Growth and the plant size distribution over the long-run

Evidence from Indonesia

Oscar Fentanes[*], Jonas Gathen[‡]

15th February 2022

Abstract

In this paper, we study the link between rapid economic growth and changes in the plant size distribution. We draw on 40 years of panel data on Indonesian manufacturing plants to show that as output grew by a factor of twenty-five, the average plant size doubled and very large plants strongly increased their shares of output and employment. To link these facts, we use a statistical accounting exercise to decompose output growth into aggregate productivity growth, input growth and selection and reallocation effects. We show that the direct contribution of aggregate productivity for output growth in Indonesia is close to zero. To explain output growth in the absence of aggregate productivity growth, we build a structural model of plant size dynamics that flexibly captures an endogenous co-evolution of aggregate output growth and changes in the plant size distribution. We directly estimate the model on the plant-level data and show that it fits remarkably well untargeted moments such as the average size and output evolution as well as the entire plant size distribution over time. In the model, output growth is driven by endogenous transition dynamics and we validate this choice via a novel empirical exercise that shows the importance of transition dynamics in Indonesian manufacturing. Growth driven by transition dynamics plays out over a long period of time because (1) it takes time to grow large plants and (2) we find that many plants are far from their optimal size.

[*]Toulouse School of Economics (TSE), Email: oscar.fentanes@tse-fr.eu

[†]Toulouse School of Economics (TSE), Email: jonas.gathen@tse-fr.eu.

[‡]We are grateful for many useful comments by Fabrice Collard, Christian Hellwig, Ben Moll, Yongseok Shin, Francisco Buera, Eugenia Gonzalez-Aguado, Miguel Zerecero, Andreas Schaab and all participants of the TSE Macro workshop. For feedback and comments on some of the econometric exercises we want to especially thank Jihyun Kim, Nour Meddahi, Thierry Magnac and Jeffrey Racine. We want to thank the World Bank Jakarta team, especially Massimiliano Calí, for gratefully sharing the manufacturing plant-level data for Indonesia. We further thank participants at the Annual STEG conference 2022, ENTER jamboree 2021 as well as the PhD workshop at TSE. All remaining errors are our own.

1 Introduction

Over the last several decades, numerous developing countries around the world have experienced long periods of rapid economic growth that were historically unprecedented both in speed and length. At the same time, there is good evidence that economic growth is accompanied by important changes in the entire firm and plant size distributions. While developing countries are characterized by an abundance of small firms and a lack of medium and large sized firms (Hsieh and Olken 2014)[1], there is a strong positive correlation between income per capita and features of the firm size distribution such as dispersion, the size of the right tail and average firm size (Poschke 2018). There is also good evidence that this holds over the course of development within countries as documented for a variety of countries such as the US (Poschke 2018), Japan (Gollin 2008) and Portugal (Cabral and Mata 2003). As countries develop, larger firms and plants become more important, holding an increasingly larger fraction of total employment. However, less is known about the dynamic interaction between the two.

In this paper, we systematically study the link between plant size dynamics and economic growth over the course of development. To this end, we build on detailed plant-level Indonesian manufacturing panel data over a time span of 40 years (1975-2015).[2] Using this exceptionally long-run micro data, we start with a growth decomposition using the following accounting identity:[3]

$$\%\Delta\ \text{GDP} \equiv \%\Delta \begin{array}{c}\text{aggregate}\\ \text{productivity}\end{array} + \%\Delta \begin{array}{c}\text{total}\\ \text{inputs}\end{array} + \%\Delta\Big[\underbrace{\begin{array}{c}\text{avg plant}\\ \text{productivity}\end{array}}_{\text{selection}} \underbrace{+ N * Cov\left(\begin{array}{c}\text{plant}\\ \text{productivity}\end{array}, \begin{array}{c}\text{share of}\\ \text{inputs}\end{array} \right)}_{\text{reallocation}} \Big] \quad (1)$$

We derive this accounting identity starting from the fact that economic growth is simply the product of aggregating up the heterogeneous growth rates in value added output across all plants (N) in the economy, a standard assumption on the separability of productivity in plant production functions and the decomposition based on Olley and Pakes (1996). We quantify the contribution of each of these terms given observed revenue and inputs and estimation of the production function parameters. As is standard, we then back out plant-level total factor productivity, commonly referred to as TFPQ in our setup (see Hsieh and Klenow 2009). In line with models of aggregate productivity growth, we assume plant-level TFPQ is the product of plant-specific productivity and a common productivity term across all plants that we call aggregate productivity.

[1]There is recent work by Buera et al (2021) that shows markedly different results to Hsieh and Olken (2014) in Indian agriculture. Specifically, that Indian agricultural plants show a missing middle (plants between 10-40 workers) and that large plants are more common in agriculture than in the rest of the economy. In this paper, we focus on manufacturing and do not speak to the plant size distribution in agriculture.

[2]Our data is at the plant-level and doesn't allow to identify firms, hence we focus on plants in this paper.

[3]The exact accounting identity and the underlying assumptions are given in Section 2.3.

We identify changes in aggregate productivity from within-plant changes in TFPQ, additionally dealing with two important biases that in practice lead to the overestimation of aggregate productivity growth: survivor bias (less productive plants are more likely to exit) and that young plants may not be at their ergodic distribution of productivity (as would be the case if they need to build up demand (e.g. Foster, Haltiwanger, and Syverson 2016)). We deal with both biases by focusing on within-plant changes in TFPQ for a selected set of older plants for whom exit is unlikely and within-plant, idiosyncratic productivity should not change in distribution.[4]

Based on the decomposition exercise, aggregate productivity growth contributed less than 4% to the roughly 25-fold increase in output over the time period 1975-2015. Instead, we find that all output growth over the period 1975 to 2015 is driven by input growth (2/3) and by selection and reallocation effects (1/3). Furthermore, selection and reallocation effects contribute in almost equal shares, while the positive reallocation effect is entirely captured by a mechanical increase in the number of plants; the contribution of the pure reallocation effect is actually negative in the data.

The statistical decomposition exercise provides systematic insights into the direct drivers of economic growth, but neither explains how to think about input growth, selection and reallocation nor quantifies potential indirect effects. To this end, we are interested in a structural model of plant dynamics that flexibly captures various drivers of aggregate economic growth and lets us quantify their respective contributions. In the absence of a large role for aggregate productivity growth, we conjecture that growth must either come from other policy changes that lead plants to grow or from initial conditions that hold the seeds for future growth but simply need time to unfold (what we call transition dynamics throughout). We find strong evidence for the role of transition dynamics, drawing on two pieces of evidence. First, based on our previous production function estimates that allow the distribution of plants to be far away from a stationary distribution, we find that plants are systematically below their optimal plant size. This provides a large source of potential future growth even in the absence of any policy changes.

Second, we conduct a reduced-form exercise that fixes the policy environment at the start of the data and asks how predictive future growth and plant dynamics are. Measuring plant size by the number of workers, we show that 70-90% of future changes in the entire plant size distribution are predictable based only on the initial plant size distribution in 1975 and observed within-plant changes in size between 1975 and 1976. The key reason for this is that conditional on plant size, average growth in plant size is relatively stable over

[4] This identification strategy holds across a variety of different models where TFPQ can be identified and idiosyncratic productivity is close to an ergodic distribution for older plants. For example, this holds in all models where idiosyncratic plant productivity follows an exogenous Markov process and plants enter with an arbitrary initial distribution. It also holds in models where idiosyncratic productivity is endogenous, but reaches an ergodic distribution for older plants, as for example in models where plants build up demand (Foster, Haltiwanger, and Syverson 2016). By "change in distribution" we mean that moments of the distribution do not change, while plants' productivities may arbitrarily change across periods within this distribution.

time: plants with 100 workers grow their number of workers at roughly similar rates over time. However, the key to understand the long-run evolution of the plant size distribution is that Indonesia started out with many more small and medium-sized plants and fewer large plants. In the data, the average plant size doubles between 1975 to 2015 and this is mostly driven by an increase in the right tail of the distribution. The empirical exercise shows that it simply takes time to grow large plants and these large plants will eventually become more important overall. Technically, the initial plant size distribution is far away from the stationary plant size distribution implied by the transition matrix from 1975 and 1976. Conditional on observing transitions in 1975 and 1976 and the initial distribution, we can predict the transition towards the stationary distribution and show that it takes more than 40 years.

Based on these insights, we build a structural model of plant size dynamics that flexibly captures an endogenous co-evolution of aggregate output growth and changes in the plant size distribution. The model replicates and fully endogenizes these plant size transition dynamics. While the initial distribution of plants is directly estimated in the data, slow plant growth that is observed in the data is explained by sizable convex factor adjustment costs in the model, preventing plants from growing large quickly. These adjustment costs - directly estimated on micro data - are large enough to feature slow aggregate transitions that play out over many decades. They likely capture a variety of real-world frictions that prevent plants from growing quickly. These adjustment costs are entirely in line with systematic evidence of slower life-cycle growth of plants in developing countries (Hsieh and Klenow 2014). Together with evidence on differences in initial distributions, adjustment costs explain the lack of large plants in developing countries (Hsieh and Olken 2014). The flipside of slowly growing plants in the model is that the supply of factors of production slowly grows over time and that they are endogenously reallocated across sectors featuring manufacturing and a rest-of-the-economy sector that allows to endogenize structural change. Specifically, our model replicates that the manufacturing employment share doubles and that the total labor supply roughly triples over the period 1975-2015. The main benefit of the structural model is to allow a full decomposition of the drivers of aggregate growth, capturing both direct and indirect effects and quantifying how frictions influence plant size distributions and aggregate growth dynamically over time. We validate with the model that aggregate productivity growth in manufacturing plays no role for aggregate growth in Indonesia. We further find a large role for convex adjustment costs:[5] abolishing these costs would lead to three times larger manufacturing output over the period 1975-2015. 1/3 of these output gains would be driven by input growth, translating into a three-times larger manufacturing employment share due to endogenous structural change. The remaining 2/3 are driven by reallocation effects as resources are shifted away from less productive and towards

[5]The following quantitative results should still be seen as preliminary and may change in future versions of the paper as we make changes to the model.

more productive plants. On the other hand, a large growth in the supply of efficiency units of labor, due to a combination of population growth and a better educated workforce, made widespread plant growth possible, but prevented stronger reallocation across plants by keeping wages fairly constant.

Literature

Our paper contributes and speaks to a number of literatures. Foremost, our paper relates to a large firm dynamics and misallocation literature that has considered firm-specific frictions and distortions in driving differences in firm size distributions (e.g. Clementi and Palazzo 2016; Hsieh and Klenow 2009; Restuccia and Bento 2015; Restuccia and Rogerson 2008). Whether considered distortions enter statically (e.g. Hsieh and Klenow 2009; Restuccia and Rogerson 2008) or dynamically (e.g. Asker, Collard-Wexler, and De Loecker 2014; Buera, Kaboski, and Shin 2011; David and Venkateswaran 2019), both the misallocation and firm dynamics literature have mostly abstracted from the study of economic growth and focused on stationary distributions. Important exceptions are Moll (2014), Midrigan and Xu (2014), Song, Storesletten, and Zilibotti (2011), Buera and Shin (2013), and Buera and Shin (2017). In contrast to Moll (2014), Buera and Shin (2013), and Buera and Shin (2017), our setup allows a much closer mapping between data and theory over the transition period. This is because we directly estimate the initial distribution on micro data and because we directly estimate model parameters over the transition period on micro data, avoiding assumptions on the initial distribution and having to ever observe a stationary economy. In contrast to Midrigan and Xu (2014), we do not rely on a balanced growth path assumption.

Our paper adds to a small growth literature that has looked at the interaction between aggregate productivity and endogenous transition dynamics due to heterogeneous production in the economy. For example, Song, Storesletten, and Zilibotti (2011) showed how China's decades of growth can be understood through a model where resources are slowly reallocated from less productive state-owned companies to more productive private firms. The closest paper to ours is Poschke (2018) who builds a model of technological change and entrepreneurial selection over the course of development that can explain the coevolution of aggregate growth and observed changes in the average firm size and dispersion both within and across countries. Similarly to our paper, Poschke (2018) gives a story of fully endogenous changes in the firm size distribution that do not rely on policy distortions. In contrast, the focus in our paper is to disentangle the role of aggregate productivity growth and transition dynamics and estimating the model directly on micro data. At this point, we do not endogenize entry and do not directly model entrepreneurial choices; Poschke (2018) is complementary in this regard.

More generally, this is to the best of our knowledge the first paper to consider stochastic long-run transition dynamics. Previous papers in macro development that have looked at transition dynamics have focused on deterministic transition dynamics (e.g. Buera and Shin 2013, 2017; Moll 2014). In our setup, deterministic transition dynamics would imply that households and establishments in 1975 would have had perfect foresight over the Asian Financial Crisis in 1998 and the subsequent economic recovery. Given this important crisis, we instead consider a standard business cycle setup with aggregate uncertainty (e.g. Clementi and Palazzo 2016). The reason we can study this in detail is that transition dynamics in our setup stay highly tractable because (1) we make more use of direct estimation and (2) only the plant size distribution is nonstationary while prices and aggregate productivity seem to be stationary. In terms of direct estimation, we estimate aggregate productivity in the data and make use of the fact that we can observe one path of equilibria in the data. This allows us to directly estimate the process of aggregate productivity (which we find is well-described by a persistent stationary process). We then solve for rational expectations competitive equilibria following the standard algorithm based on Krusell and Smith (1998). To save computational costs, we again use the single path of observed equilibria to directly estimate a plant forecasting rule for prices. This is very similar in idea to tractably solving perfect foresight equilibria taking observed prices as given as in Gopinath et al. (2017) or Caliendo, Dvorkin, and Parro (2019). To ensure that model-implied prices are in line with observed equilibria and consistent with the estimated forecasting rule, we introduce a policy wedge that may change over time and is informative about model misfit over time, similar in spirit to classic wedge-accounting approaches.

Theoretically, the importance of the initial distribution and slow transition dynamics in our model are driven by convex labor adjustment costs as in Cooper, Haltiwanger, and Willis (2015). This is similar in spirit to alternative frictions considered in the literature such as random growth processes (Luttmer 2011), capital adjustment costs (Asker, Collard-Wexler, and De Loecker 2014; Clementi and Palazzo 2016), financing constraints (Buera and Shin 2013; Moll 2014) as well as having to build consumer demand (Foster, Haltiwanger, and Syverson 2016) and learning dynamics (Atkin, Khandelwal, and Osman 2017; Berman, Rebeyrol, and Vicard 2019; Ruhl and Willis 2017). We do not try to disentangle different frictions in our paper, but contribute to the literature by considering how one specific friction plays out dynamically over the course of development and how it interacts with aggregate productivity growth. (Discuss how results relate to labor adj cost literature)

At last, our paper relates to a literature on the decomposition of productivity and output. Our paper directly builds on Olley and Pakes (1996) for decomposing a weighted average of idiosyncratic productivity across plants into a selection and reallocation effect. In this regard, it also directly relates to further statistical decomposition exercises in the literature (e.g. Melitz and Polanec 2015). These statistical decomposition exercises have been criticized for not being able to show causality, not capturing indirect effects and giving misleading interpretations of aggregate productivity (e.g. Baqaee and Farhi 2020). In our paper, we show how structural modeling and statistical decomposition exercises can be used in combination, both to interpret model findings and inform modeling choices, combining the good of both worlds: using accounting identities to focus on what variation in the data is the most important and using the structural model to infer causality.

The rest of the paper is structured as follows. In section 2 we present the main empirical evidence. Section 3 builds on this empirical evidence to build a full structural model that features an endogenous coevolution of aggregate growth and changes in the entire plant size distribution, shows how this model is identified from micro data and shows that it captures well the data. In Section 4, we use the structural model to formally decompose aggregate growth into its main drivers and reveal which policies may lead to large output gains. The last section concludes.

2 Empirical evidence

In this section, we introduce the data and key facts about aggregate growth and the evolution of the plant size distribution in Indonesia. We then go through our main growth decomposition exercise as well as further empirical exercises to reveal the main drivers of aggregate growth and how it relates to changes in the plant size distribution.

2.1 Data

Our primary data comes from the plant-level Annual Manufacturing Survey (Survei Tahunan Perusahaam Industri Pengolahan), collected by Indonesia's Central Bureau of Statistics (Badan Pusat Statistik). It covers only medium- to large-sized manufacturing plants by surveying all formal manufacturing establishments with more than 20 employees. The survey contains detailed and consistent information on industry, employment, production and other plant characteristics from 1975 to 2015, spanning more than 40 years. It covers between roughly 6,500 to more than 25,000 plants per year. Using information based on a random five percent sample of all manufacturing establishments from the Indonesian Economics Census in 2006 reported in Hsieh and Olken (2014), 99% of Indonesian manufacturing plants have less than 20 workers, which our dataset entirely misses. On the other hand, most of the plants we miss are very small so that - based on the GGDC 10-sector database for Indonesia, which captures time-consistent aggregate sectoral employment and output series over the time period 1960-2012 (Timmer, de Vries, and De Vries 2015) - our manufacturing data captures about

30% of total manufacturing employment as well as about 30% of total manufacturing value-added output.[6] Given the data limitation on small plants, the focus of this paper is on the link between aggregate economic growth and relatively large plants and changes in their size distribution. In the model part, we show how we can generalize our results to the aggregate economy and in the robustness section we discuss how sensitive our results are to this data limitation.

For the empirical parts below, we draw on a yearly measure of the number of workers at a plant, which we take as the sum of total reported paid and unpaid workers. We further use the plant's yearly total wage bill (production + non-production wage bill), reported value-added output and reported per worker wages for production workers and non-production workers. Value-added output is not reported in 1976-1978, which forces us for both the growth decomposition exercise and the structural model to restrict ourselves to the time period 1979-2015. All variables denoted in Indonesian Rupiah are deflated to real values using the aggregate CPI.

2.2 Growth and changes in the plant size distribution over the long run

Figure 1 reports the evolution of employment and output at the aggregate and manufacturing-sector level in Indonesia over the period 1975-2015. Based on the GGDC 10-sector database, aggregate GDP in Indonesia grew by a factor of more than fourteen between 1975 and 2012, the working population increased by a factor of more than 2.5 and hence GDP per worker increased more than 5-fold. Manufacturing contributed importantly to this aggregate growth: manufacturing output grew 25-fold and employment reallocated majorly to manufacturing as shown by a 250% increase in the manufacturing employment share.

On the other hand, the manufacturing sector underwent large changes in its plant size distribution over the same time period. We start out by plotting the average plant size over time in Figure 2 Panel A, where size is simply measured by the number of workers. The average number of workers in medium and large manufacturing plants in Indonesia roughly doubles between 1975 to 2015.[7] Figure 2 Panel B shows the

[6]Output and employment shares turn out to be very similar over time, indicating that relative labor productivity moves little over time. We return to this finding later in the paper. Furthermore, we verify that our micro-data is consistent with capturing all manufacturing plants with more than 20 workers. For example, based on the 2006 census sample, manufacturing plants with more than 50 workers should capture 34% of total manufacturing employment, while this figure is 32% based on the aggregate sectoral employment from the GGDC 10-sector database and employment in our micro-data. Given that the manufacturing plant panel includes new plants based on the Economics Census coverage is more complete after Economics Census years. We come back to this point in the next section. Furthermore, coverage likely increased over time. We find that the micro-data manufacturing plants account for less than 20% of plants in the first 10 years after 1975 and that this share rises quickly above 30% in the 1990s and 2000s. This is in line with an improvement in the coverage of plants with more than 20 workers and - due to the jumps in coverage - is probably less driven by an increasing importance of larger plants.

[7]The series further shows substantial fluctuations with larger drops in 1985, 1995 and 2006. The reason for this is that these years are years of the Economic Census in which the plant panel was filled up with plants that were previously missed. These plants are mostly small plants, either because they are young (and thus are newly recorded in the census), because they are small (close to the cutoff of 20 workers) or both.

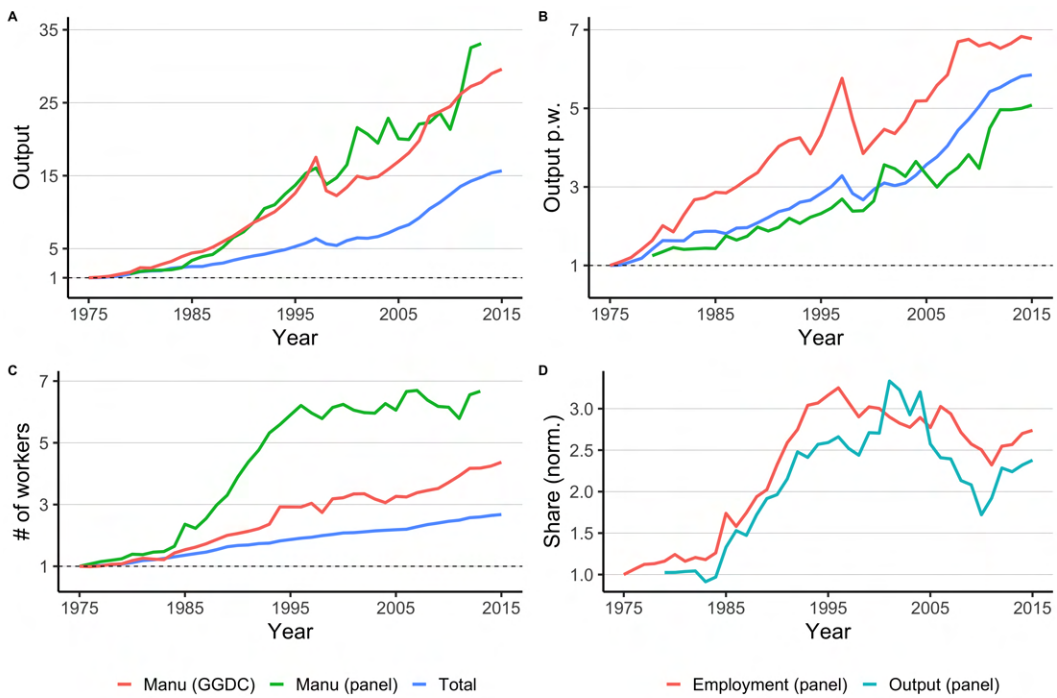

Figure 1: Evolution of aggregate and sectoral employment and output. Economy-wide data as well as one of the manufacturing series is based on the GGDC 10-sector database (1975-2012). Panel data refers to the manufacturing plants in the Indonesian manufacturing plant census (1975-2015). All series across all panels are normalized by their respective value in the first year to highlight increases over time. Panel A plots the increase in aggregate GDP and total manufacturing value-added output. Panel B gives the same as per worker versions. Panel C reports the evolution of the number of workers. Panel D gives the employment and output shares taking as totals the aggregates from the GGDC 10-sector database and the aggregates for manufacturing from the plant census.

increase in the number of plants in the panel over time. One can clearly see that the large drops in the average size of plants coincide with the inclusion of new and smaller plants around 1985, 1996 and 2006. Thus, the data gives a census of manufacturing plants with more than 20 employees around the time of the Economic Census, but fails to be a complete census in the intermediate time periods as not all new plants are included over time. However, the long-run picture of increasing average plant size and entering plants that are smaller clearly stands out, despite year-to-year fluctuations being partly driven by specific attributes of the survey sampling.

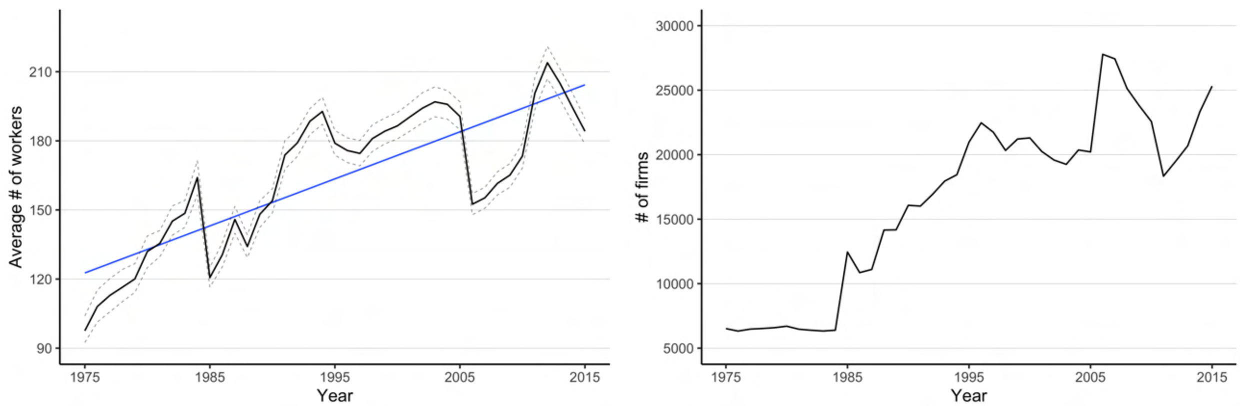

Figure 2: Panel A (left): Evolution of average plant size measured in number of workers (paid + unpaid) for medium- and large Indonesian manufacturing plants. Solid black line gives point-estimate, grey dotted lines give 95% bootstrapped confidence intervals and solid blue line gives best linear fit. Panel B (right): Evolution of number of medium- and large Indonesian manufacturing plants in the panel. Data: Based on Statistik Industri, the Indonesian manufacturing plant census capturing plants with 20 or more workers. Restricting sample to plants with at least 20 workers (enforce cutoff across survey years) and less than 15k workers. Total of 88,282 unique plants, ranging from 6,797 in 1975 to 26,311 in 2015.

Next, we consider changes in the entire plant size distribution in Figure 3. A number of features stand out. First, the plant size distribution has a heavy right tail across time, clearly visible under the log-scale in Panel A. Secondly, the plant size distribution seems to clearly move to the right. This is most clearly visible by focusing on the evolution of right tail moments as shown in Panel B. The plant at the 75th percentile almost doubles its size of 67 workers in 1975 to 125 workers in 2015. At the 90th percentile, plants more than double their size from around 175 in 1975 to around 400 in 2015. This increasingly heavier right tail is also what drives the increase in the average plant size. At last, the dispersion of the plant size distribution clearly increases over time as the 25th percentile stays roughly constant around 26 workers and the higher percentiles move further to the right.

This increasing importance of large plants over time can more clearly be seen by focusing on their total employment share. Figure 4 shows yearly employment shares in large plants by computing for each year the total employment in plants of a certain size over total employment across all plants in the panel. The

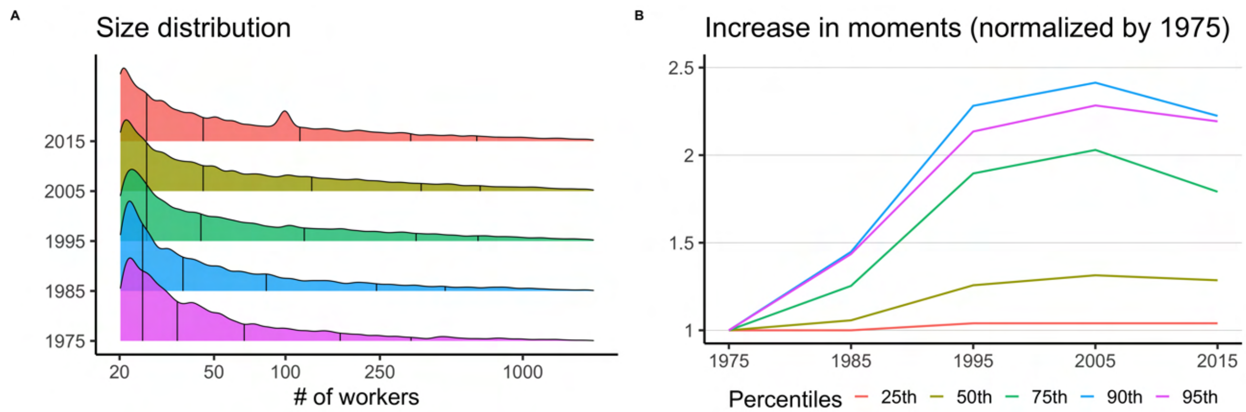

Figure 3: Evolution of distributions of total number of workers (paid + unpaid) for medium- and large Indonesian manufacturing plants. Panel A (left) gives the evolution of the entire size distribution, vertical lines give 25th, 50th, 75th, 90th and 95th percentiles. Panel B (right) plots separately the evolution of the same percentiles, normalizing each by their value in 1975. Data is based on Statistik Industri, the Indonesian manufacturing plant census capturing plants with 20 or more workers. Restricting sample to plants with at least 20 workers (enforce cutoff across survey years) and less than 15k workers as well as to data from 1975, 1985, 1995, 2005 and 2015.

employment share in plants above 500 workers increases steadily from roughly 40% to more than 60% of total employment. Similarly, for plants with at least 1,000 workers, the employment share almost doubles from around 23% to around 44%. Large plants become much more prevalent and important over time and are largely responsible for the increasing average size of manufacturing plants in Indonesia.

2.3 Decomposing aggregate growth

To better understand Indonesia's aggregate growth over the last 40 years and how changes in the plant size distribution contributed to this growth, we start out by formally decomposing aggregate manufacturing growth. This statistical exercise does not show causality, but it does reveal which features in the data are crucial to explain aggregate growth. Note that by definition, total output Y_t is simply the sum of value added output y_{it} across all plants in the economy:

$$Y_t \equiv \sum_{i \in \Omega} y_{it}$$

Next, we assume a standard plant production function of the following form:

$$y_{it} = z_t * s_{it} * f(x_{it}) \tag{2}$$

where z_t captures aggregate productivity, which is usually assumed to capture the aggregate state of technology in the economy as well as the institutional environment and the stock of public goods. s_{it} in turn

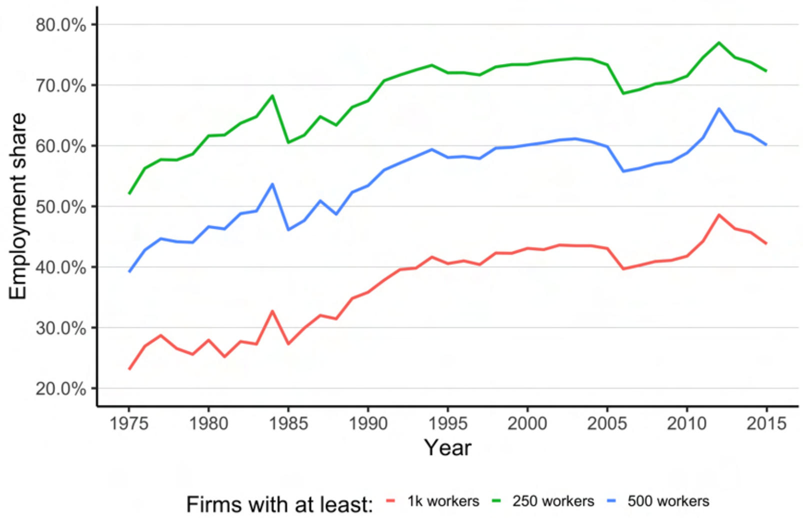

Figure 4: Evolution of employment share in large Indonesian manufacturing plants. Based on Statistik Industri, the Indonesian manufacturing plant census capturing plants with 20 or more workers. Restricting sample to plants with at least 20 workers (enforce cutoff across survey years) and less than 15k workers. Total of 88,282 plants, ranging from 6,797 in 1975 to 26,311 in 2015. Details on the variable construction in the text.

captures plants' idiosyncratic productivity. At last, $f(x_{it})$ is an increasing function in inputs x_{it}, which could for example be a vector of labor and capital. Given this general setup, we make us of the decomposition in Olley and Pakes (1996) to rewrite total output as:

$$lnY_t \equiv lnz_t + ln\sum_i f(x_{it}) + ln\left[\bar{s}_t + N_t cov\left(s_{it}, \frac{f(x_{it})}{\sum_i f(x_{it})}\right)\right] \tag{3}$$

where N_t tracks the number of active plants. Total output is the combination of aggregate productivity, the state of factor accumulation and a combination of aggregate productivity and a covariance term that captures whether resources in the economy are allocated towards the most productive plants. This last decomposition of a productivity term into an average and a covariance term is based on Olley and Pakes (1996). In comparison to the decomposition literature based on Olley and Pakes (1996), the only difference of our approach is to further decompose plant productivity into an aggregate and an idiosyncratic component and to focus on aggregate output instead of aggregate productivity. Growth in aggregate output is then given by:

$$\Delta lnY_t \equiv \Delta lnz_t + \underbrace{\Delta ln\sum_i f(x_{it})}_{\text{input growth}} + \underbrace{\Delta ln\left[\bar{s}_t + N_t cov\left(s_{it}, \frac{f(x_{it})}{\sum_i f(x_{it})}\right)\right]}_{\text{selection + reallocation effect}} \tag{4}$$

To quantify each term's contribution to growth in aggregate output, we need to separately identify each component of the plant-specific production function. First, we assume that plants produce with the same production function $f(x_{it}) = l^{\theta}$ using labor as the only production input. The reason for not considering capital is that we have a consistently defined labor series from 1975-2015, while capital series only start in 1990 and capital stocks are notoriously difficult to measure. We estimate the output elasticity θ allowing for labor to be fully dynamically chosen. As the case of fully dynamically chosen inputs is not usually considered in the production function literature (e.g. Ackerberg, Caves, and Frazer 2015; Olley and Pakes 1996), we propose a novel estimator that we discuss in more detail in the structural estimation part and in the Appendix. The idea is that in general settings where plants observe today's productivity but face dynamic frictions when making input choices (e.g. adjustment costs or financing frictions), the output elasticity in combination with first-order conditions makes sharp extensive margin predictions on whether plants should increase or decrease their inputs compared to the previous period. The estimator we propose is agnostic about the actual dynamic frictions and estimates θ only based on correctly classifying whether plants increase or decrease their number of workers. Based on this estimator, we find that $\hat{\theta} \approx 0.58$, considerably larger than the average labor share (≈ 0.45). Given $f(x_{it})$ and observable output y_{it}, total factor productivity is identified by: $TFP_{it} \equiv s_{it}z_t = \frac{y_{it}}{f(x_{it})}$. We estimate TFP using reported plant-specific value-added revenue and efficiency units of labor as based on the reported wage bill divided by an estimate of the average real wage in the economy.[8] This is TFPQ in the setup by Hsieh and Klenow (2009).

To separately identify aggregate productivity z_t from plant-specific idiosyncratic productivity, note that:

$$\Delta ln(TFP_{it}) \equiv log(s_{it}) - log(s_{it-1}) + log(z_t) - log(z_{t-1})$$

To identify changes in z_t, we solely require a subset of plants for whom on average changes in idiosyncratic productivity are approximately zero. Naively, we could take all plants for whom TFP is identified in period t and period $t-1$, compute changes in TFP for each of these plants and average across all plants to obtain an estimate of the change in aggregate productivity z_t. After normalizing the initial value of z_t to unity, we plot the resulting aggregate productivity series termed "Naive (within)" in Figure 5. As a comparison, we also plot the fully naive estimate that compares only average TFP across plants over time, highlighting the importance of using panel data. The naive within estimator is still flawed for three key reasons. First, the average within-plant change in idiosyncratic productivity is unlikely to be zero for this set of plants due to survivor bias: the naive estimator only draws on surviving plants, while exiting plants are more likely to have received negative idiosyncratic productivity shocks, leading to a positive selection on productivity. Survivor bias leads the naive estimator to overestimate the role of aggregate productivity z_t. The second reason for why the naive estimator produces biased results is that even without exit, plants may not be at their ergodic distribution of idiosyncratic productivity. For example, young plants may start with lower productivity and

[8] We discuss the identification and estimation of the wage series in the model part.

need time to build up productivity. Again, this leads to overestimates of aggregate productivity as long as the average of idiosyncratic productivities is below the average of the ergodic distribution. At last, idiosyncratic and aggregate productivity may be systematically correlated. In the case of positive correlation, this again leads to overestimates of aggregate productivity.

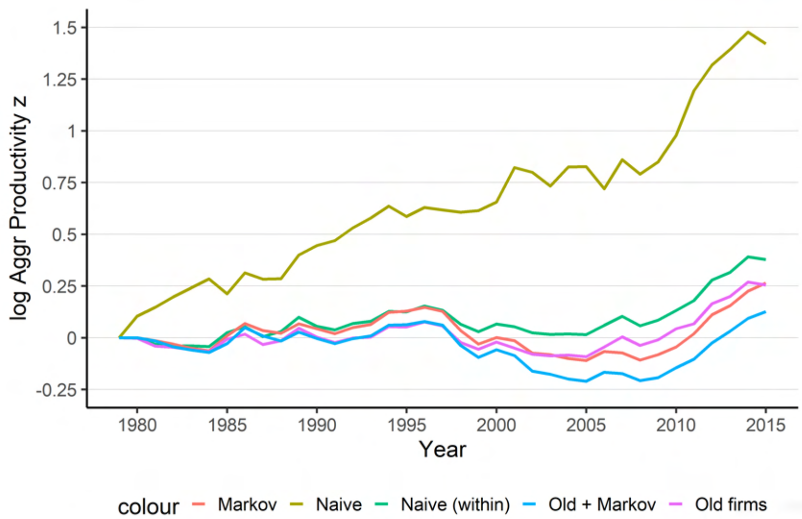

Figure 5: Evolution of aggregate productivity (in logs) based on flat-spot identification for older plants. Cutoff age for older plants is 7 years or older. Details on each of the estimator are in the text. The preferred estimator used throughout is based on Old+Markov, which restricts estimation to old plants and plants whose productivity grew positively the period before. Plant-level data based on Statistik Industri. 33,904 unique plants and 296,202 plant-year observations.

To partially solve these concerns, we focus on older plants. Clearly, older plants are a selected sample of plants for whom average idiosyncratic productivity may differ systematically from that of younger plants in the economy (e.g. they survived longer). However, our approach allows for arbitrary differences in levels of idiosyncratic productivity; we only require that the average within-plant changes in idiosyncratic productivity for our subset of plants are approximately zero. Older plants are both less likely to exit (as we show in the Appendix) and theoretically more likely to be at their ergodic distribution of idiosyncratic productivity as they had time to build productivity. We choose as an age cutoff plants that are at least seven years old (which is the plant age after which average exit rates do not decrease anymore). As seen in Figure 5, the estimator based on older plants gives lower aggregate productivity estimates as expected from survivor bias. To further ameliorate concerns for overestimating aggregate productivity changes, we can additionally assume that idiosyncratic productivity s_t follows a Markov process of order one, which is a standard assumption in models of plant dynamics and may also follow from certain models with endogenous idiosyncratic productivity. In this case, we can restrict the sample to plants who have had positive idiosyncratic shocks from $t-2$ to $t-1$, which may further reduce the likelihood of exit, while by assumption not affecting changes in idiosyncratic productivity between $t-1$ and t when controlling for s_{t-1}. Figure 5 shows the resulting productivity estimates. Estimates are lower, which is again in line with survivor bias leading to an upward bias of aggregate productivity.

We find that aggregate productivity in the Indonesian data based on our identification strategy is moving surprisingly little. This is robust to a variety of different age cutoffs and cutoffs for previous productivity increases, which all give similar dynamics for aggregate productivity over time. The result is driven by TFP of older surviving plants growing relatively little over time. Given the previous discussion, aggregate productivity is, if anything, likely overestimated and thus the contribution of aggregate productivity growth to aggregate output growth that we estimate gives an upper bound. In the Appendix, we show that this result is not sensitive to the choice of the output elasticity θ, rerunning the same estimation for widely different values of θ.

Given identification of all production function parts, we can finally fully decompose manufacturing output growth for each year in Indonesia from 1979-2015 using Equation (4). Since we are interested in the drivers of growth over the entire time period, we focus on one main summary measure: comparing the year 1979 to 2015. We find that the contribution of aggregate productivity growth to aggregate output growth in manufacturing is less than four percent (3.97%). In contrast, labor accumulation makes up around 68.5% of the total and selection plus reallocation accounts for the remaining 27.5%. Note that labor accumulation here includes not only an increase in the total efficiency units of labor, but it also captures an increase in the number of plants. Due to decreasing returns to scale in production, small is beautiful, and it is generally beneficial to have many plants, which shows up through the labor accumulation term.

To study the relative contributions of selection versus reallocation, we need to take into account that they enter non-linearly into the formula. Thus, to quantify their relative contribution, we consider the following exercise: For the contribution of the selection term between t and $t+1$, assume the reallocation term stays constant at the value at time t, then compute the log change in the sum of those two terms. Do the same exercise for the reallocation part by fixing the selection term at the value at time t. To obtain the relative contributions of the two factors, add up both contributions and compute their relative shares. Following this exercise, we find that both contribute roughly equally, with 51.3% for the selection effect and 48.7% for the reallocation effect. We then do the same non-linear decomposition for the reallocation effect, distinguishing a pure reallocation effect from the mechanical reallocation effect that comes from the increase in the number of plants. We find that the pure reallocation effect actually contributes negatively (-44%), while the increase in the number of plants has a strong positive effect (144%). The reason for the negative reallocation effect is that the covariance between plant productivity and plant output shares declined over time because of the massive entry of small but productive plants and slow plant growth leading to a slow reallocation of resources. Another likely reason that contributed to the negative reallocation effect is that labor supply increased dramatically and estimated wages remained fairly constant, ameliorating the pressure for less productive plants to let workers go and reallocate these to more productive plants.

In the Appendix, we provide further details on this exercise. First, we show in more detail how the relative contribution of each of these factors changed over time for each year in the data. Relatedly, we further decompose parts by considering entry and exit. Secondly, we study the robustness of the decomposition results by considering alternative estimates of aggregate productivity and by considering changes in the contributions from changing the output elasticity θ. The main takeaways from this exercise, and which any structural model needs to replicate, is that the contribution of aggregate productivity to aggregate growth is close to zero in the Indonesian case and that long-run growth in Indonesian manufacturing has been mostly driven by input growth and to a lesser extent by the selection of plants and reallocation of economic activity. The neat thing about the decomposition exercise is that this result holds irrespective of dynamic frictions in the economy and plant behavior; at this point, we have only made an assumption about the production function and on the identification of aggregate productivity.

2.4 The importance of initial conditions and slow transitions

The previous exercise has indicated that aggregate productivity growth did not contribute to aggregate output growth in Indonesian manufacturing over the past 40 years. The key question that remains is what drove aggregate output growth? The decomposition exercise revealed factor accumulation and reallocation effects to have driven most growth, but what exactly these are, is less clear. For example, both could have been driven by a variety of policy changes that removed distortions over time and steadily increased plant entry and input growth. In this subsection, we will provide evidence that the seeds of plant growth were already present at the beginning of our sample in 1975 and that future plant growth is highly predictive. We take this as evidence for the importance of initial conditions and slow transitions and against a large role for future policy changes.

Specifically, we consider the following exercise. Take as the starting point a discrete plant size distribution Φ_t over the number of workers in plants at a given point in time t. For this, we form X different size bins. Each bin captures the density of plants with this number of workers, which can be readily computed as the empirical fraction of plants in this bin. Additionally, we can follow individual plants and compute the probability of moving from one bin to the other between periods t and $t + 1$, which we summarize in the

Table 1: Discretized initial distribution of plants in 1975 and their average plant size measured as the number of workers.

Size bin	Share of plants	Avg # of workers
20 to 22	14.8%	21
23 to 26	15.5%	24
27 to 32	15.3%	29
33 to 44	15.7%	38
45 to 69	14.2%	54
70 to 99	7.1%	82
100 to 149	5.8%	122
150 to 249	4.7%	190
250 to 499	3.5%	356
500 to 15k	3.5%	1106

Details: N = 6,520.

transition matrix $P_{t,t+1}$ of dimension X^2. Entries of the transition matrix are again simply the observed frequencies in the data. At this point, we abstract from entry and exit behavior, that is we restrict ourselves to plants who employ workers in both period t and $t+1$, but we discuss entry and exit further below and in the Appendix.

Both factor accumulation and reallocation of resources could be driven by productive plants needing time to grow large plants. In practice, as mentioned in the introduction, a lot of constraints may prevent productive plants from directly choosing their right size including input adjustment costs and financial frictions. To study the potential importance of slow transitions, we can use the initial distribution Φ_t and premultiply the transition matrix $P_{t,t+1}$ to arrive at an estimate of Φ_{t+1}, the distribution next period. We can then iterate on this distribution by fixing the transition matrix and see how the plant size distribution evolves over time. Technically, $\hat{\Phi}_{1975+t} = (P_{1975,1976})^t * \Phi_{1975}\ \forall t = 1, 2, ...$ We use two metrics introduced previously that should capture different features of changes in the distribution. Namely, the average plant size and the share of workers working in plants with more than 500 workers. To conduct the exercise, we discretize the initial distribution into $X = 10$ bins, which gives 100 entries to estimate for the initial transition matrix $P_{1975,1976}$. We choose $X = 10$ and the exact intervals by trading off keeping the total number of entries in the transition matrix manageable, having easy to interpret bins, having many plants in each bin and overrepresenting size bins for large plants. Table 1 reports the 10 bin intervals we chose as well as the initial distribution over these bins in 1975.

Figure 6 tracks the estimated measures in Panel A and B based on three different exercises we conduct. The results for the initial exercise are given by the lines "1975 Hypothetical." The other two sets of lines give alternative exercises that additionally allow for an entry and exit state (line "Hypothetical 1975 (EE)" where

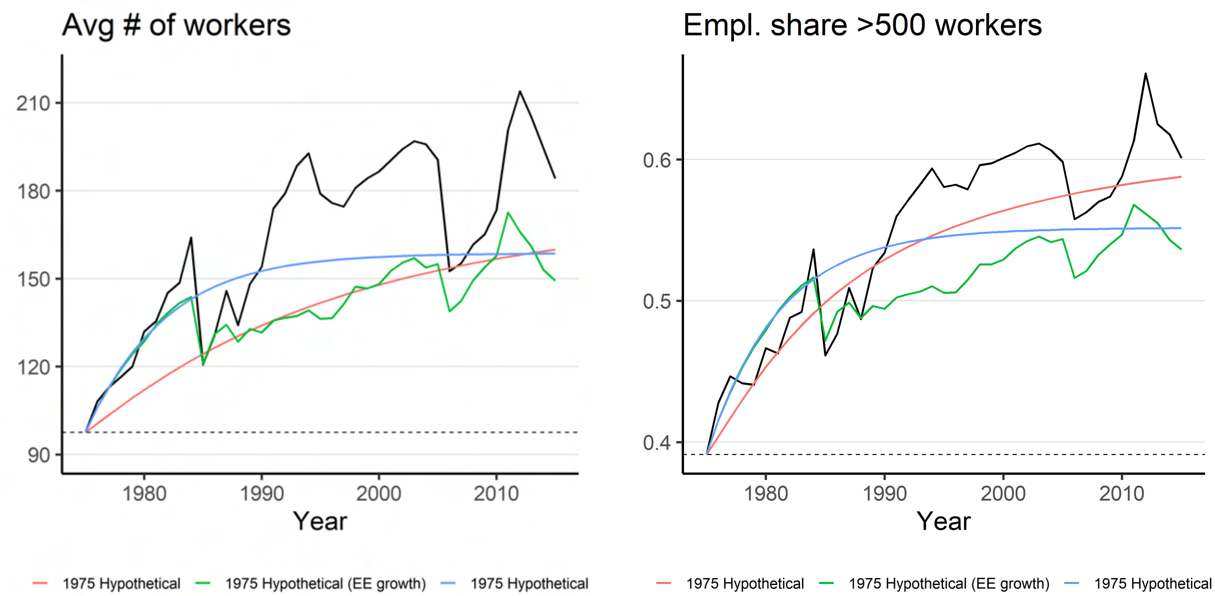

Figure 6: Key empirical exercises on transition dynamics implied by initial plant size distribution. Black line gives the data and the other three lines give predicted changes over time based on different exercises. Details are in the text.

EE stands for entry and exit) and additionally feed in observed entry and exit rates (line "Hypothetical 1975 (EE growth)").[9] Starting from the initial plant size distribution in 1975 and only feeding in information of transitions between 1975 and 1976, we get a long way in explaining changes in the entire plant size distribution over the subsequent 40 years. More than 70% of average plant size increases (A) and more than 90% of the increase in the employment share of plants with more than 500 workers (B) are explained by simply iterating on the initial distribution. Together, these two results also help explain what is driving these transition dynamics. In 1975, the distribution lacks large plants and based on the transitions observed between 1975 and 1976, we can rightly predict that the share of large plants will increase slowly over time until it converges somewhere close to 60% of total employment; the stationary share based on the transition matrix $P_{1975,1976}$. Since the initial employment share is far from this, the average plant size increases. Additionally accounting for entry and exit improves this fit and almost perfectly explains the long run distributional changes given by our two metrics. In the Appendix, we provide a battery of robustness exercises to these results, including varying initial distributions and transition matrices, averaging transition matrices over multiple years and varying assumptions on entry and exit behavior. Furthermore, we show very similar results for real value added output and the labor wage bill.

We take this as strong evidence for the importance of initial conditions in explaining changes in the plant size distribution. In combination with the previous results, this would be in line with reallocation and input

[9] In the Appendix, we provide further details on how we deal with entry and exit in these alternative exercises.

growth being mostly driven by initial conditions and some frictions in the economy that take a long time to play out. Of course, this is not to say that policy plays no role here - it might have easily driven a lot of the transitions from 1975 to 1976 - but fixing these policy effects is highly predictive of future plant size changes and there is little need for further policy changes in the future.

However, it is not clear at this point whether these plant size changes are also driving aggregate output growth. To provide suggestive empirical evidence of this, we look at changes in labor shares within and across plants over time. Following a large literature on plant dynamics, the idea is that factor revenue shares (defined as the spending bill for factor x over real total value-added revenue) are indicative of how close plants are to their (socially) optimal size. Specifically, the presence of dynamic frictions lead plants to change their factor revenue shares only slowly over time, growing them in the case their optimal size is larger and reducing them in the case their optimal size is lower. In the case of static frictions (e.g. as the wedges in Hsieh and Klenow (2009) or markups), reductions in those static frictions over time will also lead plants to grow their factor shares. Panel A of Figure 7 shows our second main empirical result, focussing again on labor as the main input in production. Here, we compute within-plant growth in labor shares over time and take the median growth across plants by plant age to highlight expected differences in plant growth by age. We take the median across plants and further exclude plants for which labor shares more than doubled or halved between any two periods to ensure that our results are not driven by outliers. Furthermore, this exercise pools data across time and computes a weighted median weighting by previous plant size to ensure that the estimates are representative at the aggregate level. We find that over time, plants slowly but persistently grow their labor shares. Importantly, this growth in the labor share does only slowly decrease over plant age and we can rule out zero or negative growth across all ages. Doing the same exercise looking at other moments of the distribution, such as the average or other percentiles gives similar results of persistently positive growth rates. We take this as suggestive evidence that most plants are below their optimal size, but face frictions that prevent them from quickly increasing their factor shares. These frictions persist even for older plants, suggesting that plants adjust very slowly. Our estimate of the production function elasticity is informed by this within-plant growth in labor shares. In Panel B, we show cross-sectional distributions in labor shares and plot them against our estimated production function elasticity. In a static and frictionless setup, the optimal plant size is given when the labor share equals the production function elasticity. Again, this plot shows that there is considerable scope for plant growth as most plants are below this optimum. In the cross-section, we do not see a systematic movement towards this optimum over time, because the composition of plants changes due to entry and exit.

Taken together, this section has provided evidence for aggregate growth being entirely driven by endogenous

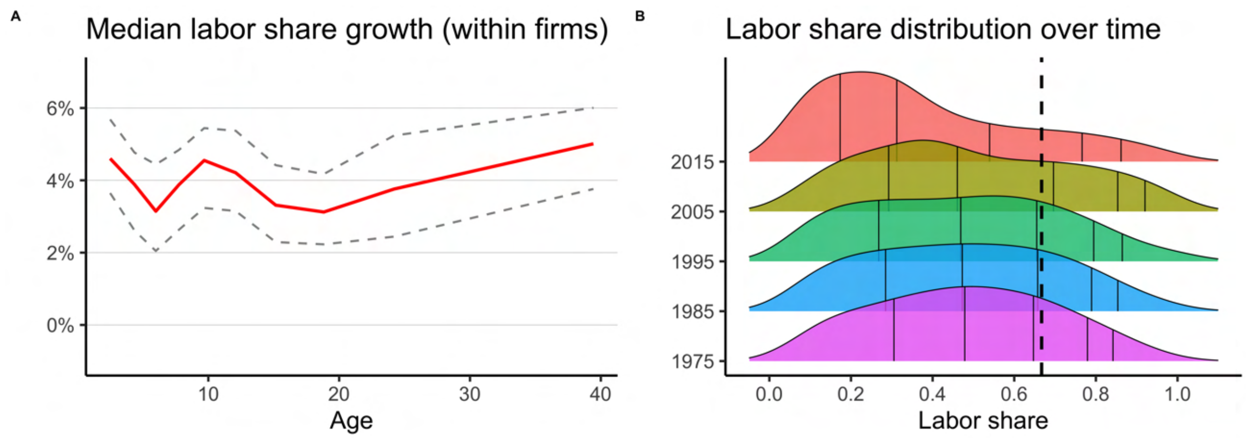

Figure 7: Within-plant and cross-sectional evidence on labor share evolution. Panel A: Evolution of within-plant labor share growth over plant age. Estimates give the weighted median across plants by plant age, weighted by the previous number of workers. Furthermore, estimates are pooled across all years (N = 394,397) and by 10 equal-sized age groups. Plot gives non-parametric bootstrapped 95 percent confidence bands. Further details in the text. Panel B: Evolution of cross-sectional labor share distributions over time. Dotted black line gives estimated output elasticity.

growth due to plants slowly transitioning closer towards their optimal size, the initial plant size distribution in Indonesia capturing a large potential for future growth and aggregate productivity growth to play no role. We now turn to a structural model that flexibly allows an important role for initial conditions and slow plant growth being driven by adjustment costs. This gives a more direct interpretation of the decomposition exercise and endogenizes the transition matrix $P_{1975,1976}$ as well as changes in this transition matrix over time.

3 Structural model of aggregate growth and plant dynamics

This section builds a structural model to analyze the coevolution of changes in the plant size distribution and aggregate growth and to decompose drivers of aggregate growth quantitatively. We consider a standard model of heterogeneous plant dynamics in the tradition of Hopenhayn (1992) featuring exogenous entry and endogenous exit. Plant size evolves endogenously due to the combination of time-varying exogenous aggregate and idiosyncratic productivity, endogenous exit and slow endogenous hiring of labor that is affected by adjustment costs. We then embed this model of plant heterogeneity into a two-sector general equilibrium model to study economy-wide growth processes and endogenous reallocation of resources across sectors.

3.1 Model Setup

Sectors

The model economy is set in discrete time indexed by $t = 1, 2, ...$ There are two sectors of production: Manufacturing (M) and a rest-of-the-economy sector (R). Both sectors produce a single consumption good. The manufacturing sector features heterogeneous plants, while the rest-of-the-economy produces the consumption good as a representative plant with a decreasing returns to scale (DRS) production function:

$$y_t^R = A_t \left(h_t^R\right)^{\theta_R} \quad \text{with} \quad \theta_R \in (0, 1)$$

where A_t is time-varying TFP, h_t^R gives labor employed in the rest-of-the-economy as measured in efficiency units and θ_R gives the output elasticity for sector R. The rest-of-the-economy sector takes as given productivity A_t and the wage rate w_t and chooses optimal labor demand maximizing per period profits: $\pi_t^R(A_t, w_t, \tau_t) = y_t^R(A_t) - w_t h_t^R$, but faces labor demand wedges τ_t^R that distort first-order conditions such that optimal labor demand is given by:

$$h_t^{R*} = \left(\frac{\theta_R A_t}{(1 + \tau_t^R) w_t}\right)^{\frac{1}{1-\theta_R}}$$

For simplicity, we assume that both A_t and τ_t^R follow deterministic exogenously given processes that flexibly capture changes in technology adoption and labor frictions over time. Labor markets in both sectors are fully competitive and there is a single wage w_t across all labor markets.

Manufacturing plants

In manufacturing, the model features a mass of heterogeneous, risk-neutral plants that produce y_{it} at time t according to the production function introduced in the previous section:

$$y_{it} = z_t s_{it} \left(h_{it}^M\right)^{\theta}$$

Production depends on exogenous aggregate manufacturing productivity z_t and exogeneous idiosyncratic manufacturing productivity s_{it}. Furthermore, efficiency units of labor h are the only factor of production and θ captures the output elasticity.[10] Both processes are potentially highly persistent. Specifically, we assume

[10] As explained previously, we consider only labor due to data limitations, observing labor for the entire time period, while capital is only observed starting in 1990 and faces more serious measurement problems. We consider the role of capital in one of the extensions of this model.

that aggregate manufacturing productivity follows a log-linear Gaussian autoregressive process: $log(z_t) = \rho^z log(z_{t-1}) + \epsilon_t^z$ with $\epsilon_t^z \sim \mathcal{N}(0, \sigma^z)$. And idiosyncratic manufacturing productivity is allowed to follow any general bounded Markov process of order one, nesting nonlinear and potentially non-Gaussian processes. Throughout, we assume that both processes are independent.

Plants choose labor partly based on previous plant size due to the presence of flexible labor adjustment costs. That is, plants pay additional adjustment costs if they want to change their workforce. Formally, we model adjustment costs AC following Cooper, Gong, and Yan (2018):

$$AC(h_{i,t-1}, h_{i,t}) = \begin{cases} F^+ + c_0^+(h_{i,t} - h_{i,t-1}) + \frac{c_1}{2}\left(\frac{h_{i,t}-h_{i,t-1}}{h_{i,t-1}}\right)^2 h_{i,t-1} - \text{if } h_t > h_{t-1} \\ 0 - \text{if } h_t = h_{t-1} \\ F^- + c_0^-(h_{i,t-1} - h_{i,t}) + \frac{c_1}{2}\left(\frac{h_{i,t}-h_{i,t-1}}{h_{i,t-1}}\right)^2 h_{i,t-1} - \text{if } h_t < h_{t-1} \end{cases}$$

where F are fixed adjustment costs that capture overhead in dealing with hiring (F^+) or firing (F^-) and c_0 captures per worker hiring and firing costs, which are also allowed to be asymmetric. Importantly, there are also convex adjustment costs whose importance is captured by c_1 and which capture costs of growing or shrinking plants quickly. These convex costs could be organizational in nature. For example, there might be limits to training new workers or the human resources department might have limited capacity and increasing this capacity has large short-run costs by having to reallocate labor, retraining workers, etc. Alternatively, there could be additional costs from collective action (in the case of firing) or it might simply be hard to locally find many suitable job candidates at once. These convex adjustment costs will be key to explain slow growth of plants over time.

Plants choose current labor taking into account adjustment costs to maximize expected discounted life-time profits that in turn depend on per period profits:

$$\pi(s_{it}, h_{i,t-1}, h_{i,t}; z_t, w_t) = y_{it}(s_{it}, z_t) - w_t h_{i,t} - AC(h_{i,t-1}, h_{i,t})$$

The presence of adjustment costs makes this a dynamic problem where plants take into account that contemporaneous changes in the workforce will influence adjustment costs in the future. Plants discount profits at rate $1/R$, where R is the exogenously given international interest rate that is assumed to be constant over time. Additionally, each period plants endogenously decide whether to exit or not. We endogenize exit in a standard way whereby plants draw each period an iid random operating cost c_F from a time-invariant distribution G after production and labor choices are realized (e.g. Clementi and Palazzo 2016).

Further assume that the outside value of exiting is zero as there is no capital that can be dismantled and sold off. It is straightforward to show that the survival decision simply reduces to the probability that the operating cost draw c_F is lower than the future expected value of an incumbent manufacturing plant V_i^M: $\mathbb{P}\big(\mathbb{E}_{s,z,w}[V_{i,t+1}^M] \geq c_F\big) = G\big(\mathbb{E}_{s,z,w}[V_{i,t+1}^M]\big)$. This gives a structural survival probability that depends on the form of G as well as all determinants of $\mathbb{E}_{s,z,w}[V_{i,t+1}^M]$ and which we denote by $\lambda(\Omega_t)$ throughout (where Ω_t denotes all possible states in period t).[11]

The maximization problem of manufacturing plant i is to optimally choose labor demand in each possible state and period t:

$$V_i^M = \max_{\{h_{i,t}(\Omega_t)\}_{\Omega_t,t=1}^{\infty}} \mathbb{E}_{s,w,z} \sum_t \left(\frac{\lambda(\Omega_t)}{R}\right)^t \pi(s_{it}, h_{i,t-1}, h_{i,t}; z_t, w_t)$$

where $\mathbb{E}_{s,w,z}$ denotes expectations over future prices, aggregate and idiosyncratic manufacturing productivity.

At last, we consider exogenous deterministic plant entry. Specifically, denote by $\mu(h_t, s_t)$ the mass of entrants for each state (h_t, s_t) in period t. With slight abuse of notation, denote by μ_t both the entire distribution and total mass of entrants at time t and by $\{\mu_t\}_{t=0}^{\infty}$ the entire sequence of plant entry in the economy. Given that endogenous labor demand choices of manufacturing plants depend on past labor demand, we assume that plants that enter in t only make a model-based decision starting in period $t+1$. Similarly, define by $m(h_t, s_t)$ the mass of producing plants for each state (h_t, s_t) in period t and by M_t the entire distribution and total mass of producing plants at time t.

Households

The economy is populated by a mass of hand-to-mouth households j who consume their labor income and who inelastically provide labor supply each period. The aggregate resource constraint is given by households inelastically consuming all final products $C_t = Y_t = \sum_{i \in M} y_{it} + Y_t^R$. Furthermore, households may have idiosyncratic efficiency units of labor h_{jt} which aggregate to total labor supply: H_t. Aggregate labor supply may vary due to exogenous population growth, exogenous changes in the composition of workers and exogenous changes in average skills and schooling, which we all treat as deterministic processes, so that we denote the exogenous sequence of aggregate labor supply by $\{H_t\}_{t=0}^{\infty}$. Importantly, households allocate their labor supply across sectors based on maximizing labor income. Thus, the labor supply elasticity that manufacturing plants face changes over time due to exogenous changes in the total labor supply and changes

[11] This setup is isomorphic to one where c_F is instead an outside opportunity for the entrepreneur running the plant. Given that we directly estimate $\lambda(\Omega_t)$ in the data, it is also observationally equivalent to a setup where the plant pays adjustment costs of dismantling the plant; one simply needs to redefine G. However, for counterfactuals, this difference matters.

in labor demand in the rest of the economy that are in part driven endogenously by the prevailing wage rate.

Equilibrium

The focus in this paper is on a (growth) path of per-period *Recursive Competitive Equilibria.* Denote by $Z_t \equiv \{A_t, \tau_t^R, H_t, \mu_t\}_t^\infty$ the sequence of exogenous deterministic processes in the economy starting at t. A *Recursive Competitive Equilibrium* is then defined by Z_t, an endogenous wage w_t and a total mass of producing plants M_t such that:

1. the rest-of-the-economy sector statically chooses optimal labor demand maximizing profits taking productivity A_t and the wage rate w_t as given and facing labor demand wedges τ_t^R.
2. manufacturing plants choose optimal labor demand and optimal output by maximizing profits, taking as given previous labor demand, s_t, z_t and w_t and forming rational expectations over future s_{it}, z_t and w_t. This means that a manufacturing plant's problem can be written recursively as:

$$V(s_{i,t}, h_{i,t-1}, z_t, M_t, Z_t) = \max_{h_{i,t}} \left\{ \pi(s_{i,t}, h_{i,t-1}, h_{i,t}; z_t, w_t) + \left(\frac{1 - \lambda(h_t, s_t)}{R} \right) \mathbb{E}_{s,w,z} \left[V(s_{i,t+1}, h_{i,t}, z_{t+1}, M_{t+1}, Z_{t+1}) \right] \right\}$$

3. Households inelastically supply total labor H_t, but optimally allocate labor across sectors to maximize labor income.
4. the aggregate wage w_t adjusts to ensure that the labor market clears: $H_t = h_t^R(w_t, A_t, \tau_t^R) + \sum_{i \in M_t} h(s_{it}, h_{i,t-1}; z_t, w_t)$
5. The mass of active plants in t is equal to surviving plants from $t-1$ plus exogenously given new entrants:

$$\forall (h_t, s_t): \quad m(h_t, s_t) = \sum_{h_{t-1}, s_{t-1}} (1 - \lambda(h_{t-1}, s_{t-1})) m(h_t, s_t | h_{t-1}, s_{t-1}) + \mu(h_t, s_t)$$

   where $m(h_t, s_t | h_{t-1}, s_{t-1})$ denotes the mass of plants that moved from $m(h_{t-1}, s_{t-1})$ to $m(h_t, s_t)$.
6. The output market clears each period: $C_t = Y_t = \sum_{i \in M} y_{it} + Y_t^R$

3.2 Identification, Estimation-Computation

The model features a flexible combination of aggregate productivity, a number of other time-varying aggregate processes (labor supply, entry, productivity and labor wedges in the rest-of-the-economy) and a path of time-varying equilibria. Parameters of the model directly influence this path of time-varying equilibria.

The “standard Macro” way of solving this would be to solve for the entire time path of equilibria for each set of parameters and find the set of parameters that brings the model closest to the data based on a set of well-chosen moments (e.g. via simulated method of moments). The main problem with this approach is that solving for each path of equilibria is computationally intensive, preventing flexibility on the amount of model parameters that are estimated. Following recent ideas in the estimation of macroeconomic models (e.g. Gopinath et al. (2017)(2019)),[12] we exploit the fact that in the data we observe one path of equilibria. The idea is to invert the model computation by not trying to solve the model for a path of prices that clears markets, but instead using the observed price path in the data and finding the set of parameters that are consistent with the observed path of equilibria. It turns out that this reformulation allows for much more flexible estimation of model parameters, because for many sets of model parameters, the model never actually has to be solved. Specifically, this allows us to non-parametrically estimate idiosyncratic productivity and exit decisions as well as flexibly estimate the labor adjustment cost function. The potential downside of this approach is that it forces us to introduce a time path of additional wedges that ensure market clearing at each point in time at observed prices. We actually find these wedges helpful, because they are informative about model-data mismatch and because we interpret these wedges as time-varying policy residuals.

The main difference to Gopinath et al. (2017)(2019) is that we consider a stochastic process for aggregate productivity in conjunction with rational expectations instead of a deterministic process with perfect foresight. Separately identifying aggregate productivity realizations from idiosyncratic productivity in the data helps us to directly estimate the process of aggregate productivity using a single path realization in the data without ever having to solve the model. To solve the *Recursive Competitive Equilibria* with heterogeneous plants that follows from rational expectations, we make use of the standard constrained rational expectations algorithm by Krusell and Smith (1998). The only difference to the standard Krusell and Smith (1998) approach is that the forecasting rules with which plants forecast prices can be directly estimated in our data using the single observed path of productivity realizations. This means that we never have to solve the fixed point problem over the parameters of the forecasting rule, the main computational difficulty of Krusell and Smith (1998). It is important to note that under the null hypothesis that all observed data is indeed generated by the model, the two approaches are equivalent up to differences in the precision of the estimates.

In the following, we go through all parameter identification and estimation steps and discuss in more detail how we practically implement Krusell and Smith (1998). We distinguish direct estimation, which allows estimation without having to solve a model first, and indirect estimation, which requires to solve the model.

[12]To be clear, more microeconomic approaches to estimation (such as most approaches in Empirical Industrial Organization) have made very productive use of this revealed price and preference argument for a long time.

Parameterization of standard parameters

We start out with setting $\beta = 0.95$ for all households in our economy as well as the rest of the world. Based on standard arguments, this implies that the international interest rate is given by $R = \frac{1}{\beta} - 1$. All the remaining parameters of the model are estimated.

Direct estimation of productivity and wage processes

Next, we proceed with the direct estimation part. We start with the manufacturing sector. In line with the analyses in Section 2, we use θ to directly identify TFP: $TFP_{it} \equiv z_t s_{it} = \frac{y_{it}}{h_{it}^{\theta}}$. The labor bill is observed in the data, which we divide by an estimate of the real wage to obtain efficiency units of labor: $h_{it} = \frac{\text{wagebill}_{it}}{w_t}$. We identify the real wage in the data by drawing on reported real per worker wages in the manufacturing survey. To capture true changes in the real wage, we need to make sure that these wages capture remuneration for the same job and the same worker skills. Ideally, we want to capture changes in the real wage for a worker whose efficiency units of labor remained constant. In the absence of worker-level data, we focus on average changes in reported real per worker wages within plants who have seen little changes in their workforce. For this, we focus on plants whose total number of workers changed by less than 5% between t and $t+1$ (for the smallest plant with 20 workers in the dataset, this allows a maximal change of one worker). Furthermore, we focus only on reported production worker wages in contrast to other workers and we exclude plants whose per worker wages increased or decreased by more than 50% between any two time periods. At last, we use the weighted average of wage growth across plants, weighting by the number of production workers within a plant. After normalizing the initial wage to unity, Figure 8 plots the estimated real wage series in the data. Estimating wages directly in the data means that to solve the model at the observed competitive equilibrium, we can directly solve for the model-implied total labor supply that is consistent with labor market clearing at this wage.[13]

Next, we estimate θ. As discussed in Section 2, we propose a novel estimator that allows labor to be fully dynamically chosen. While we explain technical details of the estimator in the Appendix, the key idea behind the estimator comes from looking at the policy functions implied by a model with adjustment costs as shown in Panel A of Figure 9. Each separate line gives the policy function for a specific idiosyncratic productivity realization in period t. Each dotted line gives the static unconstrained optimal choice of labor, which is independent of previous labor. Note that the dotted line in combination with the 45 degree line (which captures whether plant inputs grow or decline) separates the policy space into four regions. Importantly,

[13]This simplifies the computation of the observed equilibrium. For counterfactuals, we need to explicitly solve for wages that clear the labor market, taking into account model-implied labor supply and demand as described in the previous section.

Figure 8: Evolution of the real wage in Indonesian manufacturing using within-job changes in observed real wages. Estimation is based on reported wages of production workers from plant-level data based on Statistik Industri.

optimal model-implied policies only happen to be in two of the four regions, which is a general result across many dynamic frictions. In case a plant is above their optimal static unconstrained labor demand (as judged by current productivity and past labor demand), a plant always wants to weakly decrease their labor demand, while in the opposite case, a plant wants to weakly increase labor demand. The estimator exploits this geometric result by finding the θ that maximizes correctly classifying plants into these two regions and minimizing the number of plants that are classified into being in one of the other two regions. Based on this estimator, we find that $\hat{\theta} = 0.577$ with bootstrapped 95% confidence bands ranging from 0.52 to 0.626.

We separately identify aggregate productivity changes as in Section 2 and plotted in Figure 5. To see that this identification strategy works in the model here, remember that there are three potential biases in the estimation of aggregate productivity: selection on exit, changes in the distribution of idiosyncratic productivity and correlation between aggregate productivity and idiosyncratic productivity. Here, exit is indeed potentially negatively correlated with realizations of idiosyncratic productivity, but exit only happens after production by assumption, which eliminates the bias entirely. Note that this bias leads - if anything - to overestimate the role of aggregate productivity growth. Secondly, entering and young plants in the model may indeed be potentially far away from the ergodic distribution of idiosyncratic productivity, but focussing on older plants will eliminate this bias. At last, aggregate productivity and idiosyncratic productivity can be correlated in the model via entry (that is, a specific set of plants entering based on observed aggregate productivity), but the evolution of idiosyncratic productivity is assumed to be independent from the evolution of aggregate productivity. Most importantly, note that plants in the model may be far away from their optimal size, but that only shows up in the labor component and not in TFP and will thus not bias the estimation of aggregate productivity.

For aggregate productivity z_t, we identified a single path realization between 1980 to 2015. We use this path to estimate $z_t = \rho^z z_{t-1} + \epsilon_t^z$ via OLS and obtain $\hat{\rho}^z \approx 0.93$. Taking the standard error of the residuals of this regression as an estimate for σ^z, we then discretize the process of z_t via the Rouwenhorst algorithm using 10 grid points and a grid space that covers two standard deviations of the stochastic process. We use this discretized distribution of z_t and rational expectations over its evolution as captured by the transition matrix $\pi(z|z_{t-1})$ throughout. This also means that we discretize the observed path realization between 1980 to 2015 by taking the closest grid point for each yearly realization.

Having identified TFP and aggregate productivity, we back out idiosyncratic productivity s_{it} (from the non-discretized z_t), discretize it and estimate its transition matrix $\mathbb{P}(s'|s)$ non-parametrically by simply taking observed frequencies in the data, assuming solely that s_{it} follows a bounded first-order Markov process.[14]. Specifically, we choose 15 grid points for idiosyncratic productivity, which we select based on quantiles of the productivity distribution, oversampling highly productive plants to correctly capture the right tail of the plant size distribution.[15]

Direct estimation of the state space, entry and exit

Since we directly identify the state space of plants, we can also directly estimate the exit rate $\lambda(s, h, z, w)$ from the data. While this exit rate is fundamentally driven by plants being subject to random cost draws and endogeneously choosing whether to exit, this dependence is entirely captured by the state space plus the structural parameters that are fixed at the observed path of equilibria in the data. We can thus treat observed exit rates based on the state space as structural and estimate these without having to solve the model. In a second step (after all model parameters are estimated), one can then ask what cost draw distribution is in line with the estimated exit rates. In principle, we can estimate exit probabilities fully non-parametric by using observed exit frequencies over the entire state space. However, it is key that the estimated exit probabilities show sufficient smoothness to prevent plants in the model from endogeneously choosing particular states with low estimated exit probabilities. In the end, we estimate exit probabilities with a flexible logistic regression, allowing for different effects for each realization of idiosyncratic productivity and linear terms in the wage, aggregate productivity and current labor.

[14]See Ruiz-García (2019) for the quantitative importance of using a flexible productivity process rather than a standard log-normal process in structural models of plant dynamics

[15]We form bins based on the following 15 quantile ranges: 0.0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, 0.4-0.5, 0.5-0.6, 0.6-0.7, 0.7-0.8, 0.8-0.85, 0.85-0.9, 0.9-0.95, 0.95-0.975, 0.975-0.99, 0.975-0.995, 0.995-1.0. For the value of idiosyncratic productivity of each bin, we take the pooled sample average for each bin.

To identify the state space on a manageable grid for model computation, we also discretize efficiency units of labor h_{t-1} by choosing 1000 grid points that we choose based on equal spaced quantiles, ensuring that the entire labor distribution is well represented. Identification of the state space allows to directly take the initial distribution of plants from the data, which is the year 1979 in our data. We take 1979 as the initial year, because we observe labor of these plants in 1978, and 1979 is the first year for which we observe value added output, which we need to estimate idiosyncratic productivity.[16] Real wages and aggregate productivity are normalized to unity at the first year we can identify them. Another benefit of identifying the state space is that we can directly feed the model with true entry distributions of plants. These plants enter with (s_{it}, h_{it}) and thus do not make any model-based choices until the next period.

Direct estimation of the forecasting rule for prices

To solve the *Recursive Competitive Equilibria* with heterogeneous plants using the constrained rational expectations algorithm by Krusell and Smith (1998), we estimate linear forecasting rules. Specifically, we follow Clementi and Palazzo (2016) and assume the following forecasting rule:

$$log(w_{t+1}) = \beta_1 + \beta_2 log(w_t) + \beta_3 log(z_t) + \beta_4 log(z_{t+1}) + \epsilon_{t+1}$$

This means we abstract from the other time-varying processes (labor supply, entry, productivity and labor wedges in the rest-of-the-economy) to enter the forecasting rule. One can rationalize this choice by assuming that manufacturing plants find it hard to predict other time-varying processes or are ignorant about things that happen outside of manufacturing. We estimate the parameters of this forecasting rule by direct estimation using the single observed paths realization of $(z_t, w_t)_{t=1980}^{2015}$. This avoids having to solve the model in the first place.

Running the previous regression using non-discretized values for w_t and z_t, we find that only the previous wage significantly predicts the future wage with productivity realizations being statistically indistinguishable from linear independence. We find an $R^2 \approx 0.75$ of this regression with $\beta_2 \approx 0.83$.[17] At the current stage of

[16]Strictly speaking, value-added output is also observed in 1975, but not between 1976-1978. We cannot use the 1975 value-added data, because we do not observe efficiency units of labor in 1974.

[17]This is clearly much worse than what is usually estimated using model-based regressions (e.g. Clementi and Palazzo (2016) obtain $R^2 = 0.997$). This difference must be judged by: (1) We run this on 35-40 observations, while usual regressions run this on arbitrarily many paths. (2) We run this on non-discretized values of z and w. The usual model-based regressions use discretized z and then obtain implied wages, which should reduce the baseline variance meaningfully. (3) We estimate this on a transition path, while the usual approach is to throw away transition periods and only consider periods close to the stochastic steady state (which should lead to better R^2). (4) Our w and z are generated in the data, while usual regressions are model-based, which throws away any non-model variation in the data. Clearly, linear independence from aggregate productivity is problematic. However, given the results of this paper, wages are likely more driven by transition dynamics, changes in labor supply and structural change. An intermediate solution is to take these estimates only as an initial estimate and then iterate on these estimates using the standard model-based algorithm. Furthermore, model-based generation of the forecasting rule can

the project, we take this as evidence that the true forecasting rule is independent of productivity realizations and only use the forecasting rule including the previous wage. Furthermore, as is standard in using these Krusell-Smith-type forecasting rules, we ignore variance in ϵ_{t+1} and let plants deterministically predict future wages based on the forecasting rule. Since we cannot reject myopic price expectations based on the forecasting rule estimates, we use $\beta_1 = 0$ and $\beta_2 = 1$ for simplicity and tractability.[18]

Remaining direct estimation (labor supply + Rest of the Economy)

Total labor supply H_t is given exogenously and is simply the sum of aggregated labor supply in the two sectors of the economy: $H_t = h_t^R + H_t^M$. Total labor supply in manufacturing H_t^M is observed by aggregating up all observed h_{it} for each time period t. To obtain h_t^R, we use the fact that we observe the total number of workers l_t^R in the Rest of the Economy. To map from the number of workers in R to the total efficiency units of labor in R, we directly draw on recent estimates of worker selection for Indonesia by Hicks et al. (2017). Specifically, we use their estimates of wage differences and worker selection across rural agriculture and urban non-agriculture as a benchmark for our two sectors. This leads us to estimate that average efficiency units of labor are roughly two times larger in M than in R in our economy and we use this to infer total labor supply in the rest of the economy: h_t^R.[19]

For the rest-of-the-economy sector, we can directly identify θ_R and the sequences of A_t and τ_t^R. For this, take plant first-order conditions to obtain:

$$\frac{\theta_R}{(1+\tau_t^R)} = \frac{w_t h_t^R}{y_t^R}$$

We use observed y_t^R and can construct $w_t h_t^R$ to obtain the left-hand side. We assume that wedges behave such that the average of the right-hand side over time is exactly equal to θ_R. This gives $\theta_R \approx 0.226$. Wedges τ_t^R are backed out such that the previous equation holds exactly. Given θ_R and h_t^R, we can simply back out the sequence A_t using: $A_t = \frac{y_t^R}{(h_t^R)^{\theta_R}}$.

be used as a model test. We plan to do this in a later version of the paper.

[18] We are currently implementing the setup with non-myopic price expectations.

[19] Hicks et al. (2017), using worker-level panel data from Indonesia, find that non-agricultural jobs earn about 2.5 times higher income than agricultural jobs, but that around 80% of this earnings gap is explained by selection as captured by individual-specific fixed effects. Through the lens of our model, this implies that manufacturing workers have on average much more efficiency units of labor. We enforce the point estimates of Hicks et al. (2017) across all time periods.

Indirect estimation

The last part of the model estimation is based on indirect estimation. The remaining plant-side model parameters are the labor adjustment costs parameters $\{F^+, F^-, c_0^+, c_0^-, c_1\}$. Those cannot be directly inferred from the data and depend on a dynamic labor choice. This means that for estimation, the entire model has to be solved for each combination of the adjustment cost parameters. At this point, we enforce symmetry of adjustment costs so that the problem reduces to three parameters. We then estimate the parameters at the plant-level by minimizing the distance between model-implied policies and actually observed policies according to:

$$\{\hat{F}, \hat{c}_o, \hat{c}_1\} = argmin\Big\{ \sum_i \Big(\sqrt{h_{1980}^{model}(\hat{s}_{i,1980}, h_{i,1979}, ; \hat{z}_{1980}, \hat{w}_{1980})} - \sqrt{h_{1980}^{data}(\hat{s}_{i,1980}, h_{i,1979})} \Big)^2 \Big\}$$

To ensure that estimated policies are neither entirely driven by large plants for which changes in policies are larger nor by outliers, we (1) use the square root of observed policies and (2) further approximate observed policies in the data by a non-parametric Kernel fit. We visualize this approach in Figure 9 below. The Kernel estimation ensures that we estimate adjustment cost parameters based on the key features of the data rather than based on the outsized role of a few outliers.

At last, we restrict the estimation to the first two periods of observed policies. This allows us to distinguish the role of transition dynamics from changes in policy similar to the reduced form exercise where we iterate on the initial plant size distribution. Specifically, we fix the policy environment in 1979 and 1980 to estimate adjustment costs and introduce time-varying model wedges that ensure market clearing over time. We interpret these time-varying wedges as capturing a combination of IID noise and a potentially persistent "policy shock" component that is related to adjustment costs changing over time. Wedges are computed as average percentage deviations from model-implied policies that are needed to clear the labor market.[20] We show estimated wedges below.

Identification of the three adjustment cost parameters is explained in Figure 9. Panel A gives the intuition of the identification, where a model-implied optimal policy is plotted based on some combination of parameters F, c_o, c_1. Each line gives a different productivity realization. Intuitively, the inaction region given by the part of the policy function that is on the 45 degree line is governed by the fixed cost F and the hiring and firing cost c_0. In contrast, the curvature of the adjustment off the 45 degree line is captured by c_1. The higher these convex adjustment costs, the slower adjustments and the closer to linear these adjustments (the

[20] Specifically, we estimate a uniform wedge such that distorted labor demand is given by: $\tilde{h}^* \equiv (1+\tau)h^*(s_{i,t}, h_{i,t}, ; z_t, w_t)$ and we find the wedge such that the aggregated distorted labor demand clears the labor market.

less curvature). Fixed costs F are identified from observed changes in the workforce for small plants, while variable costs c_0 are identified from larger plants for whom the fixed cost plays a much smaller role in the decision-making process. Panel B in contrast highlights the actually estimated policy function by showing the fit against the data for a specific productivity shock realization. In the data, observed policies bunch around the inaction region in line with the presence of important fixed costs. Furthermore, observed policies are a lot noisier than what is implied by the simple model with labor adjustment costs. To avoid estimates being driven by outliers, we estimate parameters based on the Kernel fit shown in Panel B. What is crucial here is that the data clearly rules out strong curvature in labor adjustments. Most labor changes are close to the inaction region. In the model, this is only possible in the presence of high convex costs of adjustment.

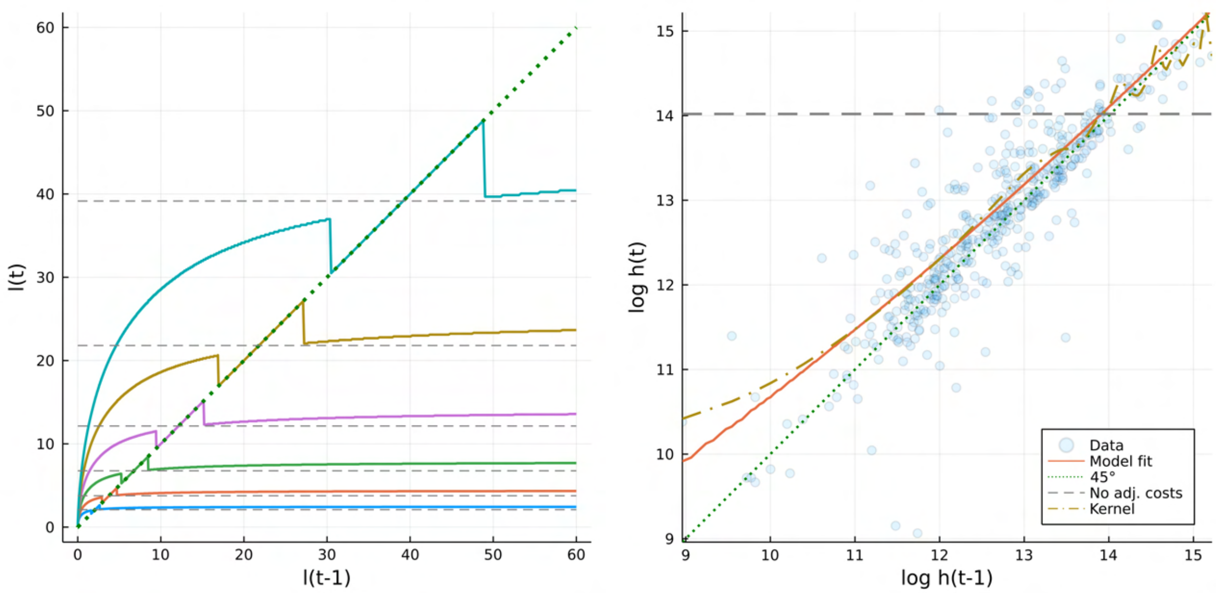

Figure 9: Identification and estimation of the adjustment cost parameters. Panel A gives a hypothetical optimal policy function based on some combination of parameters F, c_o, c_1. Each line gives a different productivity realization. The 45 degree line gives the inaction region in which plants do not change their plant size. Panel B gives the estimated policy function together with observed data for a specific productivity shock realization.

3.3 Evaluating model fit

Given the estimation and computation of the model, we want to look at how well it fits the data. For this, we focus exclusively on non-targeted aggregate moments in the data. Figure 10 shows how the model performs in predicting changes in the average plant size measured as the average efficiency units of labor (Panel A) and in aggregate manufacturing output (Panel B). The model predicts well the long-run evolution of the average plant size. To get this right, the model needs to match well the evolution of the number of plants and the total labor demand. We fit the total labor demand by construction, because we introduce wedges such that labor markets clear each period. For the number of plants, we feature exogenous exit (so fit entrants exactly), but have endogenous exit which leads to time-varying exit probabilities due to changes in aggregates and endogeneous changes in the plant-size distribution over time. At the current stage of the project, we overestimate exit, leading to too few firms over time and hence overestimating the average firm size.

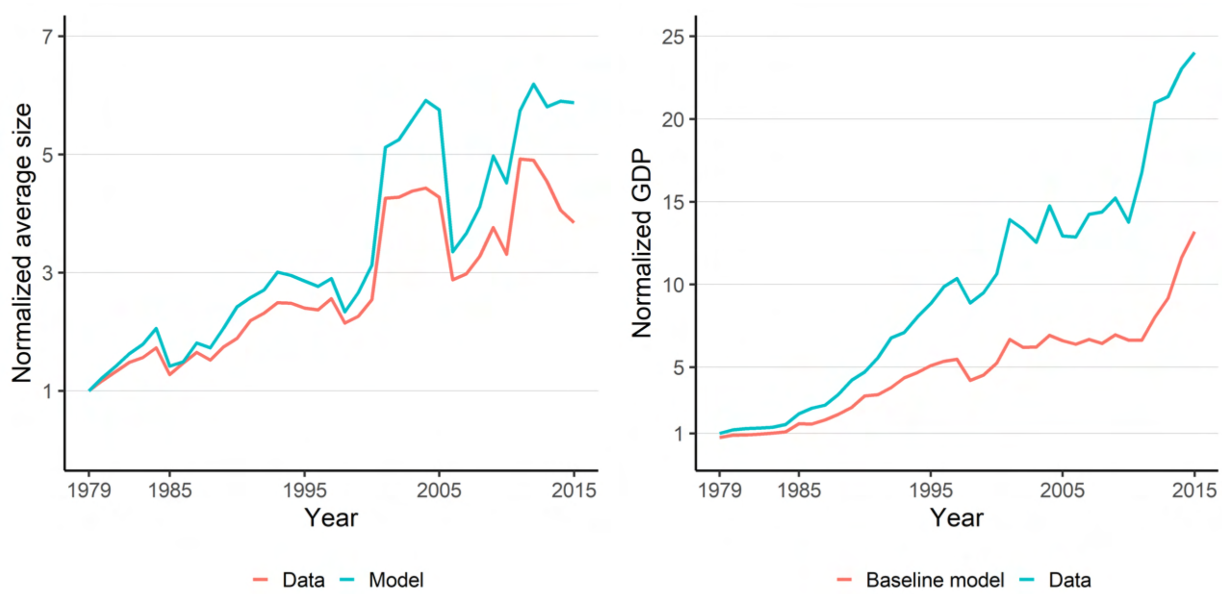

Figure 10: Model fit for the evolution of the average plant size (in efficiency units) and aggregate output in manufacturing. For interpretability, both series are normalized to 1 in the initial year.

For output, Figure 10 (panel B) shows that the model currently largely underpredicts manufacturing output growth. A large part of this is due to overpredicting exit. When we mechanically increase the measure of firms to the level seen in the data, we explain more than half of the gap in predicted output. To even better understand why we do not correctly fit output, we can replicate the accounting exercise from Equation (4) inside the model. The first two rows of Table 2 decompose aggregate growth in manufacturing from 1979 to 2012 in the data and the model. Column (1) shows the increase in log output and columns (2) to (4) its decomposition. Columns (2) and (3) show that the model fits well aggregate productivity growth (by construction) and endogenous input growth. The key reason for why the model underpredicts output growth is because it underpredicts the selection and reallocation term as shown in column (4). Columns (5) and (6) in turn show that the model separately underpredicts selection and reallocation. For selection, the model hardwires the average idiosyncratic productivity evolution through the exogenous productivity process and its transition matrix. Since new entrants are introduced with their true idiosyncratic productivity, the underestimation likely stems from too little selection of surviving plants. This again points to issues in how we currently estimate exit. This model change is also likely to have a positive effect on the reallocation part, which we currently underestimate.

Table 2: Decomposition of aggregate growth in manufacturing: 1979-2012

-(1)-(2)-(3)-(4)-(5)-(6)-(7)-(8)
-ΔlnY-Δlnz-$\Delta ln(\sum h_i^\theta)$-$\Delta ln(\bar{s} + Ncov(.))$-$\Delta ln(\bar{s})$-$\Delta ln(Ncov(.))$-$\frac{H_{2012}^M}{H_{2012}}$-$w_{2012}$
Data -3.044- -0.043-2.057-1.023-1.250- 0.906-0.067-0.956
Baseline model -2.368- -0.043-2.037-0.130-0.725–0.597-0.069-0.956
CF1: $z_t = 1$ -2.336- 0.000-2.174-0.162-0.590–0.185-0.075-1.099
CF2: $\Delta z_t = 2\%$ -3.401- 0.653-2.536-0.211-0.590–0.080-0.138-1.362
CF3: $c_1 = 0$ -3.328- -0.043-2.508-0.862-0.590- 0.985-0.208-1.900

Notes: We restrict our results to 1979-2012 because the Manufacturing survey lacks data on value added from 1976 to 1978 and our current version of the GGDC database stops in 2012. $\Delta lnX \equiv lnX_{2012} - lnX_{1979}$. (1) $\equiv$ (2)+(3)+(4). For comparison with (7) and (8), the share of (efficient) labor in the Manufacturing sector in 1979 was $\frac{H_{1979}^M}{H_{1979}} = 0.035$ and the wage $w_{1979} = 0.827$.

Next, we can look at the evolution of wedges that are needed to ensure labor market clearing each period. These wedges shown in Figure 11 are indicative not only of model fit, but also of the importance of policy changes over time. Specifically, we would expect to see systematically higher wedges over time in case model-implied adjustment behavior leads to too small factor accumulation over time. Importantly, we do not see any systematic long-run pattern in the wedge. Endogeneous growth due to plants slowly moving to their optimal plant size as predicted by adjustment cost parameters estimated on initial data in 1979/1980 explains well within-plant size growth over time. While the wedge is clearly autocorrelated, it is also well-centered around zero. We interpret this as strong evidence against adjustment costs (that drive factor accumulation in our model) or other policy parameters changing over time due to policy.

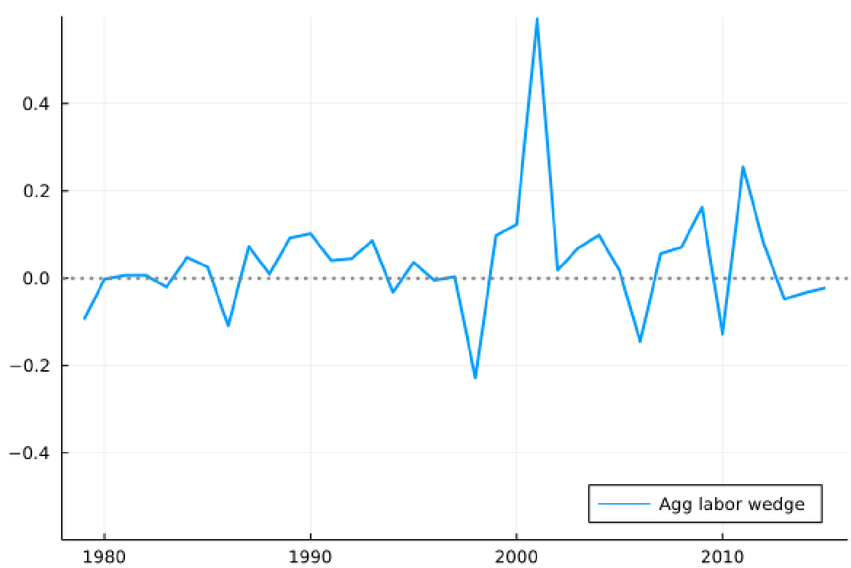

Figure 11: Estimated policy wedge that ensures labor market clearing each period.

At last, we want to look at whether the model gets right the predicted size distribution of plants over time. For this, Figure 12 plots the evolution of the entire plant size distribution for selected years. We plot both the data and the model-implied distribution over the discretized labor grid to ensure comparability. The results show that we match the evolution of the entire plant size distribution reasonably well. In 1985, we closely track the plant size distribution, which is in part driven by having fed the exact plant size distribution in 1979 and by seeing and predicting only slow transitions in the meantime. If anything, the model underpredicts the left tail of the distribution in 1985, which could be driven by overpredicting plant growth at the left tail or underpredicting declines in plant size for the left tail.[21] For 1995, the model predicts well the entire shift in the plant size distribution. In part, this is by construction as it is driven by large amounts of plant entry, on the other hand, this is because the model correctly predicts changes in the plant size across the entire distribution over time. Further over time, the model captures well not only the average shift to the right, but also a shift in the skewness of the distribution. For the end of the time period, the model captures well the plant size distribution, but underpredicts the right tail of the distribution.

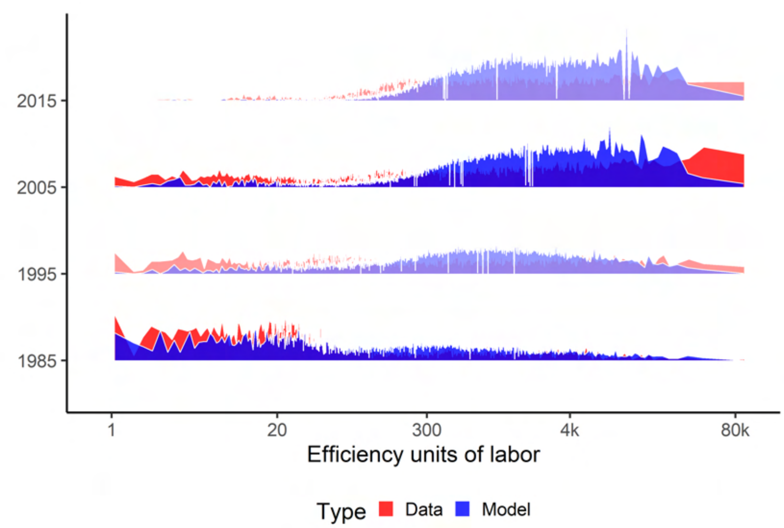

Figure 12: Model fit of the plant size distribution over time.

4 Results: Quantifying the drivers of aggregate growth

In this section, we use the structural model to better understand the main drivers of long-run economic growth and structural change in Indonesia. Specifically, we quantify the role of productivity growth and convex adjustment costs for economic growth and structural change. This section is not yet updated and

[21] This slight underprediction of the left tail is a consistent feature of the model vis-a-vis the data and could point to either too little variation in idiosyncratic productivity grid points at the left tail or model misspecification either in the adjustment costs for small plants or potentially also in forcing the same production function across all plants.

depends on results of a previous iteration of the model. Results should thus only be indicative of what we can study with the model.

The role of productivity growth in driving aggregate growth

The statistical decomposition exercise in Section 2 revealed a contribution of aggregate productivity growth to aggregate output growth that was indistinguishable from zero. Through the lens of the model, changes in aggregate productivity do not only have a direct effect on economic output - holding inputs constant, all plants will produce more - which is captured by the decomposition exercise, but also an indirect effect through incentivizing factor accumulation and reallocating economic resources across plants. To capture the role of both direct and indirect effects, we consider a counterfactual policy in which we keep aggregate productivity constant over time. Given the estimated productivity process shown in Figure 5, this means that the counterfactual smoothes out aggregate productivity, basically taking out the Asian Financial crisis and other short-run to medium-run economic shocks. Figure 13 plots counterfactual output for this scenario. We find indeed that aggregate productivity had no meaningful role in the long-run evolution of aggregate output in Indonesian manufacturing: entirely shutting down the role of aggregate productivity gives indistinguishable output until the early 1990s and then simply smoothes out the Asian Financial crisis.

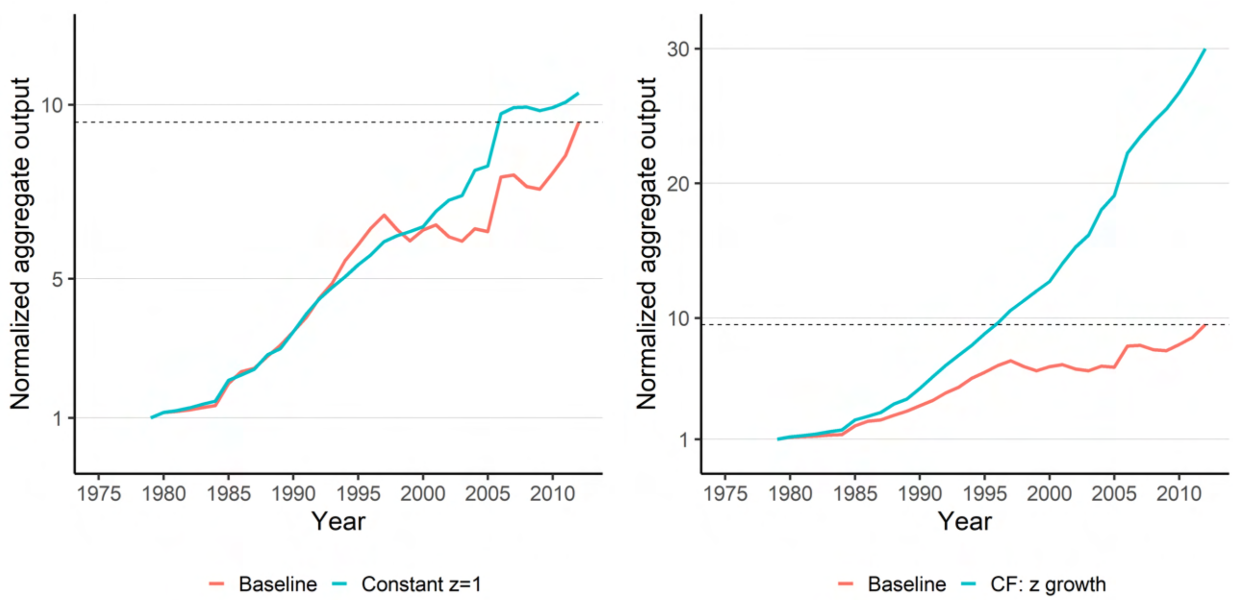

Figure 13: Counterfactual 1: Keeping aggregate productivity constant over time and Counterfactual 2: Constant aggregate productivity growth by 2 percent per year. Output is for manufacturing only.

The third row of Table 2 decomposes aggregate growth in manufacturing from 1979 to 2012 when $z_t = 1\ \forall t$. Column (1) shows, again, that smoothing out aggregate productivity fluctuations does not have a substantial impact on long-run output growth with respect to the baseline model. It is interesting, however, to notice that eliminating fluctuations has a small positive effect on output growth in the presence of adjustment costs, explained both by a small increase in factor accumulation and better reallocation.[22]

The fact that z_t played de facto no role in output growth in Indonesian manufacturing over the period 1979-2012 does not mean that aggregate productivity generally plays no role in aggregate growth. To see this point more clearly, we consider an alternative counterfactual in which aggregate productivity grows by a constant 2% per annum.[23] In Figure 13 (Panel B), we show that in this case, aggregate manufacturing output in 2012 is more than three times larger. As we show in the fourth row of Table 2 via the decomposition exercise, about 57% of this differential growth is driven by the direct effect of aggregate productivity on output. The remaining 43% are driven by indirect effects related to induced input growth (33%) and improved reallocation of resources (10%).[24] Importantly, the large indirect effect through input growth happens through induced structural change: in this counterfactual, the employment share in manufacturing increases almost fourfold between 1979 and 2012. This increase almost doubles the employment share in the manufacturing sector compared to the baseline economy. Since aggregate productivity increases the marginal productivity of workers and the price of efficiency labor is the same for manufacturing and the rest of the economy, wages increase by more than 40% for workers in all sectors.

The role of convex adjustment costs

Given the results on the role of aggregate productivity, output growth in Indonesian manufacturing is driven by slow growth of plants that are far away from their optimal plant size (as inferred from a setup without adjustment costs). Through the lens of the model, this slow growth is caused by high convex adjustment costs that prevent plants from growing faster. To quantify the role of convex adjustment costs, we consider a counterfactual in which convex adjustment costs are entirely shut down ($c_1 = 0$). Figure 14 plots output responses for this counterfactual, showing that real output in 2012 could have been already reached in 1986, 26 years earlier or in 1/5 of the time. This means that large convex adjustment costs are key for driving slow transitions as they keep plants from reaching their optimal plant size.

Another interesting feature that follows from Figure 14 is that it gives a measure of the potential of future

[22]In all our counterfactual exercises the role of selection is the same because in our current estimation of the model we suppose that entry is exogenous and exit is uniform across states and policy invariant. We are currently working on a version that generalizes exit.

[23]Technically, we are assuming that plants are myopic about this productivity growth, which will underestimate indirect effects of productivity growth. We are currently working on dropping this simplification.

[24]These contributions are computed taking differences between the baseline model contributions and the contributions for counterfactual 2. The remainders then give relative contributions where the reallocation term is computed using the fact that selection stays completely unaffected.

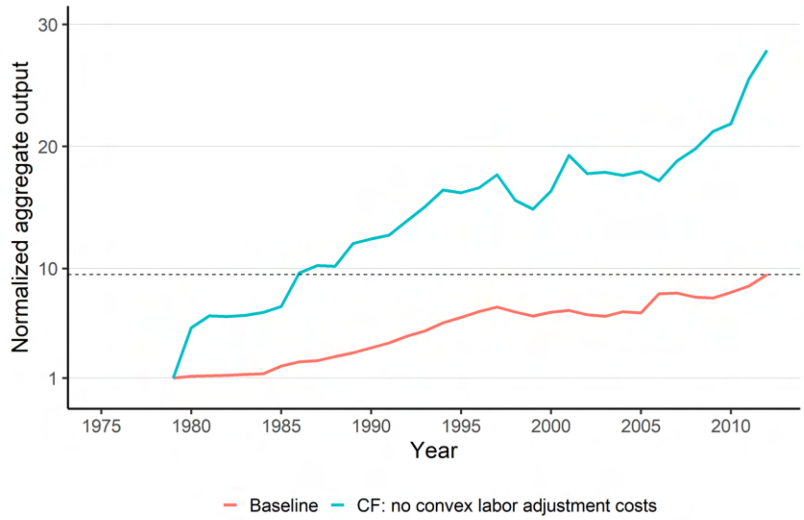

Figure 14: Counterfactual 3: Shut down convex adjustment costs. Output is for manufacturing only.

Indonesian manufacturing output growth. Faster transitions could mean that the economy reaches more quickly a steady state for total output, especially in the case where there is no aggregate productivity growth. However, due to the continuing presence of other adjustment costs and most importantly due to the large amount of entry of new plants, Indonesian manufacturing output - based on the model - has a large potential much beyond the current level in 2012. This is an important point that is relevant for current policy making as it quantifies the current costs of policies that prevent plants from reaching more quickly their optimal plant size.

The last row of Table 2 shows the results of the decomposition exercise when we shut-down convex adjustment costs. To quantify the role of convex adjustment costs, we again look at differences with respect to the baseline model. Output would be about three times larger in 2012 without convex adjustment costs and these output gains would be composed of 1/3 of input growth and 2/3 changes in reallocation of economic resources. Eliminating convex adjustment costs improves reallocation because it speeds up flows of labor h from low to high productivity plants. Moreover, it also induces a massive movement of workers from the rest of the economy to the manufacturing sector. Compared to the baseline model, the manufacturing employment share in 2012 would have been almost three times larger in the absence of convex adjustment costs. Since labor demand increases in the manufacturing sector and is better allocated in high productivity plants, the wages for all workers in the economy double in this counterfactual compared to the baseline.

5 Conclusion

This paper has shown how economic growth and changes in the plant size distribution coevolve over the course of development. By focusing on long-run evidence from Indonesian manufacturing, we showed how economic growth was driven by initial conditions and slow, but entirely predictable plant growth. The initial plant size distribution of Indonesian manufacturing in 1979 entailed the seeds of future economic growth. This is because the initial distribution featured few large plants and many medium-sized plants with a large growth potential. Output growth over the 40-year period we look at was driven mostly by general input growth due to these medium-sized plants slowly wanting to grow large and the corresponding reallocation of economic resources - mostly from other parts of the economy - towards these plants. Importantly, we find that aggregate productivity growth did not contribute to output growth in the Indonesian case. This stands in contrast to a large growth and development literature that has posited that long-run economic growth is driven by aggregate productivity growth.

Through the lens of a structural model that closely fits both plant size distributions and aggregate moments over time, we showed that Indonesian plants only grew slowly because of the presence of large convex adjustment costs, making it difficult to grow plants quickly. Together with an initial plant size distribution that we show - both through the model and a novel empirical exercise - to be far away from an ergodic distribution, this ensures persistent output growth over long periods of time. Importantly, we show that policies that reduce these constraints on plant growth imply large potentials for economic growth and there are large benefits to better understanding what drives the convexity in the adjustment costs. The adjustment costs we consider in this paper are general enough to likely capture a variety of constraints that prevent plants from growing quickly. Technically, the important lesson of our findings is that the aggregate costs of these adjustment costs depend crucially on the initial distribution of plants and focusing on steady states may be a poor approximation for reality and may greatly underestimate the true costs of adjustment costs - a lesson that has implications for various research areas such as the large literature on financial frictions (e.g. Buera and Shin 2013; Midrigan and Xu 2014; Moll 2014). At last, our results are in line with evidence of slow life-cycle growth of plants in developing countries (Hsieh and Klenow 2014) and the lack of large plants in developing countries (Hsieh and Olken 2014). The structural model allows us to view these stylized facts through the lens of the model and quantify their contribution to aggregate growth. Being able to study plant-level behavior in a structural model with flexible transition dynamics and growth offers the opportunity for studying in more detail the impact of a variety of policies and distortions and taking their study closer to the data.

6 References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6): 2411–51.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis) Allocation." *Journal of Political Economy* 122 (5): 1013–63.

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2017. "Exporting and Firm Performance: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics* 132 (2): 551–615.

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *The Quarterly Journal of Economics* 135 (1): 105–63.

Berman, Nicolas, Vincent Rebeyrol, and Vincent Vicard. 2019. "Demand Learning and Firm Dynamics: Evidence from Exporters." *Review of Economics and Statistics* 101 (1): 91–106.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review* 101 (5): 1964–2002.

Buera, Francisco J., and Yongseok Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy* 121 (2): 221–72.

———. 2017. "Productivity Growth and Capital Flows: The Dynamics of Reforms." *American Economic Journal: Macroeconomics* 9 (3): 147–85.

Cabral, Luis, and Jose Mata. 2003. "On the Evolution of the Firm Size Distribution: Facts and Theory." *American Economic Review* 93 (4): 1075–90.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro. 2019. "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock." *Econometrica* 87 (3): 741–835.

Clementi, Gian Luca, and Berardino Palazzo. 2016. "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations." *American Economic Journal: Macroeconomics* 8 (3): 1–41.

Cooper, Russell, Guan Gong, and Ping Yan. 2018. "Costly Labour Adjustment: General Equilibrium Effects of China's Employment Regulations and Financial Reforms." *The Economic Journal* 128 (613): 1879–1922.

Cooper, Russell, John Haltiwanger, and Jonathan L. Willis. 2015. "Dynamics of Labor Demand: Evidence from Plant-Level Observations and Aggregate Implications." *Research in Economics* 69 (1): 37–50.

David, Joel M., and Venky Venkateswaran. 2019. "The Sources of Capital Misallocation." *American Economic Review* 109 (7): 2531–67.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. "The Slow Growth of New Plants: Learning about Demand?" *Economica* 83 (329): 91–129.

Gollin, Douglas. 2008. "Nobody's Business but My Own: Self-Employment and Small Enterprise in Economic Development." *Journal of Monetary Economics* 55 (2): 219–33.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. "Capital Allocation and Productivity in South Europe." *The Quarterly Journal of Economics* 132 (4): 1915–67. https://doi.org/10.1093/qje/qjx024.

Hicks, Joan Hamory, Marieke Kleemans, Nicholas Y. Li, and Edward Miguel. 2017. "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata." National Bureau of Economic Research.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica: Journal of the Econometric Society*, 1127–50.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4): 1403–48.

———. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4): 1403–48.

———. 2014. "The Life Cycle of Plants in India and Mexico." *The Quarterly Journal of Economics* 129 (3): 1035–84.

Hsieh, Chang-Tai, and Benjamin A. Olken. 2014. "The Missing" Missing Middle"." *Journal of Economic Perspectives* 28 (3): 89–108.

Krusell, Per, and Anthony A. Smith. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy* 106 (5): 867–96.

Luttmer, Erzo GJ. 2011. "On the Mechanics of Firm Growth." *The Review of Economic Studies* 78 (3): 1042–68.

Maliar, Lilia, Serguei Maliar, and Fernando Valli. 2010. "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell–Smith Algorithm." *Journal of Economic Dynamics and Control* 34 (1): 42–49.

Melitz, Marc J., and Sašo Polanec. 2015. "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit." *The Rand Journal of Economics* 46 (2): 362–75.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104 (2): 422–58.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *American Economic Review* 104 (10): 3186–3221. https://doi.org/10.1257/aer.104.10.3186.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 1263–97.

Poschke, Markus. 2018. "The Firm Size Distribution Across Countries and Skill-Biased Change in Entrepreneurial Technology." *American Economic Journal: Macroeconomics* 10 (3): 1–41.

Restuccia, Diego, and Pedro Bento. 2015. "Misallocation, Establishment Size, and Productivity." In *2015 Meeting Papers*. 1391. Society for Economic Dynamics.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics* 11 (4): 707–20.

Ruhl, Kim J., and Jonathan L. Willis. 2017. "New Exporter Dynamics." *International Economic Review* 58 (3): 703–26.

Ruiz-García, J. C. 2019. "Financial Frictions, Firm Dynamics and the Aggregate Economy: Insights from Richer Productivity Processes." Working Paper.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing Like China." *American Economic Review* 101 (1): 196–233.

Timmer, Marcel, Gaaitzen J. de Vries, and Klaas De Vries. 2015. "Patterns of Structural Change in Developing Countries." In *Routledge Handbook of Industry and Development*, 79–97. Routledge.

Young, Eric R. 2010. "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell–Smith Algorithm and Non-Stochastic Simulations." *Journal of Economic Dynamics and Control* 34 (1): 36–41.

A Appendix

A.1 Further details on the estimation of aggregate productivity

To be added at a later stage.

A.2 Production Function Estimation (identification-estimation of θ)

In this section of the Appendix, we explain in more details our estimator for the production function elasticity θ. We discuss the generality of the estimator, how we estimate θ in practice as well as how we estimate standard errors. The key idea behind the estimator comes from looking at the policy functions implied by a model with adjustment costs as shown in Panel A of Figure 9. Each separate line gives the policy function for a specific idiosyncratic productivity realization in period t. Each dotted line gives the static unconstrained optimal choice of labor, which is independent of previous labor. Note that the dotted line in combination with the 45 degree line (which captures whether plant inputs grow or decline) separates the policy space into four regions. Importantly, optimal model-implied policies only happen to be in two of the four regions. In case a plant is above their optimal static unconstrained labor demand (as judged by current productivity and past labor demand), a plant always wants to weakly decrease their labor demand, while in the opposite case, a plant wants to weakly increase labor demand. The estimator exploits this geometric result by finding the θ that maximizes correctly classifying plants into these two regions and minimizing the number of plants that are classified into being in one of the other two regions.

We conjecture that the geometric result of optimal policies being concentrated entirely in the bottom left and top right region of Figure 9 holds more generally for many dynamic frictions. With flexible adjustment costs in labor, this holds for any set of economically reasonable parameters where the exact parameters only govern where plants should locate within these two regions. Similarly, the same geometric classification holds for financial frictions that prevent plants from choosing the unconstrained optimal input choice.

More formally, plants produce with:

$$y_{it} = TFP_{it} * l_{it}^{\theta}$$

They observe TFP_{it} before choosing labor demand l_{it}, but face flexible adjustment costs on labor demand l_{it} that depend on past labor demand l_{it-1} and potentially also other variables such as TFP_{it}. For any given θ, one can back out TFP_{it} using observed y_{it} and l_{it}. Furthermore, define the optimal unconstrained labor demand by

$$l_{it}^* = \left(\frac{\theta \text{TFP}_{it}}{w_t}\right)^{\frac{1}{1-\theta}}$$

where w_t captures observed real wages. Define the object $X_{it} \equiv l_{it}^* - l_{it-1}$. Then based on the economic model, plants will want to weakly increase their labor demand if $X_{it} > 0$ and weakly decrease their labor demand if $X_{it} < 0$. We can denote these indicator functions by $I_{it}^*(\theta)$ and $D_{it}^*(\theta)$ and their data counterparts by I_{it}^D and D_{it}^D. To find θ, the estimator minimizes the distance between indicators that are implied by the model and observed in the data. Specifically, the loss function we use is:

$$\text{Loss}(\theta) \equiv \sum_i \sum_t \omega_{it} \left[\mathbb{1}_{I_{it}^D=1} (I_{it}^D - I_{it}^*(\theta))^2 + \mathbb{1}_{D_{it}^D=1} (D_{it}^D - D_{it}^*(\theta))^2 \right]$$

where

$$\omega_{it} \equiv \frac{(|l_{it} - l_{it-1}|)^\varphi}{\sum_{it}(|l_{it} - l_{it-1}|)^\varphi}$$

To actually estimate θ, we take into account a number of additional constraints that complicate estimation of θ in practice and force us to make certain sample restrictions and assumptions. Based on our most preferred estimate, $\hat{\theta} \approx 0.58$. We visualize the robustness of our estimated θ with respect to these assumptions in Panels A-D in Figure 15. In each panel, we estimate θ over a grid by varying one assumption while keeping the other ones fixed at our most preferred choice. Unless otherwise noted, grey dotted lines denote the choices we make.

The first complication is that the geometric classification can fail in case plants have strong expectations about better or worse future productivity or cost realizations. In the case of labor adjustment costs, plants may anticipate that their future idiosyncratic productivity will deteriorate compared to their current one, in which case they might want to decrease their labor demand even if they should increase their labor demand based on the previous discussion. This leads to misclassification. This issue plays a role in case plants are at the upper end of the idiosyncratic productivity distribution and there is strong regression to the mean for productivity that is also anticipated by plants. Similarly, plants with low productivity may anticipate future increases in productivity and thus increase their labor demand nonetheless. This means that the issue is decreasing in the persistence of the productivity process; this issue vanishes in the limit where idiosyncratic productivity is completely persistent (unit root). By a similar argument, expectations about changes in future wages or aggregate productivity can also lead to misclassification. Hence, the estimator is likely to work much better in environments where aggregate productivity and wages are close to constant (and this is also anticipated by plants) and where idiosyncratic productivity shows high persistence. Fortunately,

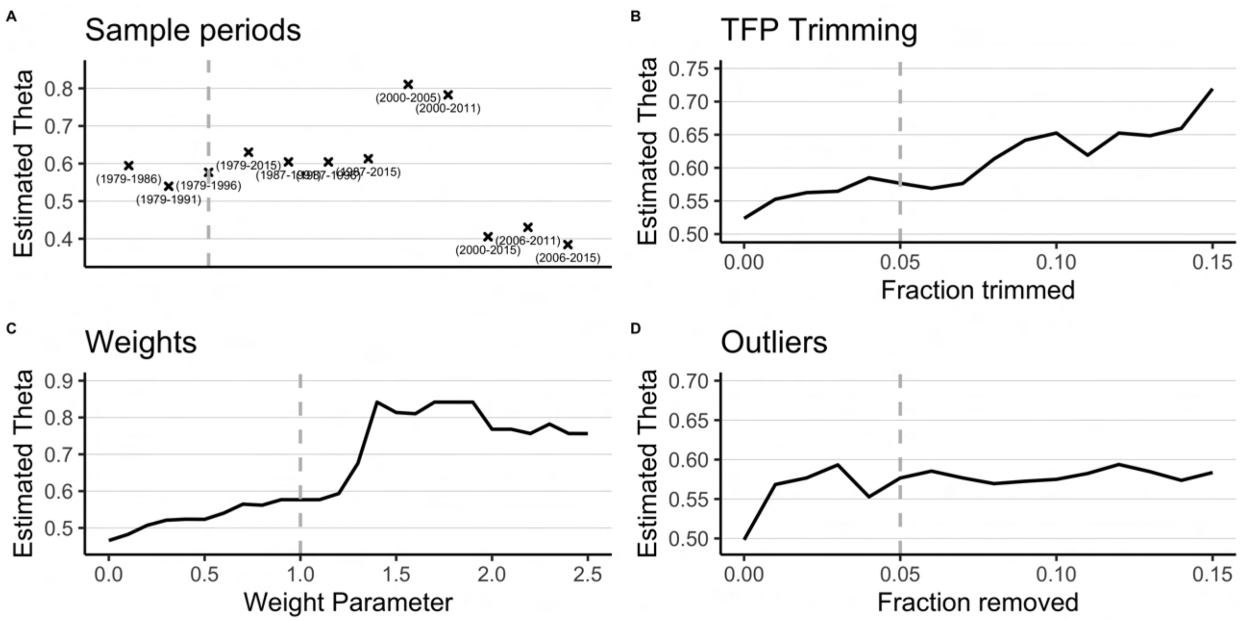

Figure 15: Robustness of production function parameter estimation. Panels A-D each vary one margin while fixing the other margins at the most preferred choice. Unless otherwise noted, grey dotted lines denote the choices we make for each margin. Panel A: Gives estimated elasticities for different sample periods (where time periods are reported in brackets below). Panel B: Considers the case of trimming extreme tails of the TFP distribution (conditional on the elasticity). Panel C: Varies the weight parameter for the weights as defined in the loss function. Panel D: Additionally considers dropping outliers by trimming a fraction of observations with the highest weights. Weights are well-defined for plants that are shrinking or growing so that such trimming will remove plants of both types.

measured real wages are fairly constant and estimated aggregate productivity (also for different values of θ) is highly persistent (see also the results in the later part of the Appendix). There is a trade-off between choosing a time period in which real wages and aggregate productivity are very stable and in choosing a long time period to maximize the number of observations. In the end, we chose as our most preferred setting the entire time period before the Asian Financial Crisis: 1979-1996. Panel A of Figure 15 varies the time-period for which θ is estimated. Estimates of θ are fairly similar for any time period before the Asian Financial Crisis, but differ when using mostly data for after the year 2000. The much lower estimates when including data from after 2006 could be caused by the issue of misclassification, by structural breaks in the production function parameters or outliers (as discussed further below).

Panel B of Figure 15 additionally trims the estimation sample by dropping x% of the top and bottom by productivity. Estimates are much more stable with respect to this change and are increasing in the amount of trimming. The reason is that plants that are dropped are more likely to be misclassified as growing their size (while they shrink in the data). We choose 5% trimming here as an intermediate or slightly conservative value.

Another complication is model-implied inaction. As shown in Figure 9, adjustment costs with fixed costs lead to inaction where plants keep labor demand constant. However, the importance of plants that are at or close to the inaction region will be determined by the weights as governed by the parameter φ. For positive φ, plants that are closer to the inaction region will be weighted less strong in the estimation of θ than plants that are further away. In Panel C, we vary φ and find that estimated θ also varies with φ. We choose the natural linear weight $\varphi = 1$, which seems to give an intermediate to slightly conservative value for θ. Additionally, we also considered dropping all plants in a specific distance to the inaction region and found that this leaves estimates for θ almost untouched. The reason is that with $\varphi = 1$, those plants will have very small weights anyway. For our most preferred estimate, we thus do not additionally drop any plants close to the inaction region. At last, we additionally drop outliers in terms of weights. The estimated θ might be disproportionately affected by a few plants who have very high weights. To obtain outlier-robust estimates, we drop x% of the plants with the highest weights (which captures both plants who grow and shrink strongly) and show results for varying x in Panel D. Again, we find stable estimates for θ and choose a value of 5% which gives an intermediate value for θ.

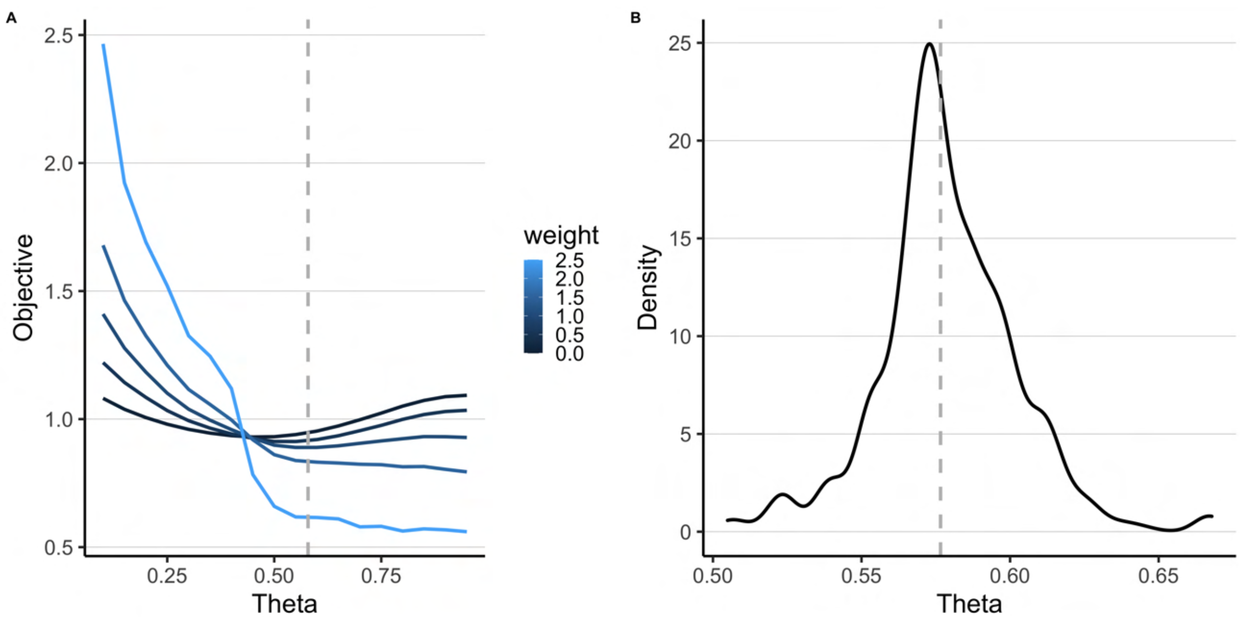

Figure 16: Visualizing identification of production function elasticity Theta. Panel A plots objective functions for different weights. The minimum of each objective function gives the respective estimate. The grey line gives the most preferred point-estimate. Panel B shows the distribution of bootstrap estimates. The grey line, again, denotes the point estimate.

In the end, our most preferred estimate is based on the time period 1979-1996, weights based on $\varphi = 1$, 5% TFP trimming and 5% of outlier trimming. The final estimation sample includes roughly 170k plant-year observations, or about 25% of all plant-year observations in the data. To quantify statistical uncertainty, we form 95% bootstrap confidence bands using the standard non-parametric bootstrap. This gives nonsymmetric confidence bands that range from 0.52 to 0.626. To show identification, Figure 16 reports two different results. Panel A formally shows identification by plotting normalized objective functions over values of θ for different weights φ. For each φ, the objective function has a unique minimum, proving identification. Furthermore, the problem is smooth and well-behaved. Panel B reports the bootstrap distribution to show the adequacy of using the nonparametric bootstrap for inference. While showing general validity of the bootstrap procedure is beyond the scope of this paper, we note that this is a smooth problem (as shown in Panel A), suggesting that bootstrapping should work well here. Looking at the distribution of bootstrap estimates of θ suggests that the distribution has well-behaved tails (based on 500 bootstrap estimates, all estimates fall between 0.5 to 0.67) and is unimodal. Based on the estimated θ, one important point needs to be highlighted: $\hat{\theta} = 0.58$ is considerably larger than the average labor share in the estimation sample (around 0.46). Through the lens of the theoretical model, this means the estimator captures a situation in which plants are systematically below their optimal plant size.

To end this section, we have also alternatively considered rewriting the estimator as a Maximum Likelihood estimator. Given the binary nature of model-predictions being entirely based on the sign of X_{it}, one could consider a Probit estimator where plants observe X_{it}, but make a decision based on $X_{it}+\varepsilon_{it}$. However, given the scaling of X_{it}, this is not a well-defined problem as long as all plants draw from the same distribution of shocks. Alternatively, one could consider rescaling X_{it} by focusing for example on growth rates. However, this gives too small weights to large plants and too large weights to small plants. Estimates based on the Maximum Likelihood procedure turned out to give unrealistic parameter estimates ($\theta < 0.05$), which led us to discard the Maximum Likelihood estimates.

A.3 Non-parametric identification of exit

In this part of the Appendix, we describe in more detail how our estimation approach allows for the non-parametric identification of endogenous plant exit. We mean by this that we can allow for an arbitrary operating cost draw function G that may also depend on other variables in the state-space including idiosyncratic productivity s and labor h_t. The idea is simple ...

our identification of plant exit

flexibly nests standard models of endogenous plant exit as used in the plant dynamics literature (cite).

In the standard setting, plants endogenously choose to exit according to: (describe setup)

Show how we can allow for any distribution of the cost draw (iid), Without any further assumptions on the

A.4 Further data

Using the Indonesian manufacturing survey data, we also draw on a measure of a plant's capital stock based on Cali-Presidente (2021). This includes using the sector-specific wholesale price indices and differentially deflating five main components of capital (land, buildings, machines, vehicles, others). The capital stock series obtained from Cali-Presidente (2021) draws primarily on self-reported capital stocks by plants, but drops observations that do not pass a battery of consistency checks. Missings are then filled up by drawing on the perpetual inventory method (PIM) using reported capital investments and assumptions on depreciation rates. We are happy to provide further details on these data cleaning and construction steps upon request.

The second dataset we draw on is at the worker level and is based on the Indonesian Family Life Survey, a large continuing longitudinal socioeconomic and health survey. We use this dataset to obtain a robustness measure of the evolution of manufacturing worker wages. The individual-level panel tracked households and individuals throughout five waves (1993, 1997-98, 2000, 2007-08, and 2014-2015), representing about 83% of the Indonesian population. This data includes detailed employment information, including questions on previous employment. As in Hicks et al. (2017), this allows to create up to a 28-year annual individual employment panel from 1988 to 2015.[25] Moreover, the survey tracks information on principal and secondary employment and includes information on both formal and informal sector employment. Throughout, we restrict ourselves to manufacturing workers. There are 3,390 unique manufacturing workers in the data and a total of 11,374 worker-year observations. As our earnings measure, we use the sum of all wages, profits, and benefits obtained in the current job (as in Hicks et al. (2017)). The data also consistently tracked how many hours individuals work per week and how many weeks they work in a year for each of their jobs, which allows to construct an annual measure of hours worked and thus the construction of hourly income, which we use in this paper as our measure of wages. We further deflate this wage series by an aggregate CPI (since the sector-specific wholesale price index cannot be readily mapped to this series).

[25]Employment status and sector of employment are available for each year, but in the fourth and fifth IFLS round, earnings were collected only for the current job.

A.5 Decomposition of aggregate growth

A.5.1 Formal derivation of main accounting identity

$$
\begin{aligned}
Y_t &\equiv \sum_i y_{it} \\
&= \sum_i z_t s_{it} f(x_{it}) = \sum_i z_t s_{it} f(x_{it}) \frac{\sum_i f(x_{it})}{\sum_i f(x_{it})} \\
&= z_t * \sum_i f(x_{it}) * \sum_i s_{it} \frac{f(x_{it})}{\sum_i f(x_{it})} \\
&= z_t * \sum_i f(x_{it}) * \sum_i (s_{it} - \bar{s}_t + \bar{s}_t) \left(\frac{f(x_{it})}{\sum_i f(x_{it})} - \frac{1}{N_t} + \frac{1}{N_t} \right) \\
&= z_t * \sum_i f(x_{it}) * \left[\bar{s}_t + N_t cov \left(s_{it}, \frac{f(x_{it})}{\sum_i f(x_{it})} \right) \right]
\end{aligned}
$$

$$
ln(Y_t) = ln(z_t) + ln \left(\sum_i f(x_{it}) \right) + ln \left(\bar{s}_t + N_t cov \left(s_{it}, \frac{f(x_{it})}{\sum_i f(x_{it})} \right) \right)
$$

A.5.2 Detailed year-to-year results

Will add details on year-to-year decomposition results in a later version of the paper.

A.5.3 Robustness with respect to output elasticity θ

As one key robustness exercise, we consider the importance of the output elasticity θ. Given that this parameter is the main input in the decomposition exercise and determines the relative role of productivity versus inputs, one might expect that the decomposition results heavily depend on θ. This is even more of a concern given the lack of a good identification strategy for θ. Fortunately, it turns out that the main results are surprisingly little affected by the actual choice of θ. Specifically, we consider a grid of values for θ between 0.4 and 0.9, nesting all commonly used estimates in the literature. For each value of θ, we repeat our identification strategy of aggregate productivity using older plants. We will add the corresponding graph in a later version of the paper.

A.6 Further details and robustness exercises for Section 2.4. (initial conditions and slow transitions)

A.6.1 Further details on main empirical exercises

While the first empirical exercise given by the lines "1975 Hypothetical" is explained in the main text, the other two empirical exercises are not fully explained. These exercises account for entry and exit. To begin with, note that entry and exit is potentially very important, especially if entering plants differ from exiting plants. Of the roughly 6,800 plants operating in 1975, less than 12% were still operating in 2015. On the other hand, as shown in Figure 2, the number of active plants increased by a factor of 3 between 1975 and 2015. This means that the vast majority of active plants in 2015 did either not exist or was not captured in the 1975 census. To capture the role of entry and exit, we amend the previous exercise by including a state-0 which captures inactive plants or potential entrants. This means that both the initial distribution is defined over an additional state-0 and the transition matrix will feature transitions into (exit) and out of state-0 (entry). To construct the new transition matrix, we can use observed entry and exit flows. Since transition matrix entries are computed as the share of flows from bin x in period t into any other bin in period $t+1$, we can readily compute transitions from an active state to an exit state. However, we cannot directly compute entries from inactivity, because the baseline is fundamentally undetermined. We do not know how many inactive or potential plants there are. This means we can also not directly compute the new initial distribution that includes the measure of plants in state-0. Since both the transition matrix and the initial distribution depend on the number of inactive plants, this number cannot be identified from observables in the first two periods alone. In theory, we can pin down the initial number of inactive plants by enforcing that the transition matrix stays constant over time and by feeding in another moment, the change in the number of plants between 1976 and 1977. However, as can be seen from Figure (?), the initial periods saw an initial decrease in the number of plants between 1975-1976 and a subsequent increase between 1976-1977. To match this pattern, we would have to enforce a negative transition matrix entry for staying inactive.

To avoid this, while giving almost indistinguishable results, we instead assume that the share of inactive plants that stay inactive is 0. This identifies the transition matrix and we then consider two additional exercises where we keep this transition matrix fixed. In the first version of the exercise with entry and exit, we simply iterate on the initial distribution and the transition matrix. This keeps the total number of plants (inactive + active) constant, while introducing interesting entry and exit dynamics that directly affect the evolution of the plant size distribution over time. Results for this exercise are given by the lines "Hypothetical 1975 (EE)" (where EE stands for entry and exit). While the long-run results are almost unchanged to the previous results, introducing entry and exit does speed up transition dynamics considerably, providing a much better out-of-sample fit for the early transition period. This is driven by observed exiting plants being smaller and less productive than observed new entrants in 1976. With a positive share of inactive plants staying inactive each period would slow these predicted transition dynamics down.

In the second version of the exercise with entry and exit, we additionally vary the number of plants that enter each period. Specifically, we exactly match the increase in the number of active plants over time as shown in Figure 2, while taking information on new entrants and exits only from 1975 and 1976. In contrast to the previous exercise with entry and exit, here we do take limited information on future plant entry and thus it does not lend as well to predicting future changes in the plant size distribution. However, this exercise gives a more complete picture of the importance of entry and exit observed in the data. Results are given by the lines denoted "Hypothetical 1975 (EE growth)." The series again behave very similarly as before, but we can more clearly see that important year-to-year fluctuations in the real data series are driven by entry shocks. For example, the inclusion of many more plants in 1985 had important medium- to long-run effects on the evolution of the size distribution.

A.6.2 Robustness results

We have robustness results for:

- varying initial distributions and transition matrices
- averaging transition matrices over multiple years
- varying assumptions on entry and exit behavior
- same exercise for real value added output and the labor wage bill

We will include these results in a later version of the paper.

A.7 Algorithm to solve for Recursive Competitive Equilibria

In this part of the Appendix, we provide detailed pseudo-code for how we compute equilibria based on our model. We first detail how we compute equilibria in the baseline economy (for which we observe one path of equilibria) and then discuss how we compute counterfactual paths of equilibria.

A.7.1 Baseline equilibria

The algorithm we propose to solve for equilibria follows closely the constrained rational expectations algorithm in Krusell and Smith (1998). There are two differences. The first main difference is with respect to the outer loop of iterating over the forecasting rule. Here, we directly estimate the forecasting rule based on the sequence of observed equilibria in the data. This can be done without ever solving a model and we do not iterate on this forecasting rule, leading to large computational gains in comparison to the classic Krusell and Smith (1998) algorithm. For this approach to be consistent, we need to ensure that the observed path of equilibria (including the same prices and evolution of other aggregate state variables) will actually be generated by our model. To ensure this, we introduce a time-varying wedge that scales model-implied policies and that, as standard in the wedge-accounting literature, are not known nor expected by agents.

The second difference is with respect to the inner loop for which any forecasting rule is fixed. Here, we follow closely the standard algorithm, but make small adjustments. To explain these small adjustments, first note that we assume a forecasting rule of the following form:

$$w_{t+1} = f(w_t, z_t, z_{t+1})$$

Prices are only a function of the past wage, past and current aggregate productivity as for example in Clementi and Palazzo (2016). We abstract from including further state variables because (1) they make individual value functions non-stationary, greatly increasing computational costs, and (2) their influence seems to cancel out as prices turn out to be relatively stable over time. As in Clementi and Palazzo (2016), this forecasting rule only enters for future expectations over prices. Hence, the value function writes:

$$V(w_t, z_t, h_{i,t-1}, s_{i,t}) = \max_{h_{i,t}} \pi(w_t, z_t, h_{i,t-1}, s_{i,t}) + \frac{1 - \lambda(w_t, z_t, h_{i,t}, s_{i,t})}{R} \mathbb{E}_{s', z', \tilde{w}'} \Big[V(\tilde{w}_{t+1}, z_{t+1}, h_{i,t}, s_{i,t+1}) \Big]$$

We follow the standard approach to discretize the price variables (here: w_t) so that all state variables are on a grid and then iterate on an initial guess of the value function over this entire grid (e.g. Maliar, Maliar, and Valli 2010; Young 2010). As in Clementi and Palazzo (2016), we interpolate between grid points for the wage forecasting, because the forecasting rule gives predictions of the future wage that may not lay on the grid of w_t. Specifically, call the predicted wage $\tilde{w}'$ which lies between two grid points w_i and w_{i+1}. Then, define $J(w_i|w, z, z') = 1 - \frac{\tilde{w}' - w_i}{w_{i+1} - w_i}$ and $J(w_{i+1}|w, z, z') = \frac{\tilde{w}' - w_i}{w_{i+1} - w_i}$ (and $J(w_j|w, z, z') = 0$ for all others). We then use $J(w'|w, z, z')$ as weights in computing the future expected value of incumbents.

Specify each step of the algorithm.