Michael Peters

Heterogeneous mark-ups, growth and endogenous misallocation

Working paper

Heterogeneous Mark-Ups, Growth and Endogenous Misallocation

Michael Peters*
London School of Economics

September 16, 2013

Abstract

The recent work on misallocation argues that aggregate productivity in poor countries is low because various market frictions prevent marginal products from being equalized. By focusing on such allocative inefficiencies, misallocation is construed as a purely static phenomenon. This paper argues that misallocation also has dynamic consequences because it interacts with firms' innovation and entry decisions, which determine the economy's growth rate. To study this link between misallocation and growth, I construct a tractable endogenous growth model with heterogeneous firms, where misallocation stems from imperfectly competitive output markets. The model has an analytical solution and hence makes precise predictions about the relationship between growth, misallocation and welfare. It stresses the importance of entry. An increase in entry reduces misallocation by fostering competition. If entry also increases the economy-wide growth rate, static misallocation and growth are negatively correlated. The welfare consequences of misallocation might therefore be much larger once these dynamic considerations are taken into account. Using firm-level panel data from Indonesia, I present reduced form evidence for the importance of imperfect output market and calibrate the structural parameters. A policy, which reduces existing entry barriers, increases growth and reduces misallocation. The dynamic growth effects are more than four times as large as their static counterpart.

JEL Codes: O11, O33, O43, D42

Keywords: Endogenous mark-ups, Imperfect product markets, Entry, TFP differences, Economic growth, Firm dynamics

---

*eMail: m.peters1@lse.ac.uk. I am especially grateful to my advisors Daron Acemoglu, Abhijit Banerjee and Rob Townsend for their invaluable guidance. I also thank Ufuk Akcigit, Marios Angeletos, Thomas Chaney, Larry Christiano, Penny Goldberg, Chang-Tai Hsieh, Felipe Iachan, Sam Kortum, Ezra Oberfield, Michael Powell, Iván Werning and particularly Susanto Basu, Arnaud Costinot and Aleh Tsyvinski and as well as numerous seminar participants whose questions and suggestions benefited the paper substantially. I am also grateful for the financial assistance of the George and Obie Shultz fund to acquire the data necessary for this project. All errors are my own.

1 Introduction

One of the major developments in the recent literature on growth and long-run economic development has been the focus on misallocation. Instead of trying to explain productivity differences across countries by differences in the set of available techniques, this literature stresses the importance of allocative efficiency and argues that aggregate TFP in poor countries is low because various frictions in the marketplace prevent marginal products from being equalized. By putting such static allocative inefficiencies at center stage, this literature suggests that misallocation is a purely static phenomenon. In this paper I argue that misallocation might also have dynamic consequences. By affecting firms' entry and innovation incentives, misallocation might not only reduce allocative efficiency statically but also have adverse growth implications. The aggregate welfare impact of misallocation might therefore be much bigger than previously appreciated.

To study this interaction between static misallocation and its dynamic ramifications, I propose a tractable endogenous growth model with heterogenous firms. Relative to existing models of misallocation, the theory is novel in two dimensions. Most importantly, productivity growth is endogenous and driven by firms' innovation and entry decisions. Hence, there is a direct link between misallocation and firms' dynamic incentives. The second novelty concerns the proximate source of misallocation, i.e. why are marginal products not equalized. Most existing theories of misallocation assume that differences in marginal products exist because some firms are constrained in their input choices. Such constraints could for example be due to credit market frictions, differences in non-market access to production factors, or preferential policies, where taxes and regulatory requirements are based on firm-specific idiosyncrasies like family ties or political conviction. While these theories stress very different causes of misallocation, they share a common economic mechanism why allocative efficiency suffers: the presence of such input barriers prevents resources from flowing from low to high productivity units whenever such constraints are binding. In this paper I take a different approach and focus instead on imperfections on output markets. I do so for empirical reasons. In an environment with imperfect output markets, static misallocation occurs because firms have monopoly power and set firm-specific mark-ups. While frictional input markets imply that high (marginal) productivity firms are constrained, imperfectly competitive output markets predict that high (marginal) productivity is indicative of market power. This difference has testable implications for the firm-level data and I provide evidence, which is supportive of misallocation driven by monopolistic power.

The main theoretical contribution of this paper is to concisely characterize the link between static misallocation and growth. The key insight is that static misallocation and productivity growth are both equilibrium outcomes given a process of firm dynamics. In particular, I consider an economy where firms have the option to invest in efficiency enhancing innovation. By doing so, they acquire a competitive edge in the market place and shield themselves from competition. Successful innovation activity will therefore allow producers to post high mark-ups. At the same time, new firms can enter the market and reduce mark-ups to arbitrage away the profits of existing producers. Hence, firms' innovation and entry behavior determines not only aggregate productivity growth but also the distribution of mark-ups, which is precisely the source of misallocation.

In fact, the theory makes precise predictions how growth and misallocation are related and stresses the importance of entry. First of all, I show that the unique equilibrium distribution of mark-ups takes a very intuitive form: it is a pareto distribution whose shape parameter is endogenous and depends only a single endogenous variable, namely the entry intensity, which is simply the share of productivity growth accounted for by entering firms. In particular, the shape of mark-up distribution is increasing in the entry intensity, reflecting the pro-competitive effect of entering firms. If an economy's aggregate productivity growth is largely accounted for by new entrants, the distribution of mark-ups is compressed because entering firms introduce sufficient churning in the economy to keep monopoly power limited. If on the other hand productivity growth is mostly generated by existing producers, the distribution of mark-ups has a fat tail because incumbent firms are unlikely to be replaced and had sufficient time to outgrow their competition. As misallocation only depends on the distribution mark-ups, it follows that the static welfare effects of misallocation are also fully parametrized by this shape parameter and hence the underlying entry intensity. In particular, entering firms reduce misallocation by continuously contesting the product market and hence keeping monopoly power in check.

Then, I exploit this equilibrium link between growth and misallocation to derive a simple expression for the dynamic gains of misallocation relative to the static impact. The key object is the correlation between the rate of entry and the aggregate growth rate. This correlation is of course endogenous and depends on the source of variation. If entry is high because entry costs low, there is limited crowding out of incumbents' innovation efforts and entry and growth are positively correlated. As entry reduces misallocation, this will also imply a negative correlation between static misallocation and equilibrium growth. If on the other hand entry is high because incumbent firms are very inefficient to improve their technology, entry and growth might be negatively related, which would imply a positive correlation between misallocation and growth.

The model makes tight predictions for the cross-sectional patterns of mark-ups and the time-series properties of firm-dynamics. I therefore apply the theory to a comprehensive panel data set of manufacturing firms in Indonesia. The empirical analysis has three parts, which follow the theory closely. I first exploit both the cross-firm and the panel dimension of the data to present evidence that the empirical pattern of inframarginal rents is consistent with models of imperfect output markets, but less easily reconciled with theories stressing constraints on input choices. In particular, I show that entering firms have relatively low marginal products in the cross-section but increasing marginal products along their life-cycle. This is consistent with the view that firms manage to charge higher prices as they successfully implement process innovations and build a customer base, but harder to square with theories where firms have dynamic incentives to alleviate input constraints. In a second step, I exploit the model's prediction that it is the ease of market entry which determines mark-ups and hence misallocation. Using the regional variation across different product markets in Indonesia, I show that seamless market entry indeed reduces monopoly power. After these reduced form exercises, I the use the structure of the model more intensely. I first conduct an accounting exercise to answer the question "How much of the dispersion in revenue productivity can the theory explain?". In the theory, the entirety of the cross-sectional productivity dispersion reflects mark-ups and is hence fully determined from the properties of firm dynamics, which I can estimate directly from the data. It is this relationship between the dynamic moments of the firm-level data and the cross-sectional distribution of revenue productivity, which puts discipline on this accounting exercise. When I perform this exercise in my data, I find a number of roughly 20%.[1] Then I calibrate the model to measure the dynamic impacts of misallocation. An increase in the extent of entry which reduces the observable dispersion of revenue productivity by 10% increases static TFP by 0.2%. This increase in entry however also affects the growth rate directly. Even though existing firms' innovation incentives are reduced by this increase in market competition, the dynamic welfare consequences are around four times as high as the static effects.

Related Literature This paper provides a link between static misallocation and aggregate productivity growth.[2] I follow Aghion and Howitt (1992) and Grossman and Helpman (1991) and construct a quality ladder model in the Schumpeterian tradition. To generate heterogeneous mark-ups, I take an otherwise standard CES demand system, but allow explicitly for limit pricing (Bernard et al., 2003; Acemoglu and Akcigit, 2012).[3] The characterization of the balanced growth path shares some similarities to Klette and Kortum (2004) and Lentz and Mortensen (2008), although I do not consider multi-product firms in my framework. My model stresses the importance of entry and the endogenous response of incumbents' innovation effort in determining the aggregate welfare gains from policies. This is similar to Atkeson and Burstein (2010, 2011) and Acemoglu et al. (2012), although these papers do not focus on the pro-competitive effect of entry.

The aggregate importance of static misallocation is pioneered in Hsieh and Klenow (2009) and Restuccia and Rogerson (2008) and has recently been applied to the cross-country data in Bartelsman et al. (2013).[4] While these papers assume misallocation to be exogenous, the reduced form of my model is isomorphic to these papers and hence provides a microfoundation for the equilibrium distribution of firm-specific wedges, which Hsieh and Klenow (2009) and Restuccia and Rogerson (2008) employ as a modeling device. As far as theories of misallocation are concerned, the vast majority of contributions focuses on frictions in firm' input choices, studying models with imperfect capital markets (Buera et al., 2011; Moll, 2010; Banerjee and Moll, 2010; Midrigan and Xu, 2010), contractual imperfections (Acemoglu et al., 2007), frictional labor markets (Lagos, 2006) or capital adjustment costs (Collard-Wexler et al., 2011). This is

[1] While larger, this finding is qualitatively consistent Yang (2012), who also shows that the observed cross-sectional variation in marginal products is much larger than the one implied by heterogeneous mark-ups.

[2] Therefore it is of course also naturally related to the large literature to understand the distribution of productivity across the world. Banerjee and Duflo (2005), Caselli (2005) and Klenow and Rodriguez-Clare (1997) show empirically that aggregate productivity is positively correlated with income per capita. A large literature aimed to explain these differences in the context of an aggregate production function and stressed the importance of social capital or institutions (Hall and Jones, 1999; Acemoglu et al., 2001), human capital externalities (Lucas, 1988, 1990), barriers of technology adoption (Parente and Prescott, 1994) or directed technological change (Gancia and Zilibotti, 2009).

[3] This is in contrast to for example Foster et al. (2008) or Melitz and Ottaviano (2008) who generate heterogeneous mark-ups from a linear demand system with product differentiation.

[4] A very insightful discussion of these approaches is contained in Hopenhayn (2012).

in contrast to this paper, which argues that the equilibrium distribution of marginal products reflects monopolistic power and not binding input constraints. The interaction between static misallocation and firm' innovation and entry incentives has also been studied in Hsieh and Klenow (2011), Fattal Jaef (2011) and Yang (2012), albeit in a very different way. While they analyze entry incentives given a particular set of exogenous distortions firms will eventually face, I argue in this paper that the causality may go the other way: it is the horse-race between innovation and entry incentives that will shape the mark-up distribution in the economy as it determines the toughness of competition.

While this is, to the best of my knowledge, the only paper that focuses on imperfect output markets in relation to the literature on misallocation,[5] there is a related literature in the field of international trade stressing the importance of mark-ups (see e.g. Bernard et al. (2003) and Atkeson and Burstein (2008)). That mark-up heterogeneity is important for the welfare gains of trade liberalization is explicitly stressed in Epifani and Gancia (2011). Edmond et al. (2011) provide evidence that the pro-competitive effects of trade-liberalization are quantitatively important. Arkolakis et al. (2012) and Holmes et al. (2013), however, find smaller effects. From an empirical point of view, there is also substantial evidence that mark-ups vary systematically in the cross-section of firms (De Loecker and Warzynski, 2012; Garcia and Voigtländer, 2013), that they respond to policy changes (De Loecker et al., 2012) and that firm-specific prices are an important source of variation in revenue-based productivity measures (De Loecker, 2011a). The dynamic model in this paper is broadly consistent with the findings of these papers.

The rest of the paper proceeds as follows. In the next section I present the model. First, I characterize the static allocations and show the relationship between mark-ups and static misallocation. Then I derive the equilibrium distribution of mark-ups from an endogenous process of firm dynamics and derive the link between misallocation and growth. Section 3 contains the empirical analysis. First I present reduced form evidence for the importance of imperfect output markets and the relationship between mark-ups and market entry. Then I calibrate the model, show that mark-ups can account for roughly one fifth of the observed variation in revenue productivity in the data and calculate the dynamic welfare implications of misallocation. Section 4 concludes.

2 The Model

Consider the following continuous-time economy. There is a measure one of infinitely lived households, supplying their unit time endowment inelastically. Individuals have preferences over the unique consumption good, which are given by

$$U = \int_{t=0}^{\infty} e^{-\rho t} ln\left(c\left(t\right)\right) dt, \tag{1}$$

with ρ being the discount rate. The final good, which I take to be the numeraire, is a Cobb-Douglas composite of a continuum of intermediate products. In particular, there is a measure one of differentiated varieties, each of which can be produced by multiple firms. Formally,

[5] The importance of the life-cycle properties of mark-ups is however also stressed in Foster et al. (2011, 2008).

$$Y(t) = exp\left(\int_0^1 ln\left(\sum_{j\varepsilon S(i,t)} y_j(i,t)\right) di\right), \tag{2}$$

where $y_j(i,t)$ is the quantity of variety i bought from producer j and $S(i,t)$ denotes the number of firms active in sector i at time t. Hence, different varieties i and i' are imperfect substitutes, whereas there is perfect substitutability between different brands within a variety. The market for intermediate goods is monopolistically competitive, so that firms take aggregate prices as given but compete a la Bertrand with producers offering the same variety. The production of intermediates is conducted by heterogeneous firms and requires capital and labor, both of which are hired on frictionless spot-markets. The only source of heterogeneity across firms is their factor-neutral productivity. In particular, a firm producing variety i with current productivity q produces output according to

$$f(k,l;q) = qk^{\alpha}l^{1-\alpha}. \tag{3}$$

Existing firms' technology is variety-specific, i.e. once a firm entered in sector i, it can only produce i-output.

Both the set of competing firms $[S(i,t)]_i$ and firms' productivities $[q_j(i,t)]_{j,i}$ evolve endogenously through entry of new firms and process innovations by incumbent firms. Along their life-cycle, firms can hire workers to increase their productivity $q_j(i,t)$. Alternatively, workers can be employed in an entry sector, where they can try to generate new firms ("blueprints"). Entry efforts are directed, i.e. targeted towards particular sectors. I will specify the details of both the innovation and the entry technology below.

2.1 Heterogeneous Mark-ups and Misallocation

To characterize the equilibrium consider variety i. Given that production takes place with a constant returns to scale technology, firms compete in prices and different brands of variety i are perceived as perfect substitutes, in equilibrium only the most productive 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. The demand function for intermediaries is given by

$$y(i,t) = \frac{Y(t)}{p(i,t)}, \tag{4}$$

where $p(i,t) \equiv \min_{j\in S(i,t)}\{p_j(i,t)\}$. As this demand function has unitary elasticity, the most efficient firm has to resort to limit pricing, so that the equilibrium price will be equal to the marginal costs of the second most productive firm, which I will refer to as the follower. Hence,

$$p(i,t) = MC(q_F,t) = \frac{\left(\frac{w}{1-\alpha}\right)^{1-\alpha}\left(\frac{R}{\alpha}\right)^{\alpha}}{q_F(i,t)} \equiv \frac{\psi(R(t),w(t))}{q_F(i,t)}, \tag{5}$$

where $q_F(i,t)$ is the follower's productivity.[6] The equilibrium mark-up is therefore given by

$$\mu(i,t) \equiv \frac{p(i,t)}{MC(q,t)} = \frac{MC(q_F,t)}{MC(q,t)} = \frac{q(i,t)}{q_F(i,t)}, \tag{6}$$

i.e. a higher quality advantage shields the current producer from competition and allows him to post a higher mark-up. Using (4) and (5) it is easy to derive the allocation of production factors as

$$k(i,t) = \frac{1}{\mu(i,t)}\frac{\alpha Y(t)}{R(t)} \text{ and } l(i,t) = \frac{1}{\mu(i,t)}\frac{(1-\alpha)Y(t)}{w(t)}, \tag{7}$$

and firms' profits as

$$\pi(i,t) \quad = \quad \left(1 - \frac{1}{\mu(i,t)}\right)Y(t). \tag{8}$$

The expressions above show that the cross-sectional variation in profits and production factors 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 aggregate misallocation. To see this, note that the equilibrium marginal revenue products are given by

$$MPK(i,t) = \alpha\frac{p(i,t)\,y(i,t)}{k(i,t)} = R(t)\,\mu(i,t) \text{ and } MPL(i,t) = \alpha\frac{p(i,t)\,y(i,t)}{l(i,t)} = w(t)\,\mu(i,t).$$

Hence, firms' marginal products are not equalized but reflect the variation in equilibrium mark-ups. Note also, that it is precisely the mark-up, which determines measured productivity as $TFPR(i,t) \equiv \frac{p(i,t)y(i,t)}{k(i,t)^{\alpha}l(i,t)^{1-\alpha}} \propto \mu(i,t)$. In the framework of Hsieh and Klenow (2009) and Restuccia and Rogerson (2008), TFPR is proportional to $\frac{(1+\tau_K(i,t))^{\alpha}}{1-\tau_Y(i,t)}$, where $\tau_K(i,t)$ and $\tau_Y(i,t)$ are exogenous firm-specific taxes on capital and output. Hence, firms charging a high mark-up have high productivity and would hence be identified as facing high distortionary taxes.

The aggregate economy has a very transparent representation. In particular, there are two intuitive sufficient statistics, which summarize the degree of misallocation. These sufficient statistics are the *efficiency*

[6] It is at this point where the assumption of the aggregate production function being Cobb-Douglas simplifies the exposition. 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. In particular, if $\sigma > 1$ was the demand elasticity, the optimal price was $p = \frac{\psi}{q_F} min\left(\frac{\sigma}{\sigma-1}\frac{q_F}{q}, 1\right)$. In the limit where $\sigma \to 1$, we get $min\left(\frac{\sigma}{\sigma-1}\frac{q_F}{q}, 1\right) = 1$, which yields (5). We will see below, that the assumption that leading firms will always set the limit price will make the dynamic decision problem of firms very tractable.

wedge $M(t)$ and the *labor wedge* $\Lambda(t)$, which are given by

$$M(t) = \frac{exp\left(\int_0^1 ln\left(\mu(i,t)^{-1}\right)di\right)}{\int_0^1 \mu(i,t)^{-1}di} = \frac{exp\left(E\left[ln\left(\mu(i,t)^{-1}\right)\right]\right)}{E\left[\mu(i,t)^{-1}\right]} \tag{9}$$

$$\Lambda(t) = \left(\int_0^1 \mu(i,t)^{-1}di\right) = E\left[\mu(i,t)^{-1}\right]. \tag{10}$$

To see why (9) and (10) are indeed the appropriate aggregate wedges, note first that aggregate output is given by

$$Y(t) = M(t)\,Q(t)\,K(t)^{\alpha}\,L_P(t)^{1-\alpha}, \tag{11}$$

where $Q(t) = exp\left(\int_0^1 ln(q(i,t))\,di\right)$ is the usual CES efficiency index, $L_P(t)$ is the mass of production workers and $K(t)$ is the total supply of capital.[7] It is also easy to verify that the statically first-best aggregate output is given by $Y^{FB}(t) = Q(t)\,K(t)^{\alpha}\,L_P(t)^{1-\alpha}$, so that $M(t)$ is indeed exactly the reduction in aggregate TFP due to monopolistic pricing. Furthermore, equilibrium factor prices satisfy

$$R(t) = \Lambda(t)\frac{\alpha Y(t)}{K(t)} \quad \text{and} \quad w(t) = \Lambda(t)\frac{(1-\alpha)\,Y(t)}{L_P(t)}, \tag{12}$$

so that $\Lambda(t)$ measures precisely the gap between equilibrium factor prices and their respective social marginal products.[8] Hence, from a reduced form perspective, $M(t)$ and $\Lambda(t)$ are exactly equal to the efficiency and the labor wedge of Chari et al. (2007).[9] This representation is not only useful to understand the aggregate implications of monopoly power, but it also stresses that different moments of the underlying distribution of mark-ups affect the two aggregate wedges differentially.

Proposition 1. *Consider the static allocations characterized above. $M(t)$ and $\Lambda(t)$ are sufficient statistics for the static losses due to misallocation and depend only on the marginal distribution of mark-ups. Furthermore, aggregate TFP is (factor prices are) homogeneous of degree zero (minus one) in μ and hence only depend on the dispersion (level) of log markups.*

Proof. Follows directly from the definition of M and Λ in (9) and (10) and (11) and (12).

□

Proposition 1 shows that the different moments of the mark-up distribution have very different impacts

[7]As workers are used in both the production and the innovation sector, $L_P(t)$ will be determined endogenously. Along the balanced growth path, $L_P(t)$ will be constant. See below.

[8]In particular, equilibrium interest rates will be depressed relative to the interest rate as imputed from the usual growth accounting calibration. From an accounting point of view this is a desirable feature, because in most calibration exercises, the implied interest rates turn out to be counterfactually high (Banerjee and Duflo (2005); Lucas (1990)).

[9]In contrast to the case of Chari et al. (2007), in my model the labor and investment wedge are identical. For simplicity I refer to it as the labor wedge but (12) shows that it applies also to the case of capital.

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. Note that the canonical case of constant mark-ups as generated by a CES demand system with differentiated products (e.g. Melitz (2003)) is a special case of Proposition 1: TFP will be identical to its efficient competitive counterpart but monopolistic power reduces factor prices. As long as factors are in fixed supply, monopolistic power will not have any effects on efficiency (Epifani and Gancia, 2011).

2.2 Growth and Endogenous Misallocation

This paper argues that misallocation is not merely a static phenomenon, but rather a corollary of the underlying process of firm dynamics with direct growth consequences. Proposition 1 and the equilibrium pricing rule show concisely why this is the case. According to Proposition 1, static misallocation is fully determined by the cross-sectional distribution of mark-ups. Mark-ups in turn are simply given by $\mu = \frac{q}{q_F}$ (see (6)), i.e. are precisely equal to the productivity advantage of leading producers. Hence, static misallocation is a by-product of the dynamic process of firm growth: if leading firms innovate faster than their competitors, they will increase their 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 static misallocation will be reduced.

To see this more formally, it is necessary to put more structure on the evolution of relative productivities. It is particularly convenient to put firms' efficiencies on a quality-ladder (Aghion and Howitt, 1992; Grossman and Helpman, 1991), so that innovations lead to proportional productivity improvements. Formally, if a firm in sector i has had $n(i,t)$ innovations in the interval $[0,t]$, its productivity is given by $q(i,t) = \lambda^{n(i,t)}$, where $\lambda > 1$ denotes the increase in productivity brought about by an innovation. Equilibrium mark-ups can then be expressed as

$$\mu(i,t) = \frac{q(i,t)}{q_F(i,t)} = \frac{\lambda^{n_L(i,t)}}{\lambda^{n_F(i,t)}} = \lambda^{\Delta(i,t)}, \tag{13}$$

where $\Delta(i,t) \equiv n_L(i,t) - n_F(i,t) \geq 0$ is the quality gap between competing firms. Hence, equilibrium mark-ups depend only on the single-dimensional state variable Δ. Intuitively: firms are able to set high prices if they managed to climb the quality ladder faster than their potential competitors.

Innovations can stem from two sources: either current producers can experience technological improvements, or new firms can enter the market and thereby replace the current quality leader.[10] If a producer with current productivity q experiences an innovation, it reaches the next step of the quality ladder so its

[10] I refer to entrants as everyone but the current producer within a particular variety. Hence, it could be either an entirely new firm, a firm that used to produce this particular product but was replaced by the current producer in the past or a firm that is currently producing another variety and adds a new variety to its product portfolio.

new productivity is given by λq. For the case of entry, I assume that leading technologies are common knowledge for the process of innovation so that the entrant in sector i enters the market with productivity $\lambda q(i,t)$, where $q(i,t)$ is the leading quality in sector i. This formulation is not only standard in most Schumpeterian models of growth, but is particularly appealing in the current context, in that it draws attention to the different allocational consequences of the two sources of productivity improvements. While both entrants and incumbents increase the frontier technology by the same amount, the implications for equilibrium mark-ups and allocational efficiency are very different. In case the innovation stems from the current producer of variety i, the equilibrium mark-up for that variety *increases* by a factor λ. However, when productivity growth is induced by entry, the equilibrium mark-up for variety i *decreases* by a factor $\lambda^{\Delta(i,t)-1}$, as the new entrant is only a single step ahead of the quality ladder.[11]

Now suppose that incumbent firms innovate at a constant rate I and each variety i experiences entry at the constant rate z. I will show below that this is the case along the unique balanced growth path equilibrium in this economy. However, it is useful to directly characterize the equilibrium allocations as a function of the two endogenous variables (I,z) to stress that whatever the particular microfoundation of innovation incentives, it is these two equilibrium outcomes which determine both the equilibrium growth rate and static efficiency by shaping the distribution of mark-ups.

To characterize the equilibrium distribution of mark-ups, it is convenient to focus directly on the distribution of quality gaps Δ (see (13)). The distribution of productivity gaps is fully characterized by the collection $\{\nu(\Delta,t)\}_{\Delta=1}^{\infty}$, where $\nu(\Delta,t)$ denotes the measure of products with quality gap Δ at time t. These measures solve the flow equations

$$\dot{\nu}(\Delta,t) = \begin{cases} -(z+I)\,\nu(\Delta,t) + I\nu(\Delta-1,t) & \text{if } \Delta \geq 2 \\ -I\nu(1,t) + z(1-\nu(1,t)) & \text{if } \Delta = 1 \end{cases}. \tag{14}$$

Intuitively, 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 enter this state. (14) is the key equation to characterize the unique stationary distribution of mark-ups, which has a very convenient closed-form representation. In particular, both the growth rate of the economy and the degree of static misallocation are fully determined from the endogenous innovation and entry rates I and z.

Proposition 2. *Consider the economy above and let I and z be the (endogenous) rates of innovation and entry. Let x be the entry intensity $x \equiv \frac{z}{I}$ and define the endogenous statistic*

$$\vartheta(x) = \frac{ln(1+x)}{ln(\lambda)}, \tag{15}$$

[11] Note 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 monotone in x. Then, the following is true:

1. *The unique stationary distribution of mark-ups is Pareto and given by*

$$F(\mu; x) \quad = \quad 1 - \mu^{-\vartheta(x)}. \tag{16}$$

2. *The efficiency and the labor wedge are given by*

$$\Lambda(x) = \frac{\vartheta(x)}{1 + \vartheta(x)} \qquad \textit{and} \qquad M(x) = e^{-1/\vartheta(x)} \frac{1 + \vartheta(x)}{\vartheta(x)}. \tag{17}$$

3. *The dispersion of log revenue labor productivity $\left(\frac{py}{wl}\right)$ is given by*

$$sd\left[ln\left(\frac{py}{wl}\right)\right] \equiv \sigma_{LP} = \vartheta(x)^{-1}. \tag{18}$$

Hence, a higher entry intensity lowers mark-ups (in a first order stochastic-dominance sense), increases allocative efficiency and reduces observable productivity differences (18)*. Furthermore:*

4. *The equilibrium growth of technology is given by*

$$\frac{\dot{Q}(t)}{Q(t)} = g = ln(\lambda)(I + z) = ln(\lambda) I e^{ln(\lambda)\vartheta(x)}. \tag{19}$$

Hence, a higher entry intensity increases the growth rate (holding the innovation rate fixed).

Proof. See Appendix. □

Proposition 2 contains the main result of this paper: both the cross-sectional distribution of mark-ups $F(\mu; x)$ and the growth rate of technology g are jointly determined from two endogenous variables, the rate of innovation and the rate of entry. In fact, the endogenous distribution of mark-ups takes a pareto form, whose shape parameter $\vartheta(x)$ is endogenous and fully determined by a *unique* equilibrium variable - the entry intensity.[12] If entry is intense, the shape parameter is large so that both mark-up heterogeneity and the average mark-up declines and allocative efficiency improves. If on the other hand entry is of little importance, the resulting distribution of mark-ups has a fat tail, both the average mark-up and their dispersion is large and the economy is plagued by static misallocation.[13] The degree of static misallocation in turn is fully pinned down by the marginal distribution of mark-ups (see Proposition 1) and summarized in the two sufficient statistics $M(t)$ and $\Lambda(t)$, which have the closed form representation given in (17)

[12]Note that Δ is not a continuous variable but only takes integer values. Hence, the distribution function in (16) is not differentiable. In the following I will treat mark-ups as continuous.

[13]The interaction between firm-level distortions and entry incentives has also been discussed in Hsieh and Klenow (2011) and Fattal Jaef (2011) albeit in a very different way. There, a given set of distortions affects equilibrium entry. Proposition 2 shows that the causality might be the other way around: in this economy equilibrium entry shapes the dispersion of productivity differences through its pro-competitive effect.

and hence also only depend on $\vartheta(x)$. The same is true for the empirical measure of marginal products, revenue productivity. Note that it is precisely the standard deviation of log revenue productivity, which is routinely used as a key statistic to measure the degree of misallocation (see e.g. Hsieh and Klenow (2009) and Bartelsman et al. (2013)). (18) therefore can be interpreted as one potential microfoundation for such differences.

Note Proposition 2, and especially (16), is very different from Bernard et al. (2003), who also generate a (truncated) Pareto distribution of mark-ups in their model: in their model, the underlying distribution of efficiencies is *exogenous* (and drawn from a Fréchet distribution) and so is the distribution of mark-ups. In particular, and in contrast to (16), it is invariant to changes in the environment. The underlying mechanism which generates the endogenous pareto tail in my model is much more akin to the city-size dynamics of Gabaix (1999). To see this, note that the properties of mark-ups (within a product line) are inextricably linked to the life-cycle of the incumbent firm. As long the firm does not exit, mark-ups stochastically increase. Once exit takes place, mark-ups are "reset" to λ and the process begins afresh. In fact, as shown in the Appendix, the distribution of quality gaps $\Delta = \frac{ln(\mu)}{ln(\lambda)}$ as a function of age conditional on survival, $\zeta_{\Delta}(t)$, is given by

$$\zeta_{\Delta+1}(t) = \frac{1}{\Delta!}(It)^{\Delta} e^{-It},$$

which is a Poisson distribution with parameter It. Hence, conditional survival, the distribution of mark-ups continuously shifts outwards. In fact, the average log mark-up of a cohort of age t conditional on survival is given by

$$E[ln(\mu)|\text{age}=t] = ln(\lambda) I \times t \equiv g_I t, \tag{20}$$

i.e. is increasing linearly in age. In a similar vain, the probability of having exited by time t is given by $1 - e^{-zt}$ as firms exit with constant flow rate z. The long-run distribution of mark-ups is shaped by the interplay of these two processes. If there is little entry in the economy, it will be easy for incumbent firms to outgrow their competition as they survive for a long time. Hence, there will be a large tail of firms with ample monopoly power. To the contrary, if entry is important there is a lot of churning in the economy and most products are produced by firms who are forced to compete with firms who are technologically similar as they did not have time yet to accumulate a productivity advantage. This will cause the distribution of mark-ups to be concentrated around the perfect competition benchmark of marginal cost pricing.

Note that Proposition 2 did not use any structure of the innovation environment but took the endogenous entry and innovation rate as parametric. This is useful for two reasons. The first concerns the empirical analysis below. As I will be able to measure entry and innovation directly in the data, I can test Proposition 2 without having to take a stand on the details of the innovation environment. Secondly, it stresses that the particular link between mark-up based misallocation and economic growth is much more general than this particular model of entry and innovation. Suppose for example that growth was driven by technology diffusion. Assume 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. This model is identical to the one above with $I = z + \theta$, so that $x = \frac{z}{z+\theta}$. If production and the adoption of new techniques are complements, we would expect that $\theta > 0$. If, on the other hand, considerations of vintage capital are important, we would expect $\theta < 0$ as current producers have capital suited to the old technology already in place. Hence, an environment where new firms have a comparative advantage in adopting frontier technologies ($\theta < 0$)) will reduce and compress equilibrium mark-ups and thereby increase static allocative efficiency.[14]

2.3 Welfare: Static Misallocation and Growth

The fact that static misallocation and aggregate productivity growth are closely linked, implies that the aggregate implications of the observed differences in the cross-sectional productivity dispersion across countries could be much larger than previously appreciated. In particular, (19) can be written as

$$\vartheta(x) = \frac{1}{ln(\lambda)} \left(ln(g) - ln(g_I) \right), \tag{22}$$

i.e. the shape parameter of the endogenous mark-up distribution, which fully parametrizes the degree of misallocation in the economy, is increasing in the economy's growth rate, holding the growth rate of incumbents (see (20)) fixed. Hence, economies where the dispersion of revenue productivity across producers is low, do not only benefit from better static allocative efficiency but also tend to have higher growth rates. Concretely, consider the experiment of Hsieh and Klenow (2009). They report that the cross-sectional dispersion in marginal products across Indian and Chinese producers is larger than the one found in the US and that the implied efficiency losses are large - holding technologies fixed, TFP in India and China could be increased by 50% if allocative efficiency were to be brought to the US benchmark.[15] Through the lens of the model of this paper, we would conclude that the US economy is statically more efficient, precisely because it is characterized by an innovation environment, which allows new firms to enter seamlessly. This however will also affect welfare directly through its effects on the rate of technological progress. Hence, the variation in static losses through misallocation might to a large degree be a symptom of more fundamental differences across countries, which affect the entire pattern of firm dynamics.

To get a sense of the the relative of magnitude of these direct, static consequences of misallocation and

[14]Note also that the pareto result for $F(\mu; x)$ does not rely on new firms always entering with a productivity gap of one. Suppose for example that incumbent firms always climb one rung of the ladder but that entrants' innovations are larger in that they enter with a blueprint of quality $q(i,t)\lambda^b$, where $b > 1$. In that environment, the stationary distribution of mark-ups will be exactly the same as before, but just with the higher minimum λ^b. Alternatively, suppose that new blueprints come in heterogeneous qualities, i.e. with probability $p(j)$ entering firms generate a blueprint of quality $q(i,t)\lambda^j$. I show in the Appendix that the stationary distribution of quality gaps Δ is given by

$$\nu(\Delta) = \left(\frac{1}{x+1}\right)^{\Delta} \left\{ \sum_{i=1}^{\Delta} p(i) x^i \left(\sum_{k=0}^{i-1} \begin{pmatrix} i-1 \\ k \end{pmatrix} x^{-k} \right) \right\}, \tag{21}$$

where x is again the entry intensity. The analysis above is a special case of (21), where $p(1) = 1$ and $p(j) = 0$ for $j > 1$. In the Appendix, I show that entry still has a pro-competitive effect if $[p_i]_{i=1}^{\infty}$ decays exponentially, i.e. $p(i) = T\kappa^i$, where T ensures that $\sum_i p(i) = 1$ and $1 > \kappa > 0$.

[15]In the notation of this paper, $\frac{M^{India}}{M^{US}} = 2/3$

the dynamic growth effects, note that, as shown in the Appendix, equilibrium welfare is given by

$$
\begin{aligned}
W &= \text{const } + \frac{1}{\rho}\left(ln\left(c\right)+\frac{1}{\rho}\frac{g}{1-\alpha}\right) \qquad (23)\\
c &= \left(\frac{\rho+\left(\delta+\frac{g}{1-\alpha}\right)\left(1-\alpha\Lambda\right)}{\alpha\Lambda}\right)\left(\frac{\alpha\Lambda M}{\rho+\frac{1}{1-\alpha}g+\delta}\right)^{\frac{1}{1-\alpha}}L_P.
\end{aligned}
$$

Here c is the level of (normalized) consumption along the BGP and L_P is the endogenous share of workers active in the production sector. Hence, welfare depends on four endogenous variables: the two aggregate wedges (M, Λ), the equilibrium growth rate g and the allocation of labor across sectors L_P. Now consider the simplified case of this model without capital $(\alpha = 0)$ and no labor reallocation between the production and the innovation sector (i.e. $L_P\left(x, I\right) = L_P$).[16] In that case, (23) simplifies to

$$W = \text{const } + \frac{1}{\rho}\left(ln\left(M\right)+\frac{g}{\rho}\right), \qquad (24)$$

i.e. welfare is additively separable in TFP, i.e. static misallocation M and the equilibrium growth rate g. Given (24) and the expressions for M and g derived in Proposition 2, we can easily compare the dynamic and static welfare effects of a given change in the productivity dispersion across firms, σ_{LP}. In particular, let $\frac{dW}{dln(\sigma_{LP})}$ be the total change in welfare due to a percentage increase in productivity dispersion and similarly $\left.\frac{dW}{dln(\sigma_{LP})}\right|_g$ the static component holding the growth rate fixed. Note that in the model without capital, the static component is simply proportional to $\frac{dln(M)}{dln(\sigma_{LP})}$, which is simply the elasticity of aggregate TFP with respect to the underlying productivity dispersion and hence the statistic calculated in Hsieh and Klenow (2009). The following proposition shows that the dynamic gains could be much larger than these static effects.

Proposition 3. *Consider the case without capital* $(\alpha = 0)$ *and without labor reallocation* (L_P *constant). Then,*

$$\frac{dW/dln\left(\sigma_{LP}\right)}{dW/dln\left(\sigma_{LP}\right)|_g} = \frac{\text{Total change in welfare}}{\text{Static effects of misallocation}} = 1+\frac{g}{\rho}ln\left(1+x\right)\vartheta\left(x\right)\left(1+\vartheta\left(x\right)\right). \qquad (25)$$

Proof. See Appendix. □

(25) is a very useful expression, because it only depends on readily observable endogenous outcomes. Note first that the average mark-up is given by $E\left[\mu\right] = \frac{\vartheta}{\vartheta-1}$. Hence, an average mark-up of $1/3$ implies $\vartheta = 4$.[17]

[16] I will come back to the full model without these simplifying assumptions in Section 3.4, when I calibrate the model.

[17] Note that $\vartheta = 4$ also implies a median mark-up of $exp\left(ln\left(2\right)/\vartheta\right) \approx 1.2$, which is exactly the number estimated in De Loecker and Warzynski (2012).

If we think of a rate of productivity growth of 2%, a discount rate of 5% and an entry intensity of $1/3$[18], (25) implies that $\frac{dW}{dx}/\frac{dW}{dx}|_g = 3.3$, i.e. the full gains are more than three times as big as their static counterpart. Hence, by only taking the static efficiency losses from misallocation into account, one could seriously underestimate the total differences in welfare.

This discussion and (25) takes incumbents' innovation effort I as parametric. I and x (and hence σ_{LP}) however are in general jointly determined in equilibrium, so that I might endogenously respond to a change in the entry rate. Letting $\omega \equiv \frac{\partial I}{\partial z}$ be the equilibrium response of incumbents' innovation effort due to a change in the entry rate, the expression akin to (25) is given by

$$\frac{dW/dln\left(\sigma_{LP}\right)}{dW/dln\left(\sigma_{LP}\right)|_g} = 1 + \frac{g}{\rho} ln\left(1+x\right)\vartheta\left(x\right)\left(1+\vartheta\left(x\right)\right)\left(\frac{1+\omega}{1-x\omega}\right). \tag{26}$$

For the case of $\omega = 0$, this expression reduces to (25). The exact value of ω depends on the details of the innovation environment and - more importantly - on the underlying source of variation. If incumbent firms and new entrants use the same scarce factor and the variation in x is due to differences in the value of entry, we would for example expect that $\omega < 0$. If on the other hand the variation in x was due to changes in supply of that factor, it is also entirely possible that $\omega > 0$. Hence, to evaluate (26) a specific model of firms' innovation incentives is required and one has to take a stand, which underlying structural parameter causes the differences in entry across countries or sectors.

2.4 Equilibrium Innovation, Entry and Balanced Growth

To analyze the joint behavior of entry and innovation as an equilibrium outcome, we have to put more structure on the entry and innovation technologies. As far as incumbents' innovation technology is concerned, I assume that if a currently producing firm with quality advantage Δ wants to achieve an innovation flow rate of I, the firm has to hire $\Gamma(I,\Delta)$ units of labor, where

$$\Gamma\left(I,\Delta\right) = \lambda^{-\Delta}\frac{1}{\varphi}I^{\gamma}. \tag{27}$$

Here φ parametrizes the productivity of the innovation technology and $\gamma > 1$ ensures that the cost function for innovation is convex so that there is a unique solution. The term $\lambda^{-\Delta}$ implies that innovations are easier the bigger the productivity advantage Δ and is similar in spirit to the assumption of knowledge capital made in Klette and Kortum (2004) or the setup in Atkeson and Burstein (2010). In particular, this assumption is necessary to make the model consistent with a balanced growth path and with Gibrat's Law, i.e. with the fact the firms' growth rates are independent of size (Sutton, 1997; Luttmer, 2010).[19]

[18] This value is for example implied by the results of Bartelsman and Doms (2000), who report that incumbents account for 75% of aggregate productivity growth, i.e. $\frac{I}{I+z} = \frac{1}{1+x} = 0.75$.

[19] Intuitively: per-period profits are given by $\left(1-\lambda^{-\Delta}\right)Y$ (see (8)) and hence concave in Δ. For innovation incentives to be constant, the marginal costs of innovation have to be lower for more advanced firms. The leading term in (27) $\left(\lambda^{-\Delta}\right)$ is exactly the right normalization to balance those effects. Note that firms only generate a high productivity gap when they have multiple innovation *in a row*. Hence, (27) effectively posits that firms can build on their own innovations of the past.

The innovation technology of entrants is assumed to be linear, i.e. to generate a unit flow rate of a new blueprints and bring this idea to the market, χ workers are required. Hence, χ is an index of entry costs, which depends both on technological aspects (e.g. the human capital of entry workers) and institutional characteristics like license requirements and entry costs, which have to be paid, before a new firm can be active on the product market (Djankov et al., 2002).

Given this innovation environment, the balanced growth path equilibrium of this economy can be characterized in closed form. A crucial object is of course the equilibrium value of a firm, as it is this value function that determines innovation and entry incentives. As neither per period profits $\pi(i,t)$, nor the innovation/entry technologies depend on productivity q (conditional on Δ), it is not a state variable for the firm's problem. Hence, let $V(\Delta,t)$ denote the value of a firm with a quality gap Δ at time t. This value function solves the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned} r(t)V(\Delta,t)-\dot{V}(\Delta,t) &= \pi(\Delta,t)-z(\Delta,t)(V(\Delta,t)-V(-1,t))+ \\ &\quad \max_I\left\{I(V(\Delta+1,t)-V(\Delta,t))-w(t)\lambda^{-\Delta}\frac{1}{\varphi}I^{\gamma}\right\}, \end{aligned} \tag{28}$$

where $z(\Delta,t)$ is the (endogenous) entry rate for a sector with quality gap Δ, which current incumbents take as given. The intuition for (28) is as follows: Given the interest rate $r(t)$, 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 $w(t)\lambda^{-\Delta}\frac{1}{\varphi}I^{\gamma}$, the firm generates a surplus of $V(\Delta+1,t)-V(\Delta,t)$ with flow rate I.[20] In the Appendix I show that the unique solution to (28) is given by

$$V(\Delta,t) = \frac{\pi(\Delta,t)+(\gamma-1)w(t)\Gamma(I,\Delta)}{\rho+z}, \tag{29}$$

where I is the optimal innovation rate, which is strictly positive and unique, z is the entry rate, which is constant along the balanced growth path and ρ is the consumers' discount rate. The solution for the value function contained in (29) is intuitive: the value of the firm is simply the risk-adjusted net present value of current profits plus the inframarginal rents of the concave innovation technology.

Given the solution for $V(\Delta,t)$, the equilibrium innovation effort I and the equilibrium entry rate z can easily be determined. The optimal innovation rate is characterized by the first order condition

$$w(t)\lambda^{-\Delta}\frac{\gamma}{\varphi}I^{\gamma-1} = V(\Delta+1,t)-V(\Delta,t), \tag{30}$$

[20] As usual, the occurrence of both a successful incumbent innovation and entry is of second order.

as $V(\Delta+1,t) - V(\Delta,t)$ is the marginal return of an innovation. The free entry condition is given by

$$V(1,t) = \chi w(t), \tag{31}$$

as new firms always enter with productivity advantage of $\Delta = 1$ conditional on entry. Finally, the labor market clearing requires that

$$1 \;=\; L_P + L_I + L_E = L_P + \Lambda(x)\frac{1}{\varphi}I^{\gamma} + \chi z, \tag{32}$$

where L_P is the number of production workers, $L_I = \int \lambda^{-\Delta}\frac{1}{\varphi}I^{\gamma}dF_{\Delta} = \Lambda(x)\frac{1}{\varphi}I^{\gamma}$ is the aggregate amount of innovators hired by incumbents and $L_E = \chi z$ is the aggregate labor requirement to generate an entry rate of z. To determine the equilibrium, note that (29), (30) and (31) determine the relative wage $\frac{w(t)}{Y(t)}$ as function of the innovation and entry rates (z, I) or equivalently the entry intensity and innovation rate (x, I). Together with (12), which showed that the marginal product of labor in the production sector was determined by $\frac{Y(t)}{w(t)} = \frac{1}{1-\alpha}\frac{1}{\Lambda(x)}L_P$, and the labor market clearing condition (32), they fully characterize the equilibrium.

Proposition 4. *Consider the economy described above. There exists a unique balanced growth path equilibrium, where the innovation rate I and the entry intensity x are constant and determined by the entry costs χ, the efficiency of the innovation technology φ, the step size λ and the discount rate ρ. In particular*

$$\begin{array}{llll} \frac{\partial I}{\partial \chi} \lesseqgtr 0, & \frac{\partial I}{\partial \varphi} > 0, & \frac{\partial I}{\partial \lambda} > 0, & \frac{\partial I}{\partial \rho} \lesseqgtr 0 \\ \frac{\partial x}{\partial \chi} < 0, & \frac{\partial x}{\partial \varphi} \lesseqgtr 0, & \frac{\partial x}{\partial \lambda} > 0, & \frac{\partial x}{\partial \rho} < 0. \end{array}$$

Proof. See Appendix. □

Proposition 4 not only proves the existence and uniqueness of an equilibrium with constant innovation and entry rates, but it also provides the link between the underlying structural parameters and the main endogenous outcomes, which determine the equilibrium growth rate and the degree of misallocation. According to Proposition 2, both measured misallocation (i.e. the dispersion of marginal products across firms) and the consequences of misallocation (i.e. the two wedges Λ and M) are monotone in the entry intensity. Hence, lower entry costs, a more radical menu of quality improvements and a lower discount rate lower misallocation through its effect on the entry intensity. The effects on the equilibrium growth rate are more involved, because entrants compete with incumbent firms for a scarce factor. While an increase in the quality step size λ e.g. increases entry and innovation efforts and hence productivity growth g, this is not necessarily the case for the other structural parameters. The joint behavior of static misallocation and equilibrium growth therefore crucially depends on the underlying source of variation.

3 Empirical Analysis

In this section I will take the theory to Indonesian plant-level data, both to test its main implications and to quantitatively assess the likely magnitude of the dynamic multiplier derived in Proposition 3. The empirical analysis has three parts, which closely follow their counterparts in the theoretical analysis. In section 3.2, I first focus on the firm-level data and provide both cross-sectional and time-series evidence that the empirical patterns of mark-ups are consistent with the theory. In particular, I try to empirically distinguish between imperfect output markets (the mechanism of this paper) and theories, where firms are constrained in their input choices. As in De Loecker and Warzynski (2012) and De Loecker et al. (2012), this analysis only requires the static first-order conditions of firms' cost minimization problem and does not have to take a stand on either the demand structure or the underlying process of firm dynamics. This analysis therefore does not require the strong theoretical assumptions of the theory. In section 3.3 I add more theoretical structure and focus on the main prediction of Proposition 2, namely the relationship between entry and the distribution of mark-ups. While Proposition 2 is dependent on the particularities of the process of firm dynamics, it does not require any assumptions on the innovation environment but solely tests the equilibrium relationship between the distribution of mark-ups and the entry intensity. Finally, in section 3.4, I exploit the entire general equilibrium structure of the theory and calibrate the underlying structural parameters of the theory. Given these parameters, I first perform an accounting exercise to ask how much of the cross-sectional dispersion in revenue productivity the theory can explain when calibrated to aggregate moments. Then I revisit the welfare calculation of Proposition 3. In particular, I study a reduction of the costs of entry and decompose the welfare consequences into their static and dynamic components to estimate the relative magnitude of the growth effects relative to the static efficiency losses of misallocation.

3.1 The Data

The empirical analysis is based on Indonesian plant-level data. The empirical analysis is 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), Blalock et al. (2008) and Yang (2012). The Statistik Industri is an annual census of all manufacturing firms in Indonesia with 20 or more employees and contains information on firms' revenue, employment, capital stock, intermediate inputs and other firm characteristics. The data has a panel dimension and I will use data from 1991 to 2000. The final sample has about 200.000 observations.[21] There are two features of the data, which makes it particularly useful for this project. First of all, in 1996, the annual census was augmented with a special supplement to learn about potential barriers to firm expansion. In particular, firms were asked about the constraints they face and wether

[21]To 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.

they willingly invest in process innovations. This information is very useful to distinguish theories based on imperfect output markets from theories stressing the importance of input constraints. Secondly, the data contains detailed information on the geographic region, the respective firm is located in. This allows me to define the appropriate product market both geographically and in the product-space.

To be able to control for regional characteristics in the regressions, I augment this data with information from the Village Potential Statistics (PODES) dataset in 1996. The PODES dataset contains detailed information on all of Indonesia's 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 and the sectoral composition. Controlling for these factors should at least alleviate the most pressing concerns about omitted variables from the equilibrium relationship between entry and the distribution of mark-ups.

3.2 Productivity and Mark-Ups

To test the theories' prediction for the cross-sectional and time-series behavior of mark-ups, it is obviously necessary to measure firm-specific mark-ups for a broad cross-section of firms. Recently, Jan De Loecker has shown in various contributions how this can be achieved using readily available data (De Loecker et al., 2012; De Loecker and Warzynski, 2012; De Loecker, 2011b). While I follow this approach closely (and hence relegate most of the details to the appendix), I am mostly interested in the cross-sectional variation of mark-ups and less so in the level of mark-ups. As will be clear below, this implies that I will not need to estimate firms' output elasticities.

To analyze firms' mark-ups, consider the firms' cost minimization problem. In particular, suppose that production is given by

$$y = qf\left(l, x\right), \tag{33}$$

where $f\left(.\right)$ is the firms' production function, l denotes the labor input and x is a vector of other inputs.[22] The firm takes the wage w as given. Furthermore, the firm is potentially subject to an input constraint, which constrains total spending as

$$wl \leq h\left(A\right), \tag{34}$$

where $h\left(A\right)$ denotes the maximum amount of spending on labor inputs and A is a vector of firm-characteristics like financial collateral or political connectedness, which induces firm-variation in the tightness of such input constraints. Letting κ be the multiplier on the constraint (34), the firm's optimality condition imply that the mark-up μ is given by

$$\mu = \theta_l \frac{1}{1+\kappa}\frac{1}{s_l}, \tag{35}$$

where $s_l = \frac{wl}{py}$ is the labor share and $\theta_l = \frac{\partial ln(f(l,x))}{\partial ln(l)}$ is the output elasticity. Rearranging terms yields the

[22] For simplicity I focus on labor as the static input. For the empirical analysis I will also report the results using materials as an input, which might be less susceptible to adjustment costs. However, see the discussion in footnote 23 below.

equation

$$ln\left(s_l^{-1}\right) = ln\left(\theta_l^{-1}\right) + ln\left(\mu\right) + ln\left(1+\kappa\right), \tag{36}$$

which simply states that - as long as firms are statically cost-minimizing - the variation in labor productivity can be due to three sources: differences in technology (or more generally non-constant output elasticities), differences in monopoly power (as measured by the mark-up) and differences in the tightness of input constraints.

While (36) does not require any assumptions on the demand and production function,[23] it is precisely this equation, which is at the heart of the identification strategy pioneered in Hsieh and Klenow (2009). For the case of the production function taking the Cobb-Douglas form and the output structure being the standard CES-monopolistic-competition framework, we get that

$$\theta_l s_l^{-1} = (1-\alpha)\frac{py}{wl} = (1-\alpha)\frac{ARPL}{w} \propto \frac{MRPL}{w},$$

where $ARPL$ and $MRPL$ denote the average and marginal revenue products respectively. Hence, $\theta_l s_l^{-1}$ is precisely a measure of firms' inframarginal rents, which can either be generated by imperfect output markets (μ) or by binding input constraints (κ). In the perfect-competition, frictionless benchmark $\mu = 1$ and $\kappa = 0$ so that firms equate their marginal product to the equilibrium wage. This is of course the basis of Hsieh and Klenow (2009): it is the dispersion of inframarginal rents which indicates dispersion in firms' marginal products.

(36) not only suggests an intuitive way to test the reduced form properties of the theory but also a way to distinguish this theory of imperfect markets from theories based in input constraints. *If* θ_l was known, we could simply run regressions of the form

$$ln\left(s_{l,it}^{-1}\right) = ln\left(\theta_{l,it}^{-1}\right) + x_{it}'\beta + u_{it}, \tag{37}$$

where x contains characteristics proxying for either market power (i.e. μ in (36)) or binding input constraints (i.e. κ in (36)) respectively. Implementing (37) however requires an appropriate control for the unknown output elasticity θ_l. Most existing contributions assume that f either takes the Cobb-Douglas or the Translog form and estimate θ_l using proxy-methods (see e.g. De Loecker and Warzynski (2012), De Loecker et al. (2012) or Garcia and Voigtländer (2013)). I take a different approach. As I am solely interested in estimating β in (37), I focus directly on (37) and treat θ_l as a nuisance parameter. In particular, I implement (37) by modeling firms' output elasticities as

$$ln\left(\theta_{l,it}^{-1}\right) = \delta_s + \delta_t + \zeta ln\left(\frac{k_{it}}{l_{it}}\right), \tag{38}$$

[23] Of course I have to assume that f is differentiable and that the respective input - here labor - is a "static' input, which is freely adjustable so that the first order condition holds. However, note that the input constraint (34) is general enough so that it covers the case of adjustment costs (see also Collard-Wexler et al. (2011)). I nevertheless also report the results when I take materials instead of labor as the input under consideration.

where δ_s is a set of 5-digit fixed effects and δ_t is a set of year fixed effect. (38) nests two leading cases. Under the assumption that f takes the Cobb-Douglas form, $\theta_{l,it}$ does not vary within industries (and years), so that $\zeta = 0$. In that case, the resulting estimates for β in (37) will be exactly the same if I had estimated θ_l in a first stage, and then included this estimate $\hat{\theta}_s$ and industry and time effects in (37).[24] The case with $\zeta \neq 0$ then corresponds to the case of a constant returns Cobb Douglas production function, where the output elasticities are firm specific.[25] It is this case, which is my preferred specification as it does not mechanically attribute the entire systematic variation in labor-shares (within 5 digit industries) to either monopoly power or binding constraints.[26]

Tables 1 and 2 contain the results. I first focus on the mechanism stressed in this paper that the variation in firms' inframarginal rents is driven by imperfect output markets and noncompetitive pricing. According to the model, mark-ups are the result of firms' past growth experiences. While entrants are predicted to have low mark-ups, old firms with successful product innovations tend to have large monopoly power and mark-ups. Table 1 therefore contains the results of estimating (37), where I independently include these different variables as proxies for firms' mark-ups. As predicted by the theory, while entering firms have 4.3% lower mark-ups, firms on average increase their mark-up by one percent a year over their life-cycle and mark-ups are particular high in growing firms.[27] Using the information contained in the aforementioned special supplement to the census in 1996, I also find that firms who report having higher profits compared to last year and firms consciously investing in product and process innovation have higher mark-ups. Finally, I also replicate two results from the literature as a consistency check. As in Hsieh and Klenow (2009) I find that exiting firms have relatively low mark-ups. And as in De Loecker and Warzynski (2012) and Garcia and Voigtländer (2013), exporters have higher mark-ups.

To explain these results in a framework with input constraints, we had to argue that entering firms face relatively *few* constraints, while old firms, exporters, and firms along a growth trajectory are characterized by binding input constraints. While not theoretically impossible, benchmark dynamic models of financing constraints imply exactly the opposite pattern, where borrowing constraints bind in the early stages of

[24]Furthermore, my data is standard in the sense that it does not contain information on firm-specific prices. Hence, without imposing much more structure on the demand system, the output elasticity θ (which corresponds to *physical* output) cannot be estimated consistently anyway.

[25]As the optimal capital labor-ratio is given by $\frac{k}{l} = \frac{\alpha}{1-\alpha}\frac{w}{R}$, the variation in $\frac{k}{l}$ will reflect variation in α.

[26]An even more flexible approach would be to model firms' output elasticities as

$$ln\left(\theta_{l,it}^{-1}\right) = \delta_s + \delta_t + \zeta_k ln\left(k_{it}\right) + \zeta_l ln\left(l_{it}\right). \tag{39}$$

The specification in (39) corresponds to the case of a translog production function, in that the resulting estimates for β are essentially identical to the ones I had obtained if I had estimated $\theta_{l,it}$ under the assumption of a translog production function but then included $ln\left(k_i\right)$ and $ln\left(l\right)$ in the regression equation (37) (as e.g. done in De Loecker and Warzynski (2012)). The reason why this added flexibility might be counterproductive, is the danger of over-controlling. In the model, (38) is of course the correct regression to run. Optimal input demands are given by (see (7)) $k = \frac{1}{\mu}\frac{\alpha Y}{R}$ and $l = \frac{1}{\mu}\frac{(1-\alpha)Y}{w}$. Hence, it is precisely the variation in mark-ups that drives the variation in factor inputs so that running (39) would give erroneous results as $ln\left(k\right)$ and $ln\left(l\right)$ would absorb the entire variation I am interested in. Note that running (38) without the restriction $\zeta = 0$ is not subject to that concern. I therefore base the main analysis in the text on (37) and (38) and include the results using (39) in the Appendix. There I also redo the analysis using materials as a dependent variable.

[27]Similar results are also reported in Foster et al. (2008) who observe plant-specific prices and show that young firms do charge lower prices.

the life-cycle and get relaxed as the firms ages (see e.g. Clementi and Hopenhayn (2006)). In fact, it is precisely the life-cycle pattern of inframarginal rents, which offer an interesting source of identification. According to the theory of this paper, a cohort of entering firms should increase their mark-up as they gradually increase their productivity relative to their competitors (see (20)). Theories of input constraints, and the above mentioned theories of financial frictions in particular, naturally predict the opposite pattern: precisely because firms increase their profits by alleviating their binding input constraint, older firms have on average had more time to do so and hence should have lower marginal products. Using the panel dimension of the data, I can look at these implications empirically. Figure 1 shows the average mark-up as a function of age for a representative cohort of firms (with the size of the dots reflecting the size of the surviving cohort). While mark-ups are relative low upon entry, mark-ups increase over the life-cycle as predicted by (20).[28]

As a final test for the importance of binding input constraints, I focus on the particular case of financing constraints as my data contains information on that margin. As for Table 2, I consider specification (37) but now include firm characteristics, which proxy for the multiplier on firms' input constraints (κ in (36)). In the first three columns, I show that firms that finance their investment at least partly through FDI, through foreign debt of through equity issuance on the Indonesian capital market have *higher* inframarginal rents, even though they have better access to the capital market and should be less capital constrained.[29] The same is true for firms, which are (at least partly) owned by the state and which are also arguably in a preferred position to finance their labor inputs. The last three columns then focus again on the special census supplement in 1996. Firms were asked wether they face a constraint regarding their expansion plans, whether this constrains was due to a shortage of capital and wether missing capital was the main reason why the firm was not expanding. All of these measures are robustly negatively correlated with firms' inframarginal rents. This is exactly the opposite of what we expect to find if the variation in labor productivity was generated from models with credit constraints. Hence, firms with high labor productivity might not be firms that are "unlucky" by being credit constrained or subject to stifling regulation, but that firms earn those rents on their production factors precisely because they were lucky enough to increase their physical productivity faster than their competitors and are rewarded for doing so with the opportunity to post a high mark-up.

While the results in Tables 1 and 2 and Figure 1 lend support to an environment with imperfect output markets,[30] I want to stress that they do not imply that financial imperfections (or other frictions to hire valuable inputs) are unimportant. To the contrary, it might be precisely such frictions, which

[28]While Figure 1 is exactly correct given the theory, it could of course reflect pure selection if intrinsically high mark-up firms are less likely to exit. In the Appendix I control for selection and show that mark-ups are also increasing for the cohort of firms that ends up surviving until the end of the sample.

[29]It might be argued that if a firm raises money on the market, the firm is likely to be constrained - if it had enough wealth to finance everything out of their own funds, it would definitely not be constrained. However, in the context of a developing economy like Indonesia, either of these measures is probably more informative about the availability of funds. See for example Blalock et al. (2008), who use these measures as proxies for the availability of funding.

[30]In the Appendix I provide various robustness checks for these results. In particular, I redo the analysis using the material share as the dependent variable and different specifications of the output elasticity (see (38)).

prevent new firms from entering. While imperfect competition on output markets might therefore be an important proximate source of static misallocation, the fundamental source of why incumbents firms are able to charge high mark-ups might be a frictional capital market at which new firms are unable to secure sufficient funds to enter the market. Hence, the variation in the financial environment across countries might be at the heart of both differences in aggregate technology (Q) and misallocation (M and Λ) even tough the static allocations of production factors are not directly influenced by binding input constraints.

3.3 Mark-Ups and Entry

While the results in Section 3.2 are purely cross-sectional, the main contribution of this paper is the endogeneity of mark-ups, especially their dependence on the underlying process of firm dynamics governing innovation and entry. As Proposition 2 showed, the distribution of mark-ups is fully parametrized by $\vartheta(x) = \frac{ln(1+x)}{ln(\lambda)}$. Hence, markets which differ in their entry intensity x are expected to have a different distribution of monopoly power. In this section I am going to test this prediction using the variation in entry across markets in Indonesia as the main source of variation. To empirically test this implication, I need a definition of a market. The logic of the model implies that a market is defined by the degree of competition across producers. Hence I define a market both geographically and in the product space. Formally, I take a market to be a province-industry cell, where I measure industries at the 3 or 4 digit level and a province, of which there are 27, by its administrative boundaries. While differentiating markets in the product space is intuitive, allowing for the geographical dimension requires some comment. There are two reasons, why the source of variation across regions is useful. The first reason is an empirical one. If goods were freely tradable across regions in Indonesia, there would be no need to treat two provinces within a product as separate markets. If trade however is not frictionless, we would want to distinguish markets geographically, as entering firms in one's home region are more of a competitive threat. Given the geography of Indonesia, where different provinces are often on different islands, I feel more comfortable to allow for this degree differentiation. The second reason is a conceptual one. While Proposition 2, and hence the analysis in this section, does *not* rely on a particular innovation environment, it of course begs the question where the variation in the endogenous entry intensity stems from. The particular microfoundation offered in Section 2.4 took a stand on this issue and stressed the importance of entry costs (χ) and the productivity of incumbents' innovation technology (φ). While these structural parameters could of course vary across products within regions[31], I think that the underlying regional variation in local policies, financial density and local labor market conditions is a valuable source of variation, which I do not want to rule out.[32]

Before presenting the formal regression analysis, consider Figure 2, which contains a simple illustration of the empirical implications of Proposition 2. Using (15) and (16), I can write the endogenous distribution

[31]One example is the well-known measure of financial dependence of Rajan and Zingales (1998), which can be interpreted as a shifter of entry costs (given some financial market imperfection) across industries.

[32]However, to show that the results are not driven by this assumption, I also report the results in a specification with a full set of region fixed effects.

of mark-ups as

$$ln\left(1-F\left(\mu;x\right)\right)=-\vartheta\left(x\right)ln\left(\mu\right). \tag{40}$$

Hence, the (logarithm of) the complementary cdf is linearly related to the (logarithm) of mark-ups and the slope is market specific and depends on the degree of entry. In Figure 2 I plot the relationship in (40) for different geographic regions in Indonesia that differ in their degree of entry.[33] In particular, I rank the different regions in Indonesia by their entry rate and call all regions with an entry rate below (above) the median the low (high) entry regions. Then I estimate the distribution of mark-ups within these two groups. Figure 2 clearly shows that these distributions are different. In particular, low entry regions display a significantly thicker tail of firms charging a large mark-up. This implication is very much in line with the theory and at the heart of the mechanism of this paper: The distribution of measured productivity is more compressed in high entry environments, as new entrants keep output markets competitive. Comparing Figure 2 with (40) also shows that while the linearity of the schedule is not borne out in the left tail of the distribution, the theoretical prediction of a linear relationship holds up quite well for a large range of the distribution.[34]

I will now test Proposition 2 more formally in a regression framework. As discussed above, from now on I will define a market at the industry-region level. According to Proposition 2, $\vartheta\left(x\right)$ is a sufficient statistic for *all* the moments of the mark-up distribution. In particular, it follows directly from (16) and (18) that $E\left[ln\left(\mu\right)\right]=sd[ln\left(\mu\right)]=\vartheta\left(x\right)^{-1}$ and $q^{\tau}\left[ln\left(\mu\right)\right]=-\vartheta\left(x\right)^{-1}ln\left(1-\tau\right),$ where $q^{\tau}\left[ln\left(\mu\right)\right]$ is the τ-th quantile of the log mark-up distribution. Hence, all these moments are decreasing in ϑ, and hence in the entry intensity. To test this prediction, I consider regressions of the form

$$\Psi_{s,r,t}=\delta_{t}+\delta_{s}+\phi Entry_{s,r,t}+T_{r}'\xi+u_{s,r,t}, \tag{41}$$

where $\Psi_{s,r,t}$ denotes these respective moments, δ_t and δ_s are a set of industry and year effects and $Entry_{s,r,t}$ is the entry rate in market (s,r) at time t. As before, mark-ups are estimated as the residual of (37). The theory predicts that $\phi<0$ and the source of identification for ϕ is the variation in entry rates across regions within industries, which is in line with the theoretical model. To control for some regional characteristics T_r is a set of (time-invariant) regional controls, which were drawn from the regional PODES files described above. Besides population size, which I include as a crude-proxy of economic development across regions, I also include the agricultural employment share to control for differences in the economic structure and measures of regional financial development and infrastructure. Finally, I run (41) also for the subset of industries, which are producing differentiated goods in the sense of Rauch (1999). For these goods, inter-

[33] While the theory has implications for the entry *intensity* $\frac{z}{I}$, I will use the rate of entry z as my measure of entry in the empirical exercise. It not only has the benefit of being conceptually easy, but it is also better measured. Some region-industry cells are quite sparsely populated so that I cannot estimate I, which is the growth rate of incumbent firms (conditional on survival), precisely. When I calibrate the model in Section 3.4, I use the theoretically required concept of the entry intensity.

[34] When I estimate the coefficient on $ln\left(m\right)$ in a simple bivariate regression, the coefficients (standard error) in the high entry region is -1.2575 (0.0007) and -1.1579 (0.001) in the low entry region.

regional trade should be less of a concern and regional entry have more predictive power for the degree of competition of local producers.

The results are contained in Table 3. In Panel A at the top I focus on the average mark-up. In column one I show that there is a strong negative relationship between the average mark-ups and the degree of entry into a market. In columns two to four I add the different regional controls. While these controls do increase the explanatory power of the regression, they leave the coefficient on the entry rate essentially unchanged. Neither the size of the respective region (when measured by its population), nor the degree of financial development or the agricultural employment share is significant. Interestingly, the one variable, which is consistently negatively related to the average mark-up is a measure of infrastructure and market access, which is the share of villages within the province, which is accessible by road the entire year. This is expected if we interpret this as a proxy for market contestability. Finally, in column five I focus only on industries, which are classified as differentiated. As expected, this strengthens the results slightly but leaves the qualitative features unchanged.

In the lower panel I report the analogous results for different quantiles and the dispersion of mark-ups.[35] Again there is a negative relationship between the rate of market entry and the different quantiles. This is expected given Figure 2, which showed that a higher rate of entry induces first-order stochastic dominance shifts in the distribution of mark-ups. Finally, in the last column, I report the results for the standard deviation of mark-ups. These show that the entry rate is essentially unrelated to the dispersion of mark-ups across regions. This is inconsistent with Proposition 2. However, from an empirical point of view, this might not be too surprising as any measurement error of firms' labor productivity will induce cross-sectional dispersion across regions, which is unrelated to the rate of entry. In the Appendix I present a simple simulation exercise, which shows that measurement error in the micro-data will bias that the results using the standard deviation as a dependent variable but neither the quantiles nor the mean. More generally, any other mechanism (of which there are plenty), which induces dispersion in firms' labor productivity within industries but is not related to the degree of entry will tend to bias the entry coefficient downward. This attenuation in the second moment is also seen in the correlation between the mean and the standard deviation across markets. According to the theory, these should be identically equal to $\vartheta(x)^{-1}$. When I regress the average (log) mark-up on the standard deviation of (log) mark-ups and a set of year and industry fixed effects, I estimate a coefficient of 0.247 with a standard error of of 0.03. Hence, there is a clear and strong positive relationship, but the coefficient is far from one.

As explained above, I consider the variation across regions as useful given the theory. Nevertheless, one might of course be concerned about persistent regional characteristics affecting both the rate of entry and the distribution of productivity without the mechanisms stresses in this paper being in place. Given the panel structure of the data, I can estimate (41) with a full set of industry, time and region fixed effects. The results are reported in Table 4. As expected from the results above, where the inclusion of different regional characteristics did not affect the coefficient of interest, the results are almost identical to the ones

[35]For brevity I only report only the specifications including the full controls. The full results are qualitatively similar and available upon request.

above: while the entry rate is strongly negatively correlated with the level of mark-ups in a market, the effect on the second moment is very weak.

Two moments of particular importance are of course the aggregate wedges Λ and M, which are directly measurable given the data on mark-ups (see (9) and (10)). While the theory implies that both these statistics are increasing in the rate of entry, the results and together with Proposition 1 already suggest where the theory will not be borne out: As M only depends on the dispersion of mark-ups and - empirically - the rate of entry did not have much explanatory power, I expect it to also be a weak predictor for M. The results of estimating (41) using these sufficient statistics as dependent variables are reported in Table 5. Consider first the case of Λ, which is contained in Panel A. According to columns one, an increase in the rate of entry of 10 percentage points (which is the interquartile range of entry rates in the sample) increases Λ by one percentage point. As Λ is exactly a labor wedge, this is akin to a reduction of labor taxes of one percentage point. In column four and five I show that this result is insensitive to either focusing solely on the subset of differentiated products or to the inclusion of regional fixed effects. The last two columns contain the results when I use the logarithm of Λ as a dependent variable. Again there is a strong positive relationship. The analogous results for M are contained in Panel B. As expected, all coefficients are essentially zero as the entry rate is not correlated with the dispersion of mark-ups across the product markets in my sample.

Finally, I want to briefly consider two alternative explanations, which could theoretically generate the variation across regions reported above. Consider first the case of input constraints, in particular capital market imperfections. As seen from (36), if the data was generated by such model, I would identify credit constrained firms as charging a high mark-up. Now suppose that the efficiency of the financial system varies across regions so that financially underdeveloped regions will see more financially constrained firms. If the regional entry rate is correlated with the regional financial development, I would conclude that entry reduces monopoly power albeit it is just the case that high entry regions are such that less firms are constrained (or firms are less constrained). The second alternative explanation I want to consider are preferential policies. In particular, the modeling device of "firm-specific taxes" used in Restuccia and Rogerson (2008) or Hsieh and Klenow (2009) is sometimes interpreted as a stand-in for actual policies like regulation, taxes or bureaucratic red tape affecting firms differentially. In particular, firms that I identify as having large monopoly power might in fact just face politically oriented barriers to expand.[36] If additionally regions with good policies are also characterized by more dynamic entry, I would observe exactly the correlations in the data, which I interpret as being informative about the pro-competitive effects of entry. To at least roughly address these concerns, I use the micro-data to calculate two product-market level measures to control for such effects. First of all, I again consider firms as being relatively *un*constrained whenever they are financed through FDI or raise funds on the national capital market of via foreign loans.[37] Similarly, I assume that firms with a government stake are relatively less constrained

[36] Note that such constraints are nested in the general formulation of the spending constraint (34).

[37] Recall however the results in Table 2, where I show that these unconstrained firms have relatively high inframarginal rents in the cross-section.

by political fiddling. In Panel A of Table 6 I first show the entry rate is indeed correlated with these measures, albeit in a somewhat unsystematic way: while high-entry markets have more firms financed by FDI, there are less firms who access the national capital market. Similarly, the share of government owned firms is negatively correlated with the rate of entry (although the coefficient is economically small). Panel B then re-estimates (41) while controlling for these additional market characteristics. Not only is the effect of entry still highly significant, the point estimates are also almost the same as before.[38]

Summarizing, I interpret the results above as supportive of the main economic mechanism stressed in this paper: Markets with a more dynamic entry environment have lower mark-ups precisely because entering firms induce churning in the market place and keep the monopoly power of existing producers in check.[39] While this relationship between the distribution of prices (and hence static misallocation) and the rate of entry is one important testable corollary of the model, it is of course the relationship between static misallocation and equilibrium growth, which I am most interested in. To study this relationship, more theoretical structure is needed. In particular, it is here that I have to take stand on the innovation environment and calibrate the structural parameters.

3.4 Calibration

So far I used the model only to qualitatively interpret the distribution of measured wedges across product markets and the cross-sectional pattern of revenue productivity across producers. In this section, I want to turn to two questions, for which (at least parts of) the theoretical structure of the general equilibrium model is essential. First of all, I want to consider the conceptually simple accounting exercise for how much of the observed productivity distribution in the micro-data the model can account for.[40] The model predicts that this dispersion should only depend on the single variable $\vartheta = \frac{ln(1+z/I)}{ln(\lambda)}$. As the entry rate z and the innovation rate I can be estimated from the micro data, one does *not* have to specify the underlying innovation environment to answer the accounting question - a simple calibration of λ is sufficient. In a second step I want to go back to the welfare comparison between the static and dynamic effects introduced in Proposition 3. There I derived a simple expression for the dynamic welfare gains under the simplifying assumptions of (a) taking incumbents' innovation rate as parametric and (b) the absence of capital and inter-sectoral labor reallocation. In this section, I will depart from these assumptions. To do so, it is indeed necessary to specify the details of the innovation environment (and to calibrate the underlying structural parameters) because the endogenous response of entry and innovation activity have to be taken into account.

[38]In the interest of brevity I only report the results for the average mark-up. The results for the quantiles and the standard deviation are qualitatively similar and are available upon request.

[39]In the Appendix I provide various robustness checks for these results. In particular, I redo the analysis using the material share as the dependent variable and different specifications of the output elasticity (see (38)). Furthermore, I redo the analysis when I define product markets at the 4-digit level (instead of the 3-digit level), when I restrict the sample to markets with at least 20 firms and when I use the average exit rate as a measure of entry (in the theory these are obviously identical).

[40]See also Song and Wu (2013) who conduct a structural estimation to distinguish capital market imperfections from measurement error, adjustment costs and technological heterogeneity.

3.4.1 Accounting

How much of the observed dispersion in labor productivity can the model explain? According to (18), the model implies that

$$sd\left[ln\left(\frac{py}{wl}\right)\right] = \sigma_{LP} = \vartheta(x)^{-1} = \frac{ln(\lambda)}{ln(1+x)} = \frac{ln(\lambda)}{ln(1+z/I)}. \tag{42}$$

(42) is useful, because it depends on parameters and endogenous outcomes, which can be directly measured or easily calibrated. z is simply the entry rate, which is directly observable. To measure I, recall that according to (20) average log mark-ups as a function of age are given by $E[ln(\mu)] = ln(\lambda) It$. Hence, $ln(\lambda) I$ is simply the coefficient in a linear regression of mean (log) mark-ups on the age of the cohort (i.e. the slope of the schedule in Figure 1). Finally, these two outcomes are related to the aggregate economy via the growth rate of TFP $g_Q = ln(\lambda)(I+z)$. Along the BGP, the misallocation wedge $M(t)$ is constant so that the growth rate of TFP can be estimated using the aggregate Indonesian data from the Penn World Tables. These three moments obviously identify the parameter vector (z, I, λ) and hence the implied distribution of labor productivity by the model, which I can contrast with the data. The results of this exercise are contained in Table 7.

The first panel contains the moments of the data. Between 1990 and 1998,[41] aggregate productivity in Indonesia grew at an annual rate of roughly 2.5%. The mean entry rate during that time is 13% and the growth rate of incumbents' labor productivity is about 1.5%. These values imply a step-size of about 1.07, i.e. each innovation increases physical efficiency by about 7%, and an innovation rate of 23%, so that incumbent firms innovate at roughly twice the intensity of entering firms.[42] Hence, the shape parameter of the mark-up distribution is given by $\vartheta \approx 6.5$, which according to Proposition 2, is sufficient to calculate all the moments of interest. Panels three and four of Table 7 contain the result. The average mark-up in the economy is slightly below 20%, which is broadly in line with the results reported in De Loecker and Warzynski (2012). While the implied TFP losses are modest (on the order of magnitude of 1%), the implicit labor (and investment) tax is quite substantial: factor prices are roughly 13% below their static social marginal product. The theory's implications for the dispersion of labor productivity are reported in column six. Compared to the micro-data, the model explains roughly 20% of the productivity dispersion across firms within industries.[43]

The implied TFP losses seem small, especially compared to the much bigger numbers reported in Hsieh and Klenow (2009). There are two reasons. First of all, by reducing misallocation in India to US standards, they consider a much bigger liberalization experiment as the empirical productivity dispersion

[41] Because of the financial crisis in East Asia in the late 90s, I only use the data prior 1998 for the calibration analysis of this balanced growth model.

[42] The step size of 1.07 is consistent with the preferred value of 1.05 in Acemoglu and Akcigit (2012).

[43] The numbers reported in Table 7 are consistent with the numbers in literature. Hsieh and Klenow (2009) for example report these dispersion measures for India as 0.69, 0.79 and 1.73 respectively. In fact, productivity dispersion is mostly a within-sector phenomenon and idiosyncratic capital-shares matter little. The dispersion of log labor productivity within industry-year cells without controlling for the capital-labor ratio is given by 0.824, 0.842 and 1.784 respectively. The dispersion of raw log labor productivity is 0.877, 0.971 and 1.933 respectively.

in India is 50% higher.[44] Secondly, they consider an elasticity of substitution across products of three, whereas I impose a unitary demand elasticity. To see that Table 7 is then directly comparable to their results, recall that in their set-up, aggregate TFP losses are approximately given by[45]

$$lnTFP = \frac{1}{\sigma - 1} ln \left(\int q(i)^{\sigma - 1} di \right) - \frac{\sigma}{2} var\left(ln\left(ARPL\right)\right).$$

For $\sigma = 1$ and $var\left(ln\left(ARPL\right)\right) = 0.15^2$, we exactly recover a TFP loss of 1% as reported in Table 7.

It is also interesting to note that it is precisely the interdependence between the static dispersion of marginal products and the dynamic properties of the economy, in particular the rate of entry, which gives the theory bite in not explaining the entire dispersion of productivity. Consider for example the exercise of calibrating the model to the *observed* productivity dispersion given a value of λ. From (42) and given the values in Table 7, the required entry intensity is given by $x = e^{ln(\lambda)\sigma_{LP}^{-1}} - 1 = 0.09$. The implied innovation and entry rates, which are consistent the observed rate of productivity growth are then given by $\hat{I} = 33\%$ and $\hat{z} = 3\%$. Hence, in order to match the high dispersion in labor productivity while not generating too much aggregate growth, the entry rate has to be counterfactually low.[46] The micro-data on firm-dynamics in conjunction with the theory therefore puts informative bounds on what fraction of the static productivity dispersion this particular mechanism can explain.

3.4.2 Productivity Differences, Misallocation and Welfare

Let me now turn to the relationship between cross-sectional productivity differences, misallocation and welfare in this economy. Proposition 3 already suggested that the welfare consequences of the dynamic channels could be much be bigger than the static effect. In fact, using the calibrated values from Table 7 (and assuming a discount rate of 5%), (25) implies a dynamic multiplier of around 12. The main simplifying assumption in deriving the dynamic multiplier in (25) concerned the endogenous response of incumbents' innovation efforts. In fact, (26) showed that any crowding out in firms' innovation incentives will reduce those dynamic gains. In this section, I will now use the full general equilibrium model to assess the likely magnitude of dynamic growth effects taking the response of firms' innovation response explicitly into account (i.e. calculating the response ω in (26)).

As in Proposition 3 I am going to consider the welfare consequences of a given reduction in the

[44] This in itself will cause the implied efficiency losses to be bigger. As noted by Jones (2013): "Small departures from the optimal allocation of labor have tiny effects on TFP (an application of the envelope theorem), but significant misallocation can have very large effects."

[45] See equation (16) in Hsieh and Klenow (2009). This equation is exact if physical productivity and wedges are independent and jointly log-normally distributed. In my model, physical productivity and mark-ups are (endogenously) independent. Given the similarity between the pareto distribution and the log-normal distribution, it is not surprising that this is a good approximation to the actual losses in this model.

[46] Another way to see that there is too much entry in the data for the theory to be able to explain the entire dispersion of marginal products is as follows: given $z = 13\%$, we could try to calibrate I and λ to match both the aggregate growth rate and the dispersion of marginal products. Using (42) and $g_Q = ln(\lambda)(I + z) = ln(\lambda)\frac{1+x}{x}z$, we get that the entry intensity x has to solve the equation $ln(1+x)\frac{1+x}{x} = \frac{g_Q}{\sigma_{LP} z}$. It is easy to show that there is no solution for x given the data (g, σ_{LP}, z).

productivity dispersion across firms. Hence, conceptually, this exercise is exactly the dynamic counterpart of Hsieh and Klenow (2009). While the static welfare losses do not depend on the underlying source of variation (and hence will be exactly identical to the Hsieh and Klenow exercise), the general equilibrium response of entry and innovation incentives is required to calculate the full dynamic gains. To do so, I have to take a stand on which structural parameter induces the change in entry and hence productivity dispersion. I will think of this variation as being caused by differences in the cost of entry χ. This is not only the parameter that - in the model - has the closest relation to the entry intensity x, but it has also been argued that entry costs indeed vary widely across countries (Bergoeing et al., 2012), that these are correlated with the distribution of living standard around the world (Barseghyan, 2008) and that they are influenced by policies like regulation or red tape (Djankov et al., 2002).

The calibration of the structural parameters is straight-forward and transparent as the model is very parsimonious and has only seven parameters $(\alpha, \rho, \delta, \lambda, \varphi, \chi, \gamma)$. As I am mostly interested in the innovation environment, I will not be concerned with calibrating (α, ρ, δ) but set them exogenously at conventional levels. The step-size λ was already calibrated above to the aggregate rate of productivity growth given the innovation and entry rate. This leaves three parameters to be identified. The curvature of the innovation technology γ is directly linked to firms' expenditure share on innovation resources. In particular, letting l_I and l_P be the number of innovators and production workers, the model implies that

$$s^I \equiv \frac{wl_I}{wl_P + wl_I} = \frac{1}{\gamma\left(1-\alpha\right)\left(\frac{\lambda}{\lambda-1}\frac{\rho+z}{I}+1\right)+\alpha}, \tag{43}$$

which uniquely identifies γ given the observed (and already calibrated) (λ, z, I). The two remaining moments for the entry costs χ and the innovation productivity φ are then simply the two equilibrium conditions for I and z (see Proposition 4). The implied structural parameters are reported in Table 8, which also shows that the model is in fact capable of replicating these moments in the data.

As in the special case above (see (24)), the implied welfare changes can be decomposed into a static and dynamic component. Using (23), we get

$$\begin{aligned} \frac{dW}{d\sigma_{LP}} &= \frac{1}{\rho}\left[\left.\frac{\partial ln\left(c\right)}{\partial ln\left(\sigma_{LP}\right)}\right|_g + \left(\frac{\partial ln\left(c\right)}{\partial g} + \frac{1}{\rho}\frac{1}{1-\alpha}\right)\frac{\partial g}{\partial ln\left(\sigma_{LP}\right)}\right] \\ &= \frac{1}{\rho}\left[\left.\frac{\partial ln\left(c\right)}{\partial ln\left(\sigma_{LP}\right)}\right|_g - \left(\frac{\partial ln\left(c\right)}{\partial g} + \frac{1}{\rho}\frac{1}{1-\alpha}\right)g\frac{ln\left(\lambda\right)}{\sigma_{LP}}\left(\frac{1+\omega}{1-x\omega}\right)\right], \end{aligned} \tag{44}$$

where $\omega = \frac{\partial I}{\partial z}$ is again the equilibrium response of incumbents' innovation effort and $\left.\frac{\partial ln(c)}{\partial ln(\sigma_{LP})}\right|_g$ is the static effect of misallocation.[47]

Empirically, I consider the welfare effects of a reduction of productivity dispersion by 10%. The results are contained in Table 9. Consider first the second row, which contains only the static component. In

[47] Note that under the assumptions of Proposition 3 we have $ln\left(c\right) = \text{const} + ln\left(M\right)$ so that $\frac{\partial ln(c)}{\partial g} = 0$ and (44) yields (26).

order to reduce the productivity dispersion by 10%, the entry intensity has to increase by 8 percentage points. This increases aggregate TFP by 0.2%. In conjunction with the reduction of the labor wedge and the increase in the capital stock, this reduction in misallocation increases total welfare by 0.6% holding the growth rate fixed. The second row corresponds to the case akin to Proposition 3, where there is neither a response in incumbents' innovation effort ($\omega = 0$) nor in the allocation of labor across activities. The increase in entry raises the growth rate by 13 percentage points. Welfare is now 3.9% higher and hence more than 6 times higher than the static gains. In the third panel I now include the endogenous innovation response by existing firms. As expected, this reduces the welfare gains as the increase in entry crowds out incumbents' innovation incentives. Still, the total welfare gains, which amount to about 3%, are almost five times as high as the static gains.[48] For comparison, I also include the case without capital in the last panel. In that case, static welfare is of course proportional to aggregate TFP so that the change in welfare is also 0.2%. While the dynamic welfare effects are smaller than before, the dynamic multiplier increases. The dynamic multiplier in the second row of 14 corresponds exactly to (25).[49] Taking into account the endogenous response of incumbents' incentives again reduces this multiplier. The dynamic gains are however still far bigger then the static ones.

(44) is not only helpful to actually perform the decomposition. It also shows, that two parameters are particularly important to understand the relative magnitude of the dynamic effects. Besides the discount rate ρ, which is akin to a weight on the dynamic component, the convexity of the innovation technology γ is crucial, because it affects the strength of the crowding out of incumbents' research effort. To see this, consider Table 10, where I redo the analysis for different values of ρ and γ (recall that according to (43) there is one-to-one mapping between s_I and γ). In the first two panels I show that - expectedly - the dynamic gains are decreasing in the discount rate but that even for a high discount rate of 10%, the dynamic gains are more than twice as big even taking the adverse innovation response into account. In the lower panel, I reduce the convexity of incumbents' cost function. This increases the strength of crowding out. For the case of $s_I = 5\%$, the dynamic multiplier reduces from 4.77 to 2.6. By reducing this convexity further, I finally show that a negative correlation between productivity dispersion and welfare is by no means an analytical result once dynamic considerations are taking into account. If incumbents' cost function is close to linear, there might be excessive crowding out in that a reduction of entry costs actually reduces the growth rate. While this is unlikely to be the empirically relevant case in this setup, it stresses the importance of taking such equilibrium adjustments into account and to take a stand on which underlying structural parameter causes the endogenous variation in misallocation.

[48] While these dynamic gains are high, I want to stress that these are comparisons of the balanced growth path and hence likely an upper bound. During the transition, I suspect the growth rate to be below its BGP value. Furthermore, the now higher capital stock has to be accumulated. Hence, the higher growth rate is only achieved in the future (which is of course discounted at rate ρ) and households have to forego consumption to save. The welfare calculation in Table 9 does not take into account either of these concerns. However, I want to stress that the same reasoning also applies to the static gains of increased TFP. It is only over time that the higher entry rate will reduce mark-ups and increase allocative efficiency. It is therefore not obvious that my reliance on the BGP increases the *relative* importance of the dynamic channels.

[49] Substituting the parameters in (25) yields a value of around 12. The difference stems from the nonlinearities in the system. When I recalculate Table 9 for a reduction of σ_{LP} of 1%, I recover the value of 12.

4 Conclusion

There is a large literature that stresses the importance of micro-misallocation for our understanding of aggregate productivity differences. This literature focuses on the allocative consequences of misallocation and hence views misallocation as solely a static phenomenon. In this paper I argued that misallocation also has dynamic consequences because it interacts with firms' innovation and entry incentives and hence aggregate growth. I proposed a tractable endogenous growth model with heterogeneous firms that provides a clear link between static misallocation and aggregate growth.

In the model, static misallocation is due to firms' monopoly power and hence depends on the underlying distribution of mark-ups. This distribution, however, is endogenous and depends on the underlying process of firm-dynamics. Static misallocation and equilibrium growth are therefore jointly determined. The theory stresses the importance of entry. As entry has a pro-competitive effect, it endogenously reduces misallocation. If entry also increases aggregate growth, i.e. as long as entrants do not fully crowd out incumbents' innovation incentives, misallocation and growth will be negatively correlated across countries. The welfare implications of the observed degree of misallocation might therefore be much larger than previously appreciated.

Using Indonesian firm-level data, I tested the predictions of the theory. At a reduced-form level, I provided evidence for the importance of imperfect output markets as a source of misallocation and for the positive effect of entry on allocative efficiency. When calibrated to the data, the model implies that the dynamic losses of misallocation are about four times as large as their static counterpart.

References

Acemoglu, D. and U. Akcigit: 2012, 'Intellectual Property Rights Policy, Competition and Innovation'. *Journal of the European Economic Association* **10**(1), 1–42.

Acemoglu, D., U. Akcigit, N. Bloom, and W. Kerr: 2012, 'Innovation, Reallocation and Growth'. *Working Paper*.

Acemoglu, D., P. Antras, and E. Helpman: 2007, 'Contracts and Technology Adoption'. *American Economic Review* **97**(3), 916–943.

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. 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.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodriguez-Clare: 2012, 'The Elusive Pro-Competitive Effects of Trade'. *Working Paper*.

Atkeson, A. and A. Burstein: 2008, 'Pricing-to-Market, Trade Costs, and International Relative Prices'. *American Economic Review* **98**(5), 1998–2031.

Atkeson, A. and A. Burstein: 2010, 'Innovation, Firm Dynamics, and International Trade'. *Journal of Political Economy* **118**(3), 433–484.

Atkeson, A. and A. Burstein: 2011, 'Aggregate Implications of Innovation Policy'. *Working Paper*.

Banerjee, A. and E. Duflo: 2005, 'Growth Theory through the Lens of Development Economics'. In: P. Aghion and S. Durlauf (eds.): *Handbook of Economic Growth*, Vol. 1A. Elsevier B.V., Chapt. 7, pp. 473–552.

Banerjee, A. and B. Moll: 2010, 'Why Does Misallocation Persist?'. *American Economic Journal: Macroeconomics* **2**(1), p. 189–206.

Barseghyan, L.: 2008, 'Entry Costs and Cross-Country Differences in Productivity and Output'. *Journal of Economic Growth* **13**(2), 145–167.

Bartelsman, E. J. and M. Doms: 2000, 'Understanding Productivity: Lessons from Longitudinal Microdata'. *Journal of Economic Literature* **38**(3), 569–594.

Bartelsman, E. J., J. C. Haltiwanger, and S. Scarpetta: 2013, 'Cross-Country Differences in Productivity: The Role of Allocative Efficiency'. *American Economic Review* **103**(1), 305–334.

Bergoeing, R., N. V. Loayza, and F. Piguillem: 2012, 'The Complementary Impact of Micro Distortions'. *Working Paper*.

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.

Buera, F., J. Kaboski, and Y. Shin: 2011, 'Finance and Development: A Tale of Two Sectors'. *American Economic Review* **101**(5), 1964–2002.

Caselli, F.: 2005, 'Accounting for Cross-Country Income Differences'. In: P. Aghion and S. Durlauf (eds.): *Handbook of Economic Growth*, Vol. 1A. Elsevier B.V., Chapt. 9, pp. 679–741.

Chari, V., P. J. Kehoe, and E. R. McGrattan: 2007, 'Business Cycle Accounting'. *Econometrica* **75**(3), 781–836.

Clementi, G. L. and H. Hopenhayn: 2006, 'A Theory of Financing Constraints and Firm Dynamics'. *The Quarterly Journal of Economics* **121**(1), 229–265.

Collard-Wexler, A., J. Asker, and J. De Loecker: 2011, 'Productivity Volatility and the Misallocation of Resources in Developing Economies'. *NBER Working Paper* **17175**.

De Loecker, J.: 2011a, 'Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity'. *Econometrica* **79**(5), 1407–1451.

De Loecker, J.: 2011b, 'Recovering Markups from Production Data'. *International Journal of Industrial Organization* **29**, 350–355.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik: 2012, 'Prices, Markups and Trade Reform'. *Working Paper.*

De Loecker, J. and F. Warzynski: 2012, 'Markups and Firm-Level Export Status'. *American Economic Review* **102**(6), 2437–2471.

Djankov, S., R. La Porta, F. Lopez-De-Silvaes, and A. Shleifer: 2002, 'The Regulation of Entry'. *Quarterly Journal of Economics* **117**(1), 1–37.

Edmond, C., V. Midrigan, and D. Y. Xu: 2011, 'Competition, Markups and the Gains from International Trade'. *Working Paper.*

Epifani, P. and G. Gancia: 2011, 'Trade, Markup Heterogeneity and Misallocations'. *Journal of International Economics* **83**(1), 1–13.

Fattal Jaef, R. N.: 2011, 'Entry, Exit and Misallocation Frictions'. *Working Paper.*

Foster, J., J. C. Haltiwanger, and C. Syverson: 2008, 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?'. *American Economic Review* **98**(1), 394–425.

Foster, J., J. C. Haltiwanger, and C. Syverson: 2011, 'The Slow Growth of New Plants: Learning about Demand'. Technical report, Working Paper.

Gabaix, X.: 1999, 'Zipf's Law for Cities: An Explanation'. *Quarterly Journal of Economics* **114**(3), 739–767.

Gancia, G. and F. Zilibotti: 2009, 'Technological Change and the Wealth of Nations'. *Annual Review of Economics* **1**, 93–120.

Garcia, A. M. and N. Voigtländer: 2013, 'Exporting and Plant-Level Efficiency Gains: It's in the Measure'. *Working Paper.*

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 Ouput Per Worker Than Others?'. *Quarterly Journal of Economics* **114**(1), 83–116.

Holmes, T. J., W.-T. Hsu, and S. Lee: 2013, 'Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade'. *Working Paper*.

Hopenhayn, H.: 2012, 'On the Measure of Distortions'. *Working Paper*.

Hsieh, C.-T. and P. Klenow: 2009, 'Misallocation and Manufacturing TFP in China and India'. *Quarterly Journal of Economics* **124**(2), 771–807.

Hsieh, C.-T. and P. Klenow: 2011, 'The Life-Cycle of Plants in India and Mexico'. Technical report, Working Paper.

Jones, C.: 2013, 'Misallocation, Economic Growth and Input-Output Economics'. In: D. Acemoglu, M. Arrelano, and E. Dekel (eds.): *Advances in Econometrics: 10th World Congress*, Vol. 2. Cambridge University Press.

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.

La Porta, R. and A. Shleifer: 2009, 'The Unofficial Economy and Economic Development'. In: *Brookings Papers on Economic Activity*. pp. 275–352.

Lagos, R.: 2006, 'A Model of TFP'. *Review of Economic Studies* **73**(4), 983–1007.

Lentz, R. and D. T. Mortensen: 2008, 'An Empirical Model of Growth through Product Innovation'. *Econometrica* **76**(6), 1317–1373.

Lucas, R. E.: 1988, 'On the Mechanics of Economic Development'. *Journal of Monetary Economics* **22**(1), 3–42.

Lucas, R. E.: 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'.

Melitz, M.: 2003, 'The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity'. *Econometrica* **71**(6), 1695–1725.

Melitz, M. and G. Ottaviano: 2008, 'Market Size, Trade and Productivity'. *Review of Economic Studies* **75**(1), 295–316.

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.

Rajan, R. G. and L. Zingales: 1998, 'Financial Dependence and Growth'. *American Economic Review* **88**, 559–586.

Rauch, J. E.: 1999, 'Networks versus markets in international trade'. *Journal of International Economics* **48**(1), 7–35.

Restuccia, D. and R. Rogerson: 2008, 'Policy Distortions and Aggregate Productivity with Heterogenous Establishments'. *Review of Economic Dynamics* **11**(4), 707–720.

Song, Z. and G. Wu: 2013, 'A Structural Estimation on Capital Market Distortions in Chinese Manufacturing'. *Working Paper.*

Sutton, J.: 1997, 'Gibrat's Legacy'. *Journal of Economic Literature* **35**, 40–59.

Yang, M.-J.: 2012, 'Micro-level Misallocation and Selection: Estimation and Aggregate Implications'. *Working Paper.*

5 Appendix

5.1 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 reduced form analysis of this paper, I mainly need information on firms' value added, wagebill, capital stock, material spending, location, 5-digit product market and entry behavior. Additionally I exploit information on the source of firms' investment to measure credit constraints and their ownership structure to identify firms with state ownership. 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. To identify entry, I rely on the panel nature of the data. My measure for the product-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, La 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 too bad a measure of when a firm starts competing within the formal sector. All nominal variables are deflated using either 5-digit product specific price deflators (for revenue, value added and material spending), 4-digit specific capital deflators (for firms' capital stock) or the CPI Index from the World Development Indicators (wages). In order to have a direct measure of credit constraints for Table 2 and R&D and growth for Table 1, 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 were also asked whether they were spending resources on productivity improvements and how their profits compared to the previous year. 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 and an indicator if the village's main source of income is agriculture. 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.

5.2 Proof of Proposition 2

In a stationary equilibrium we have $\dot{\nu}(\Delta, t) = 0$. (14) then implies that

$$\nu(\Delta) = \left(\frac{I}{z+I}\right)^{\Delta} \frac{z}{I} = \left(\frac{1}{1+x}\right)^{\Delta} x.$$

Hence,

$$F_{\Delta}(d;x) = P(\Delta \leq d;x) = \sum_{i=1}^{d} \nu(i) = x \sum_{i=1}^{d} \left(\frac{1}{1+x}\right)^{i} = x \frac{\frac{1}{1+x}\left(1 - \left(\frac{1}{1+x}\right)^{d}\right)}{1 - \frac{1}{1+x}} = 1 - \left(\frac{1}{1+x}\right)^{d} = 1 - exp\left(-ln\left(1+x\right)d\right),$$

which implies that $P(ln(\mu) \leq m) = P\left(\Delta \leq \frac{m}{ln(\lambda)}\right) = 1 - exp(-\vartheta(x)d)$, i.e. log mark-ups are exponentially distributed with parameter $\vartheta(x)$. Similarly, $F(\mu;x) = P\left(\lambda^{\Delta} \leq \mu;x\right) = 1 - \mu^{-\vartheta(x)}$ as required in (16). To derive (17), note that

$$\begin{aligned} \Lambda(x) &= \int \mu^{-1}\vartheta\mu^{-(\vartheta+1)}d\mu = \frac{\vartheta}{\vartheta+1} \\ M(x) &= exp\left(-Eln(\mu)\right)\Lambda^{-1} = exp\left(-\vartheta^{-1}\right)\frac{\vartheta+1}{\vartheta}. \end{aligned}$$

Note that these expression are derived taking the distribution of mark-ups as continuous, even though the model implies that they are discrete. Taking this discreteness explicitly into account yields

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

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

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

While M^{Dis} and M are virtually indistinguishable, Λ^{Dis} and Λ are sufficiently close for all practical purposes.[50] Furthermore, $\Lambda > \Lambda^{Dis}$ so that the approximation is conservative. The growth rate g follows directly from the fact that the frontier technology Q increases at rate $I + z$.

5.3 Proof of (20)

We have to show that the distribution of quality gaps Δ as a function of age conditional on survival, $\zeta_{\Delta}(t)$, is given by $\zeta_{\Delta+1}(t) = \frac{1}{\Delta!}(It)^{\Delta} e^{-It}$. Let $p_{\Delta}(t)$ denote the probability of having a quality gap at age t when entering at

[50] At the calibrated parameters, the difference is about 3%.

time 0. The corresponding flow equation are

$$\dot{p}_{\Delta}(t)=\begin{cases}(1-p_0(t))\,z & \text{for } \Delta=0\\ -p_1(t)(I+z) & \text{for } \Delta=1\\ p_{\Delta-1}(t)\,I-p_{\Delta}(t)(I+z) & \text{for } \Delta\geq 2\end{cases}.$$

The solution to this set of differential equations is given by

$$\begin{aligned}p_0(t) &= 1-e^{-zt}\\ p_{i+1}(t) &= \left(\frac{1}{i!}\right)I^i t^i\left(e^{-(I+z)t}\right) \text{ for } i\geq 0.\end{aligned}$$

The distribution of mark-ups conditional on survival is then

$$p_{i+1}^S(t)\equiv\frac{p_{i+1}^S(t)}{1-p_0(t)}=\left(\frac{1}{i!}\right)I^i t^i\left(e^{-It}\right).$$

Hence, this is a Poisson distribution with parameter It, so that $E[\Delta|t]=It$. (20) then follows because $ln(\mu)=ln(\lambda)\Delta$.

5.4 Proof of Proposition 3

If I and z are constant, the economy is akin to the standard neoclassical growth model with exogenous productivity growth g, TFP M and a labor and investment wedge Λ. Hence, normalized consumption and capital per capita are given by

$$c = Mk^{\alpha}L_P^{1-\alpha}-\left(\delta+\frac{g}{1-\alpha}\right)k \text{ and } k=\left(\frac{\alpha\Lambda M}{\rho+\frac{1}{1-\alpha}g+\delta}\right)^{\frac{1}{1-\alpha}}L_P.$$

Substituting k into the expression for c yields

$$c=\left(\frac{\rho+\left(\delta+\frac{g}{1-\alpha}\right)(1-\alpha\Lambda)}{\alpha\Lambda}\right)\left(\frac{\alpha\Lambda M}{\rho+\frac{1}{1-\alpha}g+\delta}\right)^{\frac{1}{1-\alpha}}L_P. \tag{45}$$

To derive the expression for welfare (23) note that $c(t)=ce^{g_Y t}=ce^{\frac{1}{1-\alpha}gt}$, where g_Y is the growth rate of output. Hence,

$$W=\int e^{-\rho t}ln(c(t))\,dt=\frac{1}{\rho}ln(c)+\frac{1}{1-\alpha}g\int e^{-\rho t}tdt=\frac{1}{\rho}ln(c)+\frac{1}{1-\alpha}\frac{1}{\rho^2}g.$$

Substituting (45) yields (23).

Expression (24) implies that

$$\frac{dW/dln(\sigma_{LP})}{dW/dln(\sigma_{LP})|_g}=1+\frac{1}{\rho}\frac{\frac{\partial g}{\partial ln(\sigma_{LP})}}{\frac{\partial ln(M)}{\partial ln(\sigma_{LP})}}=1+\frac{1}{\rho}\frac{\frac{\partial g}{\partial\sigma_{LP}}}{\frac{\partial ln(M)}{\partial\sigma_{LP}}}.$$

From (17) and (18) we get that $\frac{\partial ln(M)}{\partial \sigma_{LP}} = -\frac{\sigma_{LP}}{1+\sigma_{LP}}$. Similarly, (19) implies that $\frac{\partial g}{\partial \sigma_{LP}} = -g\frac{ln(\lambda)}{\sigma_{LP}^2}$. Hence,

$$\frac{\frac{\partial g}{\partial \sigma_{LP}}}{\frac{\partial ln(M)}{\partial \sigma_{LP}}} = gln\left(\lambda\right)\frac{1+\sigma_{LP}}{\sigma_{LP}^3} = gln\left(1+x\right)\frac{1+\sigma_{LP}}{\sigma_{LP}^2} = gln\left(1+x\right)\vartheta\left(1+\vartheta\right),$$

which yields the expression in Proposition 3. If the effect on the innovation rate is taken into account, we get that $\frac{\partial g}{\partial \sigma_{LP}} = -g\frac{ln(\lambda)}{\sigma_{LP}^2} + g\frac{\partial I}{\partial \sigma_{LP}}\frac{1}{I}$. The structural equation from the innovation equilibrium is $I\left(x\right)$, i.e. I depends on σ_{LP} only through x. Hence,

$$\frac{\partial I}{\partial \sigma_{LP}} = \frac{\partial I}{\partial x}\frac{\partial x}{\partial \sigma_{LP}} = -\frac{\partial I}{\partial x}exp\left(ln\left(\lambda\right)\sigma_{LP}^{-1}\right)ln\left(\lambda\right)\sigma_{LP}^{-2}.$$

Therefore,

$$\frac{\partial g}{\partial \sigma_{LP}} = -g\frac{ln\left(\lambda\right)}{\sigma_{LP}^2}\left(1+\frac{\partial I}{\partial x}\frac{1+x}{I}\right) = -g\frac{ln\left(\lambda\right)}{\sigma_{LP}^2}\left(\frac{1+\eta}{1-x\eta}\right),$$

where the last equality uses the change of variables $x = \frac{z}{I}$ and $\frac{\partial I}{\partial x} = \left(1 - x\frac{\partial I}{\partial z}\right)^{-1} I\frac{\partial I}{\partial z}$.

5.5 Proof of Proposition 4

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 $w_I\left(t\right)$ and the value functions $V\left(\Delta, t\right)$ also grow at rate g_Y and that innovation and entry rates are constant and equal across all sectors, i.e. $z\left(\Delta, t\right) = z$ and $I\left(\Delta, t\right) = I$. These conjectures will be verified below. To characterize the BGP, conjecture that the value function takes the form of

$$V\left(\Delta, t\right) = \kappa\left(t\right) - \phi\left(t\right)\lambda^{-\Delta}. \tag{46}$$

If V grows at rate g_Y, (46) implies that both $\kappa\left(t\right)$ and $\phi\left(t\right)$ grow at rate g_Y too. Let $V\left(\Delta\right) = exp\left(-g_Y t\right)V\left(\Delta, t\right)$ be the normalized value function, which is stationary. Using (46) and (8) and $V\left(\Delta+1\right) - V\left(\Delta\right) = \phi\lambda^{-\Delta}\frac{\lambda-1}{\lambda}$, the Bellman equation (28) reads

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

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

$$\gamma\frac{1}{\varphi}I^{\gamma-1} = \frac{\phi}{w}\frac{\lambda-1}{\lambda}, \tag{47}$$

where $\frac{\phi}{w_I}$ is constant along the BGP. Hence, I^* is independent of time and the same for all firms. In particular,

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

As this equation has to hold for all Δ, we need that

$$\phi = \frac{Y-\left(\gamma-1\right)w\frac{1}{\varphi}I^\gamma}{r+z-g_Y} = \frac{Y-\left(\gamma-1\right)w\frac{1}{\varphi}I^\gamma}{z+\rho}. \tag{48}$$

Similarly we get $\kappa = \frac{Y}{z+\rho}$, so that the value function is given by

$$V(\Delta, t) = \frac{\pi(\Delta, t) + (\gamma - 1) w(t) \Gamma(I, \Delta)}{\rho + z}, \tag{49}$$

as required in (29). That both w_I and Y grow at this rate, is seen from the free entry condition. We finally have to establish that there exists a solution for I, which is consistent with (47). Using (48) yields

$$\frac{\gamma}{\varphi} I^{\gamma-1} (\rho + z) \frac{\lambda}{\lambda - 1} + (\gamma - 1) \frac{1}{\varphi} I^{\gamma} = \frac{Y(t)}{w(t)}. \tag{50}$$

The free entry condition implies

$$\frac{Y(t)}{w(t)} = \frac{\lambda}{\lambda - 1} (\rho + z) \chi - (\gamma - 1) \frac{1}{\lambda - 1} \frac{1}{\varphi} I^{\gamma}. \tag{51}$$

From (50) and (51) we get

$$\frac{\gamma}{\varphi} I^{\gamma-1} + (\gamma - 1) \frac{1}{\varphi} \frac{I^{\gamma}}{\rho + z} = \chi, \tag{52}$$

which determines a unique level of I given z.

To finally determine the equilibrium innovation level I and the entry intensity x, consider (52), which we can write as

$$\gamma I^{\gamma-1} + (\gamma - 1) \frac{I^{\gamma}}{\rho + xI} = \varphi \chi. \tag{53}$$

It is easy to verify that (53) defines a continuous schedule $I = h(x, \rho, \varphi, \chi)$, which is increasing in all arguments. From the labor market condition, $w(t) = \Lambda(t) \frac{(1-\alpha) Y(t)}{L_P(t)}$ and (50) we get

$$1 = L_P + \Lambda(x) \frac{1}{\varphi} I^{\gamma} + \chi z = \Lambda(x)(1 - \alpha) \left(\frac{\gamma}{\varphi} I^{\gamma-1} (\rho + xI) \frac{\lambda}{\lambda - 1} + (\gamma - 1) \frac{1}{\varphi} I^{\gamma} \right) + \Lambda(x) \frac{1}{\varphi} I^{\gamma} + \chi x I, \tag{54}$$

which defines a schedule $I = m(x, \rho, \varphi, \chi, \lambda)$, where m is decreasing in x, ρ, χ and increasing in λ and φ. Furthermore, $\lim_{x \to \infty} m(x, \rho, \varphi, \chi, \lambda) = 0$ and $\lim_{x \to 0} m(x, \rho, \varphi, \chi, \lambda) = \infty$. Hence, (53) and (54) define (I, x) uniquely and imply the comparative static results given in Proposition (4). This proves the proposition.

6 Tables

	Dep. Variable: Labor productivity $ln(\frac{py}{wl})$						
	Full Sample					Census Supplement 1996	
Entrant	-0.0424***						
	(0.00700)						
Exiter		-0.0667***					
		(0.00738)					
Age			0.00680***				
			(0.00170)				
Positive Growth				0.252***			
				(0.00409)			
Exporter					0.0981***		
					(0.00642)		
Profit growth						0.0775***	
						(0.0115)	
R&D spending							0.123**
							(0.0533)
$ln(k/l)$	0.138***	0.137***	0.137***	0.136***	0.136***	0.162***	0.161***
	(0.00188)	(0.00185)	(0.00285)	(0.00176)	(0.00179)	(0.00496)	(0.00495)
N	159678	159674	73657	173863	173863	20854	20854
R^2	0.163	0.162	0.163	0.180	0.163	0.210	0.208

Notes: Robust standard errors are shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. All regressions include a full set of 5-digit product fixed effects, a set of year fixed effects and a set of province fixed effects. The first five columns use the entire sample. Columns three has less observations because it only considers firms that entered after 1990 so that firm age can be measured properly. "Entrant" and "Exiter" are dummy variables for whether the firm entered or exited in the respective year. "Positive growth" is a dummy variable for wether the firm saw its sales growing in the last year. "Profit growth" and "R\&D spending" are dummy variables for whether the firm reported an increase in profits in last year or spending on productivity enhancing activities. "$ln\left(\frac{k}{l}\right)$" denotes the firm's (log) capital-labor ratio. Capital is measured as the total value of assets reported in the industrial census.

Table 1: Determinants of labor productivity: Imperfect output markets and Mark-ups

	Dep. Variable: Labor productivity $ln(\frac{py}{wl})$						
	Full Sample				Census Supplement 1996		
FDI	0.187***						
	(0.0227)						
Foreign loans		0.208***					
		(0.0184)					
Capital market			0.189***				
			(0.0423)				
State owned				0.183***			
				(0.00519)			
Does firm face barrier to expand?					-0.103***		
					(0.0111)		
Does firm face capital constraint?						-0.0650***	
						(0.0135)	
Is missing capital the main concern?							-0.0466*
							(0.0243)
$ln(k/l)$	0.136***	0.136***	0.138***	0.127***	0.160***	0.162***	0.162***
	(0.00178)	(0.00178)	(0.00178)	(0.00181)	(0.00495)	(0.00496)	(0.00496)
N	173863	173863	173863	173863	20854	20854	20854
R^2	0.162	0.163	0.161	0.167	0.211	0.209	0.208

Notes: Robust standard errors are shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. All regressions include a full set of 5-digit product fixed effects, a set of year fixed effects and a set of province fixed effects. The first four columns use the entire sample. “FDI”, “Foreign loans”, “Capital market” and “State owned” are dummy variable indicating whether the firms finances its investment through FDI, foreign loans or funds from the Indonesian capital market or whether the firms is (partly) state owned. the last three columns use the special census supplement of the census in 1996. The survey asked whether firms were facing any barriers to expand, whether this constrains was related to missing capital and whether missing capital was the main obstacle. “$ln\left(\frac{k}{l}\right)$” denotes the firm’s (log) capital-labor ratio. Capital is measured as the total value of assets reported in the industrial census.

Table 2: Determinants of labor productivity: Imperfect input markets and borrowing constraints

	Avg. log mark-up				
	All products				Differentiated products
Entry	-0.135***	-0.145***	-0.134***	-0.120***	-0.131**
	(0.0409)	(0.0404)	(0.0414)	(0.0415)	(0.0553)
(ln) population		-0.0254	-0.0307*	-0.0147	-0.00581
		(0.0154)	(0.0174)	(0.0164)	(0.0132)
Agricultural employmment share			-0.0772	-0.112	-0.0598
			(0.0655)	(0.101)	(0.0789)
Share of villages with banks				-0.000230	0.00807
				(0.0116)	(0.00942)
Share of villages with BRI branch				-0.0237	-0.00725
				(0.0840)	(0.0928)
Share of villages with accessible markets				-0.549**	-0.607**
				(0.221)	(0.229)
N	4480	4480	4480	4480	2972
R^2	0.006	0.010	0.015	0.036	0.059

	Other moments of log mark-up distribution				
	Quantiles				Standard deviation
	25%	50%	75%	90%	
Entry	-0.126**	-0.106**	-0.123**	-0.0792	0.0300
	(0.0478)	(0.0400)	(0.0511)	(0.0521)	(0.0345)
(ln) population	-0.0441***	-0.0313*	0.00384	0.0514*	0.0243
	(0.0142)	(0.0171)	(0.0293)	(0.0291)	(0.0150)
Agricultural employmment share	0.0525	-0.128	-0.347**	-0.471***	-0.328***
	(0.0917)	(0.102)	(0.143)	(0.159)	(0.0604)
Share of villages with banks	0.00339	-0.00453	-0.00164	-0.00599	-0.0104
	(0.0101)	(0.0121)	(0.0176)	(0.0200)	(0.00823)
Share of villages with BRI branch	0.0891	-0.00191	-0.167	-0.318*	-0.237*
	(0.0845)	(0.0929)	(0.127)	(0.175)	(0.117)
Share of villages with accessible markets	-0.188	-0.607**	-1.043***	-1.208***	-0.548***
	(0.249)	(0.231)	(0.235)	(0.246)	(0.143)
N	4480	4480	4480	4480	3645
R^2	0.169	0.059	0.169	0.247	0.427

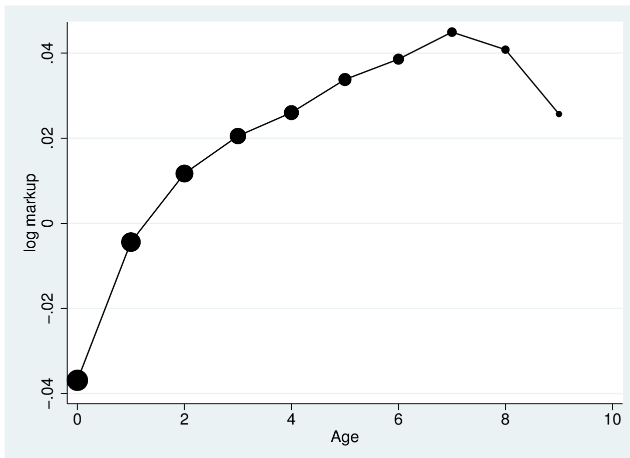
Notes: Standard errors are clustered on the regional level and shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. All regressions contain 3-digit industry fixed effects and year fixed effects. Mark-ups are measured as the residual from the regression $ln\left(py_{it}/wl_{it}\right) = \delta_s + \delta_t + \xi ln\left(k_{it}/l_{it}\right) + u_{it}$, where δ_s and δ_t are 5-digit sector and year fixed effects. "Entry" is the share of firms who enter the respective industry-region cell at year t. "ln(population)" is the log of the total population in the region in 1996. The agricultural employment share is the average share of the village population whose main income source is agricultural. The share of villages with banks/BRI branches is the share of villages within a region that report having at a bank branch or a branch of Bank Rakyat Indonesia, the main microfinance provider in Indonesia. The share of villages with accessible markets is the share of villages that can be accessed by car throughout the year. Products are differentiated in the sense of Rauch (1999).

Table 3: Entry and the Distribution of Mark-ups

	Moment of log mark-up distribution					
	Mean	25%	50%	75%	90%	Std. dev.
Entry	-0.146***	-0.140***	-0.141***	-0.164***	-0.119**	0.0140
	(0.0354)	(0.0379)	(0.0310)	(0.0503)	(0.0568)	(0.0318)
N	4480	4480	4480	4480	4480	3645
R^2	0.105	0.231	0.141	0.242	0.277	0.464

Notes: Notes: Standard errors are clustered on the regional level and shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. All regressions contain 3-digit industry fixed effects, year fixed effects and region fixed effects. Mark-ups are measured as the residual from the regression $ln\left(py_{it}/wl_{it}\right) = \delta_s + \delta_t + \xi ln\left(k_{it}/l_{it}\right) + u_{it}$, where δ_s and δ_t are 5-digit sector and year fixed effects. "Entry" is the share of firms who enter the respective industry-region cell at year t.

Table 4: Entry and the Distribution of Mark-ups: Region fixed effects



Notes: The figure shows the evolution of inframarginal rents as a function of the age of the cohort. Specifically, I calculate log labor productivity relative to the average of 5-digit product-year cells. Then I calculate the mean of this normalized log labor productivity by the age of the cohort and plot this against age. Because of attrition, the size of the cohort is declining in age. The dots reflect the size of the cohort.

Figure 1: Inframarginal rents over the life-cycle

7 Figures

	Λ				$ln(\Lambda)$		
	All products			Differentiated products	All products		Differentiated products
Entry	0.245***	0.260***	0.252***	0.271**	0.252***	0.185***	0.190**
	(0.0768)	(0.0763)	(0.0820)	(0.110)	(0.0727)	(0.0599)	(0.0800)
(ln) population		0.0396*	0.0207	0.00373		0.0380**	0.0240
		(0.0198)	(0.0195)	(0.0194)		(0.0179)	(0.0179)
Agricultural employmment share			0.0112	-0.0998		0.0430	-0.0467
			(0.128)	(0.114)		(0.108)	(0.0946)
Share of villages with banks			0.00615	-0.00644		0.00411	-0.00672
			(0.0143)	(0.0136)		(0.0122)	(0.0115)
Share of villages with BRI branch			-0.0537	-0.0807		-0.0229	-0.0367
			(0.0784)	(0.0895)		(0.0767)	(0.0866)
Share of villages with accessible markets			0.482*	0.527**		0.535**	0.565**
			(0.237)	(0.235)		(0.226)	(0.234)
N	4300	4300	4300	2857	4300	4300	2857
R^2	0.265	0.270	0.277	0.197	0.314	0.327	0.244

	M				$ln(M)$		
	All products			Differentiated products	All products		Differentiated products
Entry	0.00692	0.00522	-0.0190	-0.00707	-0.0125	-0.0309	-0.0173
	(0.0209)	(0.0205)	(0.0167)	(0.0163)	(0.0152)	(0.0217)	(0.0218)
(ln) population		-0.00496	-0.00226	-0.00115		0.00188	0.00371
		(0.00966)	(0.00658)	(0.00762)		(0.00853)	(0.0101)
Agricultural employmment share			0.140***	0.143***		0.178***	0.179***
			(0.0265)	(0.0352)		(0.0335)	(0.0457)
Share of villages with banks			0.00363	0.00377		0.00418	0.00405
			(0.00358)	(0.00434)		(0.00455)	(0.00563)
Share of villages with BRI branch			0.0751	0.0597		0.107	0.0896
			(0.0693)	(0.0636)		(0.0901)	(0.0839)
Share of villages with accessible markets			0.236***	0.262***		0.305***	0.335***
			(0.0597)	(0.0691)		(0.0784)	(0.0921)
N	3686	3686	3686	2462	3686	3686	2462
R^2	0.403	0.404	0.465	0.388	0.512	0.424	0.354

Notes: Standard errors are clustered on the regional level and shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. All regressions contain 3-digit industry fixed effects and year fixed effects. The two wedges Λ and M are calculated as $\Lambda = E\left[\mu^{-1}\right]$ and $M = exp\left(Eln\left(\mu^{-1}\right)\right)/E\left[\mu^{-1}\right]$ where $ln\left(\mu\right)$ is measured as the residual from the regression $ln\left(py_{it}/wl_{it}\right) = \delta_s + \delta_t + \xi ln\left(k_{it}/l_{it}\right) + u_{it}$, where δ_s and δ_t are 5-digit sector and year fixed effects. "Entry" is the share of firms who enter the respective industry-region cell at year t. "ln(population)" is the log of the total population in the region in 1996. The agricultural employment share is the average share of the village population whose main income source is agricultural. The share of villages with banks/BRI branches is the share of villages within a region that report having at a bank branch or a branch of Bank Rakyat Indonesia, the main microfinance provider in Indonesia. The share of villages with accessible markets is the share of villages that can be accessed by car throughout the year. Products are differentiated in the sense of Rauch (1999).

Table 5: Entry and Misallocation

	Dep. Variable: Entry rate				Dep. Variable: Avg mark-up			
Entry					-0.117***	-0.116***	-0.133**	-0.143***
					(0.0382)	(0.0378)	(0.0520)	(0.0338)
Share of firms with FDI	0.185***				0.0205	-0.0735	0.0203	0.129
	(0.0417)				(0.106)	(0.119)	(0.202)	(0.124)
Share of firms with foreign loans		0.0531			0.0319	0.0960	-0.0918	0.105
		(0.0628)			(0.210)	(0.189)	(0.170)	(0.172)
Share of firms with capital market access			-0.372**		-0.233	-0.303	-0.553	-0.306
			(0.137)		(0.418)	(0.407)	(0.377)	(0.422)
Share of firms with state ownership				-0.0239*	0.168***	0.147***	0.162***	0.130***
				(0.0118)	(0.0289)	(0.0325)	(0.0465)	(0.0264)
Regional Controls	No	No	No	No	No	Yes	Yes	No
Region FE	No	No	No	No	No	No	No	Yes
N	4745	4745	4745	4745	4480	4480	2972	4480
R^2	0.269	0.265	0.266	0.267	0.028	0.050	0.076	0.116

Notes: Standard errors are clustered on the regional level and shown in parentheses. ***, ** and * denotes significance at the 1%, 5% and 10% level respectively. All regressions contain 3-digit industry fixed effects and year fixed effects. Mark-ups are measured as the residual from the regression $ln\left(py_{it}/wl_{it}\right) = \delta_s + \delta_t + \xi ln\left(k_{it}/l_{it}\right) + u_{it}$, where δ_s and δ_t are 5-digit sector and year fixed effects. "Entry" is the share of firms who enter the respective industry-region cell at year t. The share of firms with FDI, foreign loans, capital market access and state ownership is calculated directly from the micro data. Specification with regional controls control for the log of the total population in the region in 1996, the average share of the village population whose main income source is agricultural, the share of villages within a region that report having at a bank branch or a branch of Bank Rakyat Indonesia, the main microfinance provider in Indonesia and the share of villages that can be accessed by car throughout the year.

Table 6: Entry and Misallocation: Controlling for financial and political constraints

Data		Implied parameters		Implied aggregate moments		Productivity dispersion			
						Moment	Model	Data	% Explained
g	2.47%	λ	1.071	Avg. markup	1.181	std. dev	0.153	0.789	19.39
z	13%	z	13%	M	0.989	IQR	0.168	0.789	21.29
g_{Inc}	1.58%	I	23.1%	Λ	0.867	$q^{90} - q^{10}$	0.337	1.696	19.87

Notes: g is the growth rate of aggregate productivity. I estimate g from the aggregate data. In particular, it is the average growth rate of TFP Q_t, where $Q_t = \frac{Y_t}{K_t^{\alpha} H^{1-\alpha}}$. Y_t, K_t and H_t are taken from the Penn World Tables and I take $\alpha = 1/3$. I take the data between 1989 and 1998. z is the average entry rate from the micro data. g_{Inc} is the average growth rate of mark-ups of incumbent firms conditional on survival (see (20)). First I estimate mark-ups as the residual from $ln\left(py_{it}/wl_{it}\right) = \delta_s + \delta_t + \xi ln\left(k_{it}/l_{it}\right) + u_{it}$, where δ_s and δ_t are 5-digit sector and year fixed effects. Then I calculate the average mark-up by cohort and run a regression of this average log mark-up against age. (20) implies that this coefficient is equal to $ln\left(\lambda\right) I$ and $g = ln\left(\lambda\right)\left(I + z\right)$. Hence, $\left(g, z, g_{Inc}\right)$ can be solved for $\left(z, I, \lambda\right)$. The implied moments are given in Proposition 2. The dispersion as predicted by the model is also given in Proposition 2. The dispersion in the data is simply the dispersion of the residual from $ln\left(py_{it}/wl_{it}\right) = \delta_s + \delta_t + \xi ln\left(k_{it}/l_{it}\right) + u_{it}$.

Table 7: Accounting for productivity differences

	Parameter	Value	Target	Data	Model
α	Capital share	0.35	set exogenously	-	-
δ	Depreciation rate	0.1	set exogenously	-	-
ρ	Discount rate	0.05	set exogenously	-	-
λ	Step size	1.071	Growth rate g_Q	2.47%	2.47%
γ	Convexity of innv. Tech.	2.607	R&D exp. Share s_I	0.03	0.03
χ	Entry costs	0.690	Entry rate z	13%	13%
φ	Innovation productivity	0.641	Innovation rate I	23.1%	23.1%

Notes: See the main body of the text for the calibration strategy.

Table 8: The calibrated structural parameters

				With capital ($\alpha = 0.35$)			No capital ($\alpha = 0$)		
	Productivity Dispersion	Entry intensity	Change in static TFP	Growth g_Q	Change in welfare	Dynamic multiplier	Growth g_Q	Change in welfare	Dynamic multiplier
Initial equilibrium	1.00	0.56	-	2.47	-	-	2.47	-	-
Static Misallocation	0.90	0.64	0.19	2.47	0.64	1.00	2.47	0.19	1.00
No innovation response	0.90	0.64	0.19	2.60	3.92	6.17	2.60	2.75	14.09
Innovation response	0.90	0.64	0.19	2.54	3.03	4.77	2.51	1.41	7.25

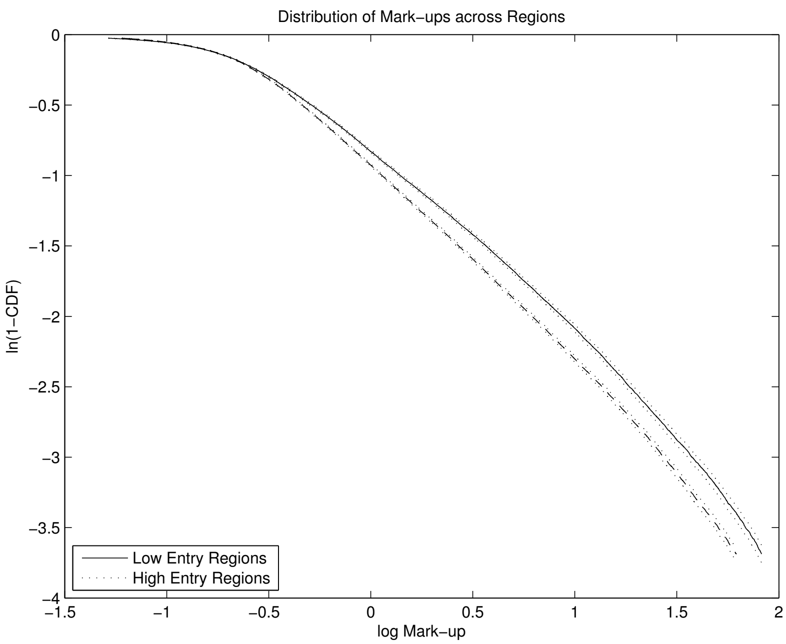
Notes: The first row contains the initial equilibrium given the parameters in Table 8. Row two contains the effects of static misallocation holding the growth rate fixed. The entry intensity in column 3 is the required entry intensity to reduce productivity dispersion (as measured by the standard deviation) by 10%. Column 4 contains the change in M (in percentage terms). Columns 5 and 8 contain the implied growth rate and columns 6 and 9 the full welfare gains (in percentage terms). These are measured in consumption equivalents holding the growth rate fixed, i.e. $exp\left[\rho\left(\hat{W} - W\right)\right]$, where $\hat{W}$ is the welfare level in the respective scenario (see (23)). The dynamic multiplier is the ratio of full welfare gains relative to the static gains. Row three contains the full effects if the innovation rate was not changed. Row four takes the endogenous change of firms' innovation incentives into account.

Table 9: Productivity differences and Welfare: Static and Dynamic Gains

	Variation in consumers' patience (ρ)					
	$\rho = 0.02$			$\rho = 0.10$		
	Growth g_Q	Change in welfare	Dynamic Multiplier	Growth g_Q	Change in welfare	Dynamic Multiplier
Initial equilibrium	2.47	-	-	2.47	-	-
Static Misallocation	2.47	0.53	1.00	2.47	0.73	1.00
No innovation response	2.60	9.95	18.62	2.60	2.13	2.89
Innovation response	2.54	6.60	12.36	2.53	1.60	2.18
	Variation in firms' innovation technology (γ)					
	R&D exp share $s_I = 0.05$			R&D exp share $s_I = 0.10$		
	Growth g_Q	Change in welfare	Dynamic Multiplier	Growth g_Q	Change in welfare	Dynamic Multiplier
Initial equilibrium	2.47	-	-	2.47	-	-
Static Misallocation	2.47	0.64	1.00	2.47	0.64	1.00
No innovation response	2.60	3.92	6.17	2.60	3.92	6.17
Innovation response	2.48	1.64	2.59	2.33	-1.62	-2.55

Notes: This table recalculates the welfare results (see Table 9) for different values of ρ (top panels) and γ (bottom panels). See the description of Table 9 for details.

Table 10: Productivity differences and Welfare: Sensitivity



Notes: The figure shows $ln\left(1-\hat{F}(\mu)\right)$, where $\hat{F}(\mu)$ is the estimated distribution function of mark-ups for "high entry" and "low entry" regions. "High entry" regions are those regions with the highest mean entry rates, which cover 50% of the population of firms (dashed line). "Low entry" regions are those regions with the lowest mean entry rates, which cover the remaining 50% of the population of firms (solid line). The dotted lines around the estimated distributions are 5% confidence intervals. Mark-ups are normalized within 5-digit product-year cells. The entry rate is the rate of entry after taking out a set of year fixed effects.

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

8 Online Appendix

8.1 More general entry process: Derivation of (21)

Suppose that incumbent firms climb one step of the quality ladder with flow rate I. In contrast, entry occurs at a flow rate z but conditional on entry, the new blueprint has a quality advantage of j steps with probability $p(j)$. In that case, the measure of firms having a quality advantage Δ, $\mu(\Delta)$, solves the flow equations

$$\begin{aligned} \dot{\mu}(\Delta, t) &= -(I+z)\mu(\Delta,t) + I\mu(\Delta-1,t) + zp(\Delta) \text{ for } \Delta \geq 2 \\ \dot{\mu}(1,t) &= -(I+z)\mu(1,t) + zp(1). \end{aligned}$$

In the stationary distribution we have $\dot{\mu}(\Delta, t) = 0$, so that

$$\mu(1) = \frac{zp(1)}{I+z} \text{ and } \mu(\Delta) = \frac{zp(\Delta) + I\mu(\Delta-1)}{I+z}.$$

It can then be verified that the expression for $\mu(\Delta)$ given in (21) solves those equations. Now consider the special case where $p(i) = T\kappa^i$. As $\sum p(i) = 1$, we need $T = \frac{1-\kappa}{\kappa}$. For that case, the stationary distribution of productivity gaps Δ (and hence log mark-ups) is given by

$$\mu(\Delta; x, \kappa) = \left(\frac{1}{x+1}\right)^{\Delta} \frac{1-\kappa}{\kappa} \left\{ \sum_{i=1}^{\Delta} (\kappa x)^i \left(\sum_{k=0}^{i-1} \binom{i-1}{k} x^{-k} \right) \right\}, \quad (55)$$

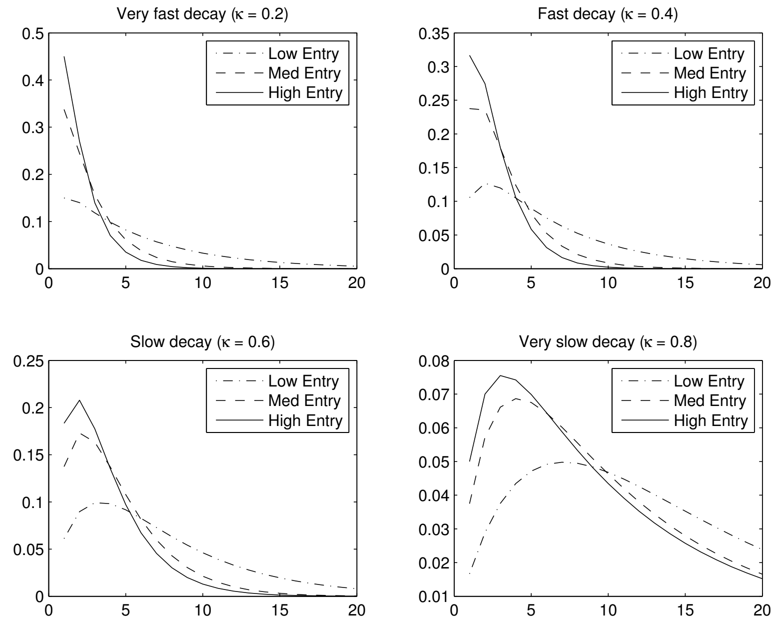
where I explicitly note the dependence on the entry intensity x and the "measure of decay" κ. In Figure 8.1 I depict (55) for different values of κ and x. It is seen that a higher level of the entry intensity induces first order stochastic dominance shifts in the distribution for different values of κ.

8.2 Robustness of main results

Tables 12 and 13 contain various robustness checks for the main results reported in Tables 1, 2, 3, 4 and 5. In particular I use different measures of the output elasticity θ_{it} and also run the results using material productivity (instead of labor productivity) as a dependent variable. Table 12 shows that the cross-sectional pattern of inframarginal rents is not sensitive to these assumptions. Table 13 shows that all these assumption also leave the negative relation between the rate of Entry and the implied distribution of mark-ups intact, but the same specifications are statistically weaker. Table 11 contains robustness results of the life-cycle pattern reported in Figure 1. In particular, I explicitly control for selection by only focusing on firms that survive until the end of the sample and I use different measures of inframarginal rents.

8.3 Attenuation of Entry coefficients

The analysis in section 3.3 showed that empirical the entry rate is only weakly correlated with the dispersion in measured productivity across markets. In Table 14 I report the results of a simple simulation exercise to show that measurement error in the microdata might very well lead to this rejection of the theory given my sample size. My strategy is as follows. To replicate the statistical properties of the regressions in section 3.3, I consider the same number of markets with the empirical distribution of firms within each market. Then I randomly draw (z_m, I_m) for



Notes: The figure shows the implied distribution of mark-ups (55) for different values of the entry intensity x and different values of the decay parameter κ. In particular, I consider $x \in \{0.2, 0.6, 1\}$ and $\kappa \in \{0.2, 0.4, 0.6, 0.8\}$.

Figure 3: Stationary distribution of mark-ups (55)

	log labor productivity ($ln(s_l^{-1})$)			log material productivity ($ln(s_m^{-1})$)	
Age	0.0229***	0.0158***	0.0150***	0.00625***	0.00463**
	(0.00283)	(0.00299)	(0.00321)	(0.00173)	(0.00187)
$ln(\frac{k}{l})$		0.154***	0.153***	0.124***	0.124***
		(0.00387)	(0.00388)	(0.00270)	(0.00271)
$ln(\frac{m}{l})$				-0.366***	-0.366***
				(0.00385)	(0.00385)
Cohort FE	No	No	Yes	No	Yes
N	51692	40759	40759	39671	39671
R^2	0.137	0.186	0.186	0.644	0.644

Notes: This table contains robustness checks for the life-cycle pattern of inframarginal rents reported in Figure 1. Columns one to three (four and five) use labor productivity (material productivity) as a measure of rents. All specifications control for selection in that I only focus on firms that survive until the end of the sample. I pool entry cohorts of the years 1991 to 1994. Hence, All firms survive at least 6 years. In columns 3 and 5 I include a set of cohort fixed effect and hence identify the age coefficient fully from the linear functional form.

Table 11: The life-cycle of inframarginal rents: Robustness

Productivity as mark-ups						
Entrant	Exiter	Age	Positive Growth	Exporter	Profit growth	R&D spending
Specification I: Labor productivity with Cobb Douglas						
-0.0431***	-0.0907***	0.0107***	0.274***	0.0326***	0.0883***	0.216***
(0.00659)	(0.00690)	(0.00154)	(0.00393)	(0.00497)	(0.0115)	(0.0514)
Specification II: Labor productivity with Translog						
-0.00464	-0.0250***	-0.00230	0.249***	0.0455***	0.0638***	0.0459
(0.00700)	(0.00741)	(0.00171)	(0.00406)	(0.00655)	(0.0115)	(0.0542)
Specification III: Material productivity, controlling for material share						
-0.0188***	-0.0583***	0.00172*	0.124***	0.0563***	0.0337***	0.217***
(0.00400)	(0.00454)	(0.000972)	(0.00247)	(0.00334)	(0.00694)	(0.0272)
Specification IV: Material productivity with Translog						
0.0162***	-0.0223***	-0.00684***	0.126***	0.00703**	0.0204***	0.140***
(0.00396)	(0.00450)	(0.000963)	(0.00241)	(0.00335)	(0.00680)	(0.0266)

Productivity as binding input constraints						
FDI	Foreign loans	Capital market	State owned	Does firm face barrier to expand?	Does firm face capital constraint?	Is missing capital the main concern?
Specification I: Labor productivity with Cobb Douglas						
0.305***	0.333***	0.300***	0.266***	-0.123***	-0.0868***	-0.0868***
(0.0218)	(0.0181)	(0.0427)	(0.00487)	(0.0110)	(0.0133)	(0.0133)
Specification II: Labor productivity with Translog						
0.106***	0.119***	0.130***	0.0960***	-0.0838***	-0.0478***	-0.0345
(0.0227)	(0.0185)	(0.0416)	(0.00569)	(0.0110)	(0.0135)	(0.0242)
Specification III: Material productivity, controlling for material share						
0.227***	0.207***	0.116***	0.201***	-0.0722***	-0.0542***	-0.0562***
(0.0105)	(0.00910)	(0.0233)	(0.00330)	(0.00676)	(0.00869)	(0.0139)
Specification IV: Material productivity with Translog						
0.158***	0.129***	0.0683***	0.125***	-0.0537***	-0.0376***	-0.0445***
(0.0104)	(0.00899)	(0.0215)	(0.00337)	(0.00660)	(0.00849)	(0.0137)

Notes: This table contains various robustness checks for the results in Tables 1 and 2. Each row contains to a different specification. Specifications I and II use labor productivity as the dependent variable but model the output elasticity as $\theta_{it} = \delta_s + \delta_t$ (the Cobb-Douglas case of (38)) or $\theta_{it} = \delta_s + \delta_t + \xi_k ln\left(k_{it}\right) + \xi_l ln\left(l_{it}\right)$ (the Translog case of (39)). Specifications III and IV use material productivity ($\frac{py}{m}$) as the dependent variable, where m is total spending on materials) and model the output elasticity as $\theta_{it} = \delta_s + \delta_t + \xi_m ln\left(\frac{m}{l}\right) + \xi_l ln\left(\frac{m}{l}\right)$ (see (38)) or $\theta_{it} = \delta_s + \delta_t + \xi_k ln\left(k_{it}\right) + \xi_l ln\left(l_{it}\right) + \xi_m ln\left(m\right)$ (the Translog case of (39)).

Table 12: Determinants of labor productivity: Robustness

	Avg mark-up	Avg mark-up	25%	50%	75%	Std. dev	Λ	M
				Specification I: Labor productivity with Cobb Douglas				
Entry	-0.117**	-0.157***	-0.137***	-0.145***	-0.0734	0.0265	0.0841***	0.0900***
	(0.0439)	(0.0394)	(0.0410)	(0.0491)	(0.0771)	(0.0406)	(0.0261)	(0.0269)
				Specification II: Material productivity, controlling for material share				
Entry	-0.0347	-0.0458*	-0.0857**	-0.0293	-0.0132	0.0648**	0.0843***	0.0835***
	(0.0361)	(0.0267)	(0.0348)	(0.0304)	(0.0417)	(0.0292)	(0.0203)	(0.0225)
				Specification III: Baseline with at least 20 firms per market				
Entry	-0.109**	-0.141***	-0.154**	-0.137**	-0.105	0.0257	0.0966***	0.0791***
	(0.0518)	(0.0439)	(0.0596)	(0.0588)	(0.0734)	(0.0486)	(0.0271)	(0.0249)
				Specification IV: Baseline with 4-digit product markets				
Entry	-0.0829**	-0.0961***	-0.118***	-0.0906***	-0.0402	0.057	0.0499***	0.0524**
	(0.0306)	(0.0282)	(0.0344	(0.0306)	(0.0306)	(0.0434)	(0.0141)	(0.0202)
				Specification V: Material productivity, controlling for material share with 4-digit product markets				
Entry	-0.0258	-0.0373**	-0.0544***	-0.0379**	-0.00181	0.0397**	0.0527***	0.851
	(0.0197)	(0.0170)	(0.0186)	(0.0169)	(0.0241)	(0.0182)	(0.0148)	(0.824)
				Specification VI: Baseline using Exit as the independent variable				
Exit	-0.148**	-0.126**	-0.118*	-0.141**	-0.183**	-0.0756*	0.0439	0.0936**
	(0.0633)	(0.0565)	(0.0621)	(0.0664)	(0.0771)	(0.0370)	(0.0283)	(0.0408)
Regional Controls	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	No	No	No	No	No

Notes: This table contains various robustness checks for the results in Tables 3, 4 and 5. Each row contains a different specification and I report only the coefficient on the markets' entry rate. Specification I uses labor productivity $\left(\frac{py}{wl}\right)$ as the dependent variable and models the output elasticity as $\theta_{it} = \delta_s + \delta_t$ (the Cobb-Douglas case of (38)). Specification II uses material productivity $\left(\frac{py}{m}\right)$ and models the output elasticity as $\theta_{it} = \delta_s + \delta_t + \xi_m ln\left(\frac{m}{l}\right) + \xi_l ln\left(\frac{m}{l}\right)$ (see (38)). Specification III contains the result of the baseline specification but drops all markets with less than 20 firms. Specification IV contains the result of the baseline specification but defines the product market as a 4-digit industry, so that a market is a year-region-4-digit-product-cell. Specification V uses materials (as in Specification II) but defines a product market as a 4-digit industry (as in Specification IV). Specification VI finally uses the baseline sample but uses the rate of Exit instead of the rate of Entry.

Table 13: Mark-ups and Entry: Robustness

	Relative scale of measurement error						
	0	1	2	3	4	5	6
	Dep. Variable: Avg log mark-up						
Entry	-1.6959	-1.5127	-1.7225	-1.8113	-1.7427	-1.4511	-1.6608
	(0.0471)	(0.0512)	(0.0727)	(0.0969)	(0.1248)	(0.1575)	(0.1778)
	Dep. Variable: standard deviation of log mark-ups						
Entry	-1.1632	-0.8359	-0.5674	-0.4637	-0.4506	-0.3762	-0.1641
	(0.0399)	(0.0438)	(0.0601)	(0.0784)	(0.1024)	(0.1268)	(0.1496)

Notes: See main body of the text for a description.

Table 14: Estimating the effect of entry: The importance of measurement error

each market m[51] , construct $\vartheta(x_m) = \frac{ln(1+z_m/I_m)}{ln(1+\lambda)}$ and draw mark-ups in each market from a pareto distribution with the market-specific tail ϑ. Then I add some measurement error drawn from a lognormal distribution with variance υ. I chose υ such that υ is a multiple of the implied variance of mark-ups and I report the analysis for different values of υ. Then I construct the average and dispersion of log measured mark-ups and regress these moments on the entry intensity. The results are contained in Table 14. In the top panel I show that in the case of the average log mark-up the estimated coefficient is unaffected by an increase in the underlying measurement error and that merely the standard error increases. In the lower panel I show that this is different for the dispersion. As the classical measurement error in the micro-data does not aggregate into classical measurement error for the market data, the Entry coefficient is biased downward. This bias becomes more severe the bigger the measurement error relative to the model-induced variation.

[51] I draw z from a uniform distribution between 0.05 and 0.3 and I from a uniform distribution with support 0.15 and 0.9.