. “share of asphalted streets” is the share of villages in the region, which are accessible by asphalt streets in 1996. “no of bank branches” is the average number of banks per village in the region as of 1996. “ln(small firms)” is the log of the average number of informal firms per village in the region in 1996. “agricultural share” is the average share of the village population whose main income source is agricultural.

Table 5: Entry and TFP

Dependent Variable:		Entry rate z				Λ		Ψ	
share of firms with foreign capital	0.192*	(0.0960)				-0.0592	-0.140	-0.0709	-0.170
						(0.248)	(0.300)	(0.198)	(0.192)
share of firms financed by foreign loans	0.0362	(0.352)				-0.193	-0.283	-0.166	-0.241
						(0.346)	(0.322)	(0.194)	(0.237)
share of firms financed by FDI	0.274**	(0.116)				0.218	0.497	0.156	0.437
						(0.331)	(0.387)	(0.313)	(0.287)
share of firms financed by the capital market	0.334	(1.716)				0.435	-0.347	0.740	-0.538
						(0.984)	(0.650)	(0.975)	(0.394)
Entry rate z						0.0990**	0.108**	0.0667**	0.0842**
						(0.0352)	(0.0337)	(0.0319)	(0.0357)
Year FE	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Regional controls	Yes	Yes	Yes	Yes		No	Yes	No	Yes
N	168	168	168	168		168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are: the share of firms who enter the manufacturing sector in region

r at time t (the entry rate) (Panel A), $\Lambda_{r,t} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i$, (Panel B) and $\Psi_{r,t} = \lambda \left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\rho_{r,t}^i) \right] / \ln(\lambda)$ (Panel

C), where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t} l_{r,t}^i$ is the wagebill of firm i ,

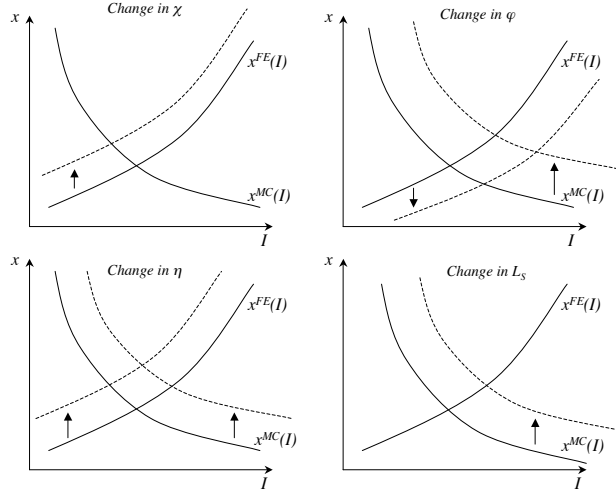
$Rev_{r,t}^i$ is the revenue of firm i and α is normalized to 0.7. The independent variables are: The share of firms in region r at time t which are at least partially owned by foreign investors and whose investment expenses are at least partially financed through foreign loans, through FDI and through the issuance of equity and/or bonds in the official capital market. The regional controls are the total population, the share of villages, which are accessible by asphalt streets, the average number of banks per village, the average number of informal firms per village and the average share of the village population whose main income source is agricultural in each region in 1996.

Table 6: The importance of credit constraints

Dependent Variable:	ρ		Entry rate z		Λ		Φ	
state owned	0.0692**							
	(0.0113)							
government stake	0.0283**							
	(0.00547)							
Entry rate z					0.0968**	0.112**	0.0616*	0.0848**
					(0.0407)	(0.0399)	(0.0355)	(0.0389)
share of firms owned by the state			-0.606		-0.430	-0.155	-0.480	-0.178
			(0.657)		(0.394)	(0.396)	(0.300)	(0.326)
share of firms with a gov. stake			0.103		0.00416	0.0177	0.00911	0.00785
			(0.126)		(0.0717)	(0.0548)	(0.0713)	(0.0416)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	-	-	Yes	Yes	No	Yes	No	Yes
N	70443	70443	168	168	168	168	168	168

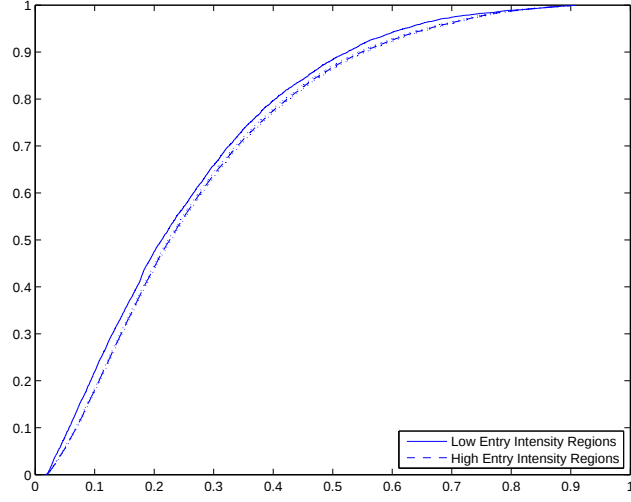
Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are: the firm-specific wage share $\rho_{r,t}^i$ (Panel A), the share of firms who enter the manufacturing sector in region r at time t (the entry rate) (Panel B), $\Lambda_{r,t} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i$, (Panel C) and $\Psi_{r,t} = \lambda \left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\rho_{r,t}^i) \right] / \ln(\lambda)$ (Panel D), where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t}^i l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i and α is normalized to 0.7. The independent variables are: Indicator variables if the firm is owned by the state or if the government (either central or local) has any financial stake in the firm and the share of firms in region r at time t which are owned by the state or where the local or central government has any financial stake. The regional controls are the total population, the share of villages, which are accessible by asphalt streets, the average number of banks per village, the average number of informal firms per village and the average share of the village population whose main income source is agricultural in each region in 1996. The firm-level regressions in columns 1 and 2 contain a full set of industry, region and type of ownership fixed effects.

Table 7: The importance of policy distortions



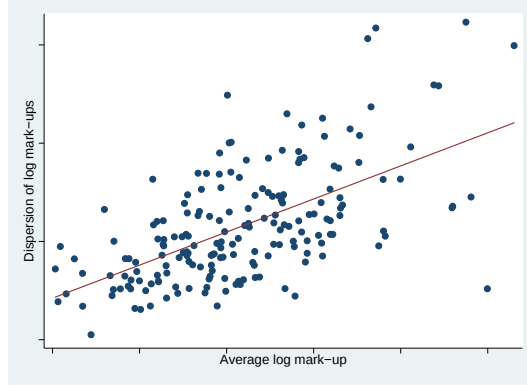
Notes: The figure shows the determination of the equilibrium innovation rate I and the equilibrium entry intensity $x = \frac{I}{z}$. The schedule $x^{FE}(I)$ stems from the free entry condition and is implicitly defined in (31). The schedule $x^{MC}(I)$ stems from the market clearing condition and is implicitly defined in (32). The comparative statics considered are: a change in the entry costs χ (Panel 1), a change in the innovation productivity of incumbents φ (Panel 2), a change in the innovation flow rate of entrants η (Panel 3) and a change in the supply of innovation resources of the economy L_S (Panel 4).

Figure 1: Comparative statics of the equilibrium entry intensity x and innovation rate I



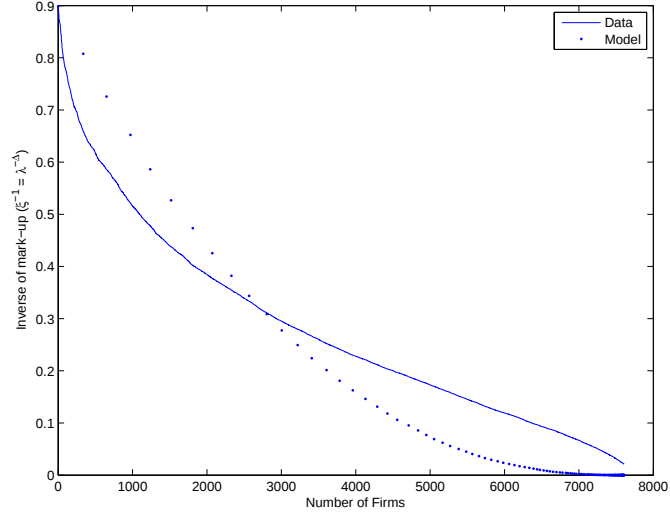
Notes: The Figure shows the estimated distribution function of firm-specific employment shares $\frac{w_i l_i}{Rev_i(1-\alpha)}$, for “high entry” and “low entry” regions. $w_i l_i$ is the wagebill and Rev_i denotes expenditure. “High entry” regions are those regions with the highest mean entry rates, which cover 25% of the population of firms (dashed line). “Low entry” regions are those regions with the lowest mean entry rates, which cover 25% of the population of firms (solid line). The dotted lines around the dashed line are 95% pointwise confidence intervals.

Figure 2: Distribution of mark-ups in high- and low-entry regions.



Notes: The figure shows the correlation between the average log mark-up $\overline{\ln(\rho)}_{r,t} = \frac{1}{N_{r,t}} \sum_i \ln(\rho_{r,t}^i)$ and the dispersion of log mark-ups $\hat{\sigma}_{r,t}$, where $\hat{\sigma}_{r,t}^2 = \frac{1}{N_{r,t}} \sum_i (\ln(\rho_{r,t}^i) - \overline{\ln(\rho)}_{r,t})^2$ and $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i(1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t}^i l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i and α is normalized to 0.7.

Figure 3: The regional correlation of mark-up dispersion and the level of mark-ups



Notes: The Figure shows the estimated distribution of firm-specific employment shares $\frac{w_i l_i}{Rev_i(1-\alpha)}$, where $w_i l_i$ is the wagebill, Rev_i denotes expenditure and α is normalized to 0.7 (solid line). According to (34), the model implies that $\frac{w_i l_i}{Rev_i(1-\alpha)} = \lambda^{-\Delta_i}$. The model's implied distribution of $\lambda^{-\Delta}$ is given in (19), is shown in dotted lines and the entry intensity x is calibrated so that $E[\lambda^{-\Delta}] = E\left[\frac{w_i l_i}{Rev_i(1-\alpha)}\right]$.

Figure 4: The distribution of mark-ups: Data versus model

7 Web Appendix [NOT FOR PUBLICATION]

In this appendix I consider various robustness checks for the results in the main text. In particular, instead of the simply entry rate, I consider the entry intensity

$$x_{r,t} = \frac{z_{r,t}}{I_{r,t}}, \quad (44)$$

which is a sufficient statistic according to the model. To calculate the incumbents' innovation rate, I use that the growth rate of TFP $Q(t)$ is given by $g_Q = \ln(\lambda)(I + z)$ (see (6.5)). The average TFP growth rate in my data is equal to 8%,³⁶ which together with $\lambda = 1.15$ from above and the observed entry rate $z_{r,t}$ allows me to calculate $I_{r,t}$. Hence, I restrict the growth rate and λ to be the same across regions. This is consistent with Proposition 2, which is stated conditionally on aggregate productivity and the level of capital. Note that $x_{r,t}$ calculated this way is a simple nonlinear transformation of the entry rate observed in the data. In Table 8 I consider various robustness checks for the main results. The results for $\Lambda(\Psi)$ are contained in upper (lower) panel). In particular, I use robust standard errors instead of the clustered standard errors in the main analysis (rows 1 and 5), estimate the model using weighted least squares, where the weights are the number of firms used to construct the regional averages from the firm level data (rows 2 and 6) and consider the entry intensity calculated in (44) as the dependent variable in rows 3 and 7. Table 8 shows that all the coefficients are positive as the theory predicts and that most of them are significant. That the weighted least squares results are generally slightly weaker than the main specification is due to the fact that the regions in Indonesia differ substantially in their population of firms so that the weighted least squares estimate puts high weights on a small number of regions. In rows 4 and 8 I include the crisis year 1998 in the sample. Doing so renders the result for the Ψ measure insignificant but the coefficients are still positive. In Figure 5 I show in what sense the crisis year 1998 is special. It is clearly seen that there is a substantial drop in both measures, i.e. the crisis episode is measured as a large decline in TFP and factor prices. This finding is reminiscent of Sandleris and Wright (2010) and Gopinath and Neiman (2010), who also find large measured declines in economic efficiency during the crisis in Argentina. In Table 9 I show more evidence that it is the crisis episode in 1998, which generates the positive but insignificant coefficient in Table 8. In particular, each column of Table 9 refers to a regression, where the respective year is dropped from the sample. In Panel A, I include the crisis year 1998 and replicate the positive but insignificant coefficient from Table 8. Panel B excludes the crisis year 1998 and shows that the effect of entry is positive and significant regardless of which other years are dropped (1995 being an exception). In Panel C, I only consider the time period before the financial crisis, which also delivers a positive and significant coefficient.³⁷ Going back to Table 8, rows 9 and 10 of Table show that the exact choice of λ is innocuous for the results. Instead of $\lambda = 1.15$ as for the benchmark results, I take $\lambda = 1.05$ and $\lambda = 2$ and the entry measure is still positive and significant.

³⁶To calculate this, I calculate aggregate output, capital and labor from the micro data by simply summing over all establishments and then calculate TFP $\ln(Q(t))$ as the residual $\ln(Y(t)) - \alpha \ln(K(t)) - (1 - \alpha) \ln(L(t))$, again assuming $\alpha = 0.3$. The average growth rate of $Q(t)$ is equal to 8%. Note that this calculation assumes that the economy is in steady state, so that the TFP distortion index $M(t)$ is constant and does not affect the growth rate. The case of Indonesia is arguably a less than perfect example, as it suffered heavily from the Asian crisis in 1997. However, the average growth rate between 1990 and 1996 is also around 8% and in the regressions below I report specifications using year fixed effects, which should at least alleviate some of the concerns.

³⁷Note that all regressions do include year fixed effects. Hence, the results in Table 9 imply that the cross-sectional correlation between the entry rate the misallocation measure Ψ differs in 1998. This is actually the case in that the entry rates are relatively high and the province with the highest entry rate has the lowest misallocation measure.

In Table 10 I show that the results are robust to the number of firms used to calculate the different dependent variables on the year-province level. In particular, recall that for example $\Lambda_{r,t} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i$, where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$ and $N_{r,t}$ denotes the number of firms in province r at year t . The larger $N_{r,t}$, the less noisy will $\Lambda_{r,t}$ be. For the main analysis I restricted attention to year-province cells satisfying $N_{r,t} \geq 25$. Table 10 replicates the main results for different values of $N_{r,t}$. Naturally the number of observations declines as the $N_{r,t}$ -cutoff increases. However, the results are robust to such changes.

Figure 6 shows in what sense the observed first order stochastic dominance between the high and low entry regions reported in Figure 2 is sensitive to the cutoff. The left panel of Figure 6 shows that first order stochastic dominance holds, when I consider the extreme 10% of the regions, the right panel shows that I reject first order stochastic dominance, when I split the sample at the median.

Tables 11 and 12 replicate the results for the different quantiles of the mark-ups distribution (Table 3) and the other measures related to the distribution of mark-ups (Table 4). In particular, I again consider the different entry measures and the different estimation methods and show that the results are robust. The coefficients conform with the theory and are mostly significant.

Table 13 contains some evidence that the firm characteristics, according to which I identify firms as being credit constraint, are informative about the financial environment. Rows 1 to 6 show that regions with a better financial infrastructure (as measured by the no of bank branches) have weakly less constraint firms as measured by foreign capital holdings and foreign direct investment. In the last four columns I exploit information from data in 1996, where firms are asked if they face credit constraints. Columns 7 and 9 show that firms with foreign capital and experiencing foreign direct investment are less likely to face credit constraints. Columns 8 and 10 show that the same results emerges, when I take the indicator if the firm answered the question what actions are undertaken to ease the constraints as the dependent variables. Hence, the firm characteristics employed in the main text are likely to be related to the credit constraints the firm faces.

Tables 14 and on page 55 contain the results concerning measurement error discussed in Section 3.4.

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: $\Lambda = E[\lambda - \Delta]$				
Robust std errors	0.0903** (0.0304)	0.102** (0.0361)	0.103** (0.0363)	0.115** (0.0356)	0.0679** (0.0251)
Weighted LS	0.0609** (0.0257)	0.0782** (0.0333)	0.0573* (0.0298)	0.0620* (0.0318)	0.0325 (0.0216)
Entry intensity $\frac{z}{i}$	0.00649** (0.00255)	0.00785* (0.00379)	0.00786* (0.00383)	0.0104** (0.00398)	0.00471** (0.00221)
Including 1998 data	0.0891** (0.0334)	0.0861* (0.0431)	0.0882* (0.0435)	0.0896* (0.0465)	0.0622** (0.0209)
	Dependent Variable: $\Psi = \lambda - E[\Delta]$				
Robust std errors	0.0711** (0.0328)	0.0671* (0.0356)	0.0686* (0.0357)	0.0871** (0.0326)	0.0343* (0.0202)
Weighted LS	0.0638** (0.0215)	0.0555* (0.0318)	0.0449 (0.0309)	0.0419 (0.0271)	0.00841 (0.0201)
Entry intensity $\frac{z}{i}$	0.00500* (0.00243)	0.00484 (0.00333)	0.00487 (0.00335)	0.00783** (0.00373)	0.00172 (0.00154)
Including 1998 data	0.0609 (0.0361)	0.0408 (0.0443)	0.0432 (0.0438)	0.0551 (0.0486)	0.0229 (0.0175)
$\lambda = 1.05$	0.0711** (0.0231)	0.0671* (0.0322)	0.0686** (0.0327)	0.0871** (0.0371)	0.0343* (0.0185)
$\lambda = 2$	0.0711** (0.0231)	0.0671* (0.0322)	0.0686** (0.0327)	0.0871** (0.0371)	0.0343* (0.0185)

Notes: Except for rows 1 and 5, standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is Λ in Panel A (see Table 1 for construction) and Ψ in Panel B (see Table 2 for construction). I report only the coefficients on the entry rate. Column 1 is the simple correlation, column 2 controls for year fixed effects, column 3 controls for the regional population and year fixed effects, column 4 contains the full set of regional controls (see again Tables 1 and 2 for details) and year fixed effects and column 5 has both year and region fixed effects. I report the results (i) using robust standard errors (row 1 and 5), (ii) using weighted least squares where the weights are the number of firms in each region-year cell to calculate the entry measure and the dependent variables (rows 2 and 6), (iii) using the entry intensity (see (44)) as the entry measure (rows 3 and 7), (iv) including the data for 1998 (rows 4 and 8) and (v) using different values for λ in the construction of Ψ (rows 9 and 10).

Table 8: Robustness of main results

Dependent Variable: $\Psi = \lambda^{-E[\Delta]}$									
	1991	1992	1993	1994	1995	1996	1997	1999	2000
Panel A: Including 1998									
Entry	0.0505 (0.0531)	0.0540 (0.0599)	0.0590 (0.0489)	0.0441 (0.0478)	0.0211 (0.0493)	0.0716 (0.0499)	0.0518 (0.0486)	0.0516 (0.0485)	0.0613 (0.0500)
N	169	169	170	169	167	166	169	168	168
Panel B: Excluding 1998									
Entry	0.0857* (0.0413)	0.0984** (0.0442)	0.0921** (0.0378)	0.0752* (0.0384)	0.0516 (0.0386)	0.112** (0.0356)	0.0829** (0.0386)	0.0852** (0.0366)	0.0991** (0.0363)
N	150	150	151	150	148	147	150	149	149
Panel C: Including only pre-crisis years (1991-1997)									
Entry	0.0973** (0.0355)								
N	130								

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is Ψ (see Table 2 for construction). Each column refers to a regression, where the respective year is dropped. All regression contain year fixed effects and the full set of regional controls (see Table 2 for details). Panel A excludes only the year given in the respective column, Panel B excludes 1998 and the year in the respective column and Panel C considers only the pre-crisis episode from 1991 to 1997.

Table 9: The crisis year 1998

$N_{r,t}$ -Cutoff	25	35	45	60	80
Dep. Var.: $\Lambda = E[\lambda^{-\Delta}]$					
Entry	0.115** (0.0372)	0.112** (0.0390)	0.0882* (0.0496)	0.107** (0.0388)	0.0695** (0.0300)
Dep. Var.: $\Phi = \lambda^{E[-\Delta]}$					
Entry	0.0871** (0.0371)	0.0852** (0.0341)	0.0509 (0.0366)	0.0755* (0.0378)	0.0503 (0.0291)
Dep. Var.: 25%-quantile					
Entry	0.0682* (0.0372)	0.0599* (0.0325)	0.0303 (0.0239)	0.0418 (0.0276)	0.0285 (0.0284)
Dep. Var.: 50%-quantile					
Entry	0.106** (0.0416)	0.0994** (0.0344)	0.0786* (0.0392)	0.0926** (0.0426)	0.0740* (0.0390)
Dep. Var.: 75%-quantile					
Entry	0.0982* (0.0528)	0.130** (0.0553)	0.110 (0.0717)	0.158** (0.0665)	0.164* (0.0786)
N	168	159	143	110	88

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is Λ in Panel A (see Table 1 for construction), Ψ in Panel B (see Table 2 for construction) and the 25, 50 and 75% quantile of the distribution of $\rho_{r,t}^i = \frac{w_{r,t} i_{r,t}^i}{Rev_{r,t}^i(1-\alpha)}$ in Panels C, D, and E respectively (see Table 3 for construction). Each column contains the results for a different cutoff for the minimum number of firms in each year-province cell ($N_{r,t}$) to construct the aggregate measures Λ and Ψ and the respective quantiles, i.e. in column 1 all year-province cells containing less than 25 firms are dropped. All regressions contain year fixed effects and the full set of regional controls (see Tables 1 and 2 for details).

Table 10: Robustness for the number of firms in year-region cells

	10%	25%	50%	75%	90%
Robust std errors	0.0460** (0.0190)	0.0682** (0.0291)	0.106** (0.0358)	0.0982* (0.0546)	0.118 (0.0769)
Entry intensity $x = \frac{z}{I}$ [no contr.]	0.00306** (0.00114)	0.00431 (0.00257)	0.00545 (0.00328)	0.00626 (0.00387)	0.00552 (0.00682)
Entry intensity $x = \frac{z}{I}$ [with contr.]	0.00409** (0.00189)	0.00641* (0.00358)	0.00947** (0.00399)	0.00932 (0.00545)	0.0100 (0.00729)
weighted LS	0.0173 (0.0150)	0.0312 (0.0221)	0.0592* (0.0343)	0.0729 (0.0521)	0.0912 (0.0533)
N	168	168	168	168	168

Standard errors in parentheses

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

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variable is τ th quantile of the distribution of $\rho_{r,t}^i$ (see Table 3) and I report the coefficient on the entry measure. All specifications contain year fixed effects and except for row 2 also control for the full set of regional controls (see again Table 3). I consider: (i) robust standard errors in row 1, (ii) the entry intensity (see (44)) as a measure of entry with and without regional controls and the weighted least squares regression, where the weights are the number of firms in each region-year cell to calculate the entry measure and the quantiles.

Table 11: Entry and the distribution of mark-ups: Robustness for quantile regressions

	(1)	(2)	(3)	(4)
Dependent Variable: $E[\Delta]$				
Robust std errors	-0.366** (0.169)	-0.329* (0.183)	-0.340* (0.183)	-0.431** (0.166)
Entry intensity $x = \frac{z}{I}$	-0.0239* (0.0118)	-0.0214 (0.0169)	-0.0217 (0.0170)	-0.0354* (0.0185)
Weighted LS	-0.299** (0.105)	-0.244 (0.151)	-0.218 (0.144)	-0.201 (0.134)
Dependent Variable: $sd(\Delta)$				
Robust std errors	-0.0664 (0.125)	-0.00251 (0.130)	-0.00599 (0.131)	-0.0958 (0.124)
Entry intensity $x = \frac{z}{I}$	-0.00501 (0.00806)	0.000914 (0.0109)	0.000766 (0.0111)	-0.00691 (0.0104)
WLS	-0.163** (0.0764)	0.0133 (0.159)	-0.0342 (0.152)	0.0374 (0.0666)
Dependent Variable: $sd(\ln(L))$				
Robust std errors	-0.355** (0.122)	-0.460** (0.139)	-0.469** (0.139)	-0.365** (0.121)
Entry intensity $x = \frac{z}{I}$	-0.0173 (0.0174)	-0.0250 (0.0208)	-0.0251 (0.0209)	-0.0204 (0.0182)
Weighted LS	-0.166** (0.0660)	-0.313** (0.133)	-0.286** (0.122)	-0.171* (0.0861)
N	168	168	168	168

Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are defined in Table 4. I consider: (i) robust standard errors in rows 1,4 and 7, (ii) the entry intensity (see (44)) as a measure of entry in rows 2,5, and 8 and the weighted least squares regression, where the weights are the number of firms in each region-year cell to calculate the entry measure and the dependent variables in rows 3,6 and 9.

Table 12: Further implications of the theory: Robustness

	Foreign capital				FDI				Probability of facing credit constraints			
no of bank branches	0.0100** (0.00233)	0.0206** (0.00745)	-0.0188 (0.0252)	0.00163 (0.00132)	0.00704 (0.00426)	-0.00933 (0.0151)						
ln(population)		-0.00248 (0.00641)	-0.00134 (0.00509)		-0.00199 (0.00298)	-0.00153 (0.00275)						
share of asphalted streets		-0.118 (0.0954)	-0.0868 (0.0816)		-0.0600 (0.0550)	-0.0631 (0.0489)						
agricultural share			-0.177 (0.142)			-0.0788 (0.0845)						
ln(small firms)			-0.00160 (0.00671)			0.00251 (0.00360)						
Foreign capital									-0.483** (0.138)	-0.464** (0.135)		
FDI											-0.414* (0.229)	-0.428* (0.229)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Only 1996 cross-section			
N	168	168	168	168	168	168	168	168	7416	7416	7416	7416

Notes: Standard errors are clustered on the regional level in columns 1 - 6. Robust standard errors in columns 7 - 10. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are: the share of firms in region r at time t that (i) have foreign capital owners (columns 1 - 3), (ii) whose investment is at least partially financed through FDI (columns 4 - 6), (iii) an indicator variable if the firm reports to be credit constraint (7 and 9) and if the firm answered the question what actions are undertaken to ease the firm experiences a lack of capital (columns 8 and 10). "Foreign Capital" is an indicator variable if the firm has foreign capital owners and "FDI" is an indicator variable if the firm's investment is partially financed by FDI. For the definition of the regional controls see Table 1. Columns 7 to 10 report probit estimates and all specifications contain year, region and sector fixed effects and control for firm-size through log total assets.

Table 13: Correlates of credit constraints

Dependent Variable:	ρ		Entry rate z		Λ		Φ	
unpaid workers	-0.00273							
	(0.00195)							
share of unpaid workers	-0.155**							
	(0.0201)							
Entry rate z					0.0995**	0.110**	0.0677**	0.0839**
					(0.0356)	(0.0381)	(0.0315)	(0.0367)
share of firms with unpaid workers			0.0557		0.0424	0.0226	0.102	0.0631
			(0.117)		(0.0831)	(0.0808)	(0.0809)	(0.0759)
average share of unpaid workers			0.820		-0.0393	0.428	-0.816	-0.217
			(1.126)		(1.146)	(0.940)	(1.160)	(0.843)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	-	-	Yes	Yes	No	Yes	No	Yes
N	70443	70443	168	168	168	168	168	168

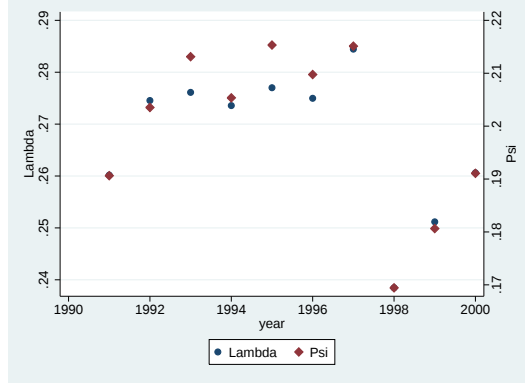
Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are: the firm-specific wage share $\rho_{r,t}^i$ (Panel A), the share of firms who enter the manufacturing sector in region r at time t (the entry rate) (Panel B), $\Lambda_{r,t} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i$, (Panel C) and $\Psi_{r,t} = \lambda \left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\rho_{r,t}^i) \right] / \ln(\lambda)$ (Panel D), where $\rho_{r,t}^i = \frac{w_{r,t} l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t} l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i and α is normalized to 0.7. The independent variables are: Indicator variable if the firm employs any unpaid workers, the share of unpaid workers employed by the firm, the share of firms in region r at time t which employing unpaid workers and the average share of unpaid workers in region r at time t . The regional controls are the total population, the share of villages, which are accessible by asphalt streets, the average number of banks per village, the average number of informal firms per village and the average share of the village population whose main income source is agricultural in each region in 1996. The firm-level regressions in columns 1 and 2 contain a full set of industry, region and type of ownership fixed effects.

Table 14: The importance of unpaid workers

Dependent Variable:	ρ	Entry rate z	Λ	Φ	Λ^{Res}	Φ^{Res}
k/l	-3.95e-09** (1.56e-09)					
$\ln(k/l)$	-0.0170** (0.00102)					
Entry			0.115** (0.0370)	0.0871** (0.0350)	0.0850** (0.0326)	0.277** (0.128)
$\log(\text{mean capital labor ratio})$		0.000323 (0.0268)	-0.00257 (0.0134)	-0.0137 (0.0119)	0.0235* (0.0126)	0.0370 (0.0495)
$\log(\text{median capital labor ratio})$		-0.0290 (0.0283)				
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Regional Controls	-	Yes	Yes	Yes	Yes	Yes
N	70443	70443	168	168	168	168
Standard errors in parentheses						
* $p < 0.10$, ** $p < 0.05$						

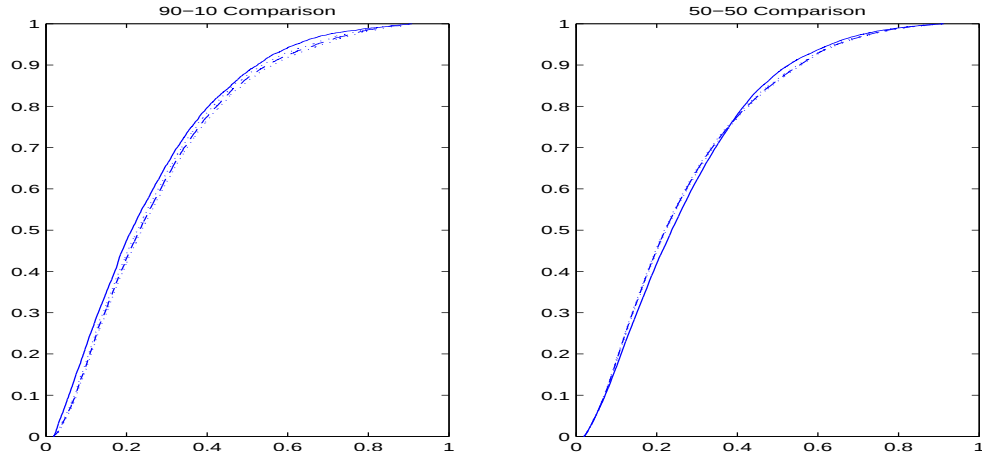
Notes: Standard errors are clustered on the regional level and shown in parentheses. ** and * denotes significance at the 5% and 10% level respectively. The dependent variables are: the firm-specific wage share $\rho_{r,t}^i$ (Panel A), the share of firms who enter the manufacturing sector in region r at time t (the entry rate) (Panel B), $\Lambda_{r,t} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \rho_{r,t}^i$ (Panel C), $\Psi_{r,t} = \lambda \left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\rho_{r,t}^i) \right] / \ln(\lambda)$ (Panel D), $\Lambda_{r,t}^{Res} = \frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \tilde{\rho}_{r,t}^i$ (Panel E) and $\Psi_{r,t}^{Res} = \lambda \left[\frac{1}{N_{r,t}} \sum_{i=1}^{N_{r,t}} \ln(\tilde{\rho}_{r,t}^i) \right] / \ln(\lambda)$, where $\tilde{\rho}_{r,t}^i = \frac{w_{r,t}^i l_{r,t}^i}{Rev_{r,t}^i (1-\alpha)}$, $N_{r,t}$ denotes the number of firms in region r at time t , $w_{r,t}^i l_{r,t}^i$ is the wagebill of firm i , $Rev_{r,t}^i$ is the revenue of firm i , $\tilde{\rho}_{r,t}^i$ is the residual of a regression of $\rho_{r,t}^i$ on a full set of sector fixed effects and the firms capital-labor ratio and α is normalized to 0.7. The independent variables are: the capital-labor ratio and the log thereof and the log of the mean (and median) capital-labor ratio in region r at time t . The regional controls are the total population, the share of villages, which are accessible by asphalt streets, the average number of banks per village, the average number of informal firms per village and the average share of the village population whose main income source is agricultural in each region in 1996. The firm-level regressions in columns 1 and 2 contain a full set of industry, region and type of ownership fixed effects.

Table 15: The importance of technological heterogeneity



Notes: The figure shows the misallocation measures Λ and Ψ (see Tables 1 and 2 for the construction) for each year in the sample.

Figure 5: Effect of the crisis in 1998



Notes: The Figure shows the estimated distribution function of firm-specific employment shares $\frac{w_i l_i}{Rev_i(1-\alpha)}$, for “high entry” and “low entry” regions. $w_i l_i$ is the wagebill and Rev_i denotes expenditure (solid line). For the figure I assume $\alpha = 0.7$. According to (34), the model implies that $\frac{w_i l_i}{Rev_i(1-\alpha)} = \lambda^{-\Delta_i}$. “High entry” regions are those regions with the highest mean entry rates, which cover 10% (left panel) or 50% (right panel) of the population of firms (dashed line). “Low entry” regions are those regions with the lowest mean entry rates, which cover 10% (left panel) or 50% (right panel) of the population of firms (solid line). The dotted lines around the dashed line are 95% pointwise confidence intervals.

Figure 6: Distribution of mark-ups in high-entry and low-entry regions: Different cut-offs