THÈSE de DOCTORAT

de l'UNIVERSITÉ TOULOUSE CAPITOLE

Présentée et soutenue par

Monsieur Jonas GATHEN

Le 2 juillet 2024

Essais en Macroéconomie du développement

École doctorale : **TSE Toulouse Sciences Economiques**
Spécialité : **Sciences Economiques - Toulouse**
Unité de recherche : **TSE-R - Toulouse School of Economics - Recherche**

Thèse dirigée par *Monsieur Fabrice COLLARD et Monsieur Stéphane STRAUB*

Composition du jury

Rapporteur : M. Francisco BUERA
Rapporteur : M. Josep PIJOAN-MAS
Examinateur : M. Matteo BOBBA
Directeur de thèse : M. Fabrice COLLARD

Essays in Macro Development

Ph.D. Thesis

Jonas Gathen[1]

Toulouse School of Economics

July 2024

[1] Address: 1, Esplanade de l'Université, 31080 Toulouse, France; email: jonas.gathen@tse-fr.eu

Acknowledgements

I started at TSE in the fall of 2017 for the M1 in Economics, close to seven years ago. Throughout the Masters and the PhD, TSE became my academic home; a home that I will dearly miss. My first gratitude thus goes out to TSE as an institution, the fantastic academic environment that it provided me over the years, including large cohorts of young, enthusiastic and bright PhD students and the many excellent faculty expanding the research frontier, all the different research groups and institutes that always left me in awe of all the things I don't (yet) know, and the excellent line-up of seminars and workshops letting me learn new things every day. I started seven years ago eager to learn as much as I could and TSE gave me an amazing environment to do just that. A big thanks thus goes to all the people that made this possible, from the Masters and PhD directors, the seminar organisers and heads of the research groups, to all the administrative staff involved in letting TSE function as a research hub.

Next, I want to thank my advisors Fabrice Collard, Stéphane Straub, Christian Hellwig and more recently also Matteo Bobba, as well as people who encouraged me throughout the PhD. In hindsight, I believe that setting me on the right path from the beginning and raising my academic aspirations throughout the PhD was instrumental for the success of my PhD. I remember Fabrice encouraging me during my ETE thesis, and Alexandre Gaillard as my Macro TA taking my half-baked ETE thesis ideas seriously, giving me invaluable feedback and pushing me to aim for more. Christian also met with me early on, raised my aspirations for my projects and proposed for me to meet with seminar speakers. Stéphane opened the door for the Indonesian plant-level data that followed me throughout my PhD and also greatly encouraged me with my early projects, two of which now show up as thesis chapters. As my PhD progressed, Fabrice and Christian both were instrumental by pushing me to submit my first paper and aiming for a top publication, and to visit LSE. Thank you so much for that.

While the PhD was a lot of work, I also want to acknowledge how lucky I feel for it having been mostly a really fun experience. Being able to spend most of the Covid time with my partner in Switzerland, spending a semester at LSE, the collegial and close-knitted academic environment that the Macro group at TSE offered, and amazing office mates, co-authors, roommates and friends in Toulouse and around the world all together ensured that I was enjoying most of my PhD. So a special thanks goes to Franzi and all my friends in Zürich and from my undergraduate, to Oscar, Jérôme and Till for having been such great roommates and friends, to Thi, Lukas, Li, Yixin, Oscar J., Johanna, Chanwoo, Pascal and all my other friends from Toulouse, to all the people and friends that made my stay at LSE such a wonderful experience, to the STEG organisers who helped me find an academic home in Macro Development, and finally to all the students and faculty in the Macro group for all their feedback and interesting discussions. Thank you.

Summary

What drives economic growth and development and how do poorer countries become richer? This fundamental question on the Macroeconomics of Development is the underlying motivation for this thesis. The three chapters focus on different aspects of the development process.

In the first chapter, co-authored with Oscar Fentanes, we ask how to quantify the effect of policies in economies where firms face adjustment frictions and the economic environment changes constantly. In such settings, at any given point in time, the economy responds to new changes in policy while still adjusting to previous changes. We show how to empirically disentangle the two using a structural model of plant dynamics and standard plant-level panel data. We apply our approach to the Indonesian Growth Miracle from 1975 to 2015, estimating the model on 40-years of micro data along the observed growth path without assuming that the economy is ever at a steady state. We find that growth from catching-up to previous changes in the economy is key and does not get less important over time, because adjustments are slow and economy-wide changes frequent.

The second chapter quantifies the aggregate costs of political connections using a general equilibrium model in which politically connected firms benefit from output subsidies and endogenously spend resources on rent-seeking activities. The model is structurally estimated using rich firm-level data for the Indonesian manufacturing sector and a firm-level measure of political connectedness based on a natural experiment. I find that subsidies to connected firms are too high and dispersed, costing the economy between 1.0-4.7% of aggregate output.

The third chapter in my thesis provides a new approach to estimate government worker skills, a key input to study how the composition and selection of government workers drives state capacity and development. The approach is applicable in settings where government output is unobserved and government wages are uninformative about skill differences. The three-step approach first estimates skills from wages in comparable jobs in the private sector, then relates these skills to skill-related observables using Machine Learning tools and finally predicts government worker skills out-of-sample. I apply the new estimation approach to rich Indonesian household-level panel data from 1988 to 2014, showing two main applications. First, I quantify that government workers are highly selected, that their skills have increased, but that their skill premium has declined over time. Second, I analyze government wage setting: the Indonesian government pays a wage premium of at least 30% conditional on skills, about 1/3 of which is driven by the large gender wage gap in Indonesia's private sector.

Résumé

Quels sont les moteurs de la croissance économique et du développement et comment les pays pauvres deviennent-ils plus riches ? Cette question fondamentale sur la macroéconomie du développement est la motivation sous-jacente de cette thèse. Les trois chapitres se concentrent sur différents aspects du processus de développement.

Dans le premier chapitre, coécrit avec Oscar Fentanes, nous nous demandons comment quantifier l'effet des politiques dans les économies où les entreprises sont confrontées à des frictions d'ajustement et où l'environnement économique change constamment. Dans ces conditions, à tout moment, l'économie réagit aux nouveaux changements de politique tout en continuant à s'adapter aux changements précédents. Nous montrons comment démêler empiriquement ces deux aspects à l'aide d'un modèle structurel de la dynamique des usines et de données de panel standard au niveau des usines. Nous appliquons notre approche au miracle de la croissance indonésien de 1975 à 2015, en estimant le modèle sur 40 ans de données microéconomiques le long de la trajectoire de croissance observée, sans supposer que l'économie n'a jamais atteint un état stable. Nous constatons que la croissance due au rattrapage des changements antérieurs dans l'économie est essentielle et ne perd pas de son importance au fil du temps, car les ajustements sont lents et les changements à l'échelle de l'économie fréquents.

Le deuxième chapitre quantifie les coûts globaux des relations politiques à l'aide d'un modèle d'équilibre général dans lequel les entreprises ayant des relations politiques bénéficient de subventions à la production et dépensent de manière endogène des ressources dans des activités de recherche de rente. Le modèle est estimé de manière structurelle à l'aide de riches données au niveau de l'entreprise pour le secteur manufacturier indonésien et d'une mesure de la connectivité politique au niveau de l'entreprise basée sur une expérience naturelle. Je constate que les subventions accordées aux entreprises connectées sont trop élevées et dispersées, coûtant à l'économie entre 1,0 et 4,7% de la production globale.

Le troisième chapitre de ma thèse propose une nouvelle approche pour estimer les compétences des fonctionnaires, un élément clé pour étudier comment la composition et la sélection des fonctionnaires influencent la capacité et le développement de l'État. Cette approche est applicable dans des contextes où la production du gouvernement n'est pas observée et où les salaires du gouvernement ne donnent pas d'informations sur les différences de compétences. L'approche en trois étapes consiste d'abord à estimer les compétences à partir des salaires versés pour des emplois comparables dans le secteur privé, puis à relier ces compétences à des éléments observables liés aux compétences à l'aide d'outils d'apprentissage automatique et enfin à prédire les compétences des fonctionnaires hors échantillon. J'applique la nouvelle méthode d'estimation à de riches données de panel indonésiennes au niveau des ménages de 1988 à 2014, et je montre deux applications principales. Premièrement, je quantifie le fait que les fonctionnaires sont hautement sélectionnés, que leurs compétences se sont accrues, mais que leur prime de compétence a diminué au fil du temps. Deuxièmement, j'analyse la fixation des salaires des fonctionnaires : le gouvernement indonésien verse une prime salariale d'au moins 30% en fonction des compétences, dont environ un tiers est dû à l'écart salarial important entre les hommes et les femmes dans le secteur privé indonésien.

Contents

List of Figures

List of Tables

Chapter 1

Making a Growth Miracle: Historical Persistence and the Dynamics of Development

Oscar Fentanes[1] & Jonas Gathen[2]

Abstract

What explains growth miracles? We argue that growth miracles are driven by a fundamental race: as the economy tries to catch-up to its steady state, changes in the economic environment move the steady state itself and provide new potential for catch-up growth. We quantify this race over the course of development using 40 years of plant-level manufacturing panel data from Indonesia and a structural model of plant dynamics. We estimate the model on the micro data along the observed growth path without assuming that the economy is ever at a steady state. While catch-up growth starting from initial conditions in 1975 accounts for 42% of Indonesia's subsequent industrialization, new changes in the economy induce new catch-up growth. In the end, the economy never catches up.

JEL Codes: D25, O11, O14, O47

Keywords: Development, industrialization, growth miracles

[1]Toulouse School of Economics; oscar.fentanes@gmail.com.

[2]Toulouse School of Economics; j.gathen@posteo.de. We are grateful for the continued support and many useful comments by our advisors: Fabrice Collard, Matteo Bobba, Christian Hellwig and Stéphane Straub. We further thank in no particular order: Ben Moll, Yongseok Shin, Francisco Buera, Ezra Oberfield, Richard Rogerson, Jihyun Kim, Eugenia Gonzalez-Aguado, Tuuli Vanhapelto, Christophe Bruneel-Zupanc, Miguel Zerecero, Alexandre Gaillard, Andreas Schaab, Gyung-Mo Kim and all participants of the TSE Macro workshop. 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 feedback during presentations at Bristol, CREi, IIES, SSE, Lausanne, Tilburg, Cergy Paris, Konstanz, Princeton, WashU St Louis, the Exeter Structural Change Workshop 2024, EWMES 2022, the Annual STEG conference 2022, EEA/ESEM 2022, ENTER jamboree 2021 as well as the PhD and Macro workshops at TSE and LSE. All remaining errors are our own.

1.1 Introduction

Over the past 50 years, extended periods of rapid economic growth in China, India and Indonesia alone lifted roughly 1 billion people out of extreme poverty (World Bank 2023). What drives such growth miracles? A common view is that growth miracles capture a transition process: a permanent policy regime change sets an initially poor and highly misallocated economy on a transition path towards a better long-run equilibrium (Buera and Shin 2013; Asturias et al. 2023). These transitions take time because frictions prevent labor and capital to quickly reallocate across firms and sectors. But if transitions take time, how does this view account for new changes in policies while the economy is still transitioning? And how can we identify the aggregate effects of new policies if the economy is still adjusting to previous policies?

In this paper, we show evidence that instead of a one-time transition, growth miracles are driven by a never-ending race: as the economy tries to catch-up to its steady state, changes in the economic environment move the steady state itself and induce new transition growth. We do so by looking at Indonesia, the fourth most populous country in the world. Indonesia provides an ideal case to study this race since we can draw on almost half a century of manufacturing plant-level panel data during which the Indonesian economy completely transformed: GDP per worker increased five-fold, the working population tripled, output in manufacturing grew 30-fold and the manufacturing employment share doubled.

Drawing on the micro data, we provide empirical evidence that motivate a model of plant dynamics in which growth is driven both by changes in the economic environment and transition phases. Empirically distinguishing these two drivers of growth is crucial because one might otherwise falsely attribute current growth to current policy changes. To do so, we estimate the model on the micro data along the observed growth path without assuming that the economy is at a steady state at any point in time. Intuitively, observed conditional choices of plants identify changes in the economic environment, while the distribution over these choices summarizes the past and reveals the potential for transition growth.

In line with the common view of growth miracles, we find that initial transition growth is important: letting the economy in 1975 transition while shutting down all future changes in the economic environment explains 42% of the manufacturing growth between 1975 and 2015. However, we also find that the Indonesian economy in 2015 is not closer to its steady state than it was in 1975, precisely because new changes in the economic environment in the meantime moved the steady state itself. As the economy is always undergoing important transition processes, one key implication is that evaluating policies without considering these adjustments is highly misleading.

We now provide further details and results. Drawing on our data, we document four main facts that help us disentangle transition growth from changes in the economic environment and motivate our subsequent model:

Fact 1: Rapid economic growth coincided with changes in the plant distribution. Average plant size doubled, the mass of plants increased four-fold and the right tail of the plant size distribution thickened.

Fact 2: Adjustment processes can account for changes in the plant distribution. These are respectively: An aging of the plant distribution together with the fact that plants enter small and grow over their life cycle, slow entry and exit dynamics, and the fact that it takes time to grow large plants.

Fact 3: The drivers of aggregate productivity differed markedly before and after the Asian Financial Crisis. Before the crisis, productivity across plants was driven entirely by the selection of more productive plants, informed by meager within-plant productivity growth. In contrast, the post-crisis growth period was characterized by strong within-plant productivity growth and little growth due to selection.

Fact 4: The allocation of resources did not improve systematically over time. This is robust to different measures of misallocation. We find evidence for volatile productivity dynamics at the plant-level, large changes in entry, and plant-level adjustment frictions – particularly in labor – from an event study design that can jointly account for this.

To quantify the race between transition growth and changes in the economic environment, we draw on a model of plant dynamics in the tradition of Hopenhayn (1992) that is motivated by the previous empirical facts. In the model, firms face risk regarding their productivity, choose to enter and exit and hire labor and capital subject to adjustment frictions that lead to the slow accumulation and reallocation of resources across sectors and firms. We embed these plant dynamics into a two-sector economy to capture the endogenous reallocation of workers across manufacturing and the rest of the economy. The main frictions in manufacturing are labor adjustment costs, in particular convex costs, that prevent plants from growing a large workforce quickly (*Facts 2 + 4*). Plants also endogenously enter and exit based on drawing entry costs and fixed costs of production. The level and dispersion of costs in turn rationalize the observed speed of entry and exit dynamics (*Facts 2*). These features imply that with an initial distribution characterized by few but productive young plants, the economy goes through a process of transition growth as plants gradually grow, more plants gradually enter over time and unproductive plants gradually exit. At any given point in time, the model economy is characterized by a set of exogenous model fundamentals and the state of the current economy as captured by the distribution of plants. Model fundamentals include all cost parameters as well as time-varying aggregates such as aggregate labor supply and technological changes in manufacturing (*Fact 3*) and the rest-of-the-economy. Policy affects growth through driving part of the changes in model fundamentals. Technically, we assume that plants make dynamic choices forming rational expectations over their future idiosyncratic risk but have perfect foresight over future aggregate changes in the economy. This introduces a computationally difficult fixed point problem: plants' dynamic choices depend on expectations over the future path of market-clearing prices, which in turn depend on the endogenous evolution of the entire distribution of plants (as in Buera and Shin 2013).

The key methodological contribution of the paper is to propose a tractable estimation strategy that allows to estimate this model economy on standard plant-level micro data along the growth path in the data without assuming that the observed economy is at a steady state at any point in time. Importantly, our model estimation allows model fundamentals to vary flexibly over time, making it particularly suited to study fast-changing economies and markets. The model estimation proceeds in three main steps that allow to distinguish transition growth from changes in fundamentals and make the computational costs of the estimation independent of the computational costs of solving for a path of model equilibria. In the first step, we identify the path of time-varying equilibrium prices – only wages in our case – along the observed growth path (e.g. as in: Gopinath et al. 2017). Given that our model can account for this equilibrium path, we can treat the path of wages as fixed throughout the estimation and thereby avoid solving for the computationally costly fixed point in the path of equilibria. In the second estimation step, we identify the distribution of plants over the state space of the model, summarizing the history of the economy. In this step, we estimate plant production functions (Ackerberg, Caves, and Frazer 2015; Demirer 2020) and propose a novel strategy to separate plant-level productivity into an idiosyncratic and a common aggregate technology component, allowing us to distinguish selection-driven from technology-driven productivity growth. In the third and last estimation step, we estimate the model parameters that govern plants' adjustment frictions. We do so by drawing on Euler equation Continuous Conditional Choice (CCC) estimation, exploiting observed conditional input and exit choices of plants and avoiding to solve a dynamic programming problem to compute model-based dynamic input choices (Hotz and Miller 1993; Bajari, Benkard, and Levin 2007). In this last step, we estimate sizable convex adjustment costs in our model, which are identified from the empirical pattern that even small but highly productive plants grow their labor force gradually over time.

Using the estimated model, we find that transition growth from starting the economy with initial conditions in 1975 and shutting down all future changes in model fundamentals explains 42% of subsequent manufacturing output growth and all of the aggregate welfare increases that are due to changes in manufacturing over time. Given an initial distribution that features young and small plants, sizable labor adjustment frictions and slow entry and exit dynamics, it takes the economy 26 years to reach 90% of the steady state manufacturing output. Importantly, transition growth remains an important driver of growth precisely because the economy's fundamentals continue to change. To quantify this point, we repeat the previous exercise to compute the transition path for each year, starting from each year's initial distribution and model fundamentals. We find that the economy does not get closer to its (time-varying) steady state. It takes on average 20 years to reach 90% of the steady state manufacturing output and – if anything – the time it takes increases over time. Based on our results, we can thus strongly reject the idea that transition growth is a transitory phenomenon.

Large changes in fundamentals are key to explain the continuing importance of transition growth. The structural model allows us to quantify the role of changes in fundamentals in Indonesia's manufacturing growth miracle and quantify how much of their effect can be explained by changes in observed government policy. We do so by focusing on two important changes in fundamentals that can be linked to development policies that the Indonesian government also pursued to varying degrees over the 40 years we study: (1) large-scale investments in education that raise the pool of skilled (and cheap) labor, and (2) the active use of FDI policy to attract manufacturing plants under foreign ownership.

We find that the manufacturing growth miracle would not have happened in the absence of the estimated doubling in human capital per worker, because labor would have been more expensive in this economy and manufacturing plants are far more sensitive to higher wages than the rest of the economy. To gauge the importance of policy in driving overall human capital increases, we then evaluate Indonesia's largest school construction program (INPRES) through the lens of the model. Building on micro-empirical evidence on the wage effects (Duflo 2001, 2004), the scale of the program (Akresh, Halim, and Kleemans 2023), and the slow labor market integration of treated cohorts, we show that by 2015, the program accounts for roughly

10% of the overall manufacturing output growth that is due to human capital per worker increases in the economy.

In contrast, for FDI, we find that manufacturing output in 2015 would have only been 8% lower in the complete absence of foreign-owned entrants, while we find that regulatory changes in FDI policy in the late 1980s may potentially account for 85% of the overall effect of FDI on manufacturing. Taken together, this still means that most growth by far stems from structural forces related to demographics. Thus, a somber conclusion based on these results – partly resonating related work on the Indian growth miracle (Bollard, Klenow, and Sharma 2013) – is that policy matters less for growth than we might think.

Related literature

We contribute to four main strands of literature. First and most importantly, we complement the growth and macro development literature by studying a model of firm dynamics where growth is driven by the combination of changes in exogenous fundamentals and transition growth. This is in contrast to a firm dynamics literature that has mostly analyzed development differences through differences in steady states.[3] Much fewer papers study transition growth with firm dynamics (e.g. Buera and Shin 2013; Moll 2014; Akcigit, Alp, and Peters 2021; Ruggieri 2022; Asturias et al. 2023; Lanteri, Medina, and Tan 2023). We add to this literature by (1) allowing for further changes in fundamentals along the transition and by (2) estimating this model on the micro data along the transition without assuming that the economy is at a steady state at any point in time. Quantitatively, we find that the combination of both sources of growth matters. Not only are initial transition growth and changes in fundamentals important for growth, but the economy is always far away from its (time-varying) steady state, questioning the usefulness of either comparing steady states or focussing on transitions in the absence of further changes in fundamentals. Apart from the methodological differences, Buera and Shin (2013) and Asturias et al. (2023) are the most closely related in their focus on understanding growth miracles. Our results mainly differ from Buera and Shin (2013) in that we find no role for reductions in frictions and misallocation – their main driver of transition growth – but rather a key role for plant and worker demographics in driving transition growth. We return to Asturias et al. (2023) below.

Second, both modeling and key results in this paper relate to the recent quantitative spatial, trade and migration literature, which focus on frictional worker mobility and trade while abstracting from firm dynamics. For example, the idea that transitions take long and the economy is persistently far away from its steady state resonates with recent findings from Allen and Donaldson (2020) and Kleinman, Liu, and Redding (2023). This literature relates to and builds on the seminal work of Caliendo, Dvorkin, and Parro (2019) who also study the combination of changes in exogenous fundamentals and transition growth. In contrast to Caliendo, Dvorkin, and Parro (2019), tractability in our case does not come from dynamic hat algebra techniques and we estimate all time-varying model fundamentals. Slow transitions in our paper and this literature share common causes: low and highly dispersed exit (moving) probabilities and slow input adjustments.

[3] While the literature on firm dynamics with a development focus is too vast to cite, overviews are for example given by Hopenhayn (2014) and Restuccia and Rogerson (2017) for misallocation, Ulyssea (2020) for informality and Alessandria, Arkolakis, and Ruhl (2021) for trade.

Third, we contribute more generally to the Quantitative Macroeconomics literature by showing how to tractably estimate the model directly on the observed transition path in the data, identify the model entirely on plant-level data and only use macro moments for model validation, an approach that we see as closely aligned to a growing literature that moves "from micro to macro" (see the overview in: Buera, Kaboski, and Townsend 2023). The "equilibrium estimation" methods that ensure tractable estimation – enforcing the observed path of equilibria throughout the estimation and Euler CCC estimation – are used in other literatures, but have not yet seen wider application in the Macroeconomic literature.[4] We find the equilibrium estimation approach to be particularly suited for studying a path of time-varying equilibria, since we can also tractably estimate entire paths of time-varying parameters. While estimation methods that require to first solve the model may offer more flexibility on the choice of moments that identify parameters, they often have to strongly restrict the parameter space.

At last, our paper also relates to the literature on growth and productivity dynamics. Our results of selection-driven aggregate productivity growth in Indonesian manufacturing mirrors similar results in Brandt, Van Biesebroeck, and Zhang (2012) for China and Asturias et al. (2023) for Chile and Korea. We add to these papers by establishing the result of selection-driven productivity growth in a non-parametric setting that nests a larger class of growth models including various endogenous growth models.

The rest of the paper is structured as follows. The next section presents the main empirical evidence. Section 1.3 develops the model and discusses identification, estimation and model validation. In Section 1.4, we quantify the main drivers of growth. The last section concludes.

1.2 Empirical evidence

In this section, we introduce the Indonesian data and key facts about the Indonesian growth experience that motivate the subsequent model.

1.2.1 Data

Our primary data comes from the plant-level Annual Manufacturing Survey, collected by Indonesia's Central Bureau of Statistics. 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 annual information on standard plant-level characteristics from 1975 to 2015, a period of 41 years. It covers between roughly 7,500 to 30,000 plants per year. Throughout, we draw on reported information on plants' age (based on birth year), industry (up to 5-digit) and ownership (including foreign ownership). On the production side, we draw on plants' capital stock, value-added revenue, and the number of workers (including paid and unpaid workers) and wage bill (including contributions and in-kind compensation) which are separately reported for production and non-production workers. Unfortunately, capital is only reported starting in 1990. All variables denoted in Indonesian Rupiah

[4]The idea of estimating models on the observed path of equilibria in Macroeconomics dates at least back to Hansen and Singleton (1982). More recent research that conditions estimation on the observed path of equilibrium prices can for example be found in the literature following Caliendo, Dvorkin, and Parro (2019). In a recent working paper, Humlum (2022) exploits these two estimation steps in a similar general model framework of growth and firm dynamics, although entirely different context of industrial robot adoption in Denmark.

are deflated to real values using the aggregate CPI and normalized to the year 2010. For aggregate data, we further combine the GGDC 10-sector and Economic Transformation databases for Indonesia, which capture time-consistent aggregate sectoral employment and output series over the time period 1960-2012 (Timmer, de Vries, and De Vries 2015) and 1990-2018 (Kruse et al. 2023) respectively. We refer to this data as the *GGDC* series throughout. In Appendix 1.A.1, we discuss in detail the data cleaning and homogenization steps we take to ensure consistency of all datasets over time.

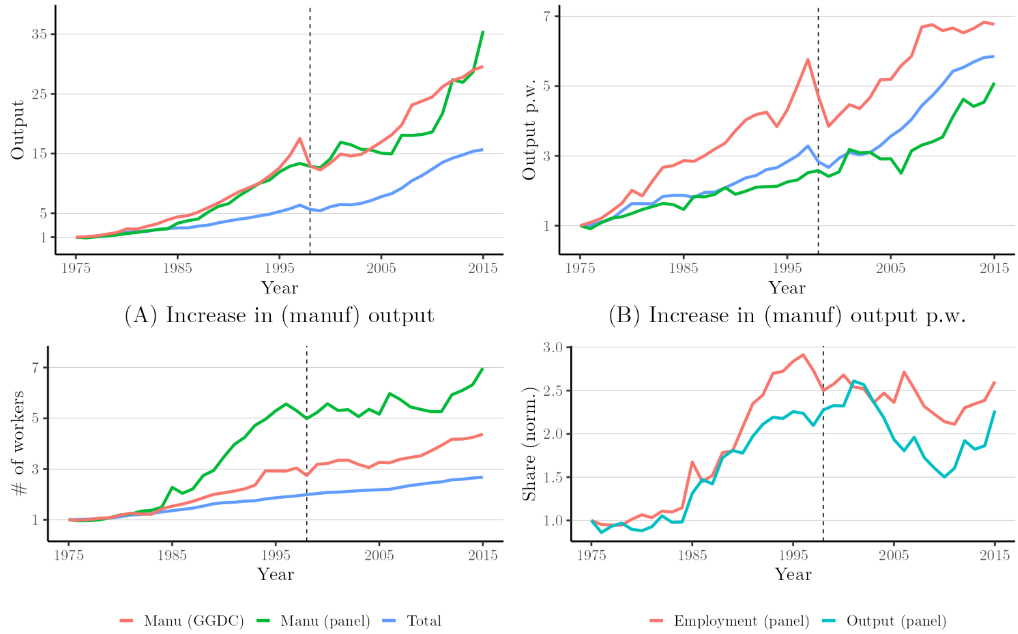
There are two important data limitations. First, while the data is in principle a census of plants with 20 or more workers, in practice the census misses plants. This shows up in discontinuous jumps in new plant entry during years of the economy-wide Economic Census in 1975, 1985, 1995 and 2006. In those years, plants are added that were previously missed, either because they were small and made the cutoff, or newly entered. For the subsequent analyses, this means that aggregate changes often show discontinuous drops in census years and should actually look more smoothed out. Given that we observe the census in 1975, the initial distribution is correctly reported. Furthermore, this does not bias results that are based on within-plant variation. Our data also misses plants because of non-reporting, either because plants miss to report in some years or because we are forced to drop a plant-year entry due to misreporting (see Appendix 1.A.1). We treat these missing entries as missing at random and specifically account for missing entries in our structural model. We correct our measurements of plant entry and exit by denoting plant entry as the first time when a plant identifier enters the panel and plant exit if we do not observe a given plant identifier at any future time period (see: Appendix 1.A.1).

The second main data limitation is with respect to the coverage of our data. As we show in Appendix 1.A.2, while our dataset misses the approximately 99% of Indonesian manufacturing plants that have less than 20 workers, most of these plants are characterized by self-employment with a modal plant size of one to two workers. After cleaning, our manufacturing data captures between 25-30% of total manufacturing employment and value-added output as based on the *GGDC* data, with shares increasing over time (Figure A.1). We think of small scale manufacturing as a separate sector given robust evidence that there are few transitions between small and larger scale manufacturing (e.g. Poschke 2013; Van Biesebroeck 2005; Schoar 2010) and most plants that enter our panel have only recently been established. For example, the median age of plants that newly enter our plant panel is only two years. The focus in this paper is thus on how relatively large plants and their dynamics drive aggregate economic growth. In the model and results parts, we explicitly model the entire economy, taking into account that our data only captures a time-varying share of output in the overall economy.

1.2.2 Four key facts of growth

We now highlight four key facts that shed light on the Indonesian growth experience. The first two facts relate to changes in the entire plant distribution and the importance of slow adjustment processes. Fact 3 and 4 document changes in the drivers of productivity growth and the absence of improvements in the allocation of resources over time.

Figure 1.1: Evolution of aggregate and sectoral employment and output



(A) Increase in (manuf) output

(B) Increase in (manuf) output p.w.

(C) Evolution of number of workers

(D) Manuf employment + output shares

Notes: (Economy-wide) Total gives the aggregate of the GGDC data. Panel refers to the Indonesian manufacturing plant census (1975-2015, 20+ workers). All series are normalized by their respective value in the first year. (A) and (B) use value-added output. Dashed vertical lines denote the onset of the Asian Financial Crisis in 1997.

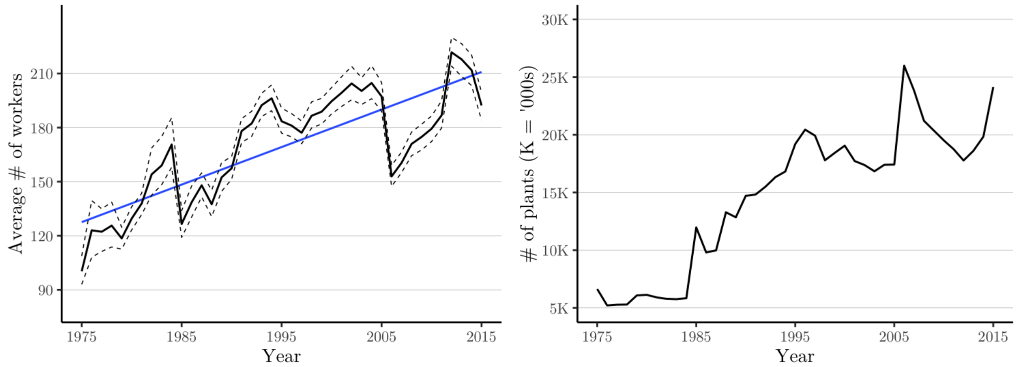
Rapid economic growth coincided with changes in the plant distribution

As the 4th most populous country in the world, Indonesia underwent a dramatic process of economic development over the past 50 years, a few key features of which are reported in Figure 1.1. Between 1975-2015, GDP per worker increased more than five-fold (Panel B), driven by a 15-fold increase in output (Panel A) and roughly a tripling of the working population (Panel C). Manufacturing contributed importantly to this aggregate growth: output grew 30-fold and the manufacturing employment share more than doubled (Panel D).

The period 1975-2015 can be divided into two main growth regimes that are separated by the Asian Financial Crisis in 1997. The pre-crisis period captures a period of rapid labor-intensive industrialization, including the period 1987-1994 that Hausmann, Pritchett, and Rodrik (2005) characterize as a growth acceleration. Most of the total worker flows into manufacturing happen before 1997 and manufacturing grows far more rapidly than the rest of the economy. Fast growth in the aggregate working population is also a defining feature of the pre-crisis period with an average annual growth rate of 3%, 70% higher than in the post-crisis period. Based on our census of medium- and large-sized plants, the rise of manufacturing rapidly takes off in the first half of the 1980s and industrialization peaks with the Asian Financial Crisis as evidenced by manufacturing employment and output shares (Panel D). After the Asian Financial Crisis, which started in 1997 and hit manufacturing mostly in 1998, the economy experienced lower total output growth that was due to substantially lower growth in plant entry and employment and – as we will show further below – by higher plant-level productivity growth.

These aggregate changes went in hand with systematic changes among manufacturing plants. As evidenced in Figure 1.2, the rapid increase in the total number of workers in manufacturing

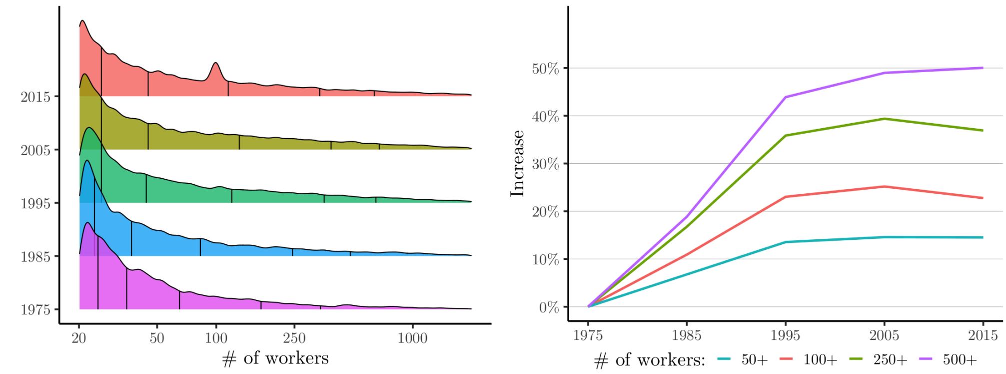
Figure 1.2: Evolution of average plant size and number of plants



(A) Avg number of workers over time (B) Number of plants over time

Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. Panel A: Workers include paid and unpaid workers. Dotted lines give 95% bootstrapped confidence intervals and solid blue line gives best linear fit. Panel B: Jumps in 1985, 1995 & 2006 are explained by Economic Census years.

Figure 1.3: Evolution of employment distribution over time



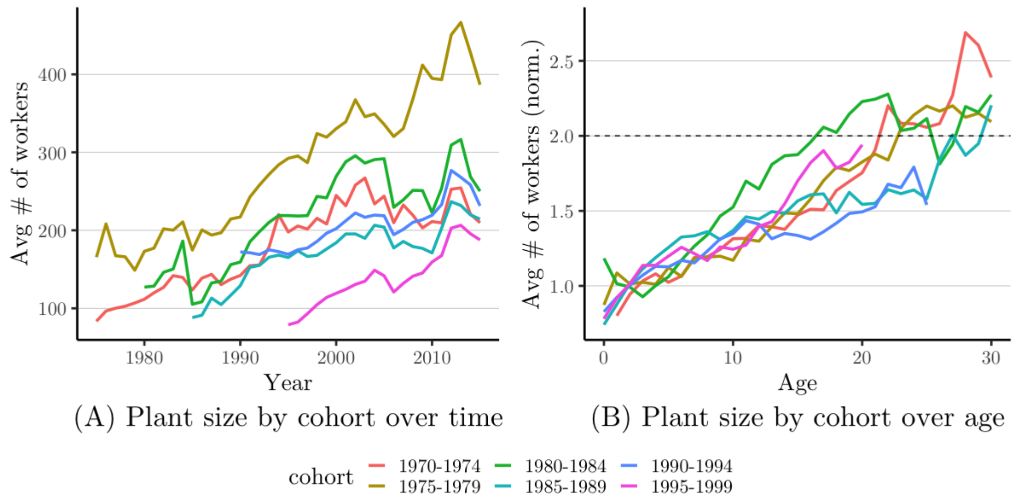
(A) Size distribution over time (B) Employment shares over time (1975 = 1)

Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. Only showing years 1975, 1985, 1995, 2005 and 2015. Panel A: Vertical lines give 25th, 50th, 75th, 90th and 95th percentiles. Bunching at 99 workers starts in 2013, after the passing of the 2012 Worker Safety Law that binds for establishments with 100+ workers. Panel B: Share of employment in plants with more than X workers.

is met with a doubling of the average number of workers in manufacturing plants (Panel A) and a four-fold increase in the number of manufacturing plants. In both cases, most of the gains were already reached by 1997. Importantly, the entire plant distribution changed systematically over time. Specifically, Figure 1.3 shows that the right tail of the plant employment distribution systematically thickened over time – a key feature of the development process that has been highlighted for firms (rather than establishments) across and within countries (Chen 2022; Choi et al. 2023; Poschke 2018). Panel B shows that the employment share in plants with more than 50 workers increased by roughly 15%, while the employment share in plants with more than 500 workers increased by more than 50%.[5] The increase in the right tail of the employment distribution is a main driver of the increase in the average plant size over time.

[5]We report this metric as it is a simple transformation of the Pareto tail, which is also robust to left-censured data. We report the corresponding secular decline in Pareto tail coefficients in Appendix 1.A.2. We find systematic changes in the Pareto coefficient, both in the cross-section (which is not in line with a common Pareto distribution) and over time (which is not in line with traveling wave equilibria).

Figure 1.4: Plant life cycle growth by birth cohort



(A) Plant size by cohort over time

(B) Plant size by cohort over age

Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. Workers include paid and unpaid workers. Plant age is based on reported year of establishment. Panel B normalizes each entry by the cohort-specific average plant size of plants below age 5 (as in Hsieh & Klenow 2014). Note that each cohort over time is an unbalanced panel as only surviving plants stay in the panel and there is (limited) plant entry as plants make the cutoff of 20 workers to be included in the census.

Adjustment processes can account for changes in the plant distribution

We now show evidence for three slow adjustment processes that can respectively account for increases in the plant size, the mass of plants and the right tail of the plant size distribution.

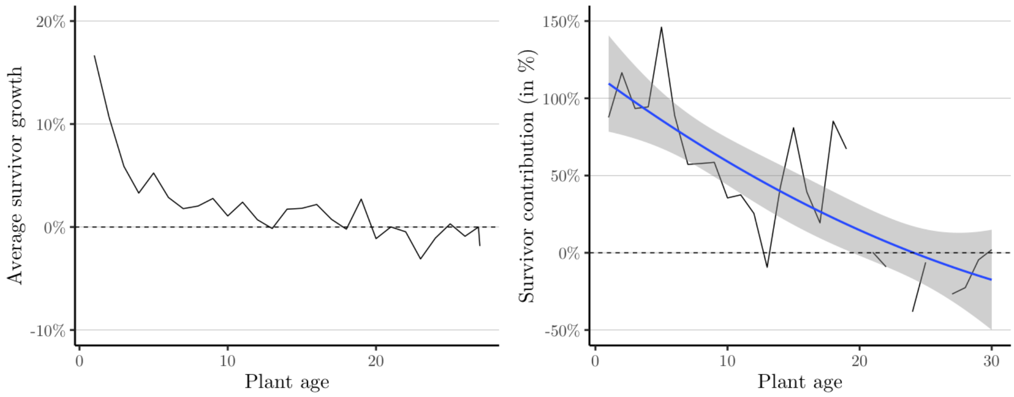
The aging of the plant distribution We start by showing that the slow aging of the plant distribution – average plant age increased by 40% since 1975 – can explain average plant size increases. The reason is that plants enter small and grow over their life cycle. Figure 1.4 plots life-cycle growth profiles across different cohorts of surviving manufacturing plants. Plants enter roughly with a similar average number of workers, which grows with plant age. Plants that survive at least 20 years have about twice as many workers as new entrants; in comparison and as documented in Hsieh and Klenow (2014), manufacturing plants in the US that survive for that long are about six times as large as new entrants.[6]

A benefit of our panel data is that we can show that the increase in the average size of surviving plants is mostly driven by within-plant growth rather than selection (larger plants being more likely to survive). Figure 1.5 Panel A shows average year-to-year within-plant growth by age. Young plants grow their employment quickly, with growth declining slowly as plants become older, average growth running out after around 20 years. Panel B reports the relative contribution of survivor growth to average plant size increases by age using the following accounting identity:

$$\underbrace{\bar{L}_a - \bar{L}_{a-1}}_{\Delta\text{avg plant size}} \equiv \underbrace{\frac{N_a^S}{N_a}(\bar{L}_a^S - \bar{L}_{a-1}^S)}_{\text{Survivor contribution}} + \underbrace{\frac{N_a^E}{N_a}\bar{L}_a^E}_{\text{Entry}} - \underbrace{\frac{N_{a-1}^X}{N_{a-1}}\bar{L}_{a-1}^X}_{\text{Exit}} + \underbrace{\bar{L}_{a-1}^S\left(\frac{1}{N_a} - \frac{1}{N_{a-1}}\right)}_{\text{Net reallocation}} \tag{1.1}$$

[6] The numbers are not perfectly comparable, because of the cutoff of 20 workers in the Indonesian data. If young plants in the US data are smaller, this overestimates the difference across countries, while plants that stay below 20 workers in the US data bias downward.

Figure 1.5: Within-plant (survivor) growth in employment



(A) Avg employment growth by age

(B) Relative contribution of survivors

Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. Sample restricted to survivors and plants for which plant age is defined (N = 543,586). Panel A: Average annual growth in total number of workers (paid + unpaid) across plants by age, weighted by a plant's previous employment. Panel B: Relative contribution of survivor growth to average plant size growth based on accounting identity in Equation 1 and dividing by the left-hand side to obtain relative contributions.

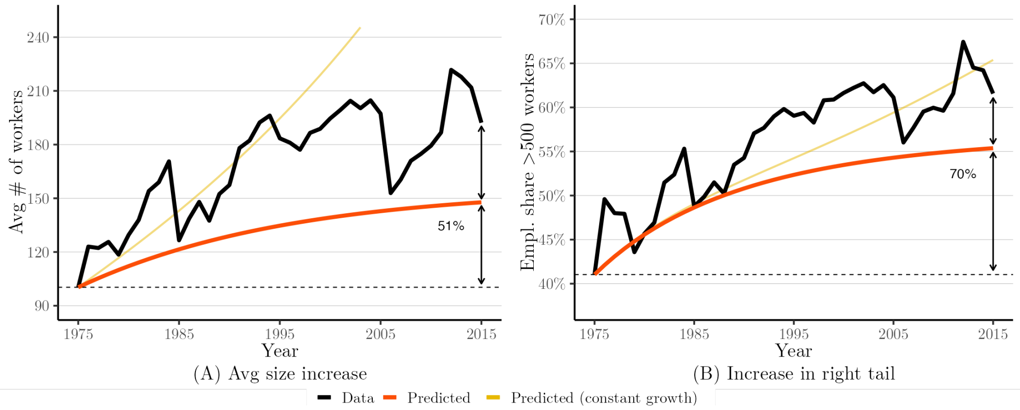
where S refers to the set of surviving plants from age $a-1$ to age a, E refers to entering plants (which exist because of the size threshold in our data) and X refers to exiting plants. The contribution of survivors, entry and exit respectively measure their average size weighted by their share in the population of plants. The net reallocation effect is driven by changes in the total number of plants over age: if exit outweighs entry (as in our case) then workers are reallocated towards fewer plants, mechanically increasing average plant size. We find that for young plants, growth by survivors explains all of the increase in average employment across plants, while selection as given by the remainder dominates the total effect after age 10-15.

Together, this evidence implies that average plant size crucially depends on where the age distribution of plants is; since the initial distribution of plants in the newly emerging manufacturing sector in 1975 featured mostly young plants, average plant size was small. Despite the rapid entry of new plants, plants became older and hence larger over time. For example, average plant age increased by 40% between 1975 to 2006 (the last year in which plant age was reported in the survey). Figure A.5 also shows how the entire age distribution shifted right over time.

It takes time for entry and exit dynamics to play out Next, we highlight a basic driver of slow entry and exit dynamics that can partially account for the observed four-fold increase in the mass of plants: low entry and exit rates. The idea is simple: the young manufacturing sector in 1975 features few plants and if only few potential entrepreneurs move into and out of entrepreneurship, it takes time to build up a mass of plants. For example, suppose there is a fixed mass of potential entrepreneurs and entry and exit rates into entrepreneurship are the same. Then the long-run (steady state) share of entrepreneurs is 1/2, but if the economy starts with no entrepreneurs, it can take many years to get close to the steady state. For entry and exit rates at 7.9% – equal to the average exit rate across all years and plants in our data[7] – it

[7] This exit rate is substantially lower than the 14-18% documented for informal establishments in Vietnam (McCaig and Pavcnik 2021) and slightly smaller than the 8.3% documented for small establishments across 12 developing countries (McKenzie and Paffhausen 2019). Within manufacturing, exit rates also seem to decline with plant

Figure 1.6: Reduced-form transition dynamics implied by initial plant size distribution



Notes: Predicted changes in distribution based on exercise taking discretized initial plant size distribution (# of workers) in 1975 and transition matrix giving conditional probabilities of moving from one plant size bin to another for years 1975-1976. Predictions iterate on initial discretized distribution with fixed transition matrix. Predicted (constant growth) instead enforces transition matrix incorrectly enforcing constant growth taking average plant size growth for 1975-1976.

already takes the economy 14 years to reach just 90% of the long-run steady state. While we do not generally observe potential entrants and thus cannot study entry rates without further assumptions, in the Appendix we show additional evidence for slow exit processes. Specifically, Figure A.8 shows that exit rates only slowly decrease with plant productivity and that exit rates do not increase with aggregate shocks such as the Asian Financial Crisis. This is in line with growing evidence that stagnant firms in developing countries tend to survive longer compared to firms in developed countries (e.g. Hsieh and Klenow 2014; Akcigit, Alp, and Peters 2021; Eslava, Haltiwanger, and Pinzon 2022).

It takes time to grow large plants At last, we show that a lack of large plants in the 1970s and the fact that it takes time to grow large plants can jointly explain the slow fattening of the right tail of the employment distribution. We do so by considering the following exercise. Take as the starting point the initial employment distribution of plants Φ_0 in 1975 and discretized in $X = 10$ different size bins. Each bin captures the fraction of plants with this number of workers. We now follow individual plants and compute the probability of moving from one bin to the other between 1975 and 1976, which we summarize in the transition matrix P of dimension X^2. We predict changes in the distribution by iterating on the initial distribution using the fixed transition matrix: $\hat{\Phi}_{t+\tau} = \Phi_0 \cdot P^{\tau}$. Figure 1.6 shows that the exercise explains 50% of average plant size increases (A) and 70% of the increases in the employment share of plants with more than 500 workers (B) over time. The reason is that in 1975, the distribution lacks large plants in comparison to the stationary distribution implied by the employment growth observed between 1975 and 1976 and it takes time to grow large plants. The exercise predicts that it takes 25 years to reach 90% of the steady state average plant size, broadly capturing the speed at which plant growth plays out over time.

The exercise is robust to a number of concerns which we address in Appendix 1.A.3.[8] Im-

size, explaining a lower exit rate of 6.2% across all US manufacturing (e.g. Clementi and Palazzo 2016) and broadly similar exit rates for all manufacturing plants in India and Mexico (Hsieh and Klenow 2014).

[8]Specifically, we show similar results when allowing for entry and exit, taking any other starting years in the 1970s and that the slow filling up of the right tail of the employment distribution holds for any other year-to-year

portantly, the exercise cannot distinguish the drivers of employment growth. The transition matrix $P_{t,t+1}$ only gives a reduced-form summary of plant employment growth subject to any frictions and growth drivers that are present between time t and $t+1$, which may include adjustment frictions, changes in wages or productivity growth. The next two subsections thus focus on two key determinants of within-plant employment growth: productivity growth and plant-level hiring frictions.

Productivity growth and selection

What is the role of productivity growth in the Indonesian growth experience and how much productivity growth is explained by the better selection of plants? In this section, we show that aggregate productivity increased roughly 3.5-fold between 1975 to 2015. However, the underlying drivers of this productivity growth differ fundamentally across Indonesia's two main growth periods: during the period of rapid labor-intensive industrialization (1975-1997), all of the aggregate productivity gains are driven by the selection of more productive plants, while aggregate productivity gains during the period after the Asian Financial Crisis are almost entirely driven by within-plant productivity gains. We obtain these results by standard production function estimation (Ackerberg, Caves, and Frazer 2015; Demirer 2020) and then separately identifying the productivity improvements that come from the better selection of plants versus within-plant productivity improvements.

Estimating productivity Following the literature, we estimate a standard value-added Cobb-Douglas production function in capital k and efficiency units of labor h:

$$y_{it} = x_{it} h_{it}^{\theta_{jt}} k_{it}^{\alpha_{jt}} \tag{1.2}$$

with $\theta_{jt} + \alpha_{jt} \leq 1$ giving the output elasticities of labor and capital in sector j at time t and x_{it} capturing plant-level productivity. As a baseline and in line with our subsequent dynamic model, we start with common output elasticities across manufacturing industries, but discuss further industry variation below. We estimate the labor and capital output elasticities separately for each year allowing for both inputs to be potentially fully dynamically chosen, which means that at this point we can remain agnostic about the frictions that determine plant input choices and changes in these frictions over time. Specifically, we draw on the control function approach in Demirer (2020), which is close in spirit to the standard control function value-added production function estimation based on Ackerberg, Caves, and Frazer (2015), but does not require intermediate inputs (also see: Gandhi, Navarro, and Rivers 2017). We provide a discussion and an identification proof adapted to our setting in Appendix 1.A.5, but the intuition of identification is as follows: exploiting the assumption that productivity follows a first-order Markov process, conditional on previous input choices and output, the ranking of current inputs identifies the ranking of productivity innovations, which can be used to construct a control function for (unobserved) productivity in the output regression.

transition matrix that can be constructed between 1975 and 2015. An important caveat that we cannot address is that due to the discretization, the exercise would also give an increase in the plant size if all plants were simply growing at a constant rate. The red lines in Figure 1.6 report changes with a counterfactual transition matrix enforcing average employment growth rates between 1975 to 1976 across all plants. While observed within-plant employment growth is far from constant, the alternative exercise still shows that it is easy to overestimate the growth in the right tail.

Results and more details on the estimation are shown in Appendix 1.A.5. Before we discuss the results, it is important to highlight that we estimate the production function on available revenue data. This comes with the standard limitations that we only identify revenue elasticities and revenue-based productivity and cannot distinguish between productivity and demand nor identify changes in markups.[9]

We find no systematic changes in estimated output elasticities over time and very standard values for the output elasticity of labor close to 2/3. Importantly, the estimated output elasticity of labor is substantially larger than the median plant-level labor share (≈ 0.45) and the aggregate labor share in manufacturing (≈ 0.25). In a frictionless model with Cobb-Douglas production, plants would equalize cost shares and revenue elasticities. In the next section, we provide evidence on frictions for labor choices that could rationalize this difference. Our subsequent model then quantitatively accounts for these large differences. We find smaller estimates for the capital output elasticity than generally found in the literature, which – as we discuss in Appendix 1.A.5 – is likely due to attenuation bias from measurement error in observed capital. A lower capital output elasticity means that observed (mismeasured) variation in capital has smaller effects on output and we further show that mismeasurement of the capital elasticity is not biasing our estimates for labor.

Selection versus plant-level productivity growth Next, we quantify how much of the productivity improvements across plants are driven by the selection of more productive plants versus within-plant productivity growth. For this, we assume that plant-level productivity is the product of a common, potentially endogenous, aggregate technology component z_t that improves the productivity of all plants and an idiosyncratic productivity shock s_{it}: $x_{it} \equiv z_t s_{it}$. This setup allows us to separate shared technology growth in z_t, selection on idiosyncratic productivity s_{it} and within-plant growth in s_{it}, and nests the productivity side of many exogenous as well as endogenous growth models in the literature.[10] In this setup, we provide a novel non-parametric identification approach that separates the path of z_t from s_{it} (up to a normalization of z_0).

To understand why separate identification of selection and technology growth is difficult in

[9]Using separate information on prices and quantities, the previous literature has highlighted the important role of demand for driving firm growth (Hottman, Redding, and Weinstein 2016; Foster, Haltiwanger, and Syverson 2016; Eslava and Haltiwanger 2020) and an important part of what we subsequently call "productivity" likely captures demand. We return to this issue when discussing model counterfactuals where the distinction between demand and productivity is key. Also, revenue-based productivity measures may be preferred in the Indonesian context where large changes in product quality bias quantity-based productivity estimates (see Atkin, Khandelwal, and Osman (2019) for the argument and Hill (2000) for a discussion of strong quality improvements in Indonesian manufacturing). Relatedly, we also do not identify changes in mark-ups, which usually requires separate information on prices and quantities (Bond et al. 2021). Studying changes in markups and its drivers over the course of development is an exciting direction of future research, but beyond the scope of this paper.

[10]Specifically, the setup nests standard neoclassical growth models that feature exogenous aggregate productivity growth and firm selection (e.g. Luttmer 2007; Clementi and Palazzo 2016) as well as endogenous growth models that feature a common endogenous growth component such as Romer (1990) or models where Gibrat's law holds and productivity growth is independent of firm size (Klette and Kortum 2004; Atkeson and Burstein 2010; Restuccia and Bento 2015; Peters 2020).

the first place, let us start by looking at log changes in average productivity over time:

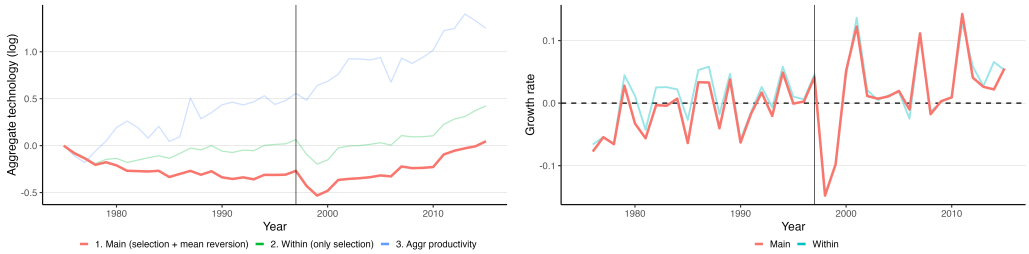
$$\frac{1}{N_t}\sum_{i\in\mathcal{N}_t}\tilde{x}_{it}-\frac{1}{N_{t-1}}\sum_{i\in\mathcal{N}_{t-1}}\tilde{x}_{it-1}=\underbrace{\tilde{z}_t-\tilde{z}_{t-1}}_{\Delta log(z)}+\underbrace{\frac{1}{N^S_{t,t-1}}\sum_{i\in\mathcal{N}^S_{t,t-1}}\Delta\tilde{s}_{it}}_{\text{Survivor }\Delta log(s)}+\underbrace{\frac{1}{N^E_t}\sum_{i\in\mathcal{N}^E_t}\tilde{s}_{it}}_{\text{Entry }\overline{log(s)}}-\underbrace{\frac{1}{N^X_t}\sum_{i\in\mathcal{N}^X_t}\tilde{s}_{it-1}}_{\text{Exit }\overline{log(s)}}$$

where $log(x)\equiv\tilde{x}$ denotes variables in logs. Changes in average productivity $\tilde{x}_{it}$ only identify growth in technology z under the special case that average changes in productivity $\tilde{s}$ among survivors as well as changes in average productivity $\tilde{s}$ between exiting and entering plants exactly cancel out. These terms capture two important selection biases. The entry and exit terms capture a "static" compositional selection bias that leads to overestimates of aggregate technology changes as long as less productive plants are more likely to exit and entering plants are positively selected on productivity. Focusing instead on within-plant changes in productivity deals with the "static" selection bias, but still leaves a term capturing average changes in idiosyncratic productivity among surviving plants. We call this term the "dynamic" selection bias, which generally biases estimated aggregate technology changes downwards if productivity s_{it} is persistent. Intuitively, if surviving plants are selected based on good histories of productivity realizations s_{it}, then they are more likely to mean revert in the future.

Non-parametric identification in our setup means that we make no functional form assumptions on the arbitrarily time-varying path of aggregate technology z_t, the productivity shock process s_{it}, nor on the plant entry and exit processes that drive endogenous selection. For expositional purposes, we provide an idea of the identification and estimation approach and relegate the precise technical assumptions, a detailed identification proof and further estimation details to Appendix 1.A.6. Identification of changes in z_t relies on two sets of assumptions. The first assumption ensures that the productivity shock process s_{it} has a stationary distribution. Technically, we assume that s_{it} follows the same underlying general first-order, ergodic Markov process across plants and time, allowing for flexible forms of error dependence. The stationary distribution is useful because if we could reweight changes in observed productivity x_{it} among the – potentially highly selected – set of surviving plants based on the stationary distribution of s_{it}, the "dynamic" selection bias exactly cancels out. That is, $\mathbb{E}_i\Delta\tilde{s}_{it}$ is exactly equal to zero at the stationary distribution of s. The second set of assumptions ensures that such a stationary distribution can always be constructed using the observed data, restricting the degree of selection at exit. Specifically, we require that (1) plants' exit decisions are not based on future productivity shock realizations, and (2) there is no sharp productivity cutoff at which all plants would exit, so that there is always common support that allows an appropriate reweighting of the distribution. The latter can be empirically tested and – as shown in Figure A.8 Panel B – finds strong support in the Indonesian data.

With the assumptions in hand, the only remaining difficulty is how to construct the weights of the stationary distribution and solve for the time path of changes in z_t. Here, we first solve "forward" for the stationary distribution by starting with equal weights over the initial distribution. Whenever a plant selectively exits, we pass on their weight to plants with similar productivity who did not exit using a standard Kernel estimator, creating a synthetic panel among surviving plants that is "representative" of the underlying process of s. For time growing large – no matter how selected the initial distribution is – one can thereby identify appropriate weights over the selected set of producing plants that recovers the stationary distribution

Figure 1.7: Aggregate manufacturing technology estimates



Notes: Panel A: Aggregate technology estimates, showing the main estimator (explained in the data) as well as the within estimator (that only controls for selection, but not mean reversion) and aggregate productivity (measured as value-added-weighted average productivity). Panel B: Corresponding growth rates in aggregate technology. The main and within estimators both use (weighted) median changes in plant-level productivity. Further details in the text.

of s (up to a common scaler z). We then move "backwards" from the last period T to identify the path of z_t: initially normalizing z_T, we start with a guess over z_{T-1}, compute the weighted changes in productivity between $T-1$ and T based on the stationary distribution of s in $T-1$ and solve for the implied z_T. This implies finding a root in z_{T-1}. Iterating on this procedure until z_0 recovers the entire path.

Figure 1.7 shows the estimated path of technology using the full sample and baseline productivity estimates. Over the entire 40-year period, technology improves little, being less than 5% higher in 2015 than in 1975. However, this masks important changes over time. Specifically, technology actually declined strongly throughout the 1970s, remained almost constant throughout the 1980s and 1990s and then saw rapid growth of roughly 4% per year since the year 2000. At the same time, aggregate productivity – measured as the value-added-weighted average productivity across plants – increased roughly 3.5-fold by 2015 and increased by 75% by the time of the Asian Financial Crisis in 1997. Together, this means that the sources of productivity growth fundamentally differed over the two Indonesian growth periods: plant selection rather than shared technology growth drove more than all of the productivity gains during Indonesia's period of rapid labor-intensive industrialization (1975-1997), while the pattern reversed after the year 2000, with common technology growth explaining more than 90% of the aggregate productivity gains. These estimates – especially after the Asian Financial Crisis – are driven by within-plant productivity growth of surviving plants. Pre-1997, true technology growth is lower than the unweighted within-plant productivity growth because of positive mean reversion that can be explained by the presence of many young and low productive plants that mean revert upwards in their productivity shocks. Post-1997, positive and negative mean reversion roughly balance out, explaining the similar growth paths of the main and within estimators. In Appendix 1.A.6 we further discuss what could be driving these large changes in the role of aggregate technology over time.

The allocation of resources did not improve over time

At last, we look at measures of the misallocation of resources over time. We start with an accounting-based decomposition of growth in Indonesian manufacturing. Using the previously assumed production structure and separation of plant-level TFP into an aggregate technology and idiosyncratic productivity component, we can write growth in manufacturing output as:[11]

$$\Delta\, ln\, Y_t \equiv \underbrace{\Delta\, ln \sum_i f(h_{it}, k_{it})}_{\text{input growth}} + \underbrace{\Delta\, ln\, z_t}_{\text{aggr techn.}} + \underbrace{\Delta\, ln \left[\bar{s}_t + N_t cov \left(s_{it}, \frac{f(h_{it}, k_{it})}{\sum_i f(h_{it}, k_{it})} \right) \right]}_{\text{selection + reallocation effect}} \tag{1.3}$$

where N_t tracks the number of active plants. Total output is the combination of the state of factor accumulation and aggregate TFP. Aggregate TFP, in turn, can be further decomposed into aggregate technology and – following Olley and Pakes (1996) – a combination of average productivity and a covariance term that captures whether resources in the economy are allocated towards the most productive plants. Since the covariance is affected by common trends in its inputs such as changes in the sample size, Figure 1.8) plots the correlation of plant-level productivity and resource shares, which is robust to common trends and simply the normalized covariance. Panel A and B show this measure of the allocation of resources for the full sample of plants and only for surviving plants that already operated in 1975, normalizing each by the first year. Each panel additionally plots the cross-sectional correlation within industries.[12] Across all plants, the allocation of resources actually deteriorated over time, while it remained stable for survivors, indicating that the deterioration is driven by the entry of small plants. As further documented in Appendix 1.A.7, this result also holds within a balanced panel and separately within each cohort of entering plants between 1970 to 1999.

In Appendix 1.A.7, we also report changes in the dispersion of marginal revenue products of capital and labor over time, which maps to changes in misallocation in the literature building on Hsieh and Klenow (2009).[13] Again, we find that, if anything, the dispersion of measured marginal revenue products (even within 5-digit industries) tends to increase over time. Thus, based on our estimates and in contrast to Buera and Shin (2013), we find little evidence for an undoing of misallocation being a feature of growth in Indonesia.

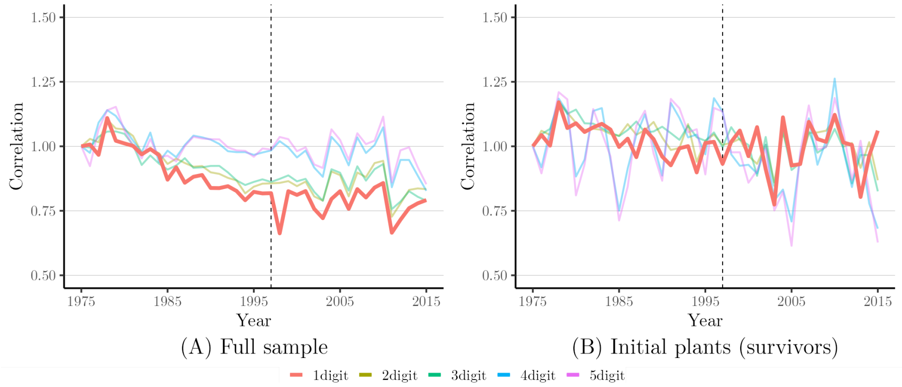
The misallocation dynamics shown in Figure 1.8) are a function of changes in plants' productivities and how inputs reallocate across plants. To shed light on these dynamics at the micro-level, we are interested in how plants' inputs respond to changes in productivity. Figure 1.9 shows how plants' labor and capital shares evolve as a response to a positive and permanent productivity shock of at least 20% – roughly equal to the 75th percentile of within-plant productivity changes. We use a standard staggered differences-in-differences design (Callaway and Sant'Anna 2021) that is particularly suited here, because it captures plant-level dynamic responses while controlling for time and plant fixed effects that ensure that results are neither driven by aggregate shocks in specific years nor by fixed differences across plants such as differences across industries or even plant-specific production functions. Given selective plant entry and exit and resulting composition biases from estimating treatment effects on un-

[11] The proof can be found in Appendix 1.B.4.

[12] Specifically, we proceed similar as in Gopinath et al. (2017): we first estimate the correlation across plants in a given industry and year and then construct the weighted average correlation across industries using the industry's average share in manufacturing value added as an industry-specific time-invariant weight. Using the same weights when aggregating across industries ensures that within-industry estimates reflect purely variation within industries over time.

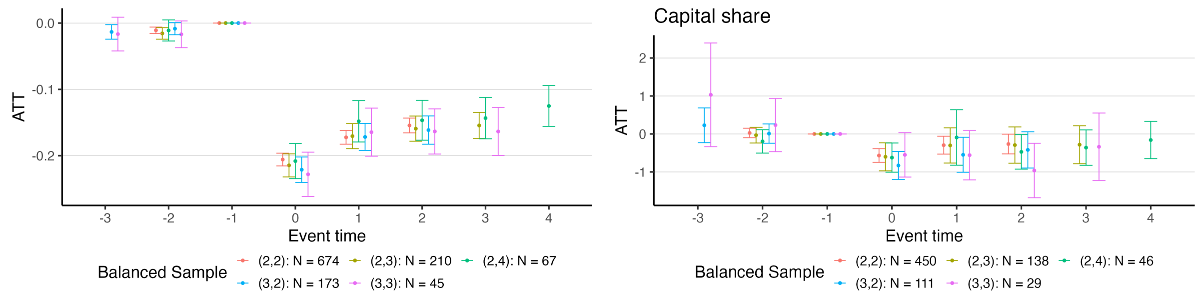
[13] There are important limitations to interpreting changes in the covariance between measured revenue productivity and input shares as changes in misallocation. For example, in Hsieh and Klenow (2009), an efficient allocation implies zero correlation between TFPR and the input share. We note that revenue productivity in our model is not equal to TFPR in Hsieh and Klenow (2009), but instead much more closely correlated with their measure of TFPQ.

Figure 1.8: Evolution of cross-sectional correlation of plant productivity and input share



Notes: Input shares are computed based on Cobb-Douglas aggregator. Within-industry results based on first estimating the correlation across plants in a given industry and year and then constructing the weighted average correlation across industries using the industry's average share in manufacturing value added as a time-invariant weight. All series are normalized by the first year.

Figure 1.9: Event study: Input share responses to a permanent productivity shock



Notes: Treatment defined by permanent productivity shock of 20 percent change that does not revert back, neither before nor after treatment. Event study following Callaway & Sant'Anna (2021) estimated on balanced panel of treated plants and with all non-treated plants as control, event time zero gives first period of treatment, effects at -1 normalized to zero by assumption. Sample for capital share is restricted to post 1990 due to data availability. Details in the text.

balanced panels, Figure 1.9 reports estimated dynamic treatment effects for different balanced panels.[14]

Plants' labor shares drop by more than 20 percentage points in the first year of treatment and recover slowly over time as plants respond by increasing hiring. In Appendix 1.A.8, we further show that the slow labor share adjustments are indeed driven by slow hiring rather than wage increases. In a model with the above production side but where plants can adjust employment without constraints, there should be no response in the labor share, indicating frictions in adjusting labor. Quantitatively, these frictions are large: For example, the average "treatment" shock is roughly 40%, so that a plant with an initial labor share of 0.7 would not adjust labor at all in the first year of a large productivity increase. In contrast, estimated dynamic responses

[14]The treatment effects are estimated with all non-treated plants as control. Event time zero gives the first period of treatment, while treatment effects in event time -1 are normalized to zero by assumption. Apart from the 20% cutoff, the treatment definition also ensures we identify a permanent productivity shock by ruling out shocks followed and preceded by productivity changes of more than 10% per year and ruling out an accumulation of shocks over multiple periods that undo the shock at period 0. E.g. if we look at an event window from -3 to 2, then productivity at -3 cannot be more than 10% apart from the pre-level at -1 and productivity at 2 cannot be more than 10% apart from productivity at event 0. This does not restrict input shares. The rapid observed declines in the sample sizes are due to the high observed volatility of productivity, making it difficult to identify permanent shocks in the data.

for capital are much noisier and we cannot reject that capital shares immediately recover after the first year. For both inputs, we find no evidence for pre-trends, pointing away from an anticipation of productivity shocks. Taken together, volatile plant-level productivity and slow labor adjustments thus offer an explanation as to why the allocation of resources did not improve over time.

1.3 Structural model

While the four empirical facts document important changes in the economy and distribution of manufacturing plants over time (Fact 1), entry and exit dynamics that shape the selection of plants (Facts 2, 3 & 4), and the importance of slow plant-level adjustment processes (Facts 2 & 4), the empirical evidence is not sufficient to quantify the drivers of the Indonesian growth miracle. In particular, it does not allow us to quantify the aggregate effects of policy changes separately from transition growth. For this, we now build a model of plant dynamics and growth in the tradition of Hopenhayn (1992). In the model, plants face idiosyncratic risk in their productivity and choose capital and labor inputs subject to labor adjustment costs and a simple financing constraint that rationalize slow plant-level adjustments (Facts 2 & 4). Plants face fixed costs that drive endogenous entry and exit, which in turn drives aggregate selection dynamics (Fact 3).

The model features a time-varying growth path which is driven by three endogenous forces: changes in the input distribution, changes in the productivity distribution due to a combination of exogenous technology growth and plant selection, and changes in (mis)allocation as given by their joint distribution. All three forces are driven by the race between transition growth and by changes in model fundamentals that induce new transition growth. Changes in model fundamentals include changes in labor supply, potential entrants, technology, adjustment frictions and taxes. We further embed the model of plant heterogeneity into a two-sector general equilibrium model to capture changes in the rest of the economy and the endogenous reallocation of labor across sectors over time. The potential of transition growth at any point in time is given by the current distribution of plants encoding the history of the economy and future growth potential as given by current model fundamentals and expectations over the future. We follow the growth literature in treating the growth path as deterministic and by assuming that agents have perfect foresight over aggregate changes in the economy.

1.3.1 Model Setup

The model economy is set in discrete time indexed by $t = 1, 2, ...$. We assume that Indonesia is a small open economy vis-a-vis the rest of the world and has access to world capital markets at interest rate r^*. There are two sectors of production: Manufacturing (M) and the rest-of-the-economy (R). Both sectors produce the same homogeneous, perfectly substitutable good, which serves as numeraire. Manufacturing features heterogeneous plants whose endogenous mass and distribution are time-varying, while we model the rest-of-the-economy as a simple representative firm whose exogenous technology and endogenous labor demand change over time. Labor is inelastically supplied by households which choose in which sector to work. Labor markets in both sectors are fully competitive. There is a government that levies a valued added and a corporate income tax, the two main corporate tax instruments in Indonesia.

We assume the government levies these taxes and redistributes revenue back to households. Similarly, we assume that all plant profits are simply transferred back to households.

Manufacturing

The manufacturing sector is composed of risk-neutral plants that are heterogeneous in their productivity s_{it}. Each period, plants choose capital k and labor h on spot markets to produce output, while facing idiosyncratic risk over their future productivity s_{it} and time-varying changes in the economy. We denote plants' payoff-relevant aggregate state of the economy, including perfect anticipation of future changes in aggregates by Ω_t and make its components more explicit below. A plant's output y_{it} and taxable profits π_{it} at time t are given by:

$$y_{it}(s_{it}, z_t, h_{i,t}, k_{i,t}) = z_t s_{it} h_{it}^{\theta} k_{it}^{\alpha} \tag{1.4}$$

$$\pi_t(s_{it}, h_{i,t}, k_{i,t}; z_t, w_t) = (1 - \tau_t^C)\big((1 - \tau_t^{VAT}) y_{it}(s_{it}, z_t, h_{i,t}, k_{i,t}) - w_t h_{i,t} - R_t k_{i,t}\big) \tag{1.5}$$

where the production function is as in Section 1.2.2, τ_t^{VAT} gives the Indonesian value-added tax and τ_t^C the corporate income tax that is levied on taxable profits. Given the frictions in this economy, both tax instruments generally distort input choices. R_t gives plants' capital borrowing rate, which is equal to the deposit rate r_t plus depreciation δ assuming competitive rental markets. Idiosyncratic risk s_{it} follows a Markov process of order one. We further assume that s_{it} is exogenous and independent of aggregate technology z_t. For z_t, we leave the path unrestricted, but assume it is exogenous.

What drives slow plant adjustments? We assume that plants face labor adjustment costs. These capture, for example, managerial time constraints that arise from the time it takes to hire, fire and reorganize production tasks, a key constraint for plant growth in developing countries (Bloom et al. 2013, 2020).[15] Following the literature, we model them as follows (e.g. Cooper, Gong, and Yan 2018; Cooper, Haltiwanger, and Willis 2015):

$$AC(h_{i,t-1}, h_{i,t}) = \begin{cases} F^+ + c_0^+(h_{i,t} - h_{i,t-1}) + \frac{c_1^+}{2}\left(\frac{h_{i,t} - h_{i,t-1}}{h_{i,t-1}}\right)^2 h_{i,t-1} & \text{if } h_t > h_{t-1} \\ 0 & \text{if } h_t = h_{t-1} \\ F^- + c_0^-(h_{i,t-1} - h_{i,t}) + \frac{c_1^-}{2}\left(\frac{h_{i,t} - h_{i,t-1}}{h_{i,t-1}}\right)^2 h_{i,t-1} & \text{if } h_t < h_{t-1} \end{cases} \tag{1.6}$$

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. Importantly, there are convex adjustment costs whose importance is captured by c_1 and which capture costs of growing (c_1^+) or shrinking (c_1^-) plants quickly. Convex adjustment costs are key to explain the slow growth of plants over time and are a key determinant of the speed of transition growth. We allow all costs to be asymmetric to accommodate that firing and hiring is often regulated differently. At last, we index all adjustment costs in terms of wages since wage indexation provides a simple way to let costs grow with the economy. Besides this indexation, in the baseline model, we

[15] In Appendix 1.B.1, we provide a simple microfoundation in terms of the costs of scarce managerial time to show how organizational changes induce convex costs. An alternative interpretation of convex adjustment costs is given in labor search models where they are rationalized via convex (reduced-form) hiring or vacancy posting costs (e.g. Bilal et al. 2022; Coşar, Guner, and Tybout 2016)). The key difference is that adjustment costs in search models become partly functions of equilibrium outcomes such as market tightness. We abstract from this general equilibrium mechanism here given that the primary focus of the paper is on longer run growth dynamics and not business cycle variation in unemployment and market tightness.

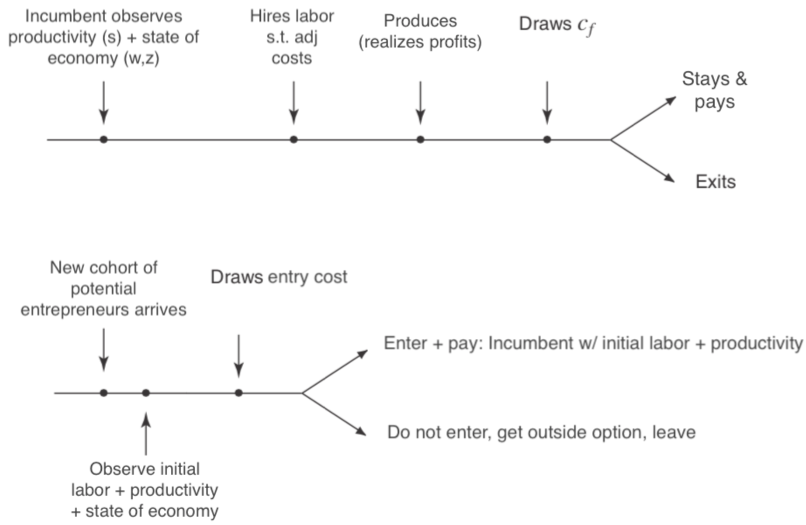


Figure 1.10: Timing in time t for manufacturing firms.

assume that adjustment cost parameters are fixed over time, but the model and estimation can accommodate time-varying parameters and we return to this point in later sections.

On top of adjustment costs, we assume that plants face a standard financing constraint that captures underdeveloped financial markets, in line with a large Macro Finance literature that emphasizes their importance (e.g. Buera and Shin 2013; Midrigan and Xu 2014; Moll 2014). Specifically, we assume the following working capital constraint:

$$w_t h_{it} \leq \kappa_t (1 - \tau_t^{VAT}) y_{it} \quad \text{(Working capital constraint)} \tag{1.7}$$

As we show in Appendix 1.B.2, the constraint can be microfounded as a simple limited enforcement problem whereby plants may want to run deficits in order to grow their workforce, but cannot commit to repaying. Limited contract enforcement due to a weak judicial system then leads to the borrowing constraint in equilibrium. κ_t then captures the probability that the judicial system will enforce the contract, with higher κ_t mapping to stronger institutions and a less binding constraint. The reason that plants may want to run deficits to grow their workforce is that in the presence of convex adjustment costs it is costly to build up a plant's workforce quickly, so that even less productive plants may want to run deficits to build up their workforce hoping for good future productivity realizations.

The timing of manufacturing production is summarized in Figure 1.10. At the beginning of a period, incumbents observe their current productivity and make production decisions. Plants' payoff-relevant aggregate state is given by: $\Omega_t \equiv \{z_t, w_t, r^*\}_t^{\infty}$. After production takes place, incumbent plants incur a fixed cost of production $c_{i,t}^F$, upon which plants decide whether they want to continue producing (and pay $c_{i,t}^F$) or permanently exit (as in Clementi and Palazzo 2016). The fixed cost is drawn from a distribution G, which we assume to be Gumbel with scale and variance parameters (μ_t^X, σ_t^X). A larger variance in the fixed costs rationalizes more overlap in the labor and productivity distributions of surviving and exiting plants. The exit decision of the plant depends on the plant's expected future value, the cost shock as well as the costs of closing down the plant (as in Hopenhayn and Rogerson 1993):

$$\max\left\{\beta\mathbb{E}[V^M(s_{i,t+1},h_{i,t};\Omega_{t+1})|s_{i,t},h_{i,t};\Omega_t]-c^F_{i,t},-C_E(h_{i,t})\right\} \tag{1.8}$$

where V^M gives the continuation value of an incumbent plant and $C_E(h_{i,t})$ the costs of closing down the plant. The above maximization problem implicitly defines plant i's survival probability that the operating cost draw $c^F_{i,t}$ is lower than its future expected continuation value: $\lambda(s_{i,t},h_{i,t};\Omega_t)\equiv\mathbb{P}\big(x\geq c^F_{i,t}\big)=G\big(x\big)$ where $x\equiv\beta\mathbb{E}[V^M(s_{i,t+1},h_{i,t};\Omega_{t+1})|s_{i,t},h_{i,t};\Omega_t]+C_E(h_{i,t})$.

The ex-ante value of an incumbent manufacturing plant can be written in recursive form according to:

$$\begin{aligned}V^M(s_{i,t},h_{i,t-1};\Omega_t)=\max_{h_{i,t}\leq\bar{h},k_{i,t}}\Bigg\{&\pi(s_{it},h_{i,t},k_{i,t};z_t,w_t)-w_tAC_t(h_{i,t-1},h_{i,t})+\lambda(s_{i,t},h_{i,t};\Omega_t)\Big\{\\&-\mathbb{E}_c[c^F_{it}|\text{stay}(s_{i,t},h_{i,t};\Omega_t)]+\beta\mathbb{E}\big[V^M(s_{i,t+1},h_{i,t},\Omega_{t+1})|s_{i,t},h_{i,t},\Omega_{t+1}\big]\Big\}\Bigg\}\end{aligned} \tag{1.9}$$

Plants have a common discount factor $\beta=1/(1+r^*)$, which is pinned down by the world interest rate. The presence of adjustment costs and financial constraints in combination with productivity dynamics makes this a dynamic problem since plants take into account that contemporaneous changes in their inputs influence adjustment and financing costs in the future.

Next, we consider endogenous plant entry. As visualized in Figure 1.10, each period there is a cohort of potential entrants (PE) of measure $|PE|_t$. Each potential entrant draws a random entry cost c^E_{it} from a distribution P, which they need to pay in case they start producing. Again, we assume that entry costs follow a Gumbel distribution with scale and variance parameters (μ^E_t,σ^E_t). Potential entrants differ in their idiosyncratic productivity s_{it} and their initial labor $h_{i,t}$, which they know when making the entry decision for producing in period t. The initial heterogeneous level of labor and productivity is key to capture that there is plant entry of many small as well as some very large plants that matter in the aggregate. It also accounts for the fact that we only observe and model plants with 20 or more workers. The exogenous distribution of potential entrants is given by $PE_t(h_t,s_t)$, which is time-varying due to exogenous reasons such as demographic changes that have been shown to be key for explaining variation in firm creation over time (Bernstein et al. 2022; Karahan, Pugsley, and Şahin 2019; Liang, Wang, and Lazear 2018). Potential entrant i with entry cost shock c^E_{it} enters if its net value is positive:

$$V_{PE}(s_{i,t},h_{i,t};\Omega_t)=\max\left\{V^M(s_{i,t},h_{i,t};\Omega_t)-c^E_{it},0\right\} \tag{1.10}$$

where we have normalized the outside option to zero. Similar to exit, this gives the following mapping $\mathbb{P}\big(V^M(s_{i,t},h_{i,t};\Omega_t)\geq c^E_{it}\big)=P\big(V^M(s_{i,t},h_{i,t};\Omega_t)\big)$. Note that in this specification the initial mass and distribution of entrants is endogenous, but entrants only start making input choices the period after they entered. A time-varying distribution of potential entrants also allows us to deal with plant entry jumps in years of the economic census as we further discuss in the estimation section. We denote the endogenous mass of entry for each state (h_t,s_t) in period t by $\mu(h_t,s_t)$, which is a function of Ω_t. Similarly, we define by $m(h_t,s_t)$ the endogenous mass of producing plants for each state (h_t,s_t) in period t. With slight abuse of notation, denote by M_t the set of producing plants at time t.

Rest-of-the-economy

We model the rest-of-the-economy parsimoniously as a representative firm with a decreasing returns to scale (DRS) production function:

$$Y_t^R = A_t \left(H_t^R\right)^{\theta_R} \quad \text{with} \quad \theta_R \in (0,1) \tag{1.11}$$

where A_t is time-varying TFP, H_t^R gives labor employed and θ_R gives the output elasticity in the rest-of-the-economy. Decreasing returns to scale ensure that economy-wide wages can be affected by changes in manufacturing. 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: $\pi_t^R(A_t, w_t, \tau_t) = Y_t^R(A_t) - (1+\tau_t^R)w_t H_t^R$, subject to labor demand wedges τ_t^R. Labor demand is then given by:

$$H_t^{R*} = \left(\frac{\theta_R A_t}{(1+\tau_t^R)w_t}\right)^{\frac{1}{1-\theta_R}} \tag{1.12}$$

Labor demand wedges in the rest-of-the-economy are a simple way to capture observed variation in the labor intensity of output and one can think of them as changes in labor frictions. A_t captures changes in technology of the rest-of-the-economy. We allow both A_t and τ_t^R to change over time, but treat them as deterministic and exogenous paths.

Households

There is a continuum of households j that are characterized by their exogenous household-specific efficiency units of labor h_{jt} and whose exogenous mass at time t is denoted by L_t. Households supply labor inelastically so that the aggregate labor supply is given by $H_t = \int_j h_{jt} dj$. Changes in L_t and $\frac{H_t}{L_t}$ capture changes in the working population and education per worker respectively. We abstract from consumption-savings decisions by assuming that households are hand-to-mouth, simply consuming their labor income y_{jt} net of transfers from the government and plant profits T_{jt}: $c_{jt} = y_{jt} + T_{jt}$.[16] Households allocate their labor supply across both sectors based on maximizing labor income: $y_{jt} = \max\{h_{jt} w_t^M, h_{jt} w_t^R\}$.

Equilibrium

We assume that the observed growth path in the data is characterized by a path of per-period perfect foresight *Recursive Competitive Equilibria*.

Definition 1 (Model fundamentals.)**.** *Model fundamentals at time t capture all exogenous model parameters, processes and distributions as given by:*

$\Theta_t^F = \{\theta, \alpha, \delta, F^-, c_0^-, c_1^-, F^+, c_0^+, c_1^+, \{A_t, \tau_t^R, H_t, PE_t, z_t, \tau_t^C, \tau_t^{VAT}, \kappa_t, \mu_t^X, \sigma_t^X, \mu_t^E, \sigma_t^E\}_t^\infty\}.$

We further denote by $\bar{\Theta}_t^F$ the modified set of model fundamentals where all fundamentals are fixed to

[16] We only make this assumption to fix ideas. Given the small open economy setup, the domestic supply of capital is inelastic to changes in domestic savings behavior such that the production side – which is the focus of this paper – would look exactly the same if heterogeneous households would instead solve a savings and consumption choice. Given inelastic labor supply, the model is also isomorphic to one with uniform labor income taxes in both sectors.

their value at time t forever. At last, denote by $\Theta_t^F \setminus \{x = \bar{x}\}$ the modified set of model fundamentals where only model fundamental x is changed to $\bar{x}$.

Definition 2 (Initial distribution.)**.** *The distribution of surviving plants from period $t-1$ gives the initial distribution at time t and is denoted by S_t.*

A path of perfect foresight *Recursive Competitive Equilibria* starting at time t is then given by model fundamentals Θ_t^F, an initial distribution S_t, and endogenous sequences of prices $\{w_t, r^*\}_t^\infty$, corresponding quantities and distribution of producing plants $\{m_t\}_t^\infty$ such that each period $\tau \in [t, \infty)$:

1. The rest-of-the-economy sector statically chooses optimal labor demand maximizing profits taking as given productivity A_τ, the wage w_τ and labor demand wedges τ_τ^R.
2. Manufacturing plants choose optimal labor and capital demand.
3. Potential entrants optimally make entry and incumbents optimally make exit decisions.
4. Households inelastically supply total labor H_τ and optimally allocate labor across sectors to maximize labor income.
5. The aggregate wage w_τ adjusts to ensure that the labor market clears: $H_\tau = H_\tau^R(w_\tau, A_\tau, \tau_\tau^R) + \sum_{i \in M_\tau} h(s_{i\tau}, h_{i,\tau-1}; \Omega_\tau)$
6. The government runs a balanced budget by levying a value added and corporate income tax and redistributing revenue back to households.
7. The capital market clears every period such that international capital supply equals domestic capital demand: $\sum_{i \in M_\tau} k(s_{i\tau}, h_{i,\tau-1}; \Omega_\tau) = K_\tau^{INT}$.
8. The mass of active plants in τ and previous aggregate state $\Omega_{\tau-1}$ is equal to surviving plants from $\tau - 1$ plus endogenous new entrants:

$$\forall (s_\tau, h_\tau): \quad m(s_\tau, h_\tau; \Omega_\tau) = \sum_{s_{\tau-1}, h_{\tau-1}} \Big(\mathbb{1}_{h^* = h_\tau} \mathbb{P}[s_\tau | s_{\tau-1}] \lambda(s_{\tau-1}, h_{\tau-1}; \Omega_{\tau-1}) \times m(s_{\tau-1}, h_{\tau-1}; \Omega_{\tau-1}) \Big) + \mu(s_\tau, h_\tau) \tag{1.13}$$

9. The goods market clears each period such that total production is either consumed or exported: $Y_\tau = \sum_{i \in M_\tau} y_{i,\tau} + Y_\tau^R = C_\tau + NX_\tau$ where $NX_\tau = EXP_\tau - K_\tau^{INT}$ are net exports.[17]

The observed growth path features a combination of changes in model fundamentals that move the economy's *steady state* and transition growth as the economy is trying to catch up to this steady state. We now formalize these concepts.

Definition 3 (Balanced Growth Path (BGP) and Steady State (SS))**.** *Along a* BGP, *underlying technology in both sectors (A_t, z_t) and the endogenous wage grow at the same constant rate, while all remaining model fundamentals and the endogenous distribution of plants stays constant (see details in Appendix 1.B.3). A* steady-state *is a BGP for which the growth rate is zero. Both* BGP *and* steady-state *are uniquely defined by* model fundamentals Θ^F *that admit a BGP/SS.*[18]

Definition 4 (Transition path)**.** *The unique perfect foresight* transition path *starting at t towards*

[17]Note that we have implicitly treated all entry costs, fixed costs and adjustment costs as shadow costs here, as they neither directly enter labor market clearing nor the aggregate resource constraint.

[18]Uniqueness depends on the uniqueness of the individual policy functions and the unique mapping between policies and prices. We treat our numeric algorithm as formally defining the equilibrium refinement conditions sufficient for uniqueness.

a BGP is defined by an initial distribution S_t and model fundamentals Θ_t^F (which admit a BGP), and gives a path of equilibrium wages over the transition.[19]

In Section 4, we use the model to separately quantify the role of *transition growth* from changes in the *steady state*.

1.3.2 Estimation

The model captures a race between changes in model fundamentals and the distribution of plants (over the state space) trying to catch up to these changes. We now take this model to the data and show how to disentangle changes in the distribution from changes in model fundamentals. Estimation proceeds in three main steps, while an additional step is needed for model counterfactuals. The first step identifies equilibrium prices – only wages in our case – in the data. We take an equilibrium estimation approach (Hotz and Miller 1993; Bajari, Benkard, and Levin 2007; Caliendo, Dvorkin, and Parro 2019), which means that we treat our model as generating the equilibrium wage path we observe in the data. We can thus treat the path of equilibrium wages as fixed throughout the estimation and only need to solve for changes in the equilibrium wage path for counterfactuals. This greatly simplifies the estimation as it avoids solving for the equilibrium path of the model during the estimation. In the second step, we identify the distribution over the entire state space of the economy over time and use this to back out related model fundamentals such as the initial distribution. The third step then solves for remaining fundamentals that are related to the dynamic input and exit choices of plants drawing on observed choices of plants conditional on the state space. In this step, we also need to make an explicit assumption about the evolution of model fundamentals beyond the time frame of our data. To conduct model counterfactuals, we further back out model fundamentals that are not needed to solve the baseline economy, such as fundamentals of the rest of the economy.

With each estimation step, we enforce more model structure and assumptions, making parameter identification very transparent. An important benefit of our approach is that we can directly draw on the production function and aggregate technology estimates discussed in Section 1.2.2 whose identifying assumptions nest our model. Table 1.1 provides an overview of the estimation steps and all model fundamentals and estimates. We now discuss each step in more detail. Given the larger number of model fundamentals, we focus throughout on the most important parts and relegate details to Appendix 1.B.5.

Step 1: Equilibrium wage estimation

In the first step, we estimate the path of equilibrium wages that clear labor markets in each period. While prices could in principle be directly observed in the data, based on our model, we only observe plants' wage bills ($w_t h_{it}$), which are a combination of the wage and the quantity of labor. We are only interested in changes over time and thus normalize the level of initial wages w_0 to unity. Ideally, we would like to capture changes in the wage by looking at wage changes for a worker whose efficiency units of labor remained constant. This would for example avoid any assumptions on how workers with different skills select across plants. In the

[19]Uniqueness of the *transition path* can be proven via contraction mapping arguments over the path of price expectations. Again, our numeric algorithm for the transition gives a unique path for perfect foresight equilibria and we treat this as formally defining the equilibrium refinement conditions sufficient for uniqueness.

Table 1.1: Overview of parameter identification and estimation

Object	Description	Type	Identification idea	Value	Details
Parameterization:					
r^*	World interest rate	F	Risk-free rate	0.04	
δ	Depreciation rate	F	Standard	0.1	
τ_t^C	Corporate tax	F	Official rate	0.2	Section A.2.4
τ_t^{VAT}	VAT	F	Official rate	0.1	Section A.2.4
Estimation:					
Step 1:					
θ	Prod function	F	Control function	0.694	Section 2.3.2
α	Prod function	F	Control function	0.03	Section 2.3.2
w_t	Wage path	E	$\Delta_i(w_t h_{it})/l_{it}$	Fig. 12	Section 3.2.2
κ_t	Borrowing constraint	F	Max labor share	1.7	Section A.2.4
Step 2:					
z_t	Techn path	F	$\Delta_i$productivity$_{it}$	Fig. 8	Section 2.3.2
$\mathbb{P}(s'\|s)$	Transition matrix	F	Obs. transitions		Section 3.2.2
Init distrib		F	Obs. survivors		Section 3.2.2
E_t	Entrants	E	Obs. entry		Section 3.2.2
Step 3:					
Adj costs		F	Euler CCC	Table 2	Section 3.2.2
Cost ratio	Fixed cost	F	Euler CCC	Table 2	Section 3.2.2
Cost level		F	Match mass 2015		Section 3.2.2
For counterfactuals:					
A_t, θ_R, τ_t^R	Rest-of-Economy	F	First-order condition		Section A.2.4
Entry costs		F			Section A.2.4
PE_t	Potential Entrants	F	Entrants + entry proba		Section A.2.4

Details: Types are: F(undamental) and E(quilibrium object). The former stay fixed in counterfactuals, the latter change endogenously. If applicable, reported standard errors correct for multi-step estimation procedure and cluster at the plant-level by using block bootstrap across all estimation steps (This is currently still work in progress).

Figure 1.11: Evolution of estimated real wage in Indonesian manufacturing

Notes: Based on within-task changes in the wage bill per worker by task (production and non-production). Data: Indonesian manufacuring census (1975-2015, 20+ workers).

absence of worker-level data that spans the entire time period, we instead draw on changes in within-plant per worker wages for similar job types, exploiting that the Indonesian data reports wages and the number of workers separately for production and non-production work. Our identification strategy for the wage allows for arbitrary sorting of workers with different skills not only across plants but also across different job types within plants, but restricts changes in the skill sorting within job types over time. Formally, we assume that plant i uses on average the same skills per worker within job types k: $h^k_{it}/l^k_{it} = \alpha^k_i \cdot \varepsilon^k_{it}$. ε^k_{it} allows job types within plants to vary in their skill intensity around α^j_i over time. With standard restrictions on ε^k_{it}, this ensures that wages are identified from:

$$\mathbb{E}_i \left[log(w_{t+1} h^k_{it+1}/l^k_{it+1}) - log(w_t h^k_{it}/l^k_{it}) \right] = log(w_{t+1}) - log(w_t)$$

As our estimate of changes in wages, we use median within-plant-worker-type changes in wages, weighting observations by the average of total workers of type k between t and $t+1$, ensuring that wages are identified from median wage changes of workers (not plants). If anything, we think this estimator overestimates wage increases, because (1) any increase in within-worker human capital (e.g. on-the-job learning) will be attributed to increases in the wage, and (2) the estimates are for surviving plants, which might see more wage growth.

Figure 1.11 plots the estimated real wage series in the data. Our preferred estimator uses wage bills for production workers only, as production workers are relatively homogeneous and thus the identifying assumption is more likely to hold. For completeness, we also report estimated wage series using non-production workers and pooling all workers. Our estimates show that wages per efficiency unit of labor increased by more than 85% over the 40-year period with important variation over time. Given that the average wage bill per manufacturing worker increased roughly 4-fold in the data, these estimates imply that the average manufacturing worker in 2015 was about 2.2x more efficient than the average manufacturing worker in 1975. We find this a reasonable estimate given the large educational gains of Indonesian workers observed over this period. As external validation, Gathen (2021) also finds a similar wage increase from separate estimates on Indonesian worker data between 1998 and 2015.

Step 2: Mapping the entire distribution over the state space

In the second step, we map any plant in our data to the state space of our model: $(h_{it-1}, s_{it}, w_t, z_t)$. While our data captures a discrete number of plants, we treat this mass as continuous for the estimation. This allows us to identify changes in the entire distribution of plants over the state space over time, which is crucial to determine the potential for transition growth. Specifically, we use this mapping to identify two key model fundamentals – the productivity process and the initial distribution – and an equilibrium object which changes in counterfactuals: the distribution of entrants.

To identify the state space, we draw on the wage estimates from Step 1, which together with a plant's wage bill ($w_t h_{it}$) identifies h_{it-1}. To obtain plant productivity and technology, we draw on the identification and estimation approach from Section 1.2.2, which hold under weak assumptions on exit and input choices. Having identified plant productivity s_{it} (up to a normalization of z_t), we estimate the dynamic process of s_{it} by discretizing s and then estimating the transition matrix $\mathbb{P}(s'|s)$ non-parametrically using (pooled) within-plant productivity changes in the data. The transition matrix is a fundamental of the economy and is identified based on within-plant changes in productivity conditional on previous productivity, and only requires that all productivity states are observed with positive probability at some point.[20]

Next, we exploit the state space mapping to identify the initial distribution of surviving plants over $(s_{i,t}, h_{i,t-1}; \mathbf{\Omega})$ in 1976. While our data starts in 1975, the first year for which we can identify $h_{i,t-1}$ is 1976. We treat this initial distribution as a model fundamental, implicitly assuming that in any counterfactual, initial survivors do not anticipate any changes to the baseline equilibrium paths of wages and technology prior to 1976. The main benefit of directly taking the initial distribution from the data is that we can remain agnostic about its origins and allow the data to reveal the initial degree of misallocation. A downside is that if the model does not capture all mechanisms of dispersion over the state space, the initial distribution may look more "misallocated" than it actually is; leading to overestimating model-implied transition dynamics.

At last, the state space mapping also allows us to identify time-varying entrant distributions $E(s_t, h_t; \mathbf{\Omega})$, which are equilibrium objects but related to the fundamental potential entrant distributions via: $PE_t(s_t, h_t; \mathbf{\Omega}) = E_t(s_t, h_t; \mathbf{\Omega})/\mathbb{P}_E(s_t, h_t; \mathbf{\Omega})$, where $\mathbb{P}_E(.)$ gives the entry probability, which is a function of the model-implied value of entering as well as the parameters of the entry cost distribution. For the model estimation along the baseline equilibrium path, we treat these equilibrium entrant distributions as fixed. Our approach implies that the baseline model exactly replicates observed plant entry. This is in contrast to plant-level exit and labor demand decisions, for which our estimation approach allows the model to fail.

Step 3: Estimating the dynamics of the model

Step 3 reveals the remaining parameters of the economy that are needed for the baseline model: fixed cost parameters that govern entry and exit decisions as well as adjustment cost parameters that govern how plants make dynamic labor choices. This step enforces more model

[20] While an imperfect ex-post test, we check ergodicity of the implied idiosyncratic productivity process s_{it} by verifying that all states in the discretized transition matrix can be eventually reached. In Appendix 1.B.5, we provide details on the discretization of productivity and labor, which we rely on for numerically solving the model and counterfactuals.

structure, particularly on how plants make dynamic input, exit and entry choices, and how plants form expectations over the future. We separate this section into two parts with the first part only exploiting optimal plant-level choices across consecutive periods, while the second part also enforces long-run expectations.

Euler equation CCC estimation We identify most remaining parameters by exploiting observed exit and labor input choices conditional on the state space, drawing on standard conditional choice probability (CCP) and continuous conditional choice (CCC) Euler estimation techniques (Hotz and Miller 1993; Bajari, Benkard, and Levin 2007). Taking first-order conditions with respect to labor from the incumbent's value function above and directly plugging in the envelope condition, we obtain the following Euler equation:

$$\begin{aligned} 0 = & \underbrace{\frac{\partial \pi(s_{i,t}, k_{it}, h_{i,t}, z_t)}{\partial h_{i,t}}}_{\text{Labor wedge}} - \underbrace{w_t \frac{\partial C_h(h_{i,t}, h_{i,t-1})}{\partial h_{i,t}}}_{\text{Current marginal adj costs}} \\ & + \underbrace{\frac{\partial \lambda(s_{i,t}, h_{i,t}, \Omega_t)}{\partial h_{i,t}} \left\{ -\tilde{g}(s_{i,t}, h_{i,t}, \Omega_t) + \beta \mathbb{E}\big[V(s_{i,t+1}, h_{i,t}, \Omega_{t+1}) | s_{i,t}, h_{i,t}, \Omega_t\big] \right\}}_{\text{Marginal benefit on survival}} \\ & + \underbrace{\lambda(s_{i,t}, h_{i,t}, \Omega_t) \left\{ -\frac{\partial \tilde{g}(s_{i,t}, h_{i,t}, \Omega_t)}{\partial h_{i,t}} + \beta \mathbb{E}\left[-w_{t+1} \frac{\partial C_h(h_{i,t+1}, h_{i,t})}{\partial h_{i,t}} | s_{i,t}, h_{i,t}, \Omega_t \right] \right\}}_{\text{Marginal benefits on future costs}} \end{aligned} \tag{1.14}$$

where we have used $\tilde{g}(s_{i,t}, h_{i,t}, \Omega_t)$ to denote the expected fixed cost conditional on surviving to emphasize that it is a function of the state space.

The Euler equation, which holds for any plant that is optimally adjusting labor, says that plants should equalize today's marginal product of labor with the marginal costs of labor and current as well as future labor adjustments. Adjustment costs give a natural explanation for why there is a "wedge" between the static marginal product and the marginal costs of labor (Hsieh and Klenow 2009). For our estimation purposes, the important features of the Euler equation are that it holds along the transition, and that it gives a nonlinear equation in observable plant-level choices (exit and input choices) and parameters that govern survival probabilities as well as adjustment costs. Specifically, marginal adjustment costs are a function of adjustment cost parameters. The tricky terms are expected and marginal expected fixed costs, marginal survival probabilities and the expected future continuation value. As we show in Appendix 1.B.6, the Gumbel distribution for the fixed costs ensures that we can analytically invert all of these terms as functions of observed survival probabilities and parameters of the Gumbel distribution.

Appendix Table B.1 presents the non-linear least squares (NLS) estimation results and Appendix 1.B.6 gives the exact estimating equation and estimation details. The Euler equation flexibly identifies marginal adjustment costs. Intuitively, linear adjustment costs are identified from the observed labor wedge across plants and the probability of switching between shrinking and growing as determined by the volatility of the estimated productivity process. Convex costs instead scale with the labor growth and are thus identified from the variation in within-plant labor demand growth across periods, again conditioned by the observed volatility of the productivity process. We find sizable adjustment costs – especially convex costs on growing

– that rationalize why even productive plants (with a high labor wedge) conditional on previous plant employment do not grow faster. Quantitatively, our estimates imply that growing a plant's workforce by 20% within a year – a growth rate slightly above the 75th percentile – leads to adjustment costs that are about 75% of the previous wage bill. Informed by faster observed shrinking conditional on productivity, convex costs on shrinking are estimated to be less than half as big. We also estimate that a plant pays almost 75% of a new worker's annual wage in the form of hiring costs, which is identified from the high observed wedge between the marginal product and wage and the high volatility of productivity that make any investments in the workforce risky. In Appendix 1.B.6, we also report time-varying estimates of adjustment costs. If anything, we find that convex adjustment costs tend to increase over time beyond what is implied by increases in the wage, pointing away from a reduction in frictions driving Indonesian growth.

Solving the baseline model The Euler equation only identifies the ratio between the level and scale of the fixed cost distribution that determine plant exit. To separately identify the level, we solve the model and match one moment in the data and model: the mass of plants in 2015, assuming that the census is complete for that year. The estimated level and scale of the fixed cost distribution rationalize average exit rates and the low but positive correlation with underlying productivity and size.

Solving the model introduces two issues that are common to equilibrium estimation approaches. First, by requiring to solve for plants' value functions, we need to make an explicit assumption on long-run expectations of plants after the year 2015 when our data ends. We assume that after 2015, plants expect to be on a balanced growth path with manufacturing technology growing at the average rate at which it grew in the preceding ten years.[21] Given that technology in manufacturing grew strongly since 2000, this assumption implies optimistic expectations, in line with low exit and strong observed plant growth in the years prior to 2015.

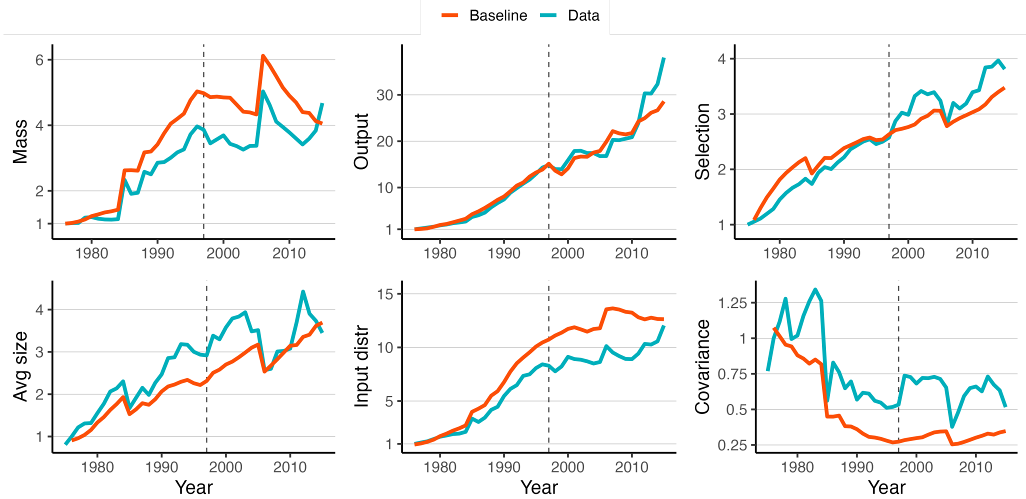
The second issue for the equilibrium estimation is that enforcing revealed equilibrium prices along the estimation of the baseline economy does not guarantee that these prices actually clear markets in our model over time. We ensure consistency – informed by our specific context and data availability – by treating the observed data as correctly revealing prices, but not necessarily correctly revealing aggregate labor demand and supply. As discussed in Section 1.2.1, our data does not correctly reveal aggregate labor demand and supply due to mis- and non-reporting and as explicitly taken into account in our measure of plant exit. We thus use the model-implied aggregate labor demand and supply for the baseline growth path and enforce the implied fundamentals that ensure market clearing for all future model-based counterfactuals. This is how our approach also ensures that counterfactuals are consistent with the baseline model economy.

1.3.3 Evaluating model fit

To assess how well the model fits the data over time, we revisit three main results from Section 1.2. Having estimated the model on micro moments, we start out by moving from "micro to

[21] We provide technical details on the BGP and how to solve for the stationarized value function in 2015 in Appendix 1.B.3. An alternative would be to solve for a continued transition towards a long-run BGP by making explicit assumptions on how all fundamentals evolve after 2015. We do not follow this approach, because it adds substantial additional computational costs while requiring similarly strong assumptions.

Figure 1.12: Baseline model fit



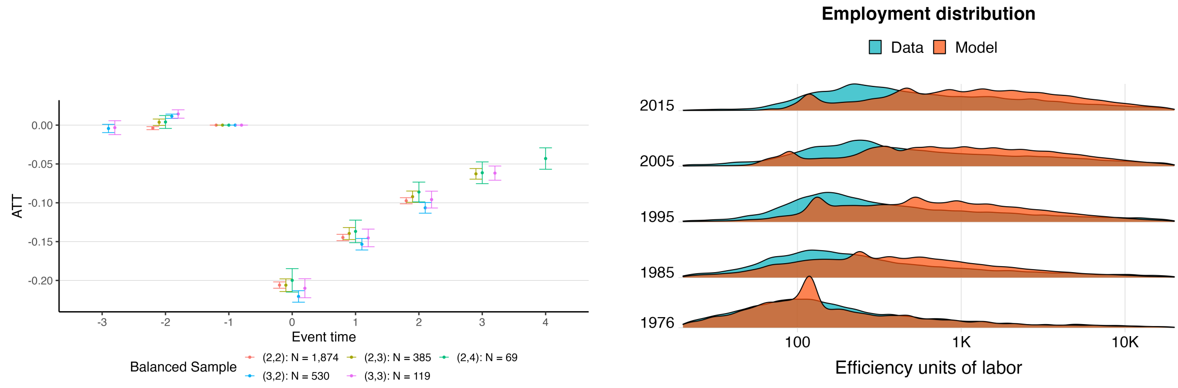
Notes: All graphs report results for manufacturing only and for each panel, both series are normalized to the initial level in the data. Average size is reported in efficiency units of labor. Selection (measuring average productivity), input distribution (the sum of the Cobb-Douglas aggregator of capital and labor) and covariance (the covariance between plant productivity and the share in inputs as given by the Cobb-Douglas aggregator) refer to the three corresponding terms in the growth accounting identity introduced in Section 2.

macro", evaluating the model's aggregate predictions. We then validate how well the model matches plants' dynamic input choices and changes in the entire plant distribution over time.

Figure 1.12 shows how the baseline model fits the mass of plants, aggregate output in manufacturing and the three main endogenous components that we used in Section 1.2.2 to formally decompose manufacturing growth. For completeness, we also plot the evolution of the average plant size as measured in efficiency units of labor. The model closely tracks the more than 30-fold increase in manufacturing output over time, including rapid growth in the absence of technology improvements until the Asian Financial Crisis, the decline during the crisis and the fast post-crisis growth. Overall, the model tends to slightly underestimate output at the plant level given the higher model-implied mass of plants.

The accounting identity helps to understand why the model fits the aggregate data well and where the model underperforms. First, as the main component of aggregate output growth, the model closely tracks the distribution of labor and capital across the endogenous plant distribution over time. Neither total labor demand, the total number of plants over time (apart from the first and last year) nor the distribution of inputs is hit by construction. The model captures the right degree of slow labor accumulation across the entire plant size distribution over time, which can be even more clearly seen from looking at the average plant size. Secondly, the model tracks well the evolution of average productivity across plants, capturing well the endogenous selection of plants over time. If anything, the model overpredicts productivity growth in the early years and underpredicts towards the end. This can be in part explained by a too fast productivity convergence implied by the estimated productivity process, stemming from frequent temporary productivity shocks that lead to overestimating productivity transitions. At last, the model also performs reasonably well on the most difficult part: the endogenous evolution of the joint distribution of productivity and inputs, as captured by the covariance term. Here, the model captures the decline in the covariance over time, as many small and productive plants enter and resources only reallocate slowly due to sizable adjust-

Figure 1.13: Model validation: Event study and changes in distribution



Notes: Left graph replicates event study exactly as in Section 2 using simulated data from model growth path. Right graph gives changes in the employment distribution (using efficiency units of labor) for model versus data. For both data and model, wage bills are used and divided by (same) estimated wage path.

ment frictions.

Next, we are interested in whether the model is also in line with the micro-level labor dynamics. Figure 1.13 Panel A shows the same estimated event study as in the data, using the exact same treatment definition and sample restrictions but now using simulated data from our baseline model from 1976 to 2015. The event study results align well. Both pre-trend, the exact magnitude of the treatment effect at impact and the slow recovery of the labor share follow the data. We also note that the balanced sample restrictions imposed in the data lead to very similar sample sizes in the simulated data indicating that we identify similarly selected sets of plants (which received a rare permanent productivity shock).

At last, Figure 1.13 Panel B compares model-implied versus observed changes in the entire employment distribution of plants over time. In 1976, the first year that our model predicts plant decisions, the distributions are still largely indistinguishable. Over time, the employment distribution moves strongly to the right with average employment increasing almost 4-fold and the mass of the distribution shifting from a strong left tail towards the right. The model tracks this overall change well, but slightly overpredicts the right tail. Importantly, we do not see a marked deterioration of the distribution even after 40 years of endogenous evolution.

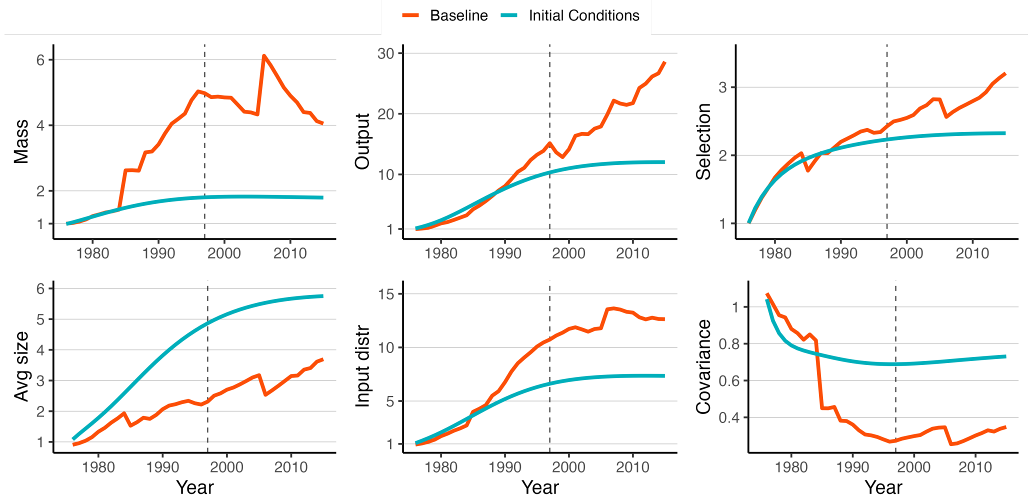
1.4 Quantifying the drivers of aggregate growth

Using the estimated model, we now quantify the drivers of growth. Specifically, we quantify (1) the importance of initial transition growth, (2) the continuing importance of transition growth over the course of development, and (3) the role of policy. We present each in turn.

1.4.1 Initial conditions and the role of transition growth

We start by quantifying the importance of transition growth based on the initial economy at the onset of the Indonesian growth miracle in 1976. How much would the 1976 economy have grown purely from transition growth in the absence of any further changes in model fundamentals? For this, we start from the initial economy with the initial distribution M_{1976}, fix initial model fundamentals $\bar{\Theta}^F_{1976}$ to their value in 1976 and solve for the perfect foresight tran-

Figure 1.14: Growth from initial conditions



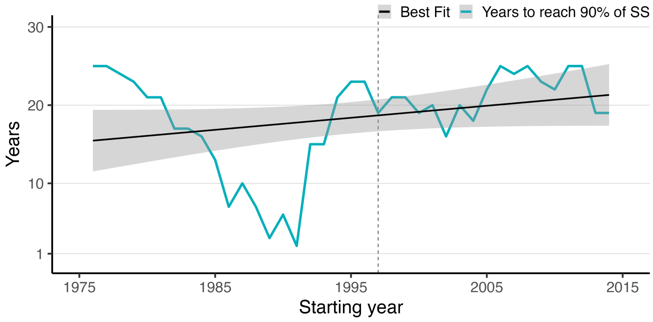
Notes: Results for counterfactual where economy evolves only based on initial conditions (all fundamentals fixed to initial level). All graphs report results for manufacturing only. Average size is reported in efficiency units of labor. Selection (measuring average productivity), input distribution (the sum of the Cobb-Douglas aggregator of capital and labor) and covariance (the covariance between plant productivity and the share in inputs as given by the Cobb-Douglas aggregator) refer to the three corresponding terms in the growth accounting identity introduced in Section 2.

sition path (see Section 1.3), which includes solving for a counterfactual path of equilibrium wages that clears the labor market over time as the initial distribution of plants transitions towards the steady state distribution defined by $\bar{\Theta}^F_{1976}$.

Figure 1.14 highlights the resulting counterfactual growth in manufacturing. Overall output in manufacturing increases roughly 12-fold over time, accounting for 42% of the overall output gains compared to the baseline (model) economy. The reason is that young and small plants – which dominate the initial distribution and new entrants – gradually hire more workers and increase their productivity through a combination of productivity convergence and the exit of less productive plants. At the same time, entry consistently exceeds exit and the mass of plants gradually doubles over time. The increase in workers across manufacturing plants is mostly driven by the reallocation of labor from the rest of the economy: in the absence of observed changes in the aggregate labor supply and technology in the rest of the economy, the model predicts that Indonesia would have seen a manufacturing miracle with manufacturing labor and output shares reaching close to 30% over 40 years (in contrast to observed shares of less than 10%). Average plant size increases far more rapidly because in the absence of productivity improvements in the rest of the economy, aggregate wages stay almost 40% lower than in the baseline economy. Hence, initial conditions explain more than all of the increase in the average plant size over time, with cheap labor being the main driving force. Looking at changes in the entire distribution, Figure C.2 highlights that the increase in the average plant size is indeed driven by a movement in the right tail, which forms slowly because it takes time to grow large plants and the initial distribution lacks large plants, in line with the empirical evidence in Section 1.2.2.

What about economy-wide effects? Aggregate output per worker increases by roughly 25% by 2015 in this counterfactual economy, explaining 5.2% of the close to 6-fold increase in aggregate output per worker observed between 1976 and 2015. However, this comparison may be

Figure 1.15: Distance to steady state over time



Notes: Years to reach 90% of the steady state manufacturing output for each transition path over 1976-2015. Each year's perfect foresight equilibrium transition path starts from that year's initial distribution and fixes fundamentals of that year over the transition. Best linear fit includes 95% CIs. Jumps are partly driven by census years in which potential entrant distributions change more strongly.

unfair given that a large part of the 6-fold increase in aggregate output per worker is driven by changes in the rest of the economy, not by manufacturing. In the end, manufacturing in our data captures less than 10% of aggregate output in the economy. We thus also consider a counterfactual in which only the rest of the economy fundamentals change as observed, but manufacturing fundamentals and the initial distribution of manufacturing plants stay fixed to their values in 1976. Using this counterfactual to "purge" the effects of changes in the rest of the economy and isolate the effects of changes in manufacturing only, we find that initial transition growth accounts for all (117%) of the aggregate output per worker gains that are due to changes in manufacturing by 2015.

1.4.2 The never-ending race: Transition growth remains important

As evidenced in the previous section, even in the absence of further changes in model fundamentals, transition dynamics – due to slow labor hiring and firing and slow plant entry and exit dynamics – take decades to play out. A key question is whether the Indonesian economy runs out of transition growth over time as new changes in model fundamentals provide new potential for transition growth. We find that the quantitative answer to this question is no. We show this result by revisiting the previous exercise but instead of computing the transition path based only on the initial distribution in 1976 and initial model fundamentals $\bar{\Theta}^F_{1976}$, we compute the perfect foresight transition paths for each year between 1976 and 2015 starting from that year's initial distribution and model fundamentals $\bar{\Theta}^F_t$. This gives a total of 40 different counterfactual transition paths with their corresponding counterfactual equilibrium wage paths and plant distributions. As a measure of the transition potential, we then calculate for each transition path the number of years it takes to reach 90% of the (time-varying) steady state manufacturing output. Figure 1.15 shows that it takes the 1976 economy 26 years to come close to the steady state if fundamentals were to remain constant at $\bar{\Theta}^F_{1976}$. On average, it takes about 20 years and, importantly, the number of years to come close to the steady state does not systematically decline and – if anything – increases over time.

This race between catching up to the steady state and changes in the steady state itself can only be studied in a model that features both transition growth and changes in fundamentals and we find strong quantitative evidence that due to the combination of large and frequent changes in fundamentals and slow transition dynamics, the Indonesian economy does not get closer to its time-varying steady state. Large demographic changes and policy changes pre-1975 also provide a simple explanation for why the initial Indonesian economy in 1976 was far away from its steady state. At last, potential for transition growth may not always provide a positive force for economic growth; in fact, after 40 years of demographic changes, the mass of plants in 2015 is above its steady state and transition growth is now negative as the mass of plants slowly declines along the transition – an important consequence of an aging population.

1.4.3 The role of policy

While the Indonesian growth experience is driven by a never-ending race of transition growth and changes in model fundamentals that induce new transition growth, the question is how government policy enters. The simple answer is: policy drives part of the changes in model fundamentals. Thus, to evaluate the effect of policy, we need to link changes in policies to changes in model fundamentals. In this section, we show how to do this by focussing on two specific but very important Indonesian government policy changes since 1975: education reform that maps to changes in human capital (and thus aggregate effective labor supply), and changes in Indonesia's foreign direct investment (FDI) policy that map to changes in the distribution of potential foreign entrants over time. In both cases, we first quantify the overall effect of changes in the specific model fundamental and then quantify the (relative) effect that can be attributed to specific policy changes. We show that overall changes in human capital were large and a necessary condition for Indonesia's manufacturing take-off, but observed education policies only explain 5% of this effect. In contrast, we find that the overall growth effects of FDI in Indonesian manufacturing were modest, but that observed changes in FDI policy explain up to 85% of its effects.

The role of cheap labor & the INPRES school construction program

What are the economy-wide and manufacturing growth effects of dramatic increases in human capital in the Indonesian economy? Over the period 1976 to 2015, our estimates suggest that human capital per worker H_t/L_t increased by 220%. To quantify the overall effects of human capital increases, we consider a counterfactual in which the Indonesian economy had not seen any human capital per worker increases over time. That is, we consider a counterfactual growth path where we start from the initial distribution in 1976 and a counterfactual path of model fundamentals with a modified path for the aggregate labor supply: $\Theta^F_{1976} \setminus \{H_t/L_t = H_{1976}/L_{1976}\}^\infty_t$. To quantify the extent to which policy contributed to the overall increases in human capital, we evaluate the effects of a particular educational policy change. Namely, we evaluate the largest school construction program in Indonesia's history and one of the largest in the world: the 1970s INPRES school construction program. The program successfully led to increases in schooling and wages (see: Duflo 2001; Akresh, Halim, and Kleemans 2023). We assume the program only affected the Indonesian economy through its effects on human capital. To quantify the effects of the INPRES program, we consider a counterfactual in which all gains in human capital materialized except the ones that were due to the INPRES program. For this we construct a counterfactual path of aggregate human capital in the absence of the INPRES program, drawing on existing estimates on the wage effects of the program (which map to marginal changes in human capital), the known scale of the program and the increasing share of treated cohorts over time (details in Appendix 1.C.2).

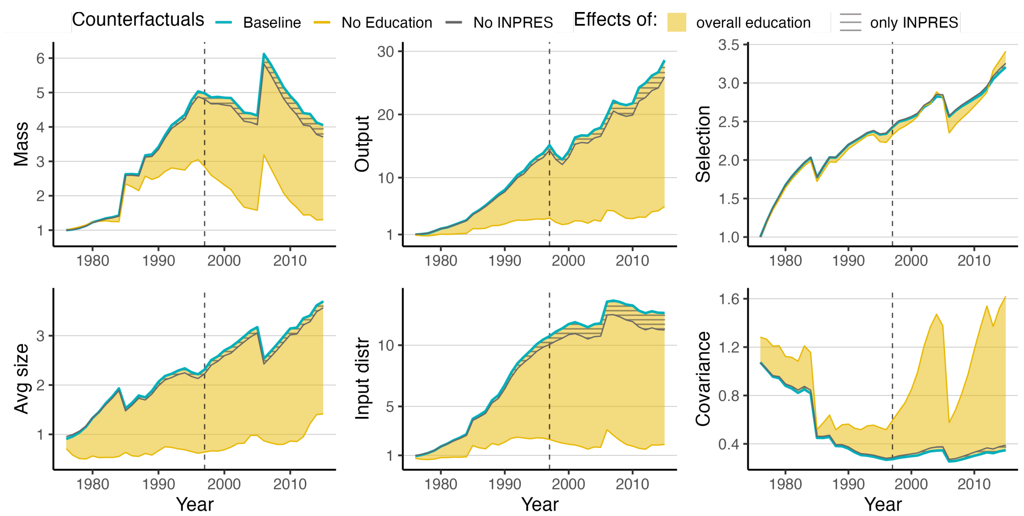


Figure 1.16: The role of education in Indonesia's manufacturing growth.

Overall, we find that the estimated 220% increase in human capital per worker increased aggregate output per worker by 26.5% in 2015. The seemingly small aggregate effect is explained by output in the rest of the economy being less dependent on labor, as captured by a low estimated labor elasticity. Figure 1.16 visualizes the quantitative effects that increases in human capital had on manufacturing growth over the period 1976-2015. Manufacturing growth in Indonesia relied heavily on the cheap labor that increases in human capital brought; in the absence of this effective increase in the supply of labor ("No Education"), wages would have roughly doubled and Indonesia would not have developed a successful manufacturing sector, with manufacturing output and its employment share not even reaching 1/4 of their historical level. Average plant size would have only increased marginally over time, far less plants would have entered and far more would have exited. The INPRES program only accounts for a small share of these effects, roughly driving 5% of the aggregate output per worker gains from increases in human capital. Figure 1.16 shows that the effect on manufacturing output, the mass of producing plants and overall hiring, is approximately twice as big as the aggregate effects. This is because manufacturing is more sensitive to labor costs than the rest of the economy. Furthermore, the positive effects of the INPRES program slowly increase over time as more cohorts of Indonesians that benefited from the new-built schools enter the labor market.

The role of foreign ownership and FDI policy

Next, we look at the role of foreign direct investment in manufacturing, which we define as the foreign ownership of manufacturing plants. Foreign-owned manufacturing plants are quantitatively important, accounting for roughly 30% of manufacturing output in 2015 (see Figure C.3 in Appendix 1.C.3) and the aggregate importance of foreign ownership increased steadily since the late 1980s. FDI policy primarily affects the entry of foreign-owned plants, since ownership shares are highly persistent and most variation in foreign ownership is across, not within plants. We thus assume that FDI only affects the Indonesian growth experience through changing the distribution of potential entrants – a model fundamental that is robust to time variation in the incentives to enter – and consider counterfactual growth paths in which we only change the path of potential entrant distributions. Again, we want to separately quantify the effects of FDI and the relative effect that changes in observed FDI policy had on FDI. For this, we separate the distribution of potential entrants at any point in time into the distribution of potential foreign entrants and potential domestic entrants enforcing model-consistent entry decisions. We then construct a counterfactual path of potential entrant distributions without foreign entrants. To capture the effect of policy, we consider important regulatory changes in FDI policy in 1987. Specifically, we exploit variation in potential foreign entrant distributions right before and after the reform to measure the effect of policy and use the estimated effect to construct a counterfactual path of potential foreign entrant distributions in the absence of the FDI policy change (details in Appendix 1.C.3).

We find that FDI helped manufacturing growth, but did not play a transformative role. Specifically, the entry of foreign-owned manufacturing plants explains 7.5% of the aggregate output per worker gains due to manufacturing growth and we estimate that manufacturing output and the manufacturing employment share would be 8% lower in 2015 in the absence of FDI. The reason for this rather small effect is that given a high estimated supply of domestic potential entrants, the downward pressure on labor demand and wages due to the disappearance of foreign entrants leads to an elastic response of domestic entry in general equilibrium that mitigates some of the negative effects of losing FDI. In contrast to the case of education policy, we find that *changes* in FDI policy potentially explain most of the overall growth effects from FDI. Specifically, changes in FDI policy potentially explain a four- to five-fold increase in potential entry and these changes in FDI policy in turn explain 85% of the overall growth effects from FDI.

1.5 Conclusion

This paper studied the drivers of growth miracles. Building on 40 years of plant-level manufacturing panel data for Indonesia, we motivated a model in which rapid growth is driven by a combination of transition growth and changes in fundamentals that are dominated by worker and plant demographics. We showed how to tractably estimate this model on the observed growth path using standard plant-level data and without assuming that the observed economy is at a steady state at any point in time. We found that transition growth is key: 42% of the observed manufacturing output growth is simply explained by initial conditions in 1975 – dominated by young and small manufacturing plants – providing ample opportunities for catch-up growth. Transition growth also does not become less important over time because important demographic changes in the economy induce further potential for transition growth.

Since our model and estimation framework maps directly to observed time-varying aggregate growth and its micro-level drivers of endogenous changes in the distribution of plants, it is particularly suited to study the dynamic growth effects of observed policy. This link is important not only to better validate macroeconomic models of growth, but also to study the aggregate growth and general equilibrium effects of policy – effects that are rarely identified in micro-empirical policy evaluations. In this paper, we only started looking at this by showing how to use the estimated model to evaluate the dynamic growth effects of two important Indonesian policies: education reform and changes in FDI policy. Based on our results, a somber conclusion – partly resonating related work on the Indian growth miracle (Bollard, Klenow, and Sharma 2013) – is that observed policy mattered less for growth than we might think. Instead, we find that Indonesian growth was mostly driven by structural forces related to demographics. This does not mean that policy necessarily plays no role. In fact, in Appendix 1.C.4 we consider two sets of reduced-form policies – a reduction in (convex) labor adjustment frictions and an increase in annual technology growth – that both would have doubled Indonesian manufacturing output by 2015, an even more remarkable manufacturing miracle closer to experiences in countries such as China and Malaysia. Future research should further unpack what drives changes in adjustment costs, technology growth or the pool of potential entering plants and link these closer to policy.

1.6 References

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

Akcigit, Ufuk, Harun Alp, and Michael Peters. 2021. "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries." *American Economic Review* 111 (1): 231–75.

Akresh, Richard, Daniel Halim, and Marieke Kleemans. 2023. "Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia." *The Economic Journal* 133 (650): 582–612.

Alessandria, George, Costas Arkolakis, and Kim J. Ruhl. 2021. "Firm Dynamics and Trade." *Annual Review of Economics* 13 (1): 253–80.

Allen, Treb, and Dave Donaldson. 2020. "Persistence and Path Dependence in the Spatial Economy." National Bureau of Economic Research.

Asturias, Jose, Sewon Hur, Timothy J. Kehoe, and Kim J. Ruhl. 2023. "Firm Entry and Exit and Aggregate Growth." *American Economic Journal: Macroeconomics*.

Atkeson, Andrew, and Ariel Tom'as Burstein. 2010. "Innovation, Firm Dynamics, and International Trade." *Journal of Political Economy* 118 (3): 433–84. https://doi.org/10.1086/653690.

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2019. "Measuring Productivity: Lessons from Tailored Surveys and Productivity Benchmarking." In *AEA Papers and Proceedings*, 109:444–49. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203. https://www.aeaweb.org/articles?id=10.1257/pandp.20191005.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin. 2007. "Estimating Dynamic Models of Imperfect Competition." *Econometrica* 75 (5): 1331–70.

Basri, M. Chatib, Mayara Felix, Rema Hanna, and Benjamin A. Olken. 2021. "Tax Administration Versus Tax Rates: Evidence from Corporate Taxation in Indonesia." *American Economic Review* 111 (12): 3827–71. https://doi.org/10.1257/aer.20201237.

Bernstein, Shai, Emanuele Colonnelli, Davide Malacrino, and Tim McQuade. 2022. "Who Creates New Firms When Local Opportunities Arise?" *Journal of Financial Economics* 143 (1): 107–30.

Bilal, Adrien, Niklas Engbom, Simon Mongey, and Giovanni L. Violante. 2022. "Firm and Worker Dynamics in a Frictional Labor Market." *Econometrica* 90 (4): 1425–62.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics* 128 (1): 1–51.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts. 2020. "Do Management Interventions Last? Evidence from India." *American Economic Journal: Applied Economics* 12 (2): 198–219.

Bollard, Albert, Peter J. Klenow, and Gunjan Sharma. 2013. "India's Mysterious Manufacturing Miracle." *Review of Economic Dynamics* 16 (1): 59–85.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation from Production Data." *Journal of Monetary Economics* 121: 1–14.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing." *Journal of Development Economics* 97 (2): 339–51.

Buera, Francisco, Joseph Kaboski, and Robert Townsend. 2023. "From Micro to Macro Development." *Journal of Economic Literature* 61 (2): 471–503.

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

Cali, Massimiliano, Marco Le Moglie, and Giorgio Presidente. 2021. "Gain Without Pain? Non-Tariff Measures, Plants' Productivity and Markups."

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.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.

Chen, Zhang. 2022. "Economic Growth and the Rise of Large Firms." *Job Market Paper*. https://zhangc-econ.github.io/files/TailGrowth.pdf.

Choi, Jaedo, Andrei A. Levchenko, Dimitrije Ruzic, and Younghun Shim. 2023. "Superstars or Supervillains? Large Firms in the South Korean Growth Miracle." https://www.freit.org/EIIT/2023/selected/choi.pdf.

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.

Coşar, A. Kerem, Nezih Guner, and James Tybout. 2016. "Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy." *American Economic Review* 106 (3): 625–63.

Demirer, Mert. 2020. "Production Function Estimation with Factor-Augmenting Technology: An Application to Markups." *UPenn. Job Market Paper*.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91 (4): 795–813.

———. 2004. "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia." *Journal of Development Economics* 74 (1): 163–97.

Eslava, Marcela, and John C. Haltiwanger. 2020. "The Life-Cycle Growth of Plants: The Role of Productivity, Demand and Wedges." National Bureau of Economic Research.

Eslava, Marcela, John Haltiwanger, and Alvaro Pinzon. 2022. "Job Creation in Colombia Versus the USA: Up or Out Dynamics Meet the Life Cycle of Plants." *Economica* 89 (355): 511–39.

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

Gandhi, Amit, Salvador Navarro, and David Rivers. 2017. "How Heterogeneous Is Productivity? A Comparison of Gross Output and Value Added." *Journal of Political Economy*.

Gathen, Jonas. 2021. "Government Jobs for Talents, Cronies or Sale? Estimating Government Worker Skills with an Application to Indonesia." *Working Paper*.

Gillis, Malcolm. 1985. "Micro and Macroeconomics of Tax Reform: Indonesia." *Journal of*

Development Economics 19 (3): 221–54.

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.

Hansen, Lars Peter, and Kenneth J. Singleton. 1982. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica: Journal of the Econometric Society*, 1269–86. https://www.jstor.org/stable/1911873.

Hastie, Trevor J. 2017. "Generalized Additive Models." In *Statistical Models in S*, 249–307. Routledge.

Hausmann, Ricardo, Lant Pritchett, and Dani Rodrik. 2005. "Growth Accelerations." *Journal of Economic Growth* 10: 303–29.

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

Hill, Hal. 2000. *The Indonesian Economy*. Cambridge University Press.

Hopenhayn, Hugo. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 1127–50.

———. 2014. "Firms, Misallocation, and Aggregate Productivity: A Review." *Annual Review of Economics* 6 (1): 735–70.

Hopenhayn, Hugo, and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy* 101 (5): 915–38.

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein. 2016. "Quantifying the Sources of Firm Heterogeneity." *The Quarterly Journal of Economics* 131 (3): 1291–1364.

Hotz, V Joseph, and Robert A Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies* 60 (3): 497–529.

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.

———. 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.

Humlum, Anders. 2022. "Robot Adoption and Labor Market Dynamics."

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2019. "Demographic Origins of the Startup Deficit." National Bureau of Economic Research.

Kleinman, Benny, Ernest Liu, and Stephen J. Redding. 2023. "Dynamic Spatial General Equilibrium." *Econometrica* 91 (2): 385–424.

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation." *Journal of Political Economy* 112 (5): 986–1018. https://doi.org/10.1086/422563.

Kruse, Hagen, Emmanuel Mensah, Kunal Sen, and Gaaitzen de Vries. 2023. "A Manufacturing (Re) Naissance? Industrialization in the Developing World." *IMF Economic Review* 71 (2): 439–73.

Lanteri, Andrea, Pamela Medina, and Eugene Tan. 2023. "Capital-Reallocation Frictions and Trade Shocks." *American Economic Journal: Macroeconomics* 15 (2): 190–228.

Lent, G., and P. Ojha. 1969. "Sales Taxes in Countries of the Far East." *IMF Staff Papers* 16.

Liang, James, Hui Wang, and Edward P. Lazear. 2018. "Demographics and Entrepreneurship." *Journal of Political Economy* 126 (S1): S140–96.

Luttmer, Erzo GJ. 2007. "Selection, Growth, and the Size Distribution of Firms." *The Quarterly Journal of Economics* 122 (3): 1103–44.

McCaig, Brian, and Nina Pavcnik. 2021. "Entry and Exit of Informal Firms and Development." *IMF Economic Review* 69: 540–75.

McKenzie, David, and Anna Luisa Paffhausen. 2019. "Small Firm Death in Developing Countries." *Review of Economics and Statistics* 101 (4): 645–57.

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.

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

Ottonello, Pablo, and Thomas Winberry. 2023. "Investment, Innovation, and Financial Frictions."

Perla, Jesse, Christopher Tonetti, and Michael E. Waugh. 2021. "Equilibrium Technology Diffusion, Trade, and Growth." *American Economic Review* 111 (1): 73–128.

Peters, Michael. 2020. "Heterogeneous Markups, Growth, and Endogenous Misallocation." *Econometrica* 88 (5): 2037–73. https://doi.org/10.3982/ECTA15565.

Poschke, Markus. 2013. "'Entrepreneurs Out of Necessity': A Snapshot." *Applied Economics Letters* 20 (7): 658–63.

———. 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. 2017. "The Causes and Costs of Misallocation." *Journal of Economic Perspectives* 31 (3): 151–74.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5, Part 2): S71–102.

Ruggieri, Alessandro. 2022. "Trade and Labour Market Institutions: A Tale of Two Liberalizations." University of Nottingham, GEP.

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

Schoar, Antoinette. 2010. "The Divide Between Subsistence and Transformational Entrepreneurship." *Innovation Policy and the Economy* 10 (1): 57–81.

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.

Ulyssea, Gabriel. 2020. "Informality: Causes and Consequences for Development." *Annual Review of Economics* 12 (1): 525–46.

Van Biesebroeck, Johannes. 2005. "Firm Size Matters: Growth and Productivity Growth in African Manufacturing." *Economic Development and Cultural Change* 53 (3): 545–83.

World Bank. 2023. *Poverty and Inequality Platform (Version 20230328_2017) [Data Set]. World Bank Group. Www.pip.worldbank.org. Accessed the 15th September, 2023.*

Appendices

1.A Data and Empirical Evidence

1.A.1 Data cleaning details

In the following subsection, we describe in detail the data cleaning steps we take to ensure that the data is consistent over time and that results are not driven by different forms of measurement error. The main data cleaning steps relate to cleaning plant-level labor (wage bill and workers) and output (value added) over time. We further discuss how we deal with "dynamic" outliers (e.g. unrealistic within-plant jumps in value added or the wage bill) and observations with extreme labor shares that pose numerous problems for the estimation and computation of the model. Besides these mentioned cleaning steps, we also drop a few clear outliers, such as when the plant ID is misreported or missing or when magnitudes of multiple reported variables are impossible. (Give details on final cleaning: How does raw data differ from cleaned data). At last, we report details on how we clean the capital series, industry codes and measure plant entry and exit.

Cleaning labor and the labor wage bill

The manufacturing census consistently reports a plant's total number of workers (including paid and unpaid) as well as separately the number of paid versus unpaid workers and the number of production and non-production (including managerial) workers. The main cleaning step we apply to ensure consistency over time is to drop all plants with less than 20 total workers (which is enforced by BPS starting in 1990, but not before), drop plants that report zero paid workers or that report more paid workers than total workers. This step drops slightly less than 2% of plant-year observations with dropped observations concentrated before 1990. We also identify a bunching at 99 workers in the years 2013-2015 (roughly 3-4% of plants), which we interpret as true bunching driven by actual policy changes and thus do not correct.[22]

For the structural model, we build on a plant's reported total wage bill. This variable is the sum of the total wage bill for other workers and for production workers. In principle, it includes all payments to labor, including in-kind transfers, overtime pay, bonuses and social contributions (e.g. pension and accident allowances). Since the survey asks about current workers and doesn't separately ask about severance pay, we treat the reported wage bill as excluding severance pay.

[22] For example, Indonesia introduced an occupational safety and health regulation in 2012 that was targeted at manufacturing firms and mandates any workplace with more than 100 workers to implement additional work safety measures.

We take three main cleaning steps for the reported wage bill. First, we correct systematic misreporting for the year 2011. Looking at the evolution of the distribution of the total wage bill across years, we find that 2011 is the only clear outlier with the only bimodal wage distribution across all years. In 2011, the bottom 20% of observations show exactly the same per worker wage at an unrealistically low level that is below the bottom 1% of observations in 2010 and 2012 and roughly 50 times lower than the average minimum wage in 2011. Given that these are well-defined misentries, the remainder of the distribution is well-behaved and we observe most plants before and after 2011, we opt for imputing misentries using within-plant averages across 2010 to 2012, enforcing linear growth for 2011.

Second, we correct for misreporting in the wages for non-production workers. Non-production workers account for roughly 16.5% of overall employment in our data. However, 17% of plant-year observations reportedly employ zero non-production workers. This also means that plants may not always report all managerial workers, which is likely if the managerial staff partly owns the establishment (and is thus not formally employed). To the extent that all payments to managerial workers should be counted as labor costs, the Indonesian data may significantly underestimate labor costs. We cannot correct for this form of underreporting. However, we can correct for the following: About 10% more plant-year observations report zero wage payments to non-production workers than plant-year observations reporting zero non-production workers. That is, there are plants that report employing non-production workers, but paying them no wages. This could be in part due to some plants reporting managerial staff who receive remuneration other than wages or due to plants simply not reporting non-production worker wages. In any case, we think it is better to impute wages here, using plants' reported wages for production workers and the average year-specific pay gap across production and non-production workers for plants that report both wages. We find that wage premia for non-production workers were around 90% in 1975 and declined to around 20% in 2015. In the end, the overall importance of this correction is small, because the correction only applies to a small number of observations.

In the third and last cleaning step, we correct the total reported wage bill across periods where the exact questions and components of the total wage bill changed. While plants were asked consistently to report total payments including cash and in-kind wages, pensions and other social contributions between 1975-1995, survey questions changed most notably between 2001-2003, 2004-2010 and 2011-2015. Looking at changes in the distribution of reported total wage bills, the 2004-2010 period appears to be the most problematic period in which total wages are systematically underreported vis-a-vis the other periods.

We correct for changes in the measurement of the total wage bill over the period 2004-2010 by exploiting within-plant changes in reported wage bills across changes in the measurement period (from 2003 to 2004 and from 2010 to 2011) and further utilizing information across different types of workers and the reporting of the number of workers by type. Plants i report x_{it}^j, the total reported wage bill for worker type j in period t. Specifically, we assume that: $x_{itm}^j = w_t h_{it}^j \epsilon_{it}^j \tau_m^j$, where w_t is the wage (in line with our model), h_{it}^j are plant-worker-type-specific efficiency units of labor, ϵ_{it}^j captures idiosyncratic measurement error and τ_m^j is systematic underreporting that is constant within a worker-type and within the measurement period m where the same questions to elicit the worker-type-specific wage bill are asked. We assume that $\tau_m^j \in (0, 1]$ for the period 2004-2010 and unity otherwise.

Our approach separately identifies wages from the measurement error τ_m^j across time and thus allows to correct for an important part of measurement error that leads to the underestimation of the total wage bill and a time inconsistent measure of the wage bill. For separate identification, we assume that over two consecutive periods, plants use the same average human capital within types of worker: $\frac{h_{it}^j}{l_{it}^j} = \frac{h_{it+1}^j}{l_{it+1}^j} = \alpha_i^j$. This allows some plants to specialize on high productive production workers or other plants to specialize on low productive managerial staff, but restricts changes in the average human capital within plant-worker-types. As long as the plant-specific number of workers by type are reported either without measurement error or with constant plant-worker-type measurement error, the assumption allows to identify:

$$\mathbb{E}_i \frac{\frac{x_{i,t+1,m}^j}{l_{i,t+1,m}^j}}{\frac{x_{i,t,m}^j}{l_{i,t,m}^j}} \equiv \mathbb{E}_i \frac{\tilde{x}_{i,t+1,m}^j}{\tilde{x}_{i,t,m}^j} = \frac{w_{t+1}}{w_t} \mathbb{E}_i \frac{\epsilon_{it+1}^j}{\epsilon_{it}^j}$$

Within measurement periods m, one can show that under realistic magnitudes for the measurement error, the following holds:[23]

$$\mathbb{E}_i \frac{\tilde{x}_{i,t+1,m}^j}{\tilde{x}_{i,t,m}^j} \approx \frac{w_{t+1}}{w_t}$$

Across measurement periods, separate identification of the change in measurement and the wage is impossible without further assumptions. To see this, write:

$$\mathbb{E}_i \frac{\tilde{x}_{i,t+1,m'}^j}{\tilde{x}_{i,t,m}^j} = \frac{w_{t+1}}{w_t} \mathbb{E}_i \frac{\epsilon_{it+1}^j \tau_{m'}^j}{\epsilon_{it}^j \tau_m^j} \approx \frac{w_{t+1}}{w_t} \frac{\mathbb{E}_i \epsilon_{it+1}^j \tau_{m'}^j}{\mathbb{E}_i \epsilon_{it}^j \tau_m^j} = \frac{w_{t+1}}{w_t} \frac{\mathbb{E}_i \tau_{m'}^j}{\mathbb{E}_i \tau_m^j}$$

where we have made use of the same approximation as above. Even then, changes in the measurement error across measurement periods cannot be separately identified from wage changes. To solve this issue, we interpolate wages from wage growth in the previous period and the next period (for which measurement does not change), assuming that wages grow smoothly over time.

In our case, we set $\tau_m^j = 1$ for all measurement periods except the period 2004-2011. To identify $\tau_{m'}^j$ for 2004-2011, we are now actually over-identified, because we can identify the measurement error from variation between 2003-2004 or from 2010-2011. We choose to use 2003-2004 because 2010-2011 featured a change in the minimum wage in Indonesia, which partly explains a large increase in plants' total wage bills and we do not know how to separate this change from a change in the measurement. Following this approach, we find that $\tau_{m'}^j \approx 0.94$, similar when restricting to production workers only or when looking at all workers. We en-

[23] Specifically, a first-order Taylor series approximation around the mean of the measurement errors gives: $\frac{w_{t+1}}{w_t} \mathbb{E}_i \frac{\epsilon_{it+1}^j}{\epsilon_{it}^j} \approx \frac{w_{t+1}}{w_t} \frac{\mathbb{E}_i \epsilon_{it+1}^j}{\mathbb{E}_i \epsilon_{it}^j}$. With a second-order Taylor series approximation, we get: $\mathbb{E}_i \frac{\epsilon_{it+1}^j}{\epsilon_{it}^j} \approx \frac{\mathbb{E}_i \epsilon_{it+1}^j}{\mathbb{E}_i \epsilon_{it}^j} \left[1 - \frac{Cov_i(\epsilon_{it+1}^j, \epsilon_{it}^j)}{\mathbb{E}_i \epsilon_{it+1}^j \mathbb{E}_i \epsilon_{it}^j} + \frac{Var_i(\epsilon_{it}^j)}{\left[\mathbb{E}_i \epsilon_{it}^j\right]^2} \right]$. Plugging in realistic measurement error, the 2nd-order correction is very small. For example, if measurement error within plants is positively correlated (which is the likely case), then the two correction terms go in opposite directions. Furthermore, both the covariance term and the variance term are close to zero for reasonable magnitudes of measurement error.

force the production worker correction across all plants for the measurement period 2004-2011 (which means that their wage bills get upward corrected by $1/\tau^j_{m'}$, a correction of roughly 6%).

Cleaning output / value-added

Throughout the paper, we use a consistent definition of value-added output. This definition coincides with how the Indonesian statistical agency (BPS) constructed value-added output for some, but not all years. Specifically,

$$\text{Value-added}_{it} \equiv \text{Gross income}_{it} - \text{Intermediates}_{it}$$

where Gross income$_{it}$ $\equiv$ Gross sales$_{it}$ + electricity sales$_{it}$ + revenue from industrial services$_{it}$ + other income$_{it}$ + Δvalue of semi-finished products$_{it}$ and Intermediates$_{it}$ $\equiv$ Raw materials$_{it}$ + Total fuel/electricity expenditures$_{it}$ + Other expenses$_{it}$. All inputs in the accounting identities are reported in their current values of Rupiah, which we deflate to 2010 constant Rupiah based on the aggregate CPI. We start by dropping observations with missing or negative gross income, which are less than 0.5% of observations.

Next, we construct a time-consistent measure of intermediates. The main issue is that intermediates are likely underreported since the survey only asks for specific categories of expenditures and intermediate expenditures have likely become more complex over time, leading plants to underreport parts of their expenditures. This leads to an overestimate of value-added output and an underestimate of capital and labor cost shares at the plant-level. We correct intermediate expenditures in two steps.

In the first step, we look at one main expenditure category of intermediate inputs for which we know that time inconsistency is an issue. Specifically, Other expenses$_{it}$ are reported inconsistently over time because not all components of other expenses are enumerated in every year. In the following, we describe the components of Other expenses$_{it}$ and how we impute them consistently across plants over time. In the years with the most detailed survey questions, Other expenses$_{it}$ (X_{it} in short) are the sum of three components (indexed by c): (1) expenses for other goods (consisting of packaging, spare parts and stationary), (2) manufacturing services, repair and maintenance, and (3) remaining other expenses (with detailed subcomponents for some years). We improve the measure of intermediate expenditures by imputing these three subcomponents in cases where they are missing. Similar to the components of the labor wage bill, we deal with underreporting of other expenses by exploiting within-plant differences in reporting around years with changes in survey questions. We separately impute missing fractions of each of the three components of other expenses, using further information on subcategories j within components c, bringing all series to the most complete level of reporting in the years 2006 and 1996/1997.

Specifically, we assume that $\forall c, j: \quad X_{icjt} = \alpha_{icj} Y_{it}$. That is, any other expense category (or subcategory) is a plant-subcategory-specific fraction of gross income Y_{it}. Since expenditures for specific subcategories are systematically missing in some years, but gross income Y_{it} is reported for all years, we impute complete missing subcategory expenditures as follows: For plants that we observe across different measurement periods, we impute their expenditure shares from average within-plant expenditure shares around the time of missing. For example, expenses for other goods are missing between 1998 and 2005, which we impute using the plant average of plant-specific expenditure shares in 1997 and 2006 together with plant-year-specific gross income Y_{it}. This ensures within-plant consistency and allows for plant-category-year-specific variation in expenditures. For plants for which we do not observe expenditures in other years, we use the aggregate category-specific expenditure share around the time of missing.

On top of this, we correct reported expenditures for the remaining other expenditures for the period 1975-1984 in which the reported series is clearly underreported in comparison to post 1985. For this correction, we again proceed separately for plants that we observe across measurement periods and for plants that we do not, using either the average within-plant difference in reported ratios or the ratio of aggregate expenditure shares across the two measurement periods as correction factors.

Overall, this first step of cleaning intermediate inputs ensures more time consistency, but does not have a sizable effect on overall intermediate expenditures. In the second step of correcting intermediate expenditures, we deal with the sizable remaining decline in the intermediate expenditure share across plants over time. For example, the aggregate intermediate expenditure share declines by more than 10 percentage points from roughly 0.65 in 1980 to 0.525 in 2015 (mostly driven by a decline in the raw material input share). We expect that a major part of this decline is in fact measurement error. One simple reason could be that plants use more processed intermediate inputs, which they do not fully report as "raw materials". To distinguish this driver from industrial composition effects (e.g. industries relying on intermediate inputs declining in relative importance over time), we construct the following correction: We regress $log(\phi_{ijt}/(1-\phi_{ijt})) = \alpha_j + \alpha_t + \epsilon_{ijt}$ where ϕ_{ijt} is the intermediate expenditure share of plant i in 5-digit industry j at time t. We use the log odds ratio to ensure that any correction we implement gives expenditure shares that are bounded between zero and one. α_j and α_t capture industry and time fixed effects. We interpret α_t as our time-varying bias term, using α_{1975} as the normalization factor (for which the bias is zero). Controlling for industry fixed effects ensures that the bias terms do not capture variation in intermediate expenditure shares from changes in the industrial composition. Corrected intermediate input shares are then given by $\tilde{\phi}_{ijt} = exp(\alpha_j + \alpha_{1975} + \epsilon_{ijt})/(1 + exp(\alpha_j + \alpha_{1975} + \epsilon_{ijt}))$. The correction maintains plant-level variation in intermediate expenditure shares and delivers both within-plant and aggregate time consistency. To ensure that the regression is well-specified, we initially drop observations with non-positive intermediate inputs or value added as well as observations with missing value added. This drops less than 2% of observations. After the correction, we recompute intermediate expenditures and value added.

Identifying problematic outliers: jumps and extreme labor shares

The last important cleaning step we take, is to identify problematic outliers that are likely misentries and would have an outsized role on the model estimation and inference.

We start with "dynamic" outliers by which we refer observations that are outliers within the time series of an individual plant. We treat a plant-year observation as a dynamic outlier if the total wage bill or value added output series (X_{it} in the following) of an individual plant sees a sizable one-time jump after which it reverts directly back. We consider two different measures: the year-to-year within-plant change $\frac{X_{it}}{X_{it-1}}$ and the year-to-year within-plant quantile difference $q_t(X_{it}) - q_{t-1}(X_{it-1})$. For both measures, we first identify a potential outlier if any of the two measures is below the 10th or above the 90th within-year percentile for the respective measure. For example, we classify the following observation a potential outlier: Between 1993 and 1994, a plant's quantile of value added output changed by more than the 90th percentile of this year's distribution of value added output quantile changes. We classify any potential outlier as an actual outlier only if the following year is also identified as a potential outlier whose change goes in the opposite direction.[24] This ensures to identify jumps, while the initial and final level do not need to coincide, allowing for plant-, time- and variable-specific drifts. Big one time changes are explicitly not counted as outliers, treating them as true shocks. We drop any dynamic outlier that we detect through this procedure. Note that this procedure also identifies observations that change back and forth multiple times in a row as outliers as long as their changes are very large. In total, this procedure drops almost 10% of plant-year observations and roughly 10% of total reported value added.

The second and last cleaning step we take is to drop observations with extreme labor shares. These are observations with reported labor shares below 5% and above 500% (roughly differing from the median labor share of 50% by a factor of 10). Extreme labor shares are likely a combination of overreported value added and underreported wage bills and these observations have a sizable impact on aggregates. They make up roughly 3% of observations, but account for 41% of total reported value added. While we think that many of these plant-year observations with large value added are correctly classified as being "granular" in their importance for output (e.g. many of these plants consistently report being large over time), their exact value added output and wage bills are likely mismeasured.

Cleaning capital

For cleaning plants' self-reported capital stock, we draw on the cleaning steps in Cali, Le Moglie, and Presidente (2021), which is the most thorough attempt at cleaning the Indonesian manufacturing plant capital series that we are aware of. The cleaning steps draw in part on the perpetual inventory method (PIM). Details can be found in Cali, Le Moglie, and Presidente (2021).

Cleaning industry codes

Industry classifications changed over time, starting with ISIC 2 in 1975 and moving to ISIC 3, ISIC 3.1 and ISIC 4 by the end of our data period. For harmonization, we start by fixing a plant's first reported 5-digit industry (the most disaggregated level reported). While plants may reasonably change industries over time, we opt for fixing industries to have a time-consistent plant-level measure of industry. We then build backward correspondences at the 2-, 3-, 4- and 5-digit industries respectively using within-plant changes in industry classifications across changes in classification systems (that is, correspondences that map from later year classifications to earlier ones). For plants that enter later, we enforce these within-plant correspondences. In the case of one-to-many mappings (e.g. the same industry code in ISIC 3.1 maps to different codes in ISIC 3), we enforce the most common one. Note that this only matters for plants that are not observed previously. In the case of no linking (e.g. a plant enters in ISIC 4 with a code that has no observed backward linkage), we check codes manually and use official crosswalks. In the cases where we think that industries are truly "new", we simply

[24] We do this separately by measure so the identification of an outlier is based on within-measure changes only.

create a new industry code. In the end, we are left with 140 unique 5-digit industry codes, 120 of which are initial ISIC 2 codes and 20 codes are new industries. At the 2-digit level, we have 9 different industries.

Plant entry and exit

In principle, classifying entry and exit should be straightforward: whenever a plant with a new panel identifier enters our panel, we would record this as plant entry and whenever that plant reports in t but does not report in $t+1$, we would record the plant as having exited at the end of t. In practice, this classification would inflate plant entry and exit because occasional non-reporting is common. This is because of actual non-reporting and – as described above – because we explicitly drop plant-year observations with misreported entries. We thus only classify a plant as having exited if we do not observe reporting by the plant at any future time period. Similarly for entry, we only count the plant as entering if it is the first time the plant identifier has entered the panel. This difference is quantitatively important: the unconditional exit rate drops almost by half from around 14% to 7.9% if we follow our classification. As we discuss in the main text, 7.9% is close to other exit rates reported for India, Mexico and the US (e.g. Hsieh & Klenow 2014).

Where and how does this matter and could this new classification bias results differentially over time? Throughout the analyses, we mostly draw on within-plant changes that are robust to compositional changes due to entry and exit. In the structural model, we explicitly continue modeling plants that are non-reporting but non-exiting, correcting for (some forms of) differential non-reporting over the state space. Bias does arise from non-reporting or misreporting that is correlated with the state space for certain estimation steps. For example, estimating conditional exit probabilities clearly suffers from bias if measured exit probabilities are biased over the state space. Two potential issues may be particularly important in our case: non-reporting due to the cutoff of 20 workers and that our measure of exit may inflate exit towards the end of our data because we cannot distinguish permanent exit from temporary non-reporting. We think that both issues likely introduce biases that are small in magnitude.

As for the cutoff of 20 workers, the issue would be particularly problematic if plants regularly moved back and forth over the threshold or if plants with more than 20 workers moved permanently below the threshold, which we would wrongly classify as plant exit. We do not think that these are important issues in the Indonesian data. For example, few plants shrink and as we show in the main text, plants with 20 workers become relatively less important vis-a-vis larger plants over time. Also, given that pre-1990, plants often continue reporting even if they move below 20 workers, we find that movements around the threshold are rare. As for the classification of exit towards the end of the sample, we note that non-reporting actually seems to decline over time and 2015 (the last year of the data) is a census year in which enumeration is most complete.

1.A.2 Further main descriptives

Figure A.1 reports the evolution of the share of employment and value-added output that is captured by the Indonesian manufacturing plant census (1975-2015) in comparison to aggregate manufacturing value-added output and employment as reported in the GGDC 10-sector

Figure A.1: Representativeness of manufacturing panel over time

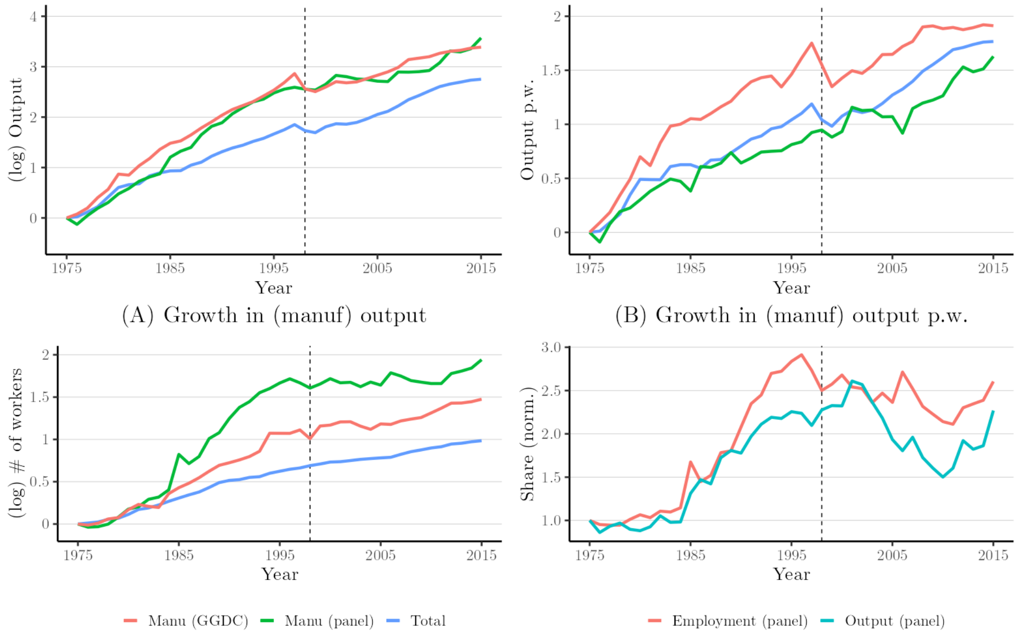
Notes: Evolution of the employment and value-added output share captured by the Indonesian manufacturing plant census (1975-2015) in comparison to aggregate manufacturing value-added output and employment as reported in the GGDC 10-sector database (1975-2012) and the Economic Transformation Database (1990-2018). Straight lines report (unweighted) time averages. The black dotted line highlights the Asian Financial Crisis.

database (1975-2012) and the Economic Transformation Database (1990-2018).[25] The series tend to increase over time, which captures the fact that medium- to large-size manufacturing plants become more important over time. However, there is important variation in the shares over time that is not well-captured by a simple secular increase in the importance of medium- to large-sized manufacturing plants. In the model, we allow for the time-varying importance of the rest-of-the-economy, and differentially so for output and employment, including the part of manufacturing that our data misses. Note that for output, the increase is much stronger if we do not clean the value-added series, because a few plant entries can have an outsized effect on total value-added (e.g. for only a minimally cleaned output series, we found that the output share of our manufacturing data can increase to up to 80% by 2015, entirely driven by a few plant entries).

In the main text, we also report that our manufacturing panel misses 99% of manufacturing plants in Indonesia. This is based on information on a random five percent sample of all manufacturing establishments from the Indonesian Economic Census in 2006 reported in Hsieh and Olken (2014). We also 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 (as reported in Hsieh and Olken (2014)), manufacturing plants with more than 50 workers should capture 34% of total manufacturing employment, while this figure is 32% based on employment in our micro-data (29.5% after cleaning) and taking the aggregate sectoral employment from the GGDC 10-sector database as denominator. Given that the manufacturing plant panel includes new plants based on the Economic Census, coverage is more complete after Economic Census years.

[25] We merge the latter two series consistently over time by enforcing the more recent vintage and aligning all series before 1990 to be consistent with the evolution after 1990. Specifically, for each variable, we take the ratio of the GGDC10 and ETD series in 1990 to be the amount that the 1990 GGDC series needs to be adjusted by. We then similarly correct each year from 1975 to 1990 by a correction factor that equals unity in 1975 (no correction in 1975), is equal to the full correction in 1990 and is taken from an equal-spaced, smoothed series in the years between.

Figure A.2: Evolution of aggregate and sectoral employment and output (in logs)



(A) Growth in (manuf) output

(B) Growth in (manuf) output p.w.

(C) Growth in number of workers

(D) Manuf employment + output shares

Notes: (Economy-wide) Total and GGDC are based on joining the GGDC 10-sector Database (1975-2012) and the Economic Transformation Database (1990-2018). Panel refers to the Indonesian manufacturing plant census (1975-2015, 20+ workers). All series are normalized by their respective value in the first year. (A) and (B) use value-added output.

Next, Figure A.2 reports the evolution of aggregates in the Indonesian economy, showing the series in logs to better visualize how growth rates changed over time.

For completeness, Figure A.3 reports the full year-to-year evolution of employment shares for different plant sizes.

We further report estimated Pareto tail coefficients for the manufacturing data in Figure A.4. We follow Chen (2022) in constructing two simple, but alternative measures of the Pareto tail. Let $F(x,t)$ be the CDF of the underlying distribution of plant employment, and f the density function. Then $\tilde{F}(x,t) \equiv 1 - F(x,t)$ denotes the fraction of plants with size greater than x, and $\tilde{F}^{emp}(x,t) \equiv \int_x^\infty y dF(y,t)$ the total employment in plants with size greater than x. In addition, let T_L be the employment size threshold for large plants and T_S for small plants. Assuming that $F(x,t)$ follows a Pareto distribution with shape parameter k_t, we have:

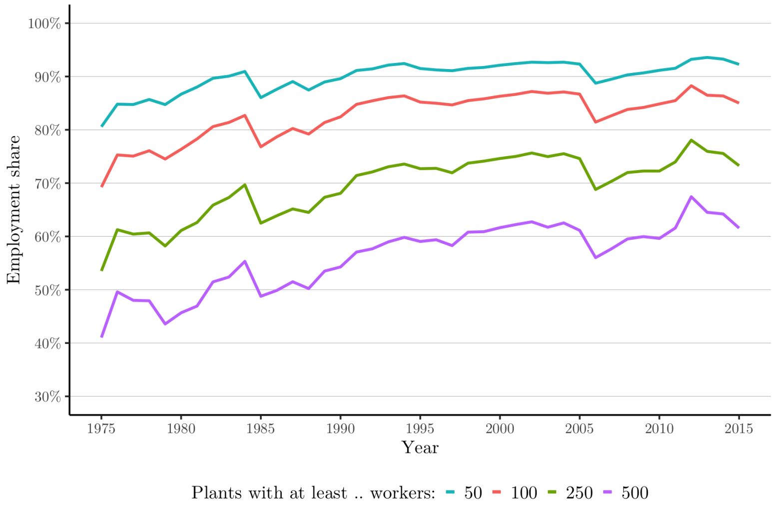
$$k_t = 1 - log\frac{\tilde{F}^{emp}(T_L)}{\tilde{F}^{emp}(T_S)}/log\frac{T_L}{T_S}$$

Alternatively,

$$k_t = -log\frac{\tilde{F}(T_L)}{\tilde{F}(T_S)}/log\frac{T_L}{T_S}$$

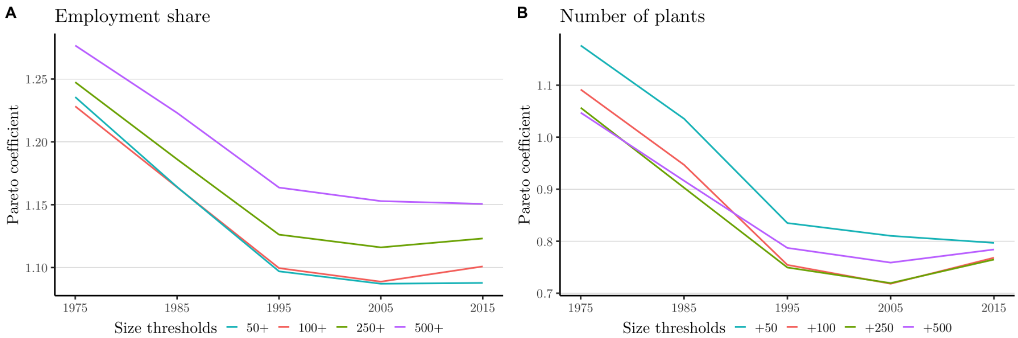
In both cases, Pareto tails can be computed in the absence of knowing the employment share or fraction of plants below 20 workers, because these shares cancel out. Panel A reports estimated Pareto coefficients for different thresholds T_L for the employment share measure, while Panel B reports the estimated Pareto coefficients for the same thresholds T_L but for the number of plants instead.

Figure A.3: Evolution of employment shares in large Indonesian manufacturing plants



Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers.

Figure A.4: Evolution of Pareto tail coefficients



Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers.

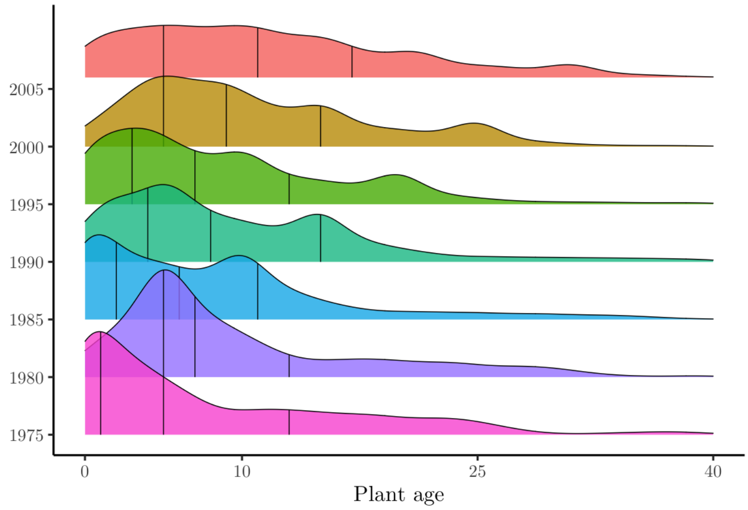
In accordance with the main text, we estimate Pareto tails by decade in 1975, 1985, 1995, 2005 and 2015. The overall trend in Pareto tails is consistent for different measures and different thresholds T_L: the tail of the employment distribution grows markedly thicker over time. However, there are important differences both across measures and across different thresholds, which is not in line with a common Pareto distribution in the cross-section. The quantitative implications are also very different for the two different measures, because Pareto tails below 1 imply that not even the mean of the distribution is defined.

Figure A.5 reports changes in the plant age distribution over time. Average plant age increased by roughly 40% between 1975 and 2006. While the 1975 plant distribution does feature very old plants, by far most plants are very young. In contrast, 30 years later, the plant age distribution is far more equally distributed, featuring more medium-old plants and relatively fewer very young plants.

1.A.3 Additional results for iterating on initial distribution

In this subsection, we provide further results on the reduced-form exercise of iterating on the discretized initial plant distribution. We start out by showing that 1975 and 1976 are good starting years, and if anything, give conservative estimates. We then show that results are very

Figure A.5: Evolution of the age distribution across plants



Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. Only showing data for years 1975, 1985, 1995,2006. The last year is 2006, because 2006 is the last year where plant age is separately asked in the survey. After 2006, we only observe plant age for surviving plants, biasing estimates of the cross-sectional age distribution.

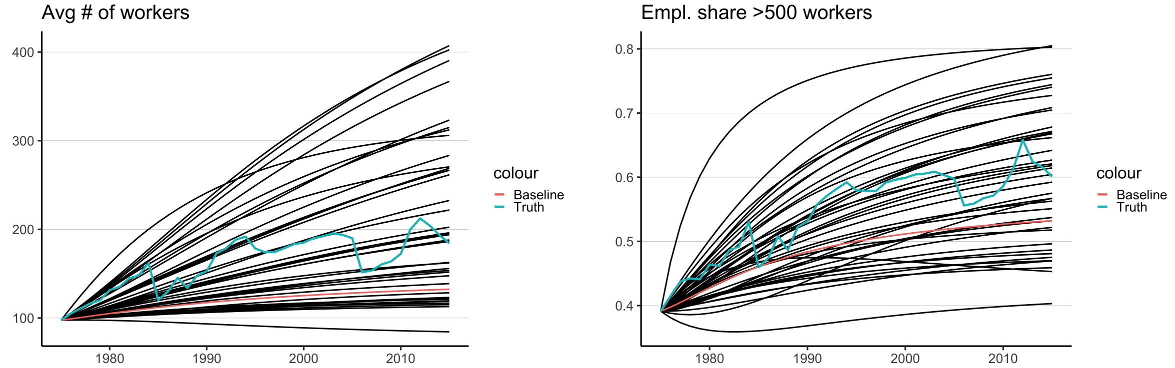
similar when accounting for entry and exit.

Results are similar when varying transition matrices and averaging transition matrices over multiple years. From an historical point of view, taking 1975 as the starting year and 1975 and 1976 as the initial years from which we construct the reduced-form transition matrix, is conservative, because (1) the two years fall in between the two periods of growth accelerations identified by Hausmann, Pritchett, and Rodrik (2005) for Indonesia (which are dated to 1967-1974 and 1987-1994 respectively), and (2) they are not affected by any notable labor or financial market reforms and predate the major tax reform of 1976 (see: Hill 2000). This is not to say that the 1970s were economically without important events. Oil prices rose dramatically in 1973 and inflation became a major macroeconomic issue that was followed by interest rate hikes and ceilings on commercial bank credit in 1974 (Hill 2000, 294). There were also important export-promoting trade policy reforms throughout the 1970s, but during a time in which Indonesia was still a very closed economy. Based on World Bank national accounts data, exports made up around 22% of GDP in 1975 whose share actually slightly decreased from 1975 to 1976, alleviating the concern that the growth between 1975 and 1976 is purely driven by trade reforms.

Figure A.6 shows that taking transition matrices for any other starting year (e.g. 1985 as starting year for 1985-1986 transitions) gives, if anything, stronger results than the ones reported for 1975 and 1976. Most years see much more growth in the average plant size and the employment share of large plants. Importantly, all years show an eventual increase in the employment share of large plants, giving credence to the idea of a tail that slowly fills up. Furthermore, any other starting year in the 1970s would have given much stronger results. E.g. taking transitions between 1976 and 1977 would have explained 67% of the average size increase and 96% of the employment share increase over time. We also considered averaging transition matrices across multiple years and obtained very similar results.

Next, we considered two variations on the exercise to account for entry and exit. To begin with,

Figure A.6: Reduced-form transition dynamics from initial conditions in 1975 and all year-to-year transition matrices

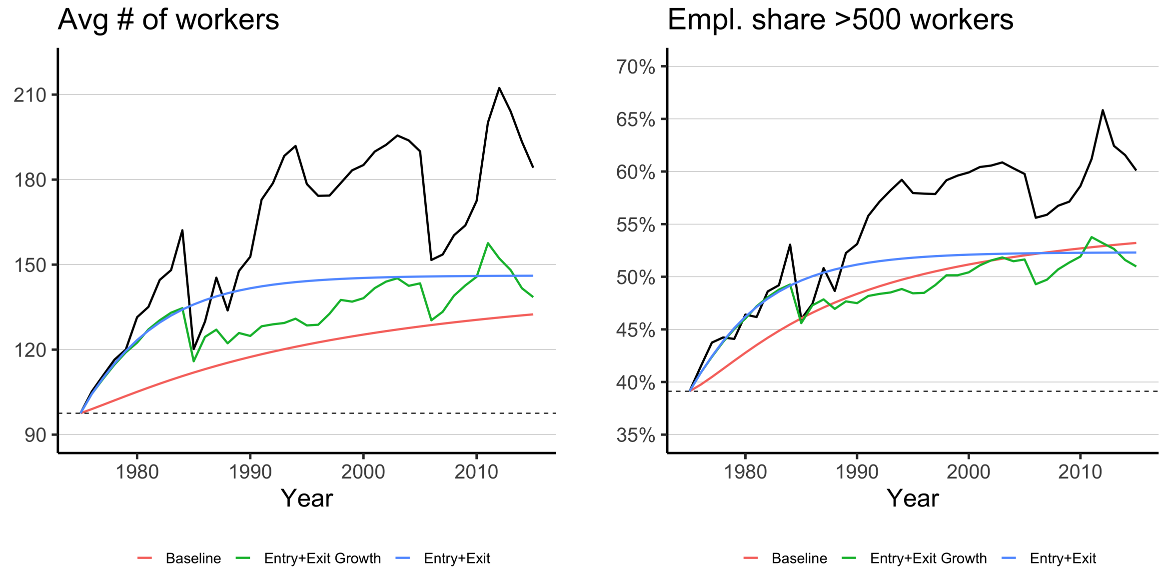


Notes: Reduced-form transition dynamics implied by initial plant size distribution in 1975, but taking transition matrices from each year-to-year pair in the data.

note that entry and exit is potentially very important, especially if entering plants differ from exiting plants. Of the roughly 6,800 plants with more than 20 workers operating in 1975, less than 12% were still operating in 2015. On the other hand, as shown previously, the number of active plants increased by a factor of 4 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, 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 "Entry + Exit" in Figure A.7. While the long-run results are almost unchanged to the previous results, introducing entry and

Figure A.7: Reduced-form transition exercise with entry and exit



Notes:

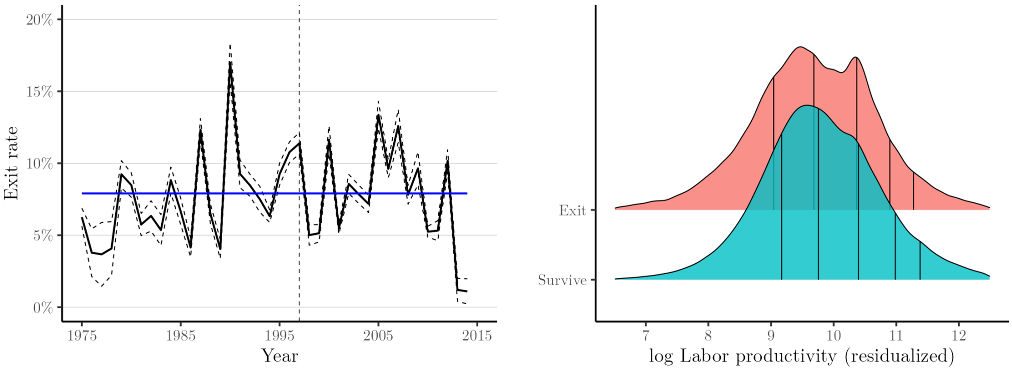
exit does speed up transition dynamics considerably, providing a much better out-of-sample fit for the early transition period. This is driven by observed exiting plants being smaller and less productive than observed new entrants in 1976. With a positive share of inactive plants staying inactive each 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 1.2, 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 "Entry + Exit Growth" in Figure A.7. 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 longer-run effects on the evolution of the size distribution.

1.A.4 Further details on exit behavior

Here, we provide evidence that exit behavior only varies little with plant productivity and does not clearly respond to aggregate shocks. Figure A.8 Panel A shows that exit rates vary quite strongly over time, but are not straightforwardly affected by measurable aggregate economic shocks. For example, exit rates actually decreased during the Asian Financial Crisis in 1998 & 1999. To focus only on within-industry variation, we residualize exit rates by 5-digit industry fixed effects here. Figure A.8 Panel B shows productivity distributions of exiting and surviving plants using value added per worker as a simple measure of (labor) productivity and only using within-industry-time variation by residualizing the measure by detailed 5-digit industry-year fixed effects. Surviving plants are more productive on average than exit-

Figure A.8: Main plant exit patterns



(A) Avg exit rates over time (B) Distribution: Survive vs Exit

Notes: Panel A: Residualized by 5-digit industry fixed effects. Standard errors are two-way clustered by industry and year. Results show 95% confidence bands and blue line gives unconditional average. Panel B: Labor productivity is measured as value added per worker, residualized by 5-digit industry-year fixed effects. Vertical lines report 25th, 50th, 75th, 90th and 95th percentiles respectively.

ing plants, but given strong overlap in the two distributions, most plants do not exit because of their productivity. A dynamic implication of this difference is that much of plant exit is not driven by productivity so that it takes time for unproductive plants to leave the economy and productivity improvements from selective exit take time to materialize.

1.A.5 Details and robustness for production function estimation

We start out by proving formal identification of the production function for the different cases (static vs. dynamic capital, time variation in production functions, industry heterogeneity). We then discuss our estimation strategy and provide detailed estimation results for the case with full flexibility on the time variation in production functions but without industry heterogeneity. In the last part, we then consider the case of industry heterogeneity.

Identification of production function

We have the following setup. Log output by firm i at time t is given by

$$y_{it} = x_{it} + f(h_{it}, k_{it})$$

where x_{it} is productivity, h_{it} is labor input, k_{it} is capital and the price of the homogeneous production good is normalized to unity throughout. We leave $f()$ unspecified here to make clear that the identification proof is non-parametric. In the estimation and model, we assume that $f()$ is Cobb-Douglas. We also suppress industry variation here, but identification extends naturally to the case with industry variation in production functions. We start with the more general case where both capital and labor are chosen dynamically and then discuss the simpler case when capital is statically chosen.

Following the literature, we assume that in the case of dynamic capital input choices, capital is pre-determined. The input choices can then be written as non-parametric functions of the relevant state-space:

$$h_{it} = f_h(h_{it-1}, k_{it}, x_{it}, \Omega_t)$$
$$k_{it} = f_k(k_{it-1}, h_{it-1}, x_{it-1}, \Omega_{t-1})$$

where we have specified in which sense capital is pre-determined. For notational simplicity, we drop the dependence on Ω_t throughout, because identification arguments are cross-sectional. At last, productivity follows a general first-order Markov process with

$$x_{it} = f_x(x_{it-1}, u_{it}) \qquad with \quad u_{it}|x_{it-1} \sim U(0,1)$$

where u_{it} is an innovation. This representation follows the Skorohod representation of random variables and is without loss of generality (see Demirer 2022). Output, labor and capital are strictly monotonic in productivity, which imposes weak regularity conditions on the productivity process x_{it}, such that each can be inverted for productivity:

$$\frac{\partial y}{\partial x} > 0 \implies x_{it} = f_y^{-1}(h, k, y, \Omega_t) \tag{1.15}$$
$$\frac{\partial h}{\partial x} > 0 \implies x_{it} = f_h^{-1}(h_{-1}, h, k, \Omega_t) \tag{1.16}$$
$$\frac{\partial k}{\partial x} > 0 \implies x_{it-1} = f_k^{-1}(k, k_{-1}, h_{-1}, \Omega_{t-1}) \tag{1.17}$$

We now adapt the identification proof by Demirer (2020):

$$h = f_h(h_{-1}, k, f_x(x_{-1}, u)) = f_h(h_{-1}, k, f_x(f_y^{-1}(h_{-1}, k_{-1}, y_{-1}), u)) = \tilde{f}_h(h_{-1}, k, k_{-1}, y_{-1}, u)$$
$$u = F_{h|h_{-1},k,k_{-1},y_{-1}}(h|h_{-1}, k, k_{-1}, y_{-1})$$

Intuitively, two firms with the same current capital, previous labor, previous capital and previous output, but different today's labor differ only in innovation to productivity. Using the identified u, we can then identify the production function using the control function $f_x(x_{-1}, u)$ for unobserved productivity x:

$$y = f(h, k) + f_x(x_{-1}, u) = f(h, k) + f_x(f_y^{-1}(h_{-1}, k_{-1}, y_{-1}), u)$$

A semi-parametric regression of y on the known function $f(h, k)$ of observables and a non-parametric term in observables/identified terms $(h_{-1}, k_{-1}, y_{-1}, u)$ identifies the output elasticities of interest.

In the case where capital is chosen statically (e.g. via a frictionless rental market), the identification approach simplifies. Specifically, dependence on k drops out in the sense that input choices are now given by:

$$h_{it} = f_h(h_{it-1}, x_{it}, \Omega_t)$$
$$k_{it} = f_k(h_{it-1}, x_{it}, \Omega_t) = f_k(h_{it}, x_{it}, \Omega_t)$$

Output and labor are strictly monotonic in productivity such that:

$$\frac{\partial y}{\partial x} > 0 \implies x_{it} = f_y^{-1}(h, y, \Omega_t) \tag{1.18}$$
$$\frac{\partial h}{\partial x} > 0 \implies x_{it} = f_h^{-1}(h_{-1}, h, \Omega_t) \tag{1.19}$$

Identification is then given by:

$$h = f_h(h_{-1}, f_x(x_{-1}, u)) = f_h(h_{-1}, f_x(f_y^{-1}(h_{-1}, y_{-1}), u)) = \tilde{f}_h(h_{-1}, y_{-1}, u)$$
$$u = F_{h|h_{-1},y_{-1}}(h|h_{-1}, y_{-1})$$

Now, without dependence on capital, two firms with the same previous labor and previous output, but different today's labor differ only in innovation to productivity. Using the identified u, we can then identify the production function using the control function $f_x(x_{-1}, u)$ for unobserved productivity x:

$$y = f(h, k) + f_x(x_{-1}, u) = f(h, k) + f_x(f_y^{-1}(h_{-1}, y_{-1}), u)$$

A semi-parametric regression of y on the known function $f(h, k)$ of observables and a non-parametric term in observables/identified terms (h_{-1}, y_{-1}, u) identifies the output elasticities of interest.

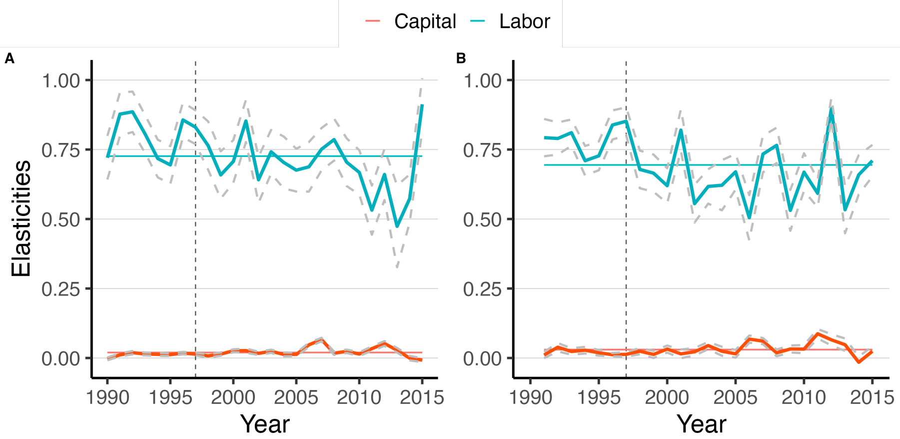
Production function estimation with time-variation but no industry variation

In the following, we report our production function estimation results. We start by showing results based on the sample from 1990 until 2015, which includes data on plant-level capital. We then discuss estimates from 1975-1990 and estimates with further industry heterogeneity. Importantly, (1) the estimated elasticities for labor are not biased by excluding capital, and (2) the estimated capital elasticities post-1990 are very low, meaning that any choice on how to model capital has only very small effects on productivity estimates.

Figure A.9 reports estimated capital and labor elasticities. Panel A gives estimates based on assuming that capital is statically chosen, while Panel B allows capital to be dynamically chosen. The estimates assume a common production function across manufacturing industries, but allow for fully flexible elasticities over time. Allowing for time-series variation is important, because policy functions are generally time-varying if the economic environment changes so that pooling estimates across years without allowing input choices to vary over time is model inconsistent. Elasticity estimates are remarkably stable and do not show a clear trend over time. Estimates are also very similar whether one assumes static or dynamic capital input choices. For example, the average estimated labor elasticity varies by less than 5% across the two different estimators (from 0.726 to 0.694). Furthermore, the estimated labor elasticity is close to 2/3, a common value in the literature. Note, however, that this is in a context where both the aggregate labor share in manufacturing (around 0.25) and the median labor share (around 0.54) are substantially below the estimated elasticities. Our model accounts for this systematic difference. Estimated capital elasticities are much lower than commonly estimated/used values in the literature and we discuss this point further below.

Elasticity estimates for each year are based on the following estimation steps: In the first step,

Figure A.9: Estimated capital and labor output elasticities for each year between 1990-2015



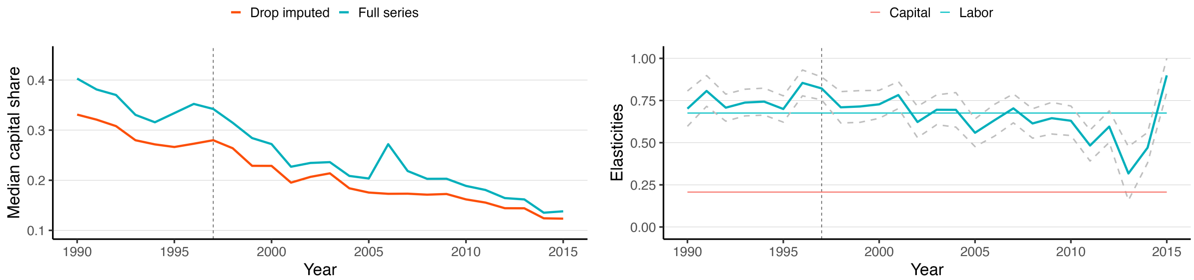
Notes: Panel A gives estimates based on a static choice of capital, Panel B gives estimates assuming capital is dynamically chosen. The latter requires to drop year 1990 in the estimation, because the estimator requires previous capital choices. Horizontal lines give average estimates over time. Grey dotted lines give 95 percent confidence bands (note that standard errors are not yet corrected for the two-stage estimation).

we flexibly estimate the rank (taking the empirical cumulative distribution function) of labor conditional on previous labor and previous output. From this estimate, we then back out a monotonic transformation of the productivity innovation u as the difference between the observed and estimated rank. In the second step, we then estimate a log-log regression in capital and labor on output, but flexibly controlling for the different components of the control function (the estimated u and previous labor, output and capital). The dynamic and static capital estimators differ only in the variables that we condition on in each of the two estimation steps. For both estimators and both estimation steps, we draw on generalized additive models (GAMs) as a flexible and robust way to estimate semi-parametric models (Hastie 2017). We obtain very similar results when choosing flexible polynomial regressions.

Given our estimation approach, why are capital elasticity estimates so low? We think the main reason here is a standard attenuation bias in the capital elasticity estimates given substantial measurement error in observed plant-level capital. Apart from the control function term, the second step estimation is a standard linear regression in capital, so that any classical measurement error in capital will attenuate the estimated capital elasticity. Why do we suspect measurement error in the capital series? As reported in Cali, Le Moglie, and Presidente (2021), one common issue in the reported plant-level capital series is misreporting in the units, which exactly shows up as a log-additive measurement error. Apart from such unit misreporting, capital – as is well known – is also more susceptible to misreporting because it is a stock that not all plants necessarily keep track of (in contrast to cost flows such as the labor bill). Inferring changes in the stock directly from reported investments and assumptions on capital-type-specific depreciation rates (as is done in perpetual inventory methods and the capital series based on Cali, Le Moglie, and Presidente (2021) that we draw on), can mitigate some of this measurement error but unlikely all.

Besides attenuation bias from noisy capital reporting, the capital series likely also suffers from more systematic biases that complicate their use. To show this, we consider the case where capital is chosen statically in which case we can directly make use of plants' first-order conditions instead of estimating the capital elasticity from the output regression. This first-order

Figure A.10: The effects of alternative estimates of capital elasticities



Notes: Left: Evolution of capital shares based on the capital series in Cali, Le Moglie, and Presidente (2021). For better comparability, using also their real value added series which deflates output by sector-specific prices. To construct capital shares, assume that rental rate is 14 percent (interest rate of 4 percent and depreciation of 10 percent) and value added tax is 10 percent. Right: Estimated labor elasticities when assuming that capital is statically chosen and its output elasticity is fixed at the median from Cali, Le Moglie, and Presidente (2021), which is 0.207.

approach does not generally suffer from attenuation bias because one can estimate capital elasticities from average or median capital shares, which are robust to log-additive measurement error. Specifically, the static capital input choice conditional on the assumed production structure implies the standard condition:

$$\alpha_{jt} = \frac{(r_t + \delta_t)k_{it}}{(1 - \tau_y)p_t y_{it}}$$

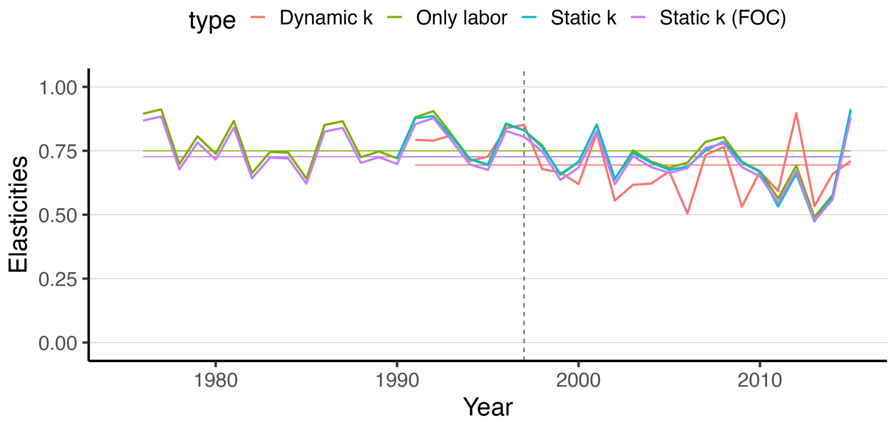
where α_{jt} gives the potentially industry-time-specific capital elasticity, $(r_t + \delta)$ gives the competitive rental rate of capital, τ_y gives a value-added tax and $p_t y_{it}$ gives plant revenue.[26] Figure A.10 (left) plots changes in median capital shares over time. Capital shares more than halved between 1990 and 2015, which would also imply a halving of capital elasticities in the case of static capital choices. This seems unlikely, among others because of global capital-biased technological change, the advent of industrial robots and a running out of labor intensive industrialization in Indonesia. Also, in the case of adjustment frictions, one would have expected an increase in the capital share as the economy is catching up, not a decline. More likely, we think plants systematically underreport new capital investments and we might overestimate depreciation of existing capital.[27] Taking noisy capital estimates together with systematic misreporting, it is hard to trust the Indonesian plant-level capital series. We thus choose to stick with the low estimated capital elasticities for the baseline model and results, which implies that observed capital variation has very small effects on plant output and labor decisions. Still, to gauge how sensitive our estimates are to low capital elasticities, Figure A.10 (right) also reports estimated labor elasticities in case where we enforce the much higher capital elasticity based on the median capital share. We find that estimated labor elasticities are almost unchanged when assuming such a higher capital elasticity.

Next, we consider production function estimation for the period 1975-1990 for which we lack data on capital. We follow two different approaches to understand whether there have been

[26] If observed capital features log-additive measurement error ($\tilde{k}_{it} = k\varepsilon$), then, $\mathbb{E}_i \frac{(r_t+\delta_t)\tilde{k}_{it}}{(1-\tau_y)p_t y_{it}} = \alpha_{jt}$.

[27] Of course, one may also explain a strong decline in estimated revenue elasticities by changes in markups. However, the required magnitude of such markup changes also seems unrealistic. Through the lens of a standard monopolistic competition model, the difference in estimated revenue elasticities would translate to a roughly 100 percentage point increase in the markup. That is, if a product is sold at 50% above marginal cost, it would now be sold at 150% above marginal cost.

Figure A.11: Estimated output elasticities of labor for each year between 1975-2015



Notes: Dynamic and static k report previously estimated labor elasticities. Only labor shows estimates based on the assumption that the production function only features labor. Static k (FOC) assumes that the capital choice is static and enforces the average estimated capital elasticity from the Dynamic k estimator (which nests the static k estimator) from 1990-2015.

important changes in production functions over time. In the first approach, we simply assume that the production function does not feature capital and estimate only labor elasticities from 1975-2015, comparing this to the estimated labor elasticities estimated with capital data from 1990-2015. In the second approach, we assume a static choice for capital and enforce the estimated capital elasticities from before. As shown in Figure A.10, we find that for both approaches estimated labor elasticities are very similar to estimates after 1990 and that they are remarkably stable and do not show a clear trend over time. We interpret this as strong evidence that production functions did not systematically change over time.

Production function estimation with industry variation

We now consider industry-level variation in production functions. (Show 2 results: 1. Test for equality across industries. 2. Check how correlated productivity estimates are)

1.A.6 Details and robustness for aggregate technology estimates

In this section, we give a formal identification proof, estimation details for separating aggregate technology from idiosyncratic productivity and a discussion of the drivers of technology growth. To simplify the exposition, in the following we will denote the logarithm of a vector by lower letter cases. Hence, in slight deviation from the exposition in the main paper, we assume that productivity for plant i at time t is given by: $Y_{it} = Z_t exp(s_{it})$. So that the log-additive form is: $y_{it} = z_t + s_{it}(s_{it-1})$. The average of within-plant changes in log productivity is then:

$$\frac{1}{N^S_{t,t-1}} \sum_{i \in \mathcal{N}^S_{t,t-1}} \Delta y_{it} = \underbrace{z_t - z_{t-1}}_{\Delta z} + \underbrace{\frac{1}{N^S_{t,t-1}} \sum_{i \in \mathcal{N}^S_{t,t-1}} \Delta s_{it}}_{\text{Avg mean reversion of survivors}}$$

Identification

Proposition 1 (Main identification result)**.** *Under the following four assumptions:*

1. *(**Common first-order stationary Markov process**) s follows the same general first-order, stationary ergodic Markov process for all i & t.*

2. *(**Selective exit**). The decision to exit after period t can flexibly depend on observables and unobservables X_{it} as well as productivity s_{it}, but may not depend on future productivity s_{it+1}. Specifically,*

$$\mathbb{P}(exit) = f(X_{it}, s_{it}, z_t) \quad \textit{with} \quad \mathbb{P}_t(exit) \perp\!\!\!\perp s_{i,t+1}|s_{i,t}$$

3. *(**No complete exit over s**) $\mathbb{P}_t(exit|s_{it}) < 1 \, \forall s \in Supp(s)$*

4. *(**Connected support in s**) For each period t, there exists at least a subset of the support of s in that period which is fully contained in the support of all s in all future periods. Formally: $\forall t, \exists S_t \subset Supp(s_{it})$ for which $S_t \subset \cup_{\tau > t} Supp(s_{i\tau})$.*

the path $z_t \, \forall t$ is identified given some normalization z_τ for some $\tau \in [0, T]$ and $\max t \equiv T \to \infty$.

Proof. To already convey the idea of a suitable estimator for the time path of z_t, let us proof Proposition 1 constructively. Identification proceeds sequentially in two fundamental steps. In the first step, I show identification of the density of the stationary distribution of s, which is identified for $t \to \infty$. In the second step, the density of the stationary distribution is used to identify the path of z_t backwards by starting at some final time T. The density of the stationary distribution is key because it can be used to construct weights under which a weighted difference Δy_{it} exactly identifies Δz_t. Specifically, there exist weights ω_s such that $\sum_{i \in \mathcal{N}^S_{T+1,T}} \omega(s_{iT})(s_{iT+1} - s_{iT}) = 0$ (where $\sum_i \omega_s(s_i) = 1$). These weights recover the stationary distribution of s. Denote by $f^{SS}(s)$ the density of the stationary distribution at s and by $f_t(s)$ the density of the distribution of s at time t. Assuming that this distribution shares the support of the stationary distribution, we have:

$$\lim_{N \to \infty} \sum_{i \in \mathcal{N}^S_{t+1,t}}^{N} \frac{f^{SS}(s_{it})}{f_t(s_{it})} \Big(log(s_{it+1}) - log(s_{it}) \Big) = 0$$

The weights are thus defined by $\omega_s(s_{it}) \equiv \frac{f^{SS}(s_{it})}{f_t(s_{it})}$ and are a function of the unknown density function of the stationary distribution of s. To identify the density $f^{SS}(s)$, start with the distribution of plants at t_0 over known y_{i0}. The idea is to follow survivors (as they follow the process for s), while replacing exiting plants with plants that stay in the panel that have similar y_{it}. More formally, denote the initial set of plants by $\mathcal{N}_0$ where each plant is given a uniform weight $\tilde{\omega}_{i0} = \frac{1}{N_0}$. We are interested in updating $\mathcal{N}$. For this, pass on the weight of each surviving plant and redistribute the weight of each plant that exits to close plants around them.[28] This gives $\mathcal{N}_1$. Updating in this way allows to eventually pass on weight to plants that have entered the economy, even if they have entered in an arbitrarily selective way. As $t \to \infty$, surviving plants will eventually populate the entire support of s and this procedure gives a synthetic sample $\mathcal{N}_\infty$ with weights $\tilde{\omega}_{i\infty}(s_{i\infty})$ that directly identify the density $f^{SS}(s)$.

The second step of the proof takes the identified density $f^{SS}(s)$ and works backwards from time T. Normalizing the final value z_T, one can show that z_{T-1} solves a fixed point problem.

[28] As $N \to \infty$ and the assumption that exiting probabilities are always strictly lower than one, there always exists a plant that is arbitrarily close to an exiting plant.

Specifically:

$$\sum_{i\in\mathcal{N}^S_{T,T-1}} \omega_{\hat{s}_{T-1}(z_{T-1})}(y_{iT}-y_{iT-1}) = z_T - z_{T-1} + \sum_{i\in\mathcal{N}^S_{T,T-1}} \omega_{\hat{s}_{T-1}(z_{T-1})}(s_{iT}-\hat{s}_{iT-1}(z_{T-1})) = z_T - z_{T-1}$$

where the last equality holds only if the guess z_{T-1} is correct. It thus gives a nonlinear equation in z_{T-1} (since the weights and the right-hand side depend on z_{T-1}). One can iterate on this procedure to identify the path of z_t backwards. At any point in time $t < T-1$, one can also alternatively guess z_{T-1} and instead of using weights at all, estimate the bias term $\sum_{i\in\mathcal{N}^S_{T,T-1}}(s_{iT} - \hat{s}_{iT-1}(z_{T-1}))$ directly using future survivors with similar s. This alternative relaxes the assumption of a common support with the stationary distribution and instead only requires that we can build a sample with similar survivors – requiring a much weaker connected support.

Estimation

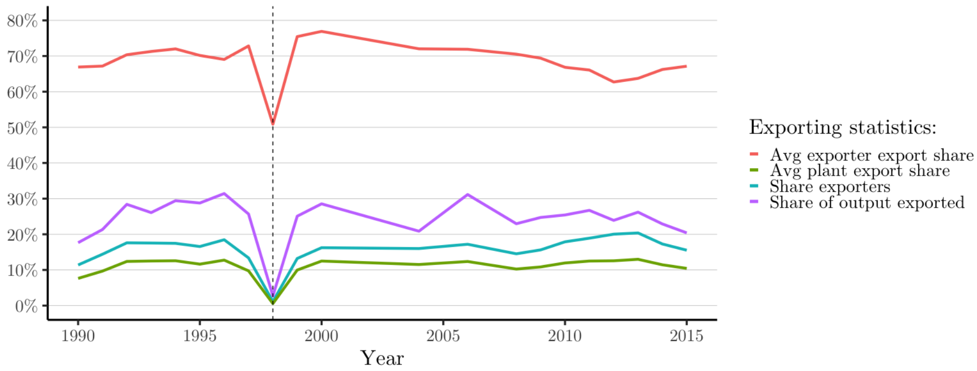
Estimation proceeds along the lines of the constructive identification proof. In the first step, one sequentially builds the synthetic panel with weights $\omega_s(s_{it})$ (which sum to 1 in each year). In principle, one can use any standard matching estimator for passing on the weight for exiting plants. We find that a Kernel matching estimator works well, because matching is only based on one variable and the Kernel estimator distributes the weight widely across multiple observations, reducing variance.[29]

One can then estimate $f^{SS}(s)$ using observed s in the last period T and constructed weights $\hat{\omega}_s(s_{iT})$. Any standard density estimator such as a Kernel density estimator works here. To reduce variance, one can also estimate $f^{SS}(s)$ on the last x periods (where x is at the discretion of the researcher). In general, for any fixed T, the bias on the estimated weights is increasing in the persistence of the process as well as in the distance of the initial distribution from the stationary distribution. That is, for large T and low persistence, one can use more periods in the end to estimate $f^{SS}(s)$.[30] Once the density is estimated, one can then proceed by sequentially estimating the path z_t. For each period t and for each guess of z_t, this means one has to estimate $f_t(s_{it}(\hat{z}_t))$. Again, any standard density estimator works here. One can then construct the weights according to: $\omega_{st}(s_{it}) \equiv \frac{f^{SS}(s_{it})}{f_t(s_{it})}$. Alternatively, one can choose not to use weights and instead directly estimate the bias from mean reversion. In that case, one can again use any kind of matching estimator to match plants in t with productivity $s(\hat{z}_t)$ to future survivors with similar s. The variance in the bias estimate reduces with the number of matched plants such that one to many matches are recommended. As before, a Kernel-based matching estimator is a natural choice here. In either the approach with weights or with an estimated bias term, one then finds z_t that solves the fixed point problem, requiring a standard root finder. We have not formally proven uniqueness of the root, but in practice, we found no issue of a multiplicity of roots. In principle, any (weighted) moment of within-plant changes in productivity that preserves scalar multiplicity can be used for the estimation. In practice, we use (weighted) median changes in productivity as the median is less susceptible to outliers. Results are similar when taking the weighted average.

[29]Note that one can readily match based on further variables such as detailed industries to minimize the risk of model misspecification.

[30]A formal treatment of optimally solving this trade-off is beyond the scope of this paper.

Figure A.12: Evolution of key exporter statistics



Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers and reported export shares (out of total value-added).

The drivers of aggregate technology

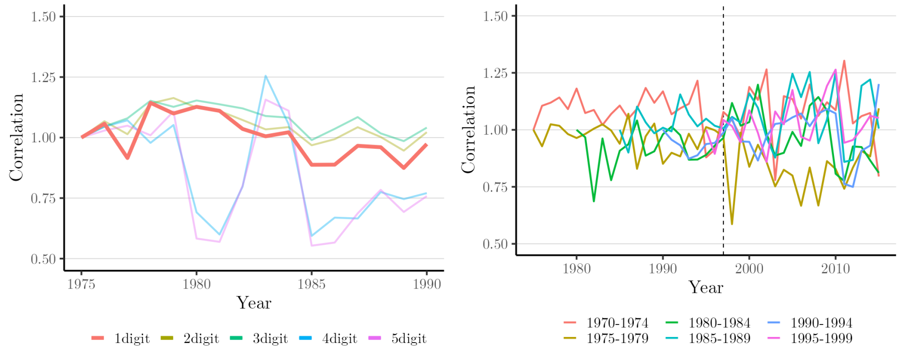
We now discuss the potential drivers of the estimated aggregate technology growth path. First, through the lens of standard endogeneous growth models such as Romer (1990), the increase in technology growth after the year 2000 could be driven by overall human capital improvements and a better integration into global markets. However, there seems to be no sharp change in human capital improvements nor in the integration into global markets. For example, Figure A.12 shows that the share of manufacturing output that is exported stays very stable around 20% and the fraction of plants that are exporting also remained flat since the mid 1980s.

Alternatively, the patterns may be in line with models of learning and imitation (e.g. Perla, Tonetti, and Waugh 2021), whereby the initial entry of many new and relatively unproductive plants lowered productivity growth and the subsequent better selection of plants increased the productivity growth from learning and imitation. While this may be an underlying driver of technology growth, we only find a very weak correlation between contemporaneous changes in aggregate technology and the evolution of average plant productivity (or other moments of the productivity distribution).

The increase in technology growth may also be in line with the recent theory in Ottonello and Winberry (2023) whereby constrained firms initially invest in factor accumulation and only later in activities that increase productivity. Given that we find little empirical evidence for capital deepening, only little plant-level labor deepening in the sense of rising labor shares and strong increases in financial access in the run-up to the Asian Financial Crisis, we are rather skeptical that this mechanism can explain large changes in aggregate technology. Still, given the limitations on the capital series and limited evidence on plant-level investments in technology, we cannot rule out that this mechanism is an important driving force of technology growth.

At last, since we cannot distinguish between productivity and demand drivers, changes in demand may also be an important driver of the patterns we observe. For example, the Asian Financial Crisis shows up as a more than 20% drop in technology, which is likely to be at least partly demand-driven. In line with this interpretation, Figure A.12 shows that plant-level exports almost completely plummet in 1998. Decreases in technology may also be partly explained by decreases in demand as the economy grows richer and consumers switch their

Figure A.13: Evolution of cross-sectional correlation of plant productivity and input share



(A) Balanced Sample (B) Survivors (by cohort)

Notes: Input shares are computed based on a Cobb-Douglas aggregator as explained in the text. For within-industry results, we first estimate the correlation across plants in a given industry and year and then construct the weighted average correlation across industries using the industry's average share in manufacturing value added as a time-invariant weight. All series are normalized by the first year.

demand towards services (e.g. Alder et al 2019, Comin et al 2021).

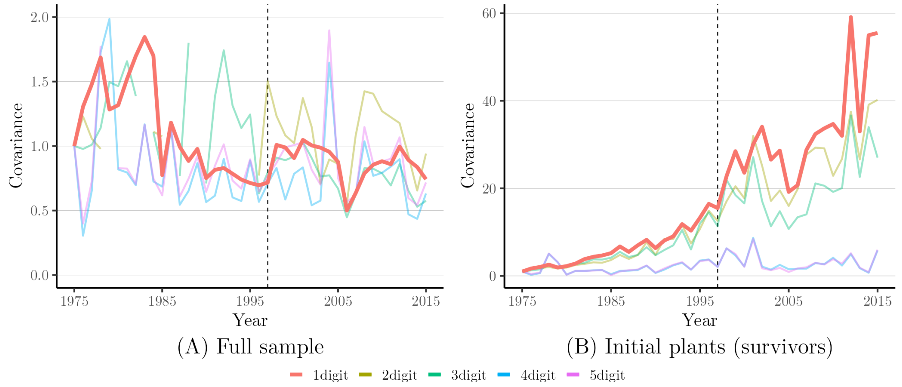
1.A.7 Additional results on changes in misallocation

In this section, we report two sets of additional results. First, we show additional evidence on the evolution of the covariance and correlation of plant-level productivity and input shares. Figure A.13 shows additional evidence on the correlation for a balanced panel of plants and for each cohort of plants between 1970 and 1999. We construct the balanced panel by selecting all plant-year observations between 1975 and 1990 for which the plant operated in 1975 and in 1990 and for which we observe more than 10 observations (to avoid dropping all plants for which individual years are missing or had to be dropped). Extending the time frame would drop too many plants.

For completeness, Figure A.14 also plots the covariance instead of the correlation. For the full sample, this covariance also does not increase over time due to the entry of small plants. For surviving plants, the covariance increases strongly. This is mechanical, because the sample shrinks over time, which naturally leads input shares to increase. Also, average productivity strongly increases among surviving plants, adding an additional trend. The correlation is robust to such common trends.

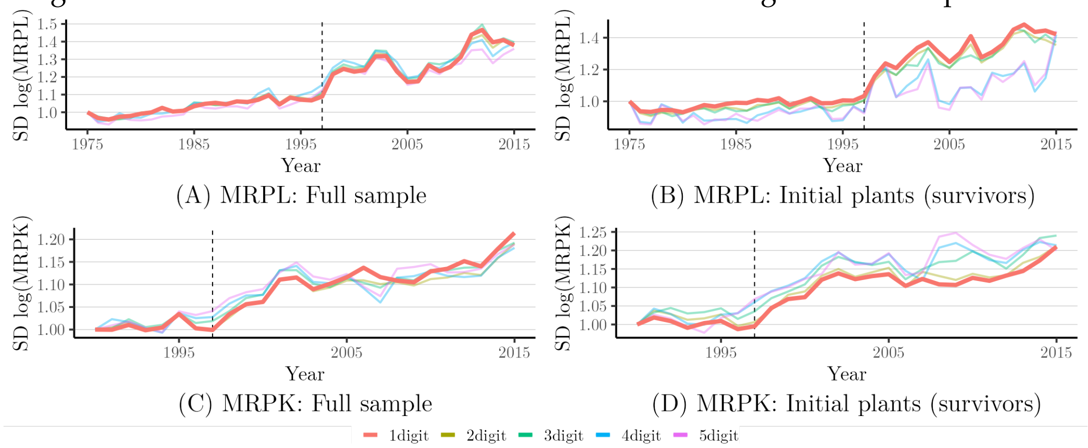
The second set of results is on an alternative measure of the allocation of resources. Figure A.15 reports changes in the dispersion of marginal revenue products of labor and capital; a sign of misallocation in the static model of Hsieh and Klenow (2009). Most importantly, we do not find any evidence for a decreasing dispersion in marginal revenue products over time, which could be linked to an "undoing of misallocation" that drives economic growth. Instead, we find evidence for an increase in the dispersion over time. Most of these increases happen after the Asian Financial Crisis in 1997. We think that the measured increases in the dispersion of marginal revenue products are at least in part driven by changes in measurement. We refer the interested reader to the discussion of measurement changes further above.

Figure A.14: Evolution of cross-sectional covariance of plant productivity and input share



Notes: Input shares are computed based on a Cobb-Douglas aggregator as explained in the text. For within-industry results, we first estimate the covariance across plants in a given industry and year and then construct the weighted average covariance across industries using the industry's average share in manufacturing value added as a time-invariant weight. All series are normalized by the first year.

Figure A.15: Evolution of cross-sectional variation in marginal revenue products

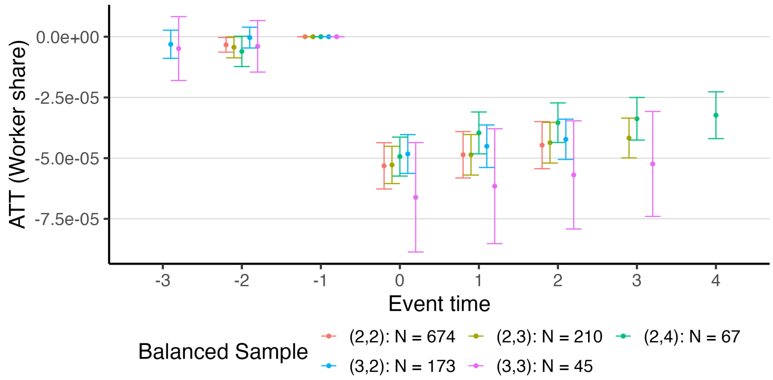


Notes: Evolution of cross-sectional standard deviation in marginal revenue products of labor and capital following Hsieh and Klenow (2009) and Gopinath et al (2017).

1.A.8 Further event study evidence

In this section, we report event study results for hiring responses (in contrast to the labor share responses). We stick to the same treatment definition as previously. To study the dynamic hiring responses of plants to a permanent positive productivity shock, we look at the *worker share* (# worker / value added) instead of the *labor share* (wage bill / value added). The concern of looking at the labor share is that a positive productivity or demand shock at the plant level may lead workers to bargain for higher wages to share in the profit gains. If these bargaining gains slowly accumulate, then we misattribute slow increases in the labor share to labor adjustment frictions. Figure A.16 shows that this concern is unwarranted. Plants actually slowly increase hiring, analogously to the labor share.

Figure A.16: Further event study results for worker shares



Notes: Worker share measured as ratio of number of workers over value added. Treatment definition as in main event study.

1.B Model and Estimation

1.B.1 Adjustment costs as costs of managerial time

In the following, we show that adjustment costs can be microfounded as costs of scarce managerial time. Our goal is to make explicit how adjustment costs can capture the time constraints of a manager working at a plant and to show why it makes sense to write adjustment costs in terms of the costs of labor w_t.

Suppose a plant owner solves the following problem:

$$V(s_{i,t}, h_{i,t-1}, \Omega_t) = \max_{h_{i,t}} \Big\{ y_{it}(s_{it}, h_{i,t}; z_t) - w_t h_t - w_t T(h_{i,t}, h_{i,t-1}) \\ + \lambda(s_{i,t}, h_{i,t}, \Omega_t) \left\{ -\mathbb{E}_c[c_F|\text{stay}] + \beta \mathbb{E}\big[V(s_{i,t+1}, h_{i,t}; \Omega_{t+1})|s_{i,t}, h_{i,t}; \Omega_t\big] \right\} \Big\}$$

where $T(h_{i,t}, h_{i,t-1})$ gives the managerial time needed to implement changes in the workforce. We denote $T(h_{i,t}, h_{i,t-1})$ in terms of the efficiency units of a worker such that we can express the manager's time cost in terms of the wage w_t. One can think of $w_t T(h_{i,t}, h_{i,t-1})$ as the actual compensation that managers receive or as a combination of compensation and opportunity costs of managerial time.

Time costs are due to two main managerial tasks: (1) the task of hiring and firing, and (2) the task of reorganizing production. Conditional on the task, we think of a single unit of the task as requiring always the same amount of time (e.g. signing one contract always takes a fixed amount of time), but the total units needed depends on the organization and the amount of hiring. Hiring and firing requires (c_F^+, c_F^-) units of time for each unit of labor hired or fired $\Delta h \equiv |h_{it} - h_{it-1}|$. This time comes from filling out paperwork, signing the contracts and adding the worker to the books. Policies that affect the paperwork that plants need to fill out, will change these costs.

Next, for any workers the plant hires or fires, managers need to assign and explain changes in worker tasks. Both for hiring and firing, we assume that managerial time to assign and explain new worker tasks is proportional to the percentage change in the workforce. We assume this is for different reasons in the case of hiring and firing and thus the unit cost of changes in the workforce for hiring and firing can differ. For hiring, the plant hires $h_{it} - h_{it-1}$ workers who they need to explain their new task. The proportional time cost in the case of hiring comes from the possibility that each new worker can also learn from their coworkers. However, in the case of relatively many new hires, each new hire can learn from relatively fewer coworkers, thus increasing the time that the manager needs to add. In the case of firing, jobs may potentially not simply disappear, but need to be reorganized. In this case, the managerial time costs of reorganizing scale with the number of lost jobs $h_{it-1} - h_{it}$ (for which replacements need to be found) and are proportional to the percentage change in the workforce because this is the amount of time the manager needs to reexplain jobs to all existing workers.

In the end, the entrepreneurial time costs T_h are then given by the functional form reported in the main text:

$$T_t(h_{i,t-1}, h_{i,t}) = \begin{cases} c^+_{0,t}(h_{i,t} - h_{i,t-1}) + \frac{c^+_{1,t}}{2}\left(\frac{h_{i,t}-h_{i,t-1}}{h_{i,t-1}}\right)^2 h_{i,t-1} & \text{if } h_t > h_{t-1} \\ 0 & \text{if } h_t = h_{t-1} \\ c^-_{0,t}(h_{i,t-1} - h_{i,t}) + \frac{c^-_{1,t}}{2}\left(\frac{h_{i,t}-h_{i,t-1}}{h_{i,t-1}}\right)^2 h_{i,t-1} & \text{if } h_t < h_{t-1} \end{cases} \tag{1.20}$$

Note that the main text also features fixed costs that can easily be rationalized by fixed time costs of managers whenever there is a change in the organization.

1.B.2 Microfoundation of working capital constraint

The microfoundation of the working capital constraint can be derived from a standard limited enforcement problem (e.g. as in Buera and Shin 2013). Assume that plant managers need to first pay their workers before being able to produce and they do so by borrowing the entire wage bill $w_t h_{it}$ with a financial intermediary. For simplicity, suppose further that the time between production and paying the wage bill is $\varepsilon \to 0$ such that the costs of borrowing go to zero. Suppose further that the plant manager – after paying their workers and producing – could decide to run away with a fraction $\frac{1}{\tilde{\kappa}_t}$ of the borrowed resources $w_t h_{it}$. Isomorphically, the plant manager runs away with all of the resources, but is caught with probability $\frac{1}{\tilde{\kappa}_t}$. We assume that the only punishment in case of successful evasion is that the financial intermediary can now sue the plant manager and claim (part of) the output of the plant in period t. We assume that the claim is proportional to plant output net of value-added tax. Importantly, the plant manager never loses access to the plant and is not excluded from any future economic activity, ensuring that the constraint remains a static problem. In equilibrium, the financial intermediary will lend $w_t h_{it}$ only to the extent that no plant manager will renege on the contract, implying the financing constraint:

$$w_t h_{it} \geq \kappa_t y_{it}$$

1.B.3 Stationarized value function and balanced growth path after 2015

After 2015, we assume that plants expect wages, all costs and aggregate productivity to rise at the same growth rate $(1 + g)$ over time. This allows to capture realistic future growth in a parsimonious way and is in line with the entire economy being on a balanced growth path after 2015. An alternative would be to only enforce constant growth in costs and productivity and then solve for the actual endogenous wage path after 2015 that clears labor markets after

2015. This would require further assumptions on how other fundamentals in the economy evolve (e.g. wedges and technology in the rest of the economy and aggregate labor supply) and feature a continued transition towards an eventual balanced growth path (as long as assumptions on the changes in future fundamentals allow for a balanced growth path). Given that the growth path after 2015 is not identified, we think that our approach strikes a good balance between realism and parsimony.

The value function in 2015 (denoted by T and suppressing dependence on Ω_T for expositional clarity) writes:

$$V_T^M(s_{i,T}, h_{i,T-1}) = \max_{h_{i,T} \in [\underline{h}, \bar{h}]} \Big\{ z_T s_{i,T} h_{i,T}^{\theta} k_{i,T}^{\alpha} - w_T h_{i,T} - (r+\delta) k_{i,T} - w_T AC(h_{i,T}, h_{i,T-1}) + \\ \lambda(s_{i,T}, h_{i,T}) \Big\{ -\mathbb{E}_c[c_F | \text{stay}_{i,T}] + \beta \mathbb{E}\big[V^M(s_{i,T+1}, h_{i,T}) | s_{i,T}, h_{i,T}\big] \Big\} \Big\} \tag{1.21}$$

In the following, we show that under the right normalization and assuming constant growth for all costs and productivity, we can write: $V_T^M(s_{i,T}, h_{i,T-1}) = \widetilde{V}^M(s_{i,T}, h_{i,T-1})\tilde{z}_T$, which implies that we can solve for $V_T^M(s_{i,T}, h_{i,T-1})$ by first solving for the stationary $\widetilde{V}^M(s_{i,T}, h_{i,T-1})$ and then renormalizing by $\tilde{z}_T$. To show this, we proceed in two steps. First, we find the normalizing factor $\tilde{z}_T$, which needs to grow at a constant rate $(1+g)$. We do so by deriving the optimal static capital choice: $k_{i,T}^* = \left(\frac{\alpha}{r+\delta} z_T s_{i,T} h_{i,T}^{\theta}\right)^{\frac{1}{1-\alpha}}$. Plugging the capital choice into the value function gives:

$$V_T^M(s_{i,T}, h_{i,T-1}) = \max_{h_{i,T} \in [\underline{h}, \bar{h}]} \Big\{ z_T^{\frac{1}{1-\alpha}} s_{i,T} h_{i,T}^{\theta} \left(\frac{\alpha}{r+\delta} s_{i,T} h_{i,T}^{\theta}\right)^{\frac{\alpha}{1-\alpha}} - w_T h_{i,T} - \\ z_T^{\frac{1}{1-\alpha}} (r+\delta) \left(\frac{\alpha}{r+\delta} s_{i,T} h_{i,T}^{\theta}\right)^{\frac{1}{1-\alpha}} - w_T AC(h_{i,T}, h_{i,T-1}) + \\ \lambda(s_{i,T}, h_{i,T}; \Omega_T) \Big\{ -\mathbb{E}_c[c_F | \text{stay}_{i,T}] + \beta \mathbb{E}\big[V^M(s_{i,T+1}, h_{i,T}) | s_{i,T}, h_{i,T}\big] \Big\} \Big\} \tag{1.22}$$

from which we can see that for output and capital to grow at a constant rate, we need that $z_T^{\frac{1}{1-\alpha}} \equiv \tilde{z}_T$ grows at a constant rate. Dividing through by $\tilde{z}_T$ gives the deflated value function $\widetilde{V}^M(s_{i,T}, h_{i,T-1})$:

$$\widetilde{V}^M(s_{i,T}, h_{i,T-1}) = \max_{h_{i,T} \in [\underline{h}, \bar{h}]} \Big\{ s_{i,T} h_{i,T}^{\theta} \left(\frac{\alpha}{r+\delta} s_{i,T} h_{i,T}^{\theta}\right)^{\frac{\alpha}{1-\alpha}} - \frac{w_T}{\tilde{z}_T} h_{i,T} - \\ (r+\delta) \left(\frac{\alpha}{r+\delta} s_{i,T} h_{i,T}^{\theta}\right)^{\frac{1}{1-\alpha}} - \frac{w_T}{\tilde{z}_T} AC(h_{i,T}, h_{i,T-1}) + \\ \lambda(s_{i,T}, h_{i,T}; \Omega_T) \Big\{ -\frac{\mathbb{E}_c[c_F | \text{stay}_{i,T}]}{\tilde{z}_T} + \beta \mathbb{E}\big[(1+g)\widetilde{V}^M(s_{i,T+1}, h_{i,T}) | s_{i,T}, h_{i,T}\big] \Big\} \Big\} \tag{1.23}$$

where we have made use of the constant growth in $\tilde{z}$:

$$\frac{V^M_{T+1}(s_{i,T+1},h_{i,T})}{\tilde{z}_T} = \frac{V^M_{T+1}(s_{i,T+1},h_{i,T})}{\tilde{z}_{T+1}}\frac{\tilde{z}_{T+1}}{\tilde{z}_T} = (1+g)\widetilde{V}^M(s_{i,T+1},h_{i,T})$$

In the second step, we need to prove that all other aggregate time-varying components also grow at the same rate. In the deflated value function, output and capital do not depend on time-varying aggregates anymore. All terms featuring wages require that wages grow at the same rate $(1+g)$, such that $\frac{w_T}{\tilde{z}_T} = \tilde{w}$ is a constant. The trickier parts are $\lambda(s_{i,T}, h_{i,T}; \Omega_T)$ and $\mathbb{E}_c[c_F|\text{stay}_{i,T}]$. We prove that $\lambda(s_{i,T}, h_{i,T}; \Omega_T)$ does not depend on time Ω_T if all costs grow by the same rate and that expected fixed costs $\mathbb{E}_c[c_F|\text{stay}_{i,T}]$ grow at the same rate $(1+g)$.

Using the analytic formula for the survival rate, it is easy to see that the survival rate does not vary with time-varying aggregates as long as μ_{xT} and σ_{xT} grow at the same rate $(1+g)$:

$$\begin{aligned}
\lambda(s_{i,T}, h_{i,T}; \Omega_T) &= exp\left(-exp\left(-\frac{\beta\mathbb{E}\big[\frac{V^M(s_{i,T+1},h_{i,T},\Omega_{T+1})}{\tilde{z}_T}|s_{i,T},h_{i,T},\Omega_{T+1}\big]-\frac{\mu_{xT}}{\tilde{z}_T}}{\frac{\sigma_{xT}}{\tilde{z}_T}}\right)\right)\\
&= exp\left(-exp\left(-\frac{\beta\mathbb{E}\big[(1+g)\widetilde{V}^M(s_{i,T+1},h_{i,T})|s_{i,T},h_{i,T},\Omega_{T+1}\big]-\frac{\mu_{xT}}{\tilde{z}_T}}{\frac{\sigma_{xT}}{\tilde{z}_T}}\right)\right)
\end{aligned} \tag{1.24}$$

At last, we use the analytic formula for the expected fixed costs to show that they indeed grow at the same rate $(1+g)$:

$$\begin{aligned}
\frac{\mathbb{E}_c[c_F|\text{stay}_{i,T}]}{\tilde{z}_T} =& \beta\mathbb{E}\big[(1+g)\widetilde{V}^M(s_{i,T+1},h_{i,T})|s_{i,T},h_{i,T},\Omega_{T+1}\big]\lambda(s_{i,T},h_{i,T})-\\
&\frac{\sigma_{xT}}{\tilde{z}_T}\Gamma\left(0, exp\left(-\frac{\beta\mathbb{E}\big[(1+g)\widetilde{V}^M(s_{i,T+1},h_{i,T})|s_{i,T},h_{i,T},\Omega_{T+1}\big]-\frac{\mu_{xT}}{\tilde{z}_T}}{\frac{\sigma_{xT}}{\tilde{z}_T}}\right)\right)
\end{aligned} \tag{1.25}$$

which does not depend on aggregate time-varying components.

1.B.4 Formal derivation of main accounting identity

We can start by giving a formal derivation of the main accounting identity that we use to validate our model.

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

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

1.B.5 Details on estimation

In this section, we provide further details on the model estimation.

Taxes

In the following, we discuss how we map Indonesian corporate income and value added taxes (VAT) over the period 1975-2015 to our model economy. In both cases, we assume that tax rates are constant over time and uniform across firms/plants. This is a very accurate mapping for the VAT, but less accurate for the corporate income tax rate. Throughout, we abstract from the very important topic of tax evasion and enforcement, but we discuss the empirical evidence on this.

We start with the simpler VAT, which we fix in the model economy to a constant 10%, the rate which was officially introduced in 1985 and has remained unchanged in place until now (see: Gillis 1985; Hill 2000; Basri et al. 2021). Officially, the only exemptions are on exports, which we do not model and thus abstract from and there were higher luxury product rates in place that we also abstract from. The VAT replaced an older sales tax that was in place between 1951-1985, whose rates varied from 5 to 20%, but with most sales subject to a 10% rate (Lent and Ojha 1969, 537). Enforcement of the sales tax before 1985 was almost absent, leading to widespread evasion (e.g. see: Gillis 1985), but the introduction of the VAT greatly improved (self-)enforcement and reduced evasion (Hill 2000; Basri et al. 2021), so that our assumption of a flat 10% rate seems reasonable (at least since 1985).

Changes in the corporate income tax rates are slightly more complicated, with major reforms in 1985 and 2009. We simplify our analysis by assuming a fixed 20% corporate income tax rate across firms and over time. Before 1985, many different tax rates were in place, including top marginal rates at 45%, which were non-enforced (Gillis 1985). With the 1985 reforms, corporate income rates were reduced and homogenized, with the maximum marginal rate capped at 35% (Gillis 1985). Between 1985 and 2009, the corporate income tax rates followed a 3-tiered schedule of different marginal tax rates defined over taxable profits (see: Gillis 1985; Basri et al. 2021). The different cutoffs and marginal rates were varied slightly over time, adjusting in part to inflation. For example, in 1985, a tax rate of 15 percent applied to the first IDR 10 million, 25 percent to the next IDR 40 million, and 35 percent on any taxable profits in excess of IDR 50 million (Gillis 1985). By 2009, as documented in Basri et al. (2021), a rate of 10 percent applied for the first IDR 50 million in taxable income; a rate of 15 percent applied for the next IDR 50 million; and a rate of 30 percent applied on all taxable profits over IDR 100 million. After 2009, the corporate income tax system moved to a flat 25 percent rate, with a more complicated schedule of discounts based on gross income that led to effective tax rates below 25 percent (see Basri et al. (2021) for details). Tax evasion and enforcement for the corporate tax rate posed a larger problem than for the VAT, especially before 1985 but also after (Hill 2000, 51f.; Basri et al. 2021). In the end, our 20% flat tax rate assumption tries to parsimoniously capture average effective corporate income tax rates, while abstracting from important temporal and cross-sectional variation.

Estimation of borrowing constraint

To identify the borrowing constraint κ_t, note that the working capital constraint writes as:

$$\frac{w_t h_{it}}{(1-\tau_t^{VAT})y_{it}} \leq \kappa_t$$

The left-hand side of this constraint is a tax-adjusted labor share, which is directly observable. The constraint gives an inequality so that in the absence of measurement error in $w_t h_{it}$ or $(1-\tau_t^{VAT})y_{it}$:

$$\kappa_t = \max_i \left(\frac{w_t h_{it}}{(1-\tau_t^{VAT})y_{it}} \right) \quad \text{for } i \to \infty$$

as long as the constraint is strictly binding for any plant. It is easy to see that the constraint will always bind for some plants as long as there is a non-zero chance for large productivity losses and plants face some positive costs of adjusting labor. The bigger problem is that any estimator based on the maximum observed adjusted labor share will be strongly influenced by outliers and individual measurement error in the labor share. As we discuss in more detail in the data cleaning Appendix, changes in the survey questions over time is one reason why we expect systematic variation in the maximum reported labor share over time that is independent of actual changes in the borrowing constraint κ_t. We thus give up on identifying variation in κ_t over time. Instead, we opt for a robust estimator of κ, taking the 95th percentile of the observed adjusted labor share. Based on this estimator, we find that $\kappa = 1.7$.

Discretization details

We discretize productivity and labor. Specifically, we choose 30 grid points for idiosyncratic productivity, which we select based on quantiles of the productivity distribution. We choose more quantiles at the right tail of the distribution as these high productivity plants are key for the aggregate economy. Non-parametrically estimating the transition matrix of idiosyncratic productivity is quantitatively important as other oft-used processes such as an AR(1) log-normal process cannot replicate empirically observed productivity dynamics (e.g. see Ruiz-García (2019)). We discretize efficiency units of labor h_{t-1} on a grid of 1000 points that we choose based on equal spaced quantiles, ensuring that the entire labor distribution is well represented. Specifically, we choose the bottom 990 grid points based on quantiles (ensuring that all plants in the data can be mapped to the grid) and then use the last 10 grid points to extend the upper bound for labor to allow plants in the model to grow beyond what we observe in the data.

Fundamentals needed for model counterfactuals

Two key sets of model fundamentals are not needed for solving the baseline model, because they are linked to reduced-form statistics that are treated as fixed along the baseline equilibrium path: the path of potential entrant distributions and the fundamentals of the rest-of-the-economy including aggregate labor supply. For model counterfactuals, however, all fundamentals are needed, so we now discuss their identification.

The potential entrant distributions can be related to objects of the baseline equilibrium path: $PE_t(s_t, h_t; \mathbf{\Omega}) = E_t(s_t, h_t; \mathbf{\Omega}) / \mathbb{P}_E(s_t, h_t; \mathbf{\Omega})$, where $E_t(s_t, h_t; \mathbf{\Omega})$ is the identified path of entrant distributions and $\mathbb{P}_E(s_t, h_t; \mathbf{\Omega}) = P\big(V^M(s_{i,t}, h_{i,t}; \Omega_t)\big)$ gives the path of entry probability distributions. The latter is a function of the incumbent's value function, which we directly obtain from the baseline model computation, and the entry cost distribution P. Since the potential entrant distributions and the entry cost distribution are not separately identified, we make the identifying assumption that the entry cost distribution is the same as the fixed cost distribution governing plant survival.[31]

The time path of aggregate labor supply is given by 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 identified from aggregating up plant-level labor demand h_{it} over the computed equilibrium path. To obtain H_t^R, we use the total observed number of workers l_t^R in the Rest of the Economy and map this to the total efficiency units of labor in R accounting for differential worker selection across sectors in Indonesia.[32]

For the rest-of-the-economy, 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. Labor 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: $A_t = y_t^R / \left(H_t^R\right)^{\theta_R}$.

[31]Given that plant entry and survival likely depend on similar economic forces (e.g. similar outside options for not running a plant), we think this gives a reasonable estimate. We also think this gives a conservative estimate of potential entry because most plants survive, implying that most potential entrants also enter. The assumption is a form of normalizing the distribution of potential entrants and is more general than normalizing the total number of potential entrants as often done in entry models (see Aguirregabiria 2021, Chp. 5).

[32]Specifically, we use the estimates of wage differences and worker selection across rural agriculture and urban non-agriculture from Hicks et al. (2017) for Indonesia. This leads us to estimate that average efficiency units of labor are roughly two times larger in M than in R. 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 more efficiency units of labor. We enforce the point estimates of Hicks et al. (2017) across all time periods.

1.B.6 Details on Euler estimation

In this subsection, we provide more details on the Euler estimation procedure we use and derive all main results.

Derivations for Gumbel distribution

We start out by showing that the Gumbel distribution for fixed costs allows closed-form expressions for the survival probability and the conditional expectation of fixed costs. For expositional clarity, we suppress dependence on the aggregate state Ω_t, but note that all objects generally depend on the aggregate state.

$$\lambda(s_{i,t}, h_{i,t}) = exp\left(-exp\left(\frac{-(x(s_{it}, h_{it}) - \mu_t^x)}{\sigma_t^x}\right)\right) \tag{1.26}$$

$$\mathbb{E}_c[c_F|\text{stay}] \equiv \tilde{g}(s_{i,t}, h_{i,t}) = x(s_{it}, h_{it})\lambda(s_{i,t}, h_{i,t}) - \sigma_t^x \Gamma\left(0, exp\left(\frac{-(x(s_{it}, h_{it}) - \mu_t^x)}{\sigma_t^x}\right)\right) \tag{1.27}$$

where $x(s_{it}, h_{it}) \equiv \beta\mathbb{E}\big[V(s_{i,t+1}, h_{i,t})|s_{it}, h_{it}\big]$ and $\Gamma()$ gives the incomplete Gamma function. That is, in principle, $\mathbb{E}_c[c_F|\text{stay}]$ depends not only non-linearly on the parameters $\{\mu_t^x, \sigma_t^x\}$, but also depends directly on the unknown expected future value x. However, given that the continuation value x is simply an invertible function of (observable) $\lambda(s_{i,t}, h_{i,t})$, we can rewrite the term to substitute for x:

$$\begin{aligned}
\tilde{g}(s_{i,t}, h_{i,t}) =& \mu_t^x \lambda(s_{i,t}, h_{i,t}) - \sigma_t^x \Big\{ ln\left(-ln\left(\lambda(s_{i,t}, h_{i,t})\right)\right)\lambda(s_{i,t}, h_{i,t}) + \Gamma\left(0, -ln\left(\lambda(s_{i,t}, h_{i,t})\right)\right) \Big\} \\
\equiv& \mu_t^x \tilde{g}_1(s_{i,t}, h_{i,t}) - \sigma_t^x \tilde{g}_2(s_{i,t}, h_{i,t})
\end{aligned}$$

We will use these equations and the invertibility of exit rates for continuation values throughout.

Identification details

In the following, we derive the estimating Euler equation and then discuss identification. To derive the estimating Euler equation, we exploit the invertibility of exit rates as shown above and simplify terms to rewrite the Euler equation only in terms of observables and model parameters:

$$\begin{aligned}
0 =& \frac{\partial y(s_{i,t}, k_{it}, h_{i,t}, z_t)}{\partial h_{i,t}} - w_t - w_t \frac{\partial C_h(h_{i,t}, h_{i,t-1}; w_t)}{\partial h_{i,t}} + \\
& \lambda(s_{i,t}, h_{i,t})\beta\mathbb{E}\left[-w_{t+1}\frac{\partial C_h(h_{i,t+1}, h_{i,t}; w_{t+1})}{\partial h_{i,t}}|s_{i,t}, h_{i,t}\right] \Bigg\{ \\
& 1 - \lambda(s_{i,t}, h_{i,t}) + ln\big(\lambda(s_{i,t}, h_{i,t})\big)\left[\frac{\mu_t^x}{\sigma_t^x}(2\lambda(s_{i,t}, h_{i,t}) - 1) - \Gamma\left(0, -ln\left(\lambda(s_{i,t}, h_{i,t})\right)\right) - \right. \\
& \left. ln\big(-ln\big(\lambda(s_{i,t}, h_{i,t})\big)\big)(2\lambda(s_{i,t}, h_{i,t}) - 1) - \lambda(s_{i,t}, h_{i,t})\frac{\partial\Gamma\left(0, -ln\left(\lambda(s_{i,t}, h_{i,t})\right)\right)}{\partial\lambda(s_{i,t}, h_{i,t})}\right] \Bigg\}
\end{aligned}$$

Given the estimating Euler equation, we can now discuss the identification of the parameters. We discuss each of the three sets of parameters in turn.

Linear and convex adjustment cost parameters: Without giving a full identification proof, one can see that the Euler equation generally identifies marginal adjustment costs $\frac{\partial C_h(h_{i,t},h_{i,t-1};w_t)}{\partial h_{i,t}}$ non-parametrically. Given our functional form assumption on adjustment costs, linear costs c_0 & convex costs c_1 are identified as follows: c_0 adds as a fixed wedge between the marginal product and the marginal costs of labor, but any adjustments today save on adjustments tomorrow. Thus, linear costs are pinned down by the observed labor wedge across plants and the probability of switching between shrinking and growing as determined by the volatility of the productivity process. Asymmetric linear costs are identified from the differential behavior of growing and shrinking plants. The convex costs c_1 instead scale with labor growth and are thus identified from the variation in within-plant labor demand growth across periods, again conditioned by the observed volatility of the productivity process. Low labor demand growth despite a high labor wedge will point to strong convex adjustment costs. Again, asymmetry here is identified from differential growth and shrinking (conditional on the productivity process and the state).

Fixed adjustment costs: The Euler equation does not identify fixed costs F^+ & F^- since they do not enter marginal adjustment costs. However, we note that fixed costs are identified from the (time-varying) distribution of plants that are not adjusting and for whom the Euler equation does not hold. The idea is that the more plants choose to not change their labor inputs (as we condition on previous labor and vary productivity), the higher the implied fixed costs. In the data, driven by the choice of focussing on efficiency units of labor, we do not see any plant that remains strictly inactive. We can thus not rule out that fixed costs are zero and fix them to zero throughout. We also note that in a previous version of the paper, we estimated fixed costs indirectly by solving for the model equilibrium path and also found them to be noisily estimated around zero. In cases where one might be particularly interested in fixed costs of adjustment that induce inaction – such as the sluggish responses to aggregate shocks – we think it is best to either work directly with the number of workers or at least work with the nominal wage bill.

Cost parameters (exit): One can immediately see that the Euler equation only identifies the ratio $\frac{\mu_t^x}{\sigma_t^x}$. The reason is that the Euler equation captures the marginal effect of changes in labor demand on the survival probability, which only depends on the ratio of the level and dispersion of costs. What variation in the data identifies this cost ratio? While the dependence in the Euler Equation looks daunting, the cost ratio is jointly disciplined by the size of the labor wedge, the dispersion in survival probabilities and the size of marginal adjustment costs next period over current labor demand. Given empirically estimated survival probabilities, one can see that a higher cost ratio generally increases the labor wedge.[33] Thus, high observed labor wedges push towards lower cost ratios.

Estimation details

The estimation proceeds in two stages. In the first stage, we estimate reduced-form survival probabilities $\lambda(s_{i,t}, h_{i,t}, \Omega_t)$ and dynamic labor input choices $h(h_{it-1}, s_{it}, \Omega_t)$ conditional on the state space. In the second stage, we enforce these reduced-form objects to estimate the Euler

[33]This is as long as marginal adjustment costs tomorrow are negative, since $\lambda(s_{i,t}, h_{i,t})ln(\lambda(s_{i,t}, h_{i,t}))\,(2\lambda(s_{i,t}, h_{i,t}) - 1)$ is generally negative (since survival probabilities are generally higher than 0.5 and $ln(\lambda(s_{i,t}, h_{i,t})) < 0$).

equation for the structural parameters of interest.

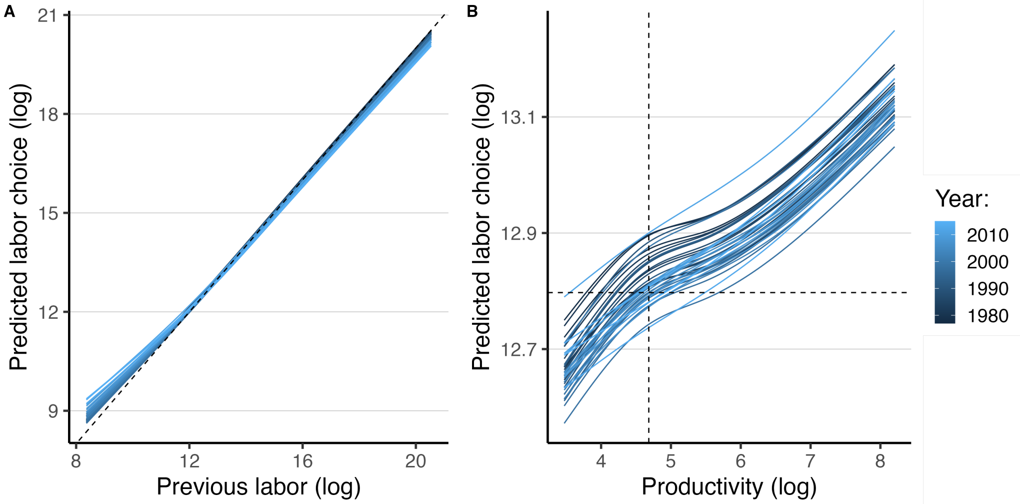
We start with the estimation of conditional survival probabilities and labor input choices. To flexibly estimate both, we draw on generalized additive models for their combination of flexibility and robustness in estimating semi-parametric functional forms. However, subsequent parameter estimates are very similar when using flexible polynomial regressions in the first stage instead. We start with survival probabilities, which – through the lens of the model – are a nonlinear, time-varying function in current labor and productivity. They are time-varying because exit decisions depend on the aggregate state space through current wages and aggregate productivity as well as through perfect foresight over future wages and aggregate productivity. We estimate survival probabilities as the combination of year fixed effects and semi-parametric functions in labor, productivity and their interaction. In general, we find that the estimated survival probabilities make sense while improving on a simple linear model in labor, productivity and year fixed effects. Specifically, the GAM achieves an adjusted R^2 that is roughly 13% higher than the linear model and produces survival probabilities that are increasing in productivity.

For labor input choices, we instead estimate the GAM as the combination of year fixed effects and semi-parametric functions in previous labor, productivity and their interaction.[34] The adjusted R^2 of our GAM is around 96%, driven by the huge explanatory power that previous labor has for current labor. In fact, a simple linear regression of previous labor on current labor already reaches $R^2 = 0.95$ with a coefficient of auto-correlation close to unity. Hence, a policy function that simply says to stick with past labor already explains observed labor choices extremely well. Through the lens of the model and the Euler equation, in the presence of sufficient productivity variation, this already points to large adjustment costs. However, to identify adjustment cost parameters, we require that the estimated policy functions also show variation across productivity, since changes in productivity conditional on previous labor vary the returns to adjusting labor. Figure B.1 plots the policy functions implied by our estimated GAM. Panel A varies previous labor and fixes productivity at the median, while Panel B varies productivity but fixes previous labor at the median. We find that policy functions are monotonic in productivity (in line with the model), but nonlinear such that labor is declining more strongly for low productivities and increasing more strongly for high productivities.

In the second stage, we enforce these reduced-form objects to estimate the Euler equation for the structural parameters of interest. For this stage, we follow the CCP literature in imposing our model structure. This means that we impose the same discretized state space as in our model and the same process of idiosyncratic productivity to ensure that the Euler estimation gives parameter estimates that are fully consistent with our model. We also impose CCC-estimated input choices for plants at time t and for all states in $t+1$ conditional on the plant's state in t, greatly reducing the noise in the estimation: this means that for a plant with labor h_{it-1} and s_{it}, we do not use observed h_{it} and h_{it+1}, but instead the predicted values $\hat{h}_{it}(s_{it}, h_{it-1}, \Omega_t)$ and $\hat{h}_{it+1}(s_{it+1}, \hat{h}_{it}, \Omega_{t+1})$, where the latter is over all possible s to be able to

[34] We also looked at more flexibility in how policy functions could vary in labor and productivity over time (beyond a simple year fixed effect), however, we found this to give less robust results. The reason is that such higher flexibility is more likely to predict more erratic differences in input choices over consecutive years, which "violate" the smoothing motive implied by the Euler equation and which can then only be rationalized by extreme adjustment cost parameters. We believe that if the underlying data has less measurement issues (that are driven by common aggregate components) and is maybe taken from an economic environment where there are less changes over time, it is more feasible to allow for more flexibility in how policy functions can change over time.

Figure B.1: Step 1 of CCC estimation: flexibly estimated labor input choices



Notes: Panel A gives labor choices over previous labor (fixing productivity at the unconditional sample median). Panel B gives labor choices over productivity (fixing previous labor at the unconditional sample median). Dotted lines give unconditional sample medians.

compute the expectation term.[35]

We think it is important to briefly mention the computational gains here. An important step in the estimation is that we follow Bajari, Benkard, and Levin (2007) in exploiting the fact that adjustment cost parameters enter linearly in the problem, which is due to the linearity of the marginal adjustment cost specification and due to plants' risk neutrality. Practically, this means that we can compute all expectation terms outside the parameter loop, greatly speeding up the parameter estimation. It is important to highlight the computational gains from this step alone: We can estimate parameters from both steps in less than 2 minutes on a standard personal computer and moving to annually estimated parameters (increasing the number of parameters by a factor of 40) can sometimes even be faster – the reason is that the pre-computed terms all stay the same and instead of estimating X parameters on N data points, we now separately estimate X parameters on N/T data points T times, which one can even parallelize.

We estimate structural parameters via nonlinear least squares (NLS). We do so by assuming that the Euler equation can be written as: $f(\Theta) + \eta_{it} = 0$, where η_{it} is model misspecification error or additive measurement error and Θ is the vector of parameters. Table B.1 reports estimated results. Table B.2 separately estimates parameters for the period before the Asian Financial Crisis and for the period after. One can see that estimated adjustment cost parameters are considerably higher post-1999 than before. At last, we also estimate adjustment costs at an annual level. Figure B.2 shows the yearly estimated convex adjustment costs and shows that they tend to increase over time.

[35] An alternative to our approach would be to directly use observed plant-level future labor adjustments as noisy realizations of expected labor adjustments, without enforcing model-based expectations (e.g. Hall 1979). Our approach is closer to our model, greatly reducing noise in the estimation, which is particularly problematic for the estimation of convex costs that disproportionately react to outliers. However, this makes our estimation approach – as other CCC/CCP estimators – more susceptible to model misspecification.

Table B.1: Main Euler estimation results

Parameters	Estimates	Std error	95% CI
c_0^+	0.735	0.010	[0.715,0.755]
c_1^+	36.656	0.059	[36.54,36.772]
c_0^-	0.000	0.011	[-0.022,0.022]
c_1^-	12.593	0.073	[12.45,12.736]
Cost ratio	-0.366	0.004	[-0.374,-0.358]

<u>*Details:*</u> Pooled across all consecutive plant-year observation pairs (N = 358,240). Adjustment cost parameters are restricted to be (weakly) positive and the cost ratio is bounded between -0.577 and -0.366 to ensure that median and mean costs are sufficiently far apart and rationalize dispersion in exit probabilities. Inference for corner solutions should be treated with care. Standard errors are not yet corrected for the multi-stage estimation.

Table B.2: Euler estimation results: Pre vs. Post Crisis

Parameters	Estimates		Std error		95% CI	
	Pre-1997	Post-1999	Pre-1997	Post-1999	Pre-1997	Post-1999
c_0^+	0.635	0.731	0.008	0.013	[0.619,0.651]	[0.706,0.756]
c_1^+	28.590	38.495	0.098	0.069	[28.398,28.782]	[38.356,38.63]
c_0^-	0.000	0.000	0.006	0.013	[-0.012,0.012]	[-0.025,0.025]
c_1^-	7.206	15.605	0.167	0.088	[6.879,7.533]	[15.433,15.777]
Cost ratio	-0.366	-0.366	NaN	0.004	[NaN,NaN]	[-0.374,-0.358]

<u>*Details:*</u> Pooled across all consecutive plant-year observation pairs (N = 136,918 for pre, N = 209,075 for post). Adjustment cost parameters are restricted to be (weakly) positive and the cost ratio is bounded between -0.577 and -0.366 to ensure that median and mean costs are sufficiently far apart and rationalize dispersion in exit probabilities. Inference for corner solutions should be treated with care. Standard errors are not yet corrected for the multi-stage estimation.

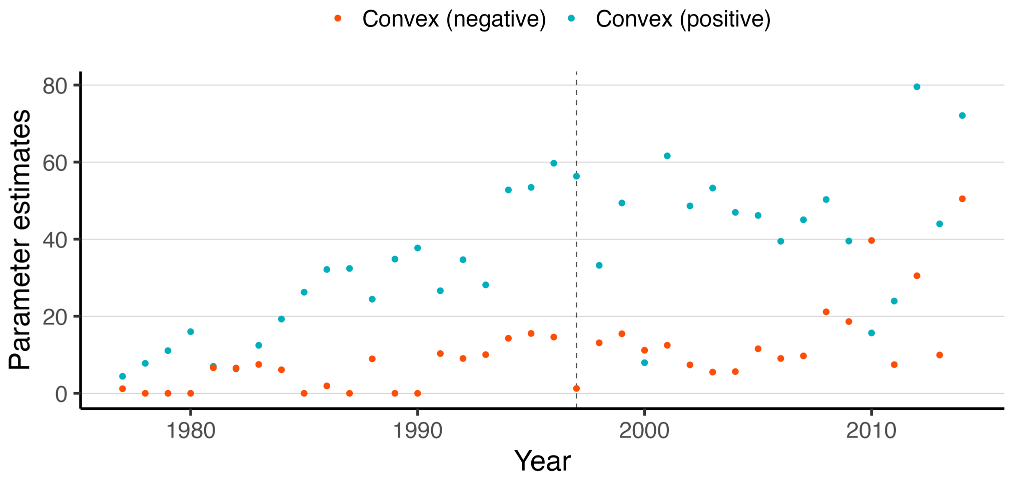


Figure B.2: Annually estimated convex adjustment costs. Details in text.

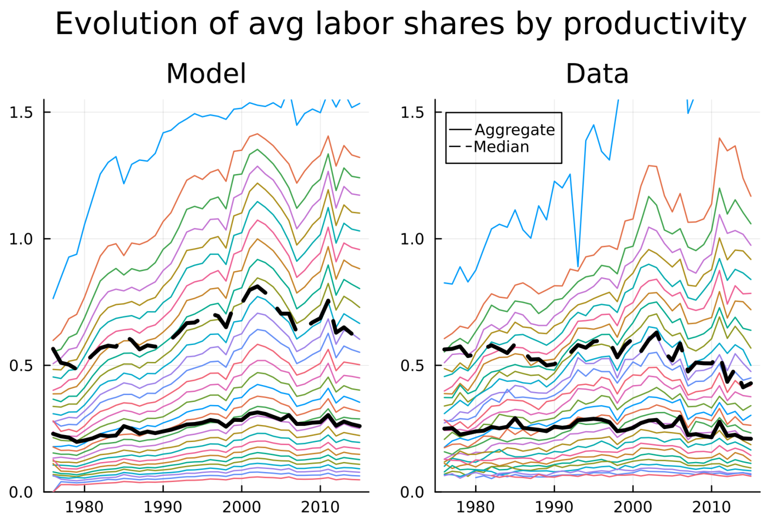


Figure C.1: Changes in the entire distribution: baseline model versus data.

1.B.7 Further model validation exercises

In this section, we show further model validation results. Specifically, Figure C.1 shows the distribution of labor shares. A key feature of the data, which the model captures, is that while average labor shares increase when holding productivity constant, the large observed shifts in productivity due to selection and productivity convergence imply that increasingly more production is concentrated in more productive plants. These productive plants, however, have substantially lower labor shares, in part because they were surprised by positive productivity shocks and adjust labor only slowly, and in part because they avoid large labor increases anticipating future mean reversion in productivity. Together, this implies that the aggregate labor share is low and remains low over time, with the median labor share even declining in the data. Apart from the decline in the median labor share, the model correctly captures these distributional changes.

1.C Counterfactuals and results

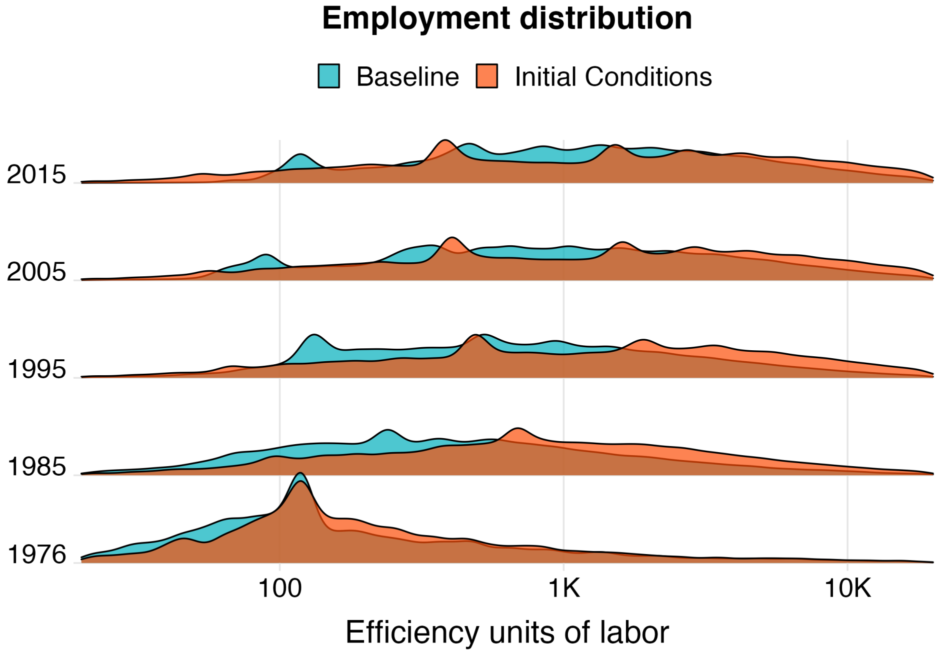
1.C.1 Further results on decomposing the drivers of growth

Figure C.2 reports the evolution of the employment distribution in the baseline (model) economy versus the counterfactual (model) economy where only initial conditions in 1976 play out over time (and all other fundamentals in 1976 are fixed).

1.C.2 Details on INPRES evaluation

This section provides further details on the model-based evaluation of the INPRES school construction program. We start out with more details on how we interpret the program through the lens of the model, how the program's effects map into changes in model fundamentals and how we identify counterfactual fundamentals of the economy had the INPRES school construction program not been implemented. In the second part, we then provide more detailed results on the model-based evaluation of the INPRES program.

Figure C.2: Evolution of employment distribution: baseline model versus initial conditions counterfactual



Notes: Details in the text.

We assume that the program's direct effect only goes through improving children's education as measured by human capital h in the model. We can be agnostic about how schools raised human capital, capturing a combination of changes along the extensive margin (some children are induced into going to school in the first place), intensive margin (some children stay longer in school) and quality margin (more schools and teachers meant smaller classroom sizes and closer proximity that may provide additional time for learning). We assume that overall demographic changes are not affected. The direct effects of changes in human capital then induce a number of endogenous changes in the model. Specifically, the increase in human capital puts downward pressure on wages and drives up labor demand both in the rest of the economy and across all manufacturing firms. Increases in human capital also have an effect on the endogenous entry and exit of firms, but we assume that this only happens through input costs. Specifically, we rule out that increases in education may have a direct effect on entrepreneurial choices and the distribution of potential entrants. Given that the policy only had measurable effects on primary and secondary schooling outcomes and that entrepreneurs in larger manufacturing firms are more likely to have tertiary education, we think this is a reasonable assumption.

Formally, we model the INPRES program as changing individual-level human capital, which aggregates up to aggregate human capital H_t over time. While the model allows for individuals being differentially affected by the school construction program and also differentially select into different sectors, the general equilibrium results only depend on the change in aggregate human capital. Denoting by L_t the evolution of the number of workers (which we assume to be unaffected by the program), we need to know the average effect of the program on human capital per worker. Through the lens of the model, the average treatment effects estimated via differences-in-differences exactly capture the average differences in human capital

h induced by the INPRES program for workers who were treated by the program, netting out aggregate changes in the wage. We can rewrite the effect of the overall program on wages as the combination of three separate effects that have been estimated in the literature: the effect of the program on school construction, school construction on years of education and years of education on wages. We further simplify the setup by assuming that we can treat the three terms as separate expectations (which is true under homogeneous treatment effects or when the shocks driving variation in the treatment effects are independent):

$$\begin{aligned}\mathbb{E}\left[\frac{\partial \text{wage}}{\partial \text{program}}\right] &= \mathbb{E}\left[\frac{\partial \text{wage}}{\partial \text{years of schooling}}\frac{\partial \text{years of schooling}}{\partial \text{no. of schools}}\frac{\partial \text{no. of schools}}{\partial \text{program}}\right] \\ &= \underbrace{\mathbb{E}\left[\frac{\partial \text{wage}}{\partial \text{years of schooling}}\right]}_{\approx 0.1 \text{ (Chaisemartin \& d'Haultfoeille '18)}} \underbrace{\mathbb{E}\left[\frac{\partial \text{years of schooling}}{\partial \text{no. of schools}}\right]}_{\approx 0.25 \text{ (Akresh et al '21)}} \underbrace{\mathbb{E}\left[\frac{\partial \text{no. of schools}}{\partial \text{program}}\right]}_{1.98 \text{ (per 1k children; direct measure)}}\end{aligned}$$

Following Akresh, Halim, and Kleemans (2023), we assume from this that the program on average increased years of schooling by half a year for individuals of any treated cohort. We further follow Akresh, Halim, and Kleemans (2023) by assuming that individuals join the workforce at age 18 and that differences in human capital induced by the program are constant over a person's life, in line with one-time educational gains. The primary schools built by the INPRES program between 1973-1979 are for children between the ages of 7-12 years, such that children fully treated by the INPRES program first joined the labor force by 1984. As in the existing literature, we assume that all cohorts born after 1968 benefit from the INPRES program. This assumes that the last cohort that we observe in 2015 still benefited from INPRES schools in 2009 (their last year of primary school).[36] To avoid having to deal with partial treatment, we further assume that cohorts before 1968 did not benefit from the program. Through the lens of our model, the INPRES program thus led to variation in aggregate human capital over the period 1986 to 2015.

We use the following steps to construct counterfactual paths of aggregate human capital in the absence of the INPRES program:

1. Start from aggregate human capital H_t given by model
2. For each year t between 1975 and 2015:
   - count share of working population affected by INPRES treatment (ϕ_t^T) & get aggregate human capital without INPRES: $\tilde{H}_t = H_t * (1 - \phi_t^T) + H_t * \phi_t^T * \frac{1}{\tilde{\beta}}$ where $\tilde{\beta}$ is the corresponding average treatment effect of the program (here: assume that this is 1.05 given as above)

We thus implicitly assume that all cohorts have the same average human capital. This is unlikely, but in the absence of better worker-level estimates of human capital, this is the best we can do. Still, we think that there are two biases that push in opposite directions so that we think the overall bias may not be too strong. First, younger cohorts likely have more human capital, which means we overestimate human capital in the absence of the INPRES program $\tilde{H}_t$. At the same time, individuals likely experience human capital increases over their life cy-

[36]The program initially planned for the INPRES-built schools to last for 20 years, however, Akresh, Halim, and Kleemans (2023) note that most even exist 40 years later. Our assumption implies that the maximum age for an INPRES school in our data is 36 years, well in line with the age range of INPRES schools.

cle such that young cohorts have less experience and less human capital, biasing our results in the opposite direction.

To construct the path of ϕ_t^T we draw on representative and harmonized population census data that we retrieve via IPUMS. For each available census wave $t \in \{1980, 1985, 1990, 1995, 2000, 2005, 2010\}$, we construct the share of working age individuals (between 18-65) who have been born in 1968 or later, which gives ϕ_t^T. For the years in between, we extrapolate from the previous census wave assuming there is no differential mortality risk. The treated share in the working population is zero before 1986 and then - due to baby-boom cohorts - increases rapidly to almost 20% by 1990, 50% by 2000 and 75% by 2015.

Applying this time path, we find that the INPRES program raised the annual economy-wide level of human capital by 3.6% by 2015. While important, this effect only accounts for less than 2% of the more than doubling of human capital per worker that we estimated over the entire time period from 1975-2015. These numbers also explain the model-based aggregate effects of the INPRES program that we find.

1.C.3 Details on FDI policy counterfactual

In this section, we provide further details on the model-based evaluation of attracting more foreign-owned plant entrants in Indonesian manufacturing. We start out with more details on the entry of foreign-owned plants, how we interpret this variation and map it to the model, how regulatory changes on FDI map to changes in model fundamentals and how we identify counterfactual fundamentals of the economy had Indonesia's FDI policy been different. In the second part, we then provide more detailed results on the model-based evaluation.

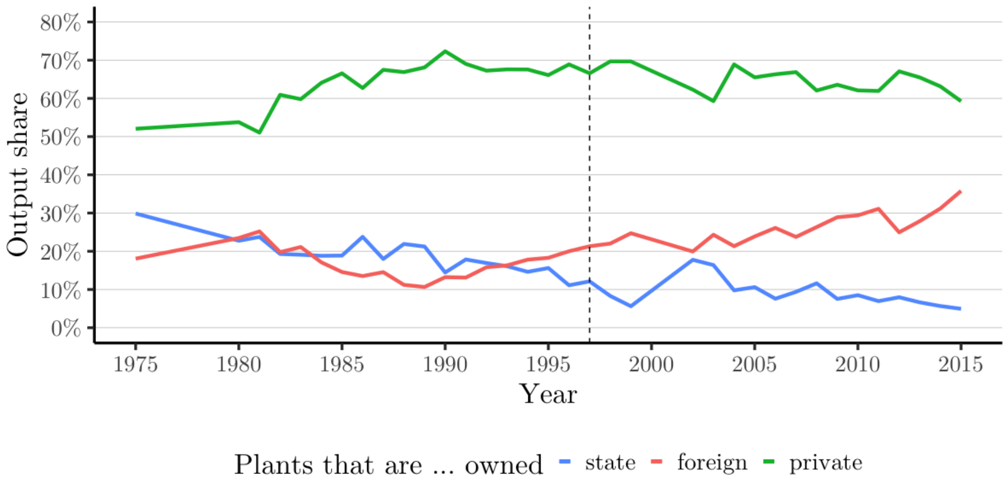
We start out by showing variation in ownership across Indonesian manufacturing plants over time. In Figure C.3 we plot the share of total manufacturing output that is owned either by the state (central + local govt), private domestic or foreign owners. We construct this by summing up all value-added output across plants but taking plants' reported ownership shares as weights. The main movement can be observed in the decline of state ownership from around 30% in 1975 to around 5% in 2015 and the rise of foreign ownership. The domestic private sector is by far the largest actor and owns between 60-70% of all manufacturing production. If we were instead to look at the share of plants, we find that more than 90% are fully domestically owned, which is stable over time. Again, we find that state ownership declines over time and foreign ownership increases, making up almost the entire remainder of 10% by 2015.

For the role of FDI policy, we are specifically interested in the effect on plant entry. Plant entry is particularly important, because most variation in foreign ownership shares is across and not within plants as plant-level ownership shares are relatively constant.[37] Figure C.4 thus reports evidence on the importance of foreign ownership among new entering plants.[38] Within a given year, foreign-owned plants (those that are majority foreign owned) make up

[37]For example, the variation in foreign ownership explained by plant fixed effects is 78% and the persistence in foreign ownership as measured by an AR(1) regression is $\rho \approx 0.9$. Restricting only to plants that were ever foreign owned gives slightly lower numbers with the $R^2 \approx 0.6$ and $\rho \approx 0.77$.

[38]We define new entering plants as plants that enter the panel for the first time. While foreign-owned plants are larger and unlikely to not make the cutoff of 20 workers, we want to make sure to compare foreign- and non-foreign-owned plants correctly. To this end, we further impose the restriction that the plant has to be younger than 10 years (which is the spacing of the censuses). The sample of entrants without the age restriction looks very similar.

Figure C.3: Evolution of ownership shares for Indonesian manufacturing plants

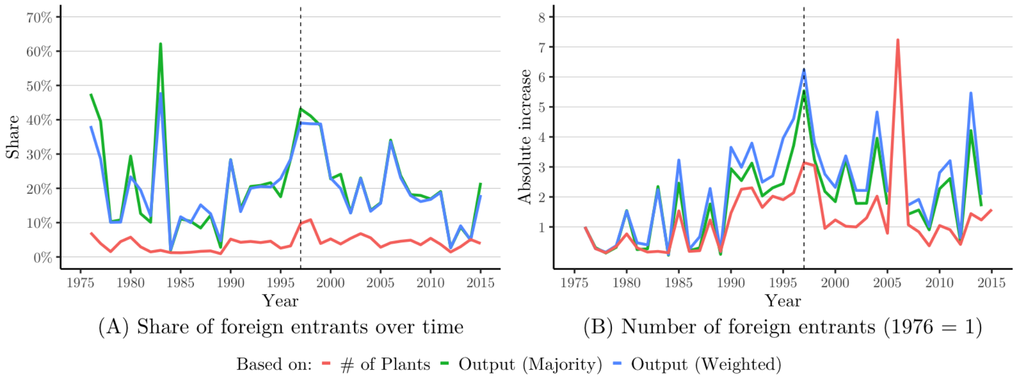


Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. Figure reports the fraction of total manufacturing output that is owned by the state (either local or central government), domestic private owners or foreign owners measured by summing up all value-added output of plants and weighting plants' output by their respective reported ownership shares.

around 3.8% of entering plants, but they account for 18-19% of total output among entrants. Entrants with some foreign ownership are almost always close to fully foreign-owned with average ownership shares around 80% and the median at 95%. Figure C.4 also documents important variation across time - variation that we exploit for identifying the effect of FDI policy. Specifically, the Indonesian FDI regulatory regime turned increasingly restrictive throughout the 1970s, forbidding 100% foreign ownership and banning FDI entirely in some sectors of the economy (see: Hill 2000). This policy regime reverted only in the second half of the 1980s with simplifications and more transparency over existing restrictions introduced in 1987. In 1992, 100% foreign ownership was permitted again and the 1990s saw increasing attempts at luring foreign manufacturing plants. We can see some of these changes in observed entry, including a marked increase in the absolute and relative weight of foreign entrants between 1990 and the Asian Financial Crisis in 1997. We also see a correlation of these policy changes with the aggregate ownership series, with the importance of foreign ownership increasing steadily from 10% in the late 1980s up to 35% by 2015.

We exploit this policy variation to study the influence of changes in foreign direct investment policy in the Indonesian growth experience. We proceed as follows: We first identify the distribution of actual foreign entrants, which - through the lens of the model - is a reduced form object that masks the underlying distribution of potential foreign entrants. To take out the variation in foreign entry that is purely explained by changes in the economic conditions that make entry more or less attractive, we proceed as before and use the model-identified, time-varying entry probabilities to invert for potential foreign entry distributions. In the next step, we are interested in whether FDI policy changes can account for changes in these potential foreign entry distributions over time. To do so, we compare the period of "restrictive FDI" from 1975-1986 with the period of "FDI promotion" from 1987-1997. Given the known data limitations of annual variation in entry (e.g. the bunching of entry around census years), we aggregate the potential entry distributions within each of the two time periods. Luckily, the census waves do not fall in a year between, so that there is no ambiguity in how to attribute entrants. We

Figure C.4: Evolution of the entry of plants with foreign ownership



(A) Share of foreign entrants over time

(B) Number of foreign entrants (1976 = 1)

Notes: Data based on Indonesian manufacturing plant census (1975-2015) with 20+ workers. In both Panels, the output measures aggregate up value-added across plants, with Majority summing up output for entering plants who are at least 50 percent foreign-owned, and Weighted summing up output based on the respective ownership share. Number of plants instead constructs the share and absolute increase based on the total number of majority-foreign-owned plants.

then compare changes in these two aggregated potential entrant distributions. For simplicity, we measure the effect of changes in the FDI policy on changes in potential foreign entrants by comparing the weighted mass of potential entrants across the two periods, taking as weights either the plants' value added or employment at entry.[39]

1.C.4 Making a (more impressive) Growth Miracle

In this part of the Appendix, we move away from Indonesia's historical growth experience and ask whether and how Indonesia could have experienced a more impressive manufacturing growth miracle, closer in comparison to the experiences of countries such as China or Malaysia. For this, we study two important policy levers that both have sizable growth effects, but play out differently over time. Specifically, we look at reduced-form policy changes that either reduce (convex) labor adjustment frictions or increase the annual growth in aggregate technology in manufacturing, but deliver the same long-run growth in manufacturing output.[40]

Figure C.5 shows how the manufacturing miracle would have played out differently in the two alternative scenarios that both see a doubling of manufacturing output by 2015 compared to the baseline miracle economy. As expected, lower adjustment costs lead to faster hiring

[39]We thus only use the time series variation and not a differences-in-differences identification design. The model-based trend correction should ensure the validity of the approach and we do not think that it is credible to compare foreign entrants to domestic entrants in a differences-in-differences design, because changes in domestic entrants (e.g. due to changes in demographics) may likely show very different trends. An alternative with a DiD design would be to compare potential entrant distributions across different countries or across different industries that were differentially treated by the regulatory regime. This is an interesting approach that we leave for future work.

[40]To reduce labor adjustment frictions, we consider a policy package that reduces both linear and convex adjustment costs. For the linear hiring and firing costs, we consider a hiring subsidy for each new hire of around 25% of the annual wage bill. Changes in convex costs are harder to map directly to tangible policies. Given our microfoundation in terms of scarce managerial time and talent, we think of them as policies that improve managerial quality in the economy such as training programs. We consider a feasible policy mix that halves the estimated convex cost parameters (for hiring and firing) – in line with the lower end of annually estimated adjustment cost parameters that we find in the data. For aggregate technology in manufacturing, z_t, we consider a policy that raises its annual growth by a constant rate to the degree that manufacturing output in 2015 is the same as in the adjustment cost counterfactual.

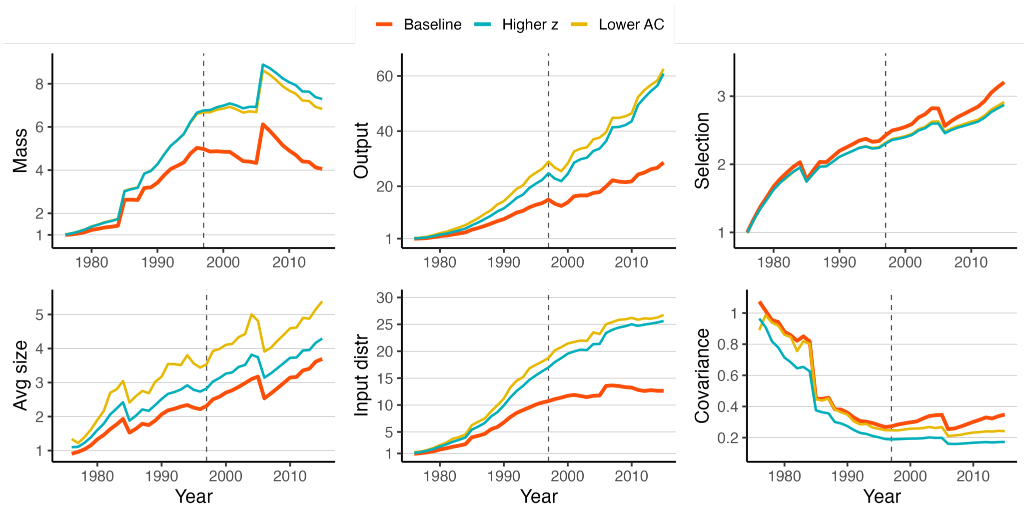


Figure C.5: Main miracle economy counterfactuals.

and thus faster transitions such that output growth is initially higher. With lower adjustment costs, far more large manufacturing plants emerge, driving up the average plant size. Growth in manufacturing technology, on the other hand, makes all plants more productive, leading to more entry and less exit of small plants, a stronger left tail and a much lower covariance of idiosyncratic productivity and resources.

Chapter 2

The Aggregate Costs of Political Connections

Jonas Gathen[1]

Abstract

This paper quantifies the aggregate costs of political connections using a general equilibrium model in which politically connected firms benefit from output subsidies and endogenously spend resources on rent-seeking activities. The model is structurally estimated using rich firm-level data for the Indonesian manufacturing sector and a firm-level measure of political connectedness based on a natural experiment from the authoritarian rule of Suharto at the end of the 1990s. A major innovation is to non-parametrically identify the output subsidy from differences in distributions of revenue-based total factor productivity (TFP) across connected and non-connected firms. In general equilibrium, both the distribution and the level of subsidies to connected firms matter. I find that subsidies to connected firms are too high and dispersed, costing the economy between 1.0-4.7% of aggregate output. At most, 45% of these output costs are due to the misallocation of factors of production towards connected firms. The large remainder is explained by the costs of subsidizing connected firms instead of putting saved subsidies to more productive use.

[1]Toulouse School of Economics; j.gathen@posteo.de. I am indebted to my advisors Fabrice Collard, Christian Hellwig and Stéphane Straub for their support throughout this project. I further thank, in no particular order, Ben Moll, Beatriz Simon Yarza, Felipe Brugués, Tim Lee, Nicolas Werquin, Mathias Reynaert, Matthias Meier, Patrick Fève, Ferenc Szucs, Mounu Prem and all participants at the TSE Macro workshop, the EEA/ESEM conference 2021 and the Mannheim Internal Seminar for helpful comments. I want to thank Bob Rijkers for sharing data, feedback and helpful information on using the measure of political connections for Indonesia used in this paper and Mushfiq Mobarak for kindly sharing their measure of political connections. I want to thank the World Bank Jakarta team, especially Massimiliano Calí, for gratefully sharing the manufacturing firm-level data for Indonesia including much-improved capital series and providing helpful feedback and suggestions on the project. This paper has been funded by the Agence Nationale de la Recherche under grant ANR-17-EURE-0010 (Investissements d'Avenir program). I am also especially grateful to Alexandre Gaillard and Oscar Fentanes for many discussions. All remaining errors are my own.

2.1 Introduction

What are the economy-wide costs of a few corrupt elites? There is strong anecdotal and quantitative evidence that autocrats and their inner circles obtain special economic privileges for their businesses to amass large fortunes. For example, wealth in excess of one-quarter of GDP was attributed to Putin's inner circle in Russia (Aslund 2019) and Tunisia's former dictator Ben Ali (Rijkers, Freund, and Nucifora 2017). This accumulation of wealth in the hands of a few politically connected elites and their businesses comes, among others, from corruption, unfair competition and systematic property rights violations and therefore is the sign of larger distortions that matter for the aggregate economy.

This paper systematically quantifies the costs that a few connected firms can pose for the entire economy. A motivating example makes the costs of political connections that this paper quantifies more explicit. In 1996, the Indonesian government decided to promote its national car industry by offering a generous combination of various tax and tariff exemptions to selected firms. Seemingly by coincidence, one day before the policy announcement, Suharto's son created a local car manufacturing company that ended up becoming the sole beneficiary of the government tax exemptions. These tax exemptions were awarded despite the company not operating a single car assembly line. Eventually, another presidential decree by Suharto allowed his son's company to import cars instead and sell these at an effective tax rate that was about 90% lower than that faced by competitors (for details, see Hale 2001). Additionally, the government further supported the company by directly buying its cars. This example illustrates two main economy-wide costs of political connections. First are *misallocation* costs: direct and indirect subsidies led the connected car manufacturer to increase its operations and demand more inputs, pushing up input prices and crowding out productive capital and labor from other firms in the economy. These *misallocation* costs depend crucially on (1) how the government selects connected firms, (2) the extent of the subsidies and (3) whether the subsidies alleviate other distortions in the economy. The second main costs of political connections are *opportunity costs of public funds*: direct and indirect subsidies to connected firms are costly because these resources could be spent on more efficient development objectives.

In Indonesia, only 1% of firms are connected, but they are disproportionately large, making up 15% of total (value-added) revenue. The average connected firm is around twelve times larger than the average non-connected firm, which also holds within narrowly defined industries. I show this by drawing on detailed annual firm-level manufacturing census data and previous micro-empirical work by Mobarak and Purbasari (2006), who identify connected firms in Indonesia under the authoritarian rule of Suharto at the end of the 1990s using a natural experiment.[2] A key question to quantify economy-wide distortions from political connections is how much of this size difference is due to political connections and how much is due to other firm fundamentals that we may simply call *productivity*. I use a structural model to disentangle the role of selection from the benefits of political connections and quantify the costs of favors to connected firms. In the model, firms flexibly spend resources on rent-seeking activities to obtain an output subsidy that can be seen as a reduced-form net transfer from the govern-

[2]The natural experiment follows Fisman (2001) and identifies all stock-listed firms that benefit from connections by looking at stock-price fluctuations in response to plausibly exogenous shocks to the health of dictator Suharto. Mobarak and Purbasari (2006) then find the remaining connected firms by exploiting a highly concentrated ownership network and link all connected firms to the micro-data.

ment.[3] The major methodological innovation of this paper is to non-parametrically identify the unobserved output subsidy from differences in distributions of revenue-based total factor productivity (TFP) across connected and non-connected firms. I do so using a type of quantile treatment effect (QTE) estimator with endogenous selection into treatment. Finally, combining estimated subsidies with a full general equilibrium model allows to quantify the aggregate costs of political connections.

The identification of subsidies is difficult because they do not just enter as "wedges" that distort model-based first-order conditions as usually studied in the misallocation literature (Hsieh and Klenow 2009), but they also directly distort observed revenue. A key contribution of this paper is to show how to identify subsidies non-parametrically, that is, without making a functional form assumption on how rent-seeking activities buy government benefits. Identification only draws on a revenue-based measure of total factor productivity (TFP) for connected and non-connected firms. This measure of TFP captures a combination of subsidies and actual productivity. To separately identify them, I crucially rely on two main assumptions. The first is a monotonicity restriction that ensures that firms with the highest measured TFP also have the highest productivity. This restriction does not mean that subsidies need to increase in productivity, only that subsidies cannot decline too fast with firm productivity. The assumption can also be indirectly tested. The second main assumption is on the selection of politically connected firms, and is more restrictive. Specifically, I parameterize how connected firms are selected, making estimated subsidies dependent on the degree of selection.[4] The benefit of this assumption is that it gives intuitive (and estimable) bounds on subsidies, spanning from the case of no selection - connected firms being a representative sample of all firms - to the case of maximal positive selection where most TFP is actual productivity.[5]

The structural estimation is based on a matching procedure that exploits observed non-connected firms that do not receive subsidies. The idea is to selectively sample firms from the population of non-connected firms and use the monotonicity restriction to order and then match them to the observed sample of connected firms to back out their productivities. Estimated subsidies reveal a high degree of selection, especially at the bottom of the productivity distribution. Based on both bounds, the least productive politically connected firm is still more productive than 40% of non-connected firms. However, despite connected firms being selected, estimated subsidies are sizable. For the average connected firm, the government subsidizes at least 40% of output or, equivalently, pays a price markup of at least 65%.

With the estimated subsidies in hand, I show that a structural model of endogenous rent-seeking can almost perfectly explain them quantitatively. The structural model is needed to

[3]This subsidy captures many of the channels through which political connections matter, such as lower taxes due to tax avoidance and evasion (Johnson and Mitton 2003; Do, Nguyen, and Tran 2017), output and input subsidies, preferential access to government contracts, state-owned land and natural resources (Brugu'es, Brugu'es, and Giambra 2018; Chen and Kung 2018; Schoenherr 2019; Straub 2014; Szucs 2017) as well as preferential access to institutions and infrastructure (Fisman and Wang 2015). While the identification of benefits allows for any combination of these factors, subsequent welfare estimates rely on the government paying for the benefits and them entering through revenue, as is the case for tax evasion, government subsidies and government demand.

[4]This differs from standard approaches in the quantile treatment effect (QTE) literature who either feature the standard QTE without selection into treatment (Doksum 1974; Lehmann and D'Abrera 1975), selection based on observables (Firpo 2007) or instrumental variable approaches with general selection on unobservables (Chernozhukov and Hansen 2005).

[5]There is good evidence that connected firms are negatively selected within subsets of large firms, such as listed firms or firms eligible for government contracts (e.g. Gonzalez and Prem 2019; Schoenherr 2019; Szucs 2017). My results are in line with this since, conditional on firm size, politically connected firms are less productive.

infer unobserved rent-seeking activities, which affect input prices in general equilibrium and hence the aggregate costs of political connections. In the model, rent-seeking activities of a firm are organized within a department that is in charge of lobbying, tax evasion, legal affairs and bribery. Connected firms then endogenously choose the size of this rent-seeking department. The estimated model shows decreasing returns to rent-seeking activity and convex costs that increase both in rent-seeking activity and firm size. Model-implied subsidies explain more than 95% of the variation in estimated subsidies for both bounds. The economic intuition is that firms optimally trade-off investing in rent-seeking activities that they use to buy subsidies with trying to stay below the radar of opposing interest groups and public scrutiny. In the data, connected firms with intermediate levels of productivity receive the largest subsidies. Through the lens of the model, these firms are at a sweet spot where they are productive enough at rent-seeking while being small enough to receive little public scrutiny. Based on the estimates, connected firms differ widely in how much they spend on rent-seeking activities. The largest connected firms spend less than 2% of their input costs on rent-seeking, while smaller and less productive firms benefit from receiving less attention and spend up to 30% of their input costs on rent-seeking. As a validation of these estimates of unobserved rent-seeking activities, I show that they align with recent quantitative evidence on high-level rent-seeking activities, as evidenced in the Odebrecht corruption scandal (see Campos et al. 2021).

The last step to quantify the aggregate costs of political connections is to consider firms' decisions in a simple general equilibrium model with competitive capital and labor markets. Even though connected firms are more productive than the average firm, and there is some rationale to subsidize them, estimated subsidies are far too large and dispersed to be efficient. According to the baseline estimates, the total annual output costs of political connections are 1.0-4.7%. Despite the focus in the literature on *misallocation costs*, I find that at least 55% of the costs of political connections are driven by *opportunity costs of public funds* as subsidies would be more efficiently spent on reducing distortionary taxes for all firms in the economy. The remainder is driven by capital and labor being *misallocated*. Connected firms end up much larger than is socially optimal, crowding out resources from other, non-connected firms in the economy. Almost all misallocation of resources happens across and not within firms because rent-seeking activities are concentrated in a few connected firms. I find even narrower welfare bounds and higher output costs ranging from 4-6.5% when considering the detrimental effects of political connections on the provision of public goods and increasing market power. These results are robust to further heterogeneity in industry- and connections-type and different forms of measurement error.

Given that a large part of the costs come from inefficiently high subsidies to connected firms that do not receive enough public scrutiny, I find high returns to increase monitoring of rent-seeking and corruption in the economy. Conservative estimates suggest policies that double all existing monitoring activities as long as extra monitoring costs are less than 0.1% of GDP - a realistic assumption considering that such extra spending amounts to 10x the global annual budget of Transparency International.[6] In summary, a few connected firms can pose high societal costs, and curbing their influence can have large returns.

The structure of the paper is as follows: Below, I discuss the related literature and contribution.

[6] This is at Indonesia's 1997 GDP. See: https://www.transparency.org/en/the-organisation/our-operating-budget. Accessed on May 12th, 2022.

Section 2 discusses the measure of political connections in the Indonesian data and how connected firms differ from non-connected firms. In section 3, I present and structurally estimate a model that can explain size differences between connected and non-connected firms and endogenizes subsidies. Section 4 quantifies the economy-wide costs of political connections. Key extensions of the baseline model and various robustness results are in section 5, while the last section concludes.

Literature

The key contribution of this paper is to provide quantitative estimates of the aggregate costs of political connections in general equilibrium. A growing micro-empirical literature has documented how favors to connected firms drain government resources[7] and lead to large allocative inefficiencies.[8] However, quantifying the aggregate costs of political connections has remained an elusive quest.[9]

Garcia-Santana et al. (2020) consider costs of political connections in general equilibrium but do not have firm-level evidence of political connections, forcing them to draw on sectoral estimates of corruption. The firm-level data allows to estimate firm-level subsidies directly. To the best of my knowledge, this paper is the first to propose a method that allows to estimate subsidies without imposing a functional form for the returns of rent-seeking.[10] Non-parametric identification matters; I find robust evidence for hump-shaped subsidy schedules in firm productivity, which rejects parametric rent-seeking technologies used in the rest of the literature (e.g. Akcigit, Baslandze, and Lotti forthcoming; Garcia-Santana et al. 2020; Arayavechkit, Saffie, and Shin 2018; Huneeus and Kim 2021). I show that standard models of rent-seeking need to feature increasing costs of rent-seeking activities to rationalize this empirical pattern and propose a simple functional form that can explain estimated subsidies, while being in line with observed spending on rent-seeking activities in other contexts (e.g. Campos et al. (2021)).

Arayavechkit, Saffie, and Shin (2018) and Huneeus and Kim (2021) both study the aggregate

[7]For example, Chen and Kung (2018) find that connected firms in China pay between 55-60% less for state-owned land.

[8]Haselmann, Schoenherr, and Vig (2018) show extensive misallocation of bank credit between connected firms and banks in Germany and Schoenherr (2019) finds that politically connected firms in Korea win a larger number of government contracts and that they execute these contracts systematically worse and at higher costs than non-connected firms. Schoenherr (2019) estimates that three quarters of the costs of contract misallocation are due to selecting the wrong firms to give contracts to. Similarly, Brugu'es, Brugu'es, and Giambra (2018) find that connected firms are more likely to win discretionary government procurement contracts in Ecuador and that these firms charge higher prices and are less efficient. Szucs (2017) shows that connected firms in Hungary sort into government procurement contracts that are allocated with higher bureaucratic discretion and finds evidence that these connected firms are of lower productivity. In contrast, Bertrand et al. (2018) does not find evidence that connected firms receive higher benefits from the state in France.

[9]Few papers looked at welfare, e.g. Faccio (2006); Fisman (2001); Gonzalez and Prem (2019); Martinez-Bravo, Mukherjee, and Stegmann (2017); Straub (2014); Gonzalez, Prem, and Urz'ua (2018); Chen and Kung (2018); Fisman and Wang (2015); Haselmann, Schoenherr, and Vig (2018); Schoenherr (2019). Notable recent exceptions are Akcigit, Baslandze, and Lotti (forthcoming), Bai, Hsieh, and Song (2020), Garcia-Santana et al. (2020), Arayavechkit, Saffie, and Shin (2018) and Huneeus and Kim (2021). Brugu'es, Brugu'es, and Giambra (2018), Szucs (2017) and Koren et al. (2015) also look at costs of rent-seeking focussing exclusively on partial equilibrium effects.

[10]This rent-seeking technology is similar to the idea of a "concealment technology" (Cremer and Gahvari 1994) or evasion technology (e.g. Slemrod and Yitzhaki 2002) used in the tax evasion literature. It is closer to the idea of tax avoidance (see Slemrod and Yitzhaki 2002; Slemrod 2001) in that I model political connections without risk, firms know how much taxes they have to pay this period and are only uncertain about future tax payments as political connections may change. This seems to be more in line with how connections work in developing countries (e.g. see Hoang 2018; Chen and Kung 2018).

costs of lobbying in the US, which they infer from firm-level lobbying expenditures and size distortions. The novelty in both papers is that they directly observe lobbying activity. I study the aggregate costs of rent-seeking in a context that is corruption-rife and non-democratic and where lobbying data is not available and would only capture a small portion of overall rent-seeking behavior. To quantify the aggregate costs of rent-seeking, one requires knowledge on the returns from rent-seeking as well as the extent of rent-seeking. My approach - using only information on standard firm inputs and output as well as whether a firm is connected or not - allows to flexibly identify returns and infer unobserved rent-seeking activities with additional assumptions in a second step. This second step is not needed if data on rent-seeking activities is directly available. Given the lack of data on rent-seeking activities - especially in corruption-rife contexts where rent-seeking activities are likely the most harmful - the approach in my paper to study the welfare costs of rent-seeking is applicable across many different contexts.

The paper is also complementary to Bai, Hsieh, and Song (2020) and Akcigit, Baslandze, and Lotti (forthcoming) in that I provide quantitative estimates on the costs and benefits of political connections that help to better understand welfare implications. Bai, Hsieh, and Song (2020) show how bureaucrats in China favor firms to help them avoid bad institutions and growth distorting regulation. I find that costs greatly outweigh benefits on aggregate. Akcigit, Baslandze, and Lotti (forthcoming) show important evidence for dynamic losses from political connections through a lack of innovation. I abstract from such dynamic losses because the data, unfortunately, does not allow me to study how connections change at the firm-level over time. Since I abstract from dynamic losses, my estimates should be seen as lower bounds on the costs of political connections.

The paper also strongly relates to the misallocation literature. Economically, I find that most costs from political distortions are due to real opportunity costs of public funds and not due to the misallocation of factors of production - a forceful reminder of the limitations from focussing only on misallocation losses. Methodologically, the paper speaks to direct and indirect approaches in the misallocation literature (e.g. Restuccia and Rogerson 2017). An interesting feature of this paper is that it combines the direct and indirect approaches by flexibly capturing different distortions when estimating subsidies with minimal structural assumptions and only later linking them to a full structural model. The paper also contributes separately to both strands of this literature. On the direct side, it adds to the literature by focussing on political connections as one particular friction.[11]

On the indirect side, most quantitative empirical work has followed Hsieh and Klenow (2009) in inferring general distortions from wedges in first-order conditions that lead to observed variation in factor shares (this also includes Garcia-Santana et al. (2020) and Huneeus and Kim (2021)). By assumption, the wedge approach captures distortions that only indirectly affect output and inputs via sub-optimal decisions. This paper takes the opposite and neglected approach of identifying "direct" distortions from differences in TFP distributions across connected and non-connected firms.[12] I term this a subsidy approach, a quantitative version of

[11]For other papers following this direct approach, see the literature cited in Restuccia and Rogerson (2017). Among papers that follow a direct approach, Guner, Ventura, and Xu (2008) is similar in that it also considers a distortion that firms internalize and that directly affects the measured productivity distribution.

[12]TFP in this paper is TFPQ in the setup by Hsieh and Klenow (2009), and they treat it s as a fundamental, while I show that it includes a combination of subsidies and fundamental *productivity*. Hsieh and Klenow (2009) and wedge approaches in their spirit assume that one observes undistorted (pre-wedge) output, while I instead assume that we can only observe distorted (post-subsidy) output.

Restuccia and Rogerson (2008) where subsidies are endogenous and identified using micro-data. Given that subsidies are identified from fundamentally different variation in the data, I show in a key extension in Section 5 how one can combine the subsidy and wedge approaches, allowing for different wedge distributions for connected and non-connected firms. In this case, a further benefit of my estimation approach is that it directly gives a control group from which one can infer counterfactual wedges - a key issue in wedge approaches. In an application, I show how these counterfactual wedges can quantify additional costs of political connections that stem from market power. A limitation of the subsidy approach is that it cannot distinguish between an output subsidy and input subsidies, and I capture all subsidies under a single output subsidy throughout.

Compared to the literature, the estimated costs of a few politically connected firms are high. I find that 1% of firms explain up to 20% of the total costs of misallocation found in Hsieh and Klenow (2009).[13] In contrast to Restuccia and Rogerson (2008), this is in a setup where subsidies are even positively correlated with productivity. The quantitative results are also in line with recent micro empirical research that has found productivity improvements and reductions in frictions (Abeberese et al. 2021) as well as positive competition effects (Hallward-Driemeier, Kochanova, and Rijkers 2021) in the wake of Indonesia's democratisation process and the fall of previously connected firms. More importantly, the structural setup in this paper shows how a decline in political connections leads to a reallocation of government spending and productive resources, reduces frictions and increases competition that in turn drive growth and development. Compared to other estimates on the costs of political rent-seeking, I find around 50% higher costs than Huneeus and Kim (2021)'s estimated costs of lobbying in the US. One might have expected even larger costs given the focus on the overall effect of rent-seeking in a highly corrupt regime. However, it is important to note that the costs of political connections studied here do not only depend on the extent of rent-seeking but also on the efficiency of public spending and the level of baseline distortions that rent-seeking may potentially ameliorate. A lower efficiency of public spending and higher baseline distortions in Indonesia explain why costs of rent-seeking can be of a similar magnitude in developing and developed countries.

2.2 Political Connections in Indonesia

The starting point is a good measure of political connections for which I draw entirely on Mobarak and Purbasari (2006). I first introduce their measure and the firm data and then briefly highlight key empirical regularities that will inform subsequent modelling choices.

Identifying connected firms in Indonesia

Indonesia under the rule of dictator Suharto at the end of the 1990s was characterized by a vast patronage network that extended from the capital city of Jakarta down to the village level (Fisman 2001; Martinez-Bravo, Mukherjee, and Stegmann 2017). By allocating public contracts, concessions, credit, and extra-budgetary revenues, a network of elites closely connected to the

[13] I take the market power and wedge estimates in this paper that are closest to models in the misallocation literature. I find aggregate productivity costs between 7.7-9.9%, computing aggregate productivity as $Z_t = \frac{Y_t}{K_t^\alpha L_t^\beta}$, which is comparable, but not identical to the setup in Hsieh and Klenow (2009).

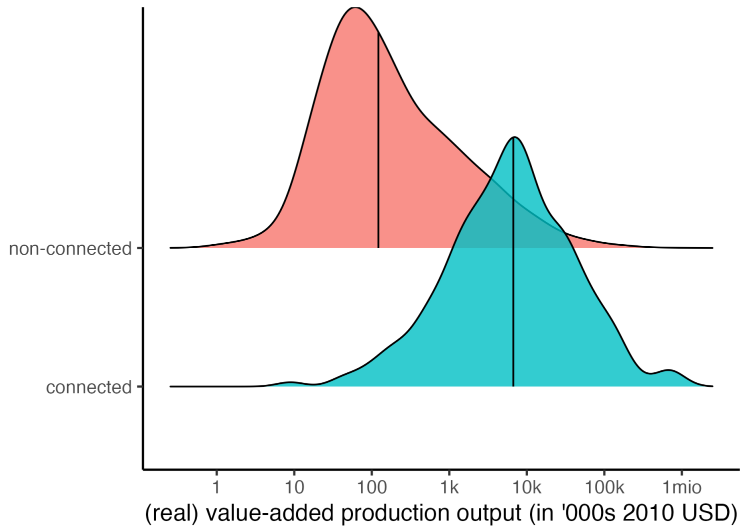
state administration was able to amass large amounts of wealth (see Hadiz and Robison 2013; Robison and Hadiz 2004). Such economic systems of patronage are, unfortunately, still widely prevalent around the world (e.g. Aslund 2019; Chen and Kung 2018; Diwan, Malik, and Atiyas 2019). Based on comparative statistics such as Transparency International's Corruption Index, today's Indonesia is similarly corrupt as countries such as Russia, Vietnam, Mexico and Brazil.

There is also strong evidence that political and economic elites held onto power after the fall of the Suharto regime in the aftermath of the Asian Financial Crisis in 1997 (see Robison and Hadiz 2004; Martinez-Bravo, Mukherjee, and Stegmann 2017). Still, recent empirical work finds that the eventual democratisation process led to productivity improvements and reductions in frictions among firms (Abeberese et al. 2021) and that this was at least in part driven by an increase in competition after the fall of previously connected firms (Hallward-Driemeier, Kochanova, and Rijkers 2021). In this paper, I will be able to quantify a number of economic mechanisms through which these effects played out.

At the same time, Indonesia is exceptional for providing several rich data sources that have allowed scholars to identify politically connected firms and link these to detailed annual firm-level panel data. Specifically, this paper draws on the Annual Manufacturing Survey (Survei Tahunan Perusahaam Industri Pengolahan) collected by Indonesia's Central Bureau of Statistics (Badan Pusat Statistik), which covers all formal manufacturing establishments with more than 20 employees. Based on the GGDC 10-sector database, these account for about 30% of all value-added manufacturing output in Indonesia (Fentanes and Gathen 2022). The survey contains detailed industry information (up to 5-digit), employment, production, and other firm characteristics and has been used extensively in the Economics literature (e.g. Amiti and Konings 2007). I combine this with the measure of political connections from Mobarak and Purbasari (2006), who identified politically connected firms and already linked these to firms in the survey.

Mobarak and Purbasari (2006) identify connected firms in two different ways. In this paper, I use the union of the two sets of firms as my main measure of whether a firm is politically connected. The first set of firms is identified by tracing firms that were directly owned and founded by blood relatives of Suharto. This set excludes firms whose owners might have strategically married into the Suharto family. For the second and more comprehensive set of firms, Mobarak and Purbasari (2006) draw on the natural experiment in Fisman (2001). Fisman (2001) uses news about plausibly exogenous health issues of dictator Suharto in various periods between 1995-1996 and looks at responses to firms' stock prices on the Jakarta Stock Exchange around these events. The idea is that news about the deteriorating health conditions of the dictator should negatively affect the stock price of firms that benefit from being politically connected to the dictator. The added benefit of the Indonesian context is that the Indonesian regime was highly centralized around the dictator, so shocks to the dictator's health should affect any listed connected firm. Mobarak and Purbasari (2006) then link the identified listed connected firms to non-listed connected firms by tracing all other firms that share ownership and management through conglomerate structures. As Claessens, Djankov, and Lang (2000) and Carney and Child (2013) show, most firms belong to larger conglomerate structures owned by specific families and ownership and control is rarely separated in Southeast Asian firms, including Indonesia. At last, Mobarak and Purbasari (2006) link the set of identified connected firms to the manufacturing census, leaving a sample of 241 firms, of which 89 firms

Figure 2.1: Output distribution: Connected vs. non-connected firms



Notes: Distributions of firm-specific real value-added output (in 2010 USD in 000's) for cross-section of Indonesian firms in 1997 based on Statistik Industri, the Indonesian manufacturing firm census. Connected vs. non-connected firms are identified as in Mobarak & Purbasari (2006). Non-connected firms: N = 18,317. Connected firms: N = 241.

are identified as being owned and founded by blood connections of Suharto.

I provide more detailed information on each of the steps in Appendix A.1 and discuss the role of measurement error in Section 5. However, it is important to highlight three key features of this data. First, this definition of political connections captures "high-level" political connections and does not capture more local connections of firms to local authorities in the bureaucracy or police. The reason is that the approach only captures firms linked to conglomerate structures that either belong to Suharto's blood family or include at least one listed firm that is identified via the natural experiment. Second, the measure of political connections is different from state-owned enterprises, but there is some overlap. About 15% of connected firms in the data can be classified as at least partly state-owned, while the remaining 85% of connected firms see no state ownership. I further discuss the role of state ownership in Section 5. At last, the approach identifies a snapshot of the connected manufacturing firms in 1994-1997, shortly before the Asian Financial crisis in 1997/8. Throughout, I consider only data before the Asian Financial Crisis, because I do not observe changes in connections after the crisis.

Differences between connected and non-connected firms

Figure 2.1 shows the firm-size differences in value-added output between connected and non-connected firms for the cross-section of Indonesian manufacturing firms in 1997, the year before the crisis. The average connected firm is about twelve times larger than the average non-connected firm, but there is also considerable overlap in output across the two distributions. In fact, there exist non-connected firms that are smaller than the smallest connected firm and non-connected firms that are larger than the largest connected firm. The size distribution of non-connected firms is visibly more skewed and more dispersed. One way to see the dispersion is that the coefficient of variation is more than four times larger for the size distribution of non-connected than for connected firms (13 vs 3).

Table 2.1 documents the average size differences between connected and non-connected firms

Table 2.1: Within-industry size ratios of average connected over average non-connected firms

		Within industry			
	unconditional	2-digit	3-digit	4-digit	5-digit
Ratio	11.96	14.19	13.51	17.74	19.93
# industries	1	9	31	115	302
# industries w/ connected firm	1	9	26	62	103

Details:
Data is real value-added output data for the cross-section of Indonesian manufacturing firms in 1997 based on Statistik Industri. Size ratios are computed based on the ratio of the average size for connected vs. the average size of non-connected firms within each considered industry and then averaged across industries using the number of connected firms in each industry as weight. Non-connected firms: N = 18,317. Connected firms: N = 241.

within industries. I separately compute the average output of all connected and non-connected firms and compute their ratios by industry. I then take these size ratios and average over them to derive an economy-wide size ratio, using as weights the number of connected firms in each industry. Column 1 reports the average size ratio without industry heterogeneity, and Columns 2-5 report ratios looking respectively within 2-, 3-, 4- and 5-digit industries. I find that connected firms are not only considerably larger on average than non-connected firms, but these size differences are just as large or even larger within industries. Even within narrowly defined industries, the average connected firm is more than 12 times and up to 20 times larger than the average non-connected firm. Outliers do not drive this pattern. As for the distribution of connected firms across industries, connected firms are widely dispersed. Only about 1.3% of all firms are connected, but connected firms still show up in all nine 2-digit industries, 26 out of 31 3-digit industries and about one-third of all roughly 300 5-digit industries. This dispersion across industries suggests that size differences are not driven by selection into specific industries.[14]

2.3 Quantifying the role of connections: A structural approach

This section develops a simple model of heterogeneous firms similar to Restuccia and Rogerson (2008) without entry and exit, where firms make static choices of production inputs. Size differences between connected and non-connected firms in the model are driven by fundamental differences in idiosyncratic productivity and differences in benefits from political connections. Benefits from political connections are modelled as idiosyncratic output subsidies, which in contrast to Restuccia and Rogerson (2008), are endogenized via strategic spending on rent-seeking activities similar to Garcia-Santana et al. (2020) and Akcigit, Baslandze, and Lotti (2018). Based on this model, I show how to identify and estimate benefits from political connections flexibly. The last part shows how these estimates align with a rich model of endogenous rent-seeking behavior, which is subsequently used for partial and general equilibrium counterfactuals.

[14] Further empirical results and robustness exercises are reported in Appendix A.2.

2.3.1 Modeling political connections

Household & government

The household side of the model is kept as simple as possible, featuring a representative household maximizing life-time discounted utility:

$$\sum_{t=0}^{\infty} \beta^t U(C_t)$$

subject to a CRRA intertemporal utility function $U(C_t)$ and a per period budget constraint:

$$A_{t+1} + C_t = (1 + r_t - \delta)A_t + w_t L_t + \Pi_t + T_t$$

where households face no risk, provide labor supply L_t inelastically at potentially time-varying wage w_t, rent capital to firms at the potentially time-varying interest rate r_t, face depreciation of capital at rate δ and demand consumption goods fully elastically. Households further receive net profits by firms and net revenue T_t from the government. The household's optimal consumption-savings choice is then characterized by the standard Euler Equation:

$$1 = \beta(1 + r_{t+1} - \delta)\frac{U'(C_{t+1})}{U'(C_t)}$$

In the baseline model, the government levies taxes, subsidizes connected firms and balances its budget each period by paying any net revenues T_t lump sum to households.

Firms

The economy is populated by a fixed and discrete number of heterogeneous firms indexed by i whose after-tax value-added revenues R_i are given by:

$$R_i = (1 - \bar{\tau})p_i y_i = (1 - \tau_i)y_i = (1 - \tau_i)z_i k_i^{\alpha} l_i^{\beta} \ \text{ with } \ \alpha + \beta \leq 1 \tag{2.1}$$

z_i captures firm-specific productivity, k_i and l_i denote the firm's input choices of capital and labor and α and β give the revenue elasticities of capital and labor respectively.[15] Crucially, as in Restuccia and Rogerson (2008), firms face idiosyncratic taxes τ_i that are directly paid to the government. Specifically, all non-connected firms face a constant *de jure* revenue tax $\tau_i = \bar{\tau}$ that is set to 25% to mimic Indonesia's flat corporate income tax rate.[16] Political connections in the model solely enter by distorting these baseline taxes. As the key object of interest, define the differential subsidy $(1+\tilde{\tau}_i) \equiv \frac{1-\tau_i}{1-\bar{\tau}}$, so that connected firms for whom $\tau_i < \bar{\tau}$ are subsidized and non-connected firms face $\tilde{\tau}_i = 0$. In this setup, taxes are like firm-specific output prices given by: $p_i = (1 + \tilde{\tau}_i)$ (where prices of non-connected firms are constant and normalized to unity). The differential subsidy $\tilde{\tau}_i$ captures in a reduced-form way many of the channels

[15] One can think of this setup as heterogeneous firms producing a single output or as I show in Appendix A.3, isomorphically as an economy with CES demand and differentiated inputs. In the latter case, z_i flexibly captures both productivity and demand. I do not separately identify the role of demand vs productivity. As is standard in models of firm-size dynamics, both factors influence firm dynamics in the same way. For ease of exposition, I simply refer to them as productivity in the following.

[16] There are reduced tax rates for small enterprises as well as public enterprises. The former do not play a role in my model and do not show up in the Manufacturing census data. I ignore the latter or implicitly capture them if they are counted as connected firms.

mentioned in the introduction through which connected firms benefit from interactions with the government. The direct identification and estimation of $\tilde{\tau}_i$ in this paper allow for any of these channels. However, to be consistent with the subsequent quantification of the costs of political connections, I interpret distortions throughout as either direct tax distortions (tax cuts or evasion) or as higher output prices that the government pays for - a public mark-up.

To allow the firm size of connected firms to flexibly depend on their benefits from political connections, I endogenize the differential subsidy $\tilde{\tau}_i()$. To fix ideas and in line with the little existing systematic evidence on high-level rent-seeking (Campos et al. 2021), think of rent-seeking activities of a firm as being organized within a department in charge of lobbying, tax evasion, legal affairs and bribery. The endogenous subsidy can then be thought of as the output of the rent-seeking department. To make this clear, call $\tilde{\tau}$ the *Political Connections Technology*, which depends on four inputs: $\tilde{\tau}(l_P, k_P, \varepsilon_i, z_i)$. First, connected firms endogenously choose the size of their rent-seeking department by choosing the amount of capital, k_P, and labor, l_P, engaged in rent-seeking activities. Labor employed in rent-seeking activities captures lawyers who renegotiate contracts and find tax loops, lobbyists who push for favorable legislation and preferential contracts, management and other workers who are involved in managing rent-seeking projects, meeting and cultivating political contacts and labor used by third parties who specialize in facilitating rent-seeking and corruption (Hoang 2018). Rent-seeking capital captures equipment and machines that are used for rent-seeking activities. These endogenous rent-seeking activities are similar to Garcia-Santana et al. (2020) but more general than the fixed cost of political connections as in Akcigit, Baslandze, and Lotti (2018).[17] Furthermore, the output of rent-seeking activities can directly depend on firm productivity z_i to capture the idea that more productive firms are also more productive at rent-seeking or that the government may interact differently with more productive firms.

At last, whether the firm takes part in rent-seeking activities depends on an exogenous binary variable ε_i that captures access to political connections. With $\varepsilon = 0$, individual firms are not currently matched to a politician in power, do not have the ear of the political elite or have a distaste for political connections, so subsidies are zero ($\tilde{\tau}_i = 0$). Similar to Akcigit, Baslandze, and Lotti (2018), this captures the idea that political connections depend partly on luck, evidenced by the fact that despite political connections being profitable, most firms - including some of the largest Indonesian firms - are not connected.[18] I model the process of ε_i in a way that nests the endogenous entry into political connections via a fixed cost (e.g. as in Huneeus and Kim 2021).

Given the *Political Connections Technology* $\tilde{\tau}(l_P, k_P, \varepsilon_i, z_i)$, how do connected and non-connected firms choose inputs? All firm choices are static.[19] Firms buy workers on the spot market at a common wage w and rent assets at a rental rate r from households. Both assets and labor can

[17]Specifically, the fixed cost in Akcigit, Baslandze, and Lotti (2018) simply captures the remuneration of one politician and is thus nested by rent-seeking activities in labor l_P in the model above. Spending on political connections in Garcia-Santana et al. (2020) captures direct in-kind bribes that are captured by total rent-seeking activities in my model.

[18]Later on, I show how to generalize ε to allow for a finite number of different types of connections. Types could then be industries or different groups of connected firms, such as connected firms that are blood-connected to the dictator Suharto versus connected firms who do not have this special link, nesting corruption-specific productivity as in Garcia-Santana et al. (2020). While not specifically modeled, this setup also captures endogenous entry into political connections via a fixed cost because of how the process of ε_i is specified.

[19]In the conclusion, I briefly discuss how the identification approach may also work in a dynamic setup.

be used instantaneously either as productive inputs or for rent-seeking activities. ε_i may vary over time (following a first-order Markov process). However, the realization of ε_i is known to the firm at the beginning of a period before it makes any other production decisions. A firm is thus fully characterized by productivity z_i as well as the realization of ε_i and solves the following static problem each period:

$$\max_{k,l,l_P,k_P} \left\{ \pi(z_{it}, \varepsilon_{it}) \equiv \big(1 + \tilde{\tau}(l_P, k_P, \varepsilon_{it}, z_{it})\big)(1-\bar{\tau}) z_{it} k^{\alpha} l^{\beta} - w_t(l + l_P) - r_t(k + k_P) \right\} \tag{2.2}$$

This problem gives simple and intuitive static FOCs that say that firms should equalize the marginal benefits and the marginal costs (as captured by the rental prices of capital and labor) for both production and rent-seeking activities:

$$\begin{aligned} r_t &= \alpha \frac{R_{it}(k^*, l^*, l_P^*, k_P^*)}{k^*} = \frac{\partial \tilde{\tau}(l_P^*, k_P^*, \varepsilon_{it}, z_{it})}{\partial k_P}(1-\bar{\tau}) z_{it} k^{*\alpha} l^{*\beta} \\ w_t &= \beta \frac{R_{it}(k^*, l^*, l_P^*, k_P^*)}{l^*} = \frac{\partial \tilde{\tau}(l_P^*, k_P^*, \varepsilon_{it}, z_{it})}{\partial l_P}(1-\bar{\tau}) z_{it} k^{*\alpha} l^{*\beta} \end{aligned} \tag{2.3}$$

Based on a revealed-preference argument, firms show up as non-connected if they optimally choose rent-seeking labor and capital such that $\tilde{\tau} = 0$.

Equilibrium

The aggregate resource constraint is given by: $Y_t = \sum_i z_{it} k_{it}^{\alpha} l_{it}^{\beta} = C_t + I_t$. The focus in this paper is on a *steady state competitive equilibrium* that is described by competitive prices (r^*, w^*) that households and firms take as given and competitive allocations such that:

- the exogenously given set of firms all produce by optimally choosing capital, labor and rent-seeking activities based on their realizations of $(z_{it}, \varepsilon_{it})$
- the household optimally chooses consumption and savings based on the Euler Equation given above and consumption and savings are constant over time
- the government levies taxes and subsidizes connected firms and balances its budget each period by transferring net revenue lump-sum to the household
- prices adjust such that capital demand and supply and labor demand and supply equalize each period and these aggregates are constant over time
- the distribution of firms over $(z_{it}, \varepsilon_{it})$ is at its stationary distribution

2.3.2 Identification of political connections

Taking a step back, it is important to highlight that the welfare implications of subsidies to connected firms are ex-ante unclear in this setup. Given the baseline distortion of revenue taxes that all firms face, there is a welfare argument for subsidizing connected firms.[20] In the end, whether subsidies to connected firms are harmful in comparison to no subsidies to connected firms depends on at least three key margins: (i) How many firms become connected, (ii) the distribution of subsidies as governed by the shape of $\tilde{\tau}(.)$ and (iii) the extent of socially wasteful spending on rent-seeking activities that directly depends on the selection of connected firms

[20] As I show formally in Appendix A.4, in a setup with heterogeneous firms, decreasing returns to scale in production and distortive output taxes, it is optimal to subsidize firms at constant rates (up to small general equilibrium corrections) and distribute subsidies as widely as possible.

as governed by ε.

To estimate the costs of political connections, it is important to capture all three margins flexibly. The first margin is directly pinned down by observing the number of connected firms in the data. In contrast, the other two margins must allow for considerable flexibility in how connected firms are selected and how rent-seeking activities by connected firms translate into firm-specific subsidies. The approach in this paper, as formally stated in Proposition 2, is to impose functional form restrictions on the selection of connected firms, pose a weak assumption on the shape of $\tilde{\tau}(.)$ and then back out $\tilde{\tau}(.)$ non-parametrically:

Proposition 2 (Main identification result)**.** *Given the previous setup and conditional on having identified total factor productivity (TFP) defined as $TFP_i \equiv (1-\tau_i)z_i$ one can separately identify τ_i and z_i based on the following two assumptions:*

1. *(**Selection**). Firms with access to political connections have been drawn from a known population of productivities z_i according to:*

$$\mathbb{P}(\varepsilon \neq 0) = \begin{cases} cz_i^{\rho}, & \text{if } z_i \geq \bar{z} \\ 0, & \text{otherwise} \end{cases}$$

where c is a normalizing constant to ensure well-defined sampling.

2. *(**Monotonicity of TFP**) The connections technology $\tilde{\tau}(.)$ is such that there is a monotonic mapping between TFP_i and productivity z_i for firms with access to political connections. Formally, $\frac{\partial TFP(z,\tilde{\tau})}{\partial z} = (1+\tilde{\tau})(1-\bar{\tau}) + \frac{\partial \tilde{\tau}}{\partial z}(1-\bar{\tau})z > 0$ for $\bar{\tau} \in (0,1)$ given and all $\tilde{\tau} \in supp(\tilde{\tau})$.*

The proof of this proposition is straightforward. The assumption on selection guarantees that we can link an identified distribution of productivities to an observed distribution of TFP for connected firms. In contrast, the monotonicity assumption allows linking moments of this identified productivity distribution to moments of the observed distribution of TFP of connected firms. We can then identify the entire subsidy distribution using $\tau(q) = TFP(q)/z(q) - 1$ for any quantile q.

There are several components in the proposition that are important to unpack. Beginning with the monotonicity assumption on TFP, this assumption states that the ranking of connected firms by TFP is identical to the underlying ranking of their productivities. This restriction on the underlying *Political Connections Technology* $\tilde{\tau}(q)$, namely that $(1+\tilde{\tau}) + \frac{\partial \tilde{\tau}}{\partial z}z > 0$, is naturally given for functions that are strictly increasing in rent-seeking activities, but exhibit any form of decreasing, constant or increasing returns to scale. More generally, it also allows $\tilde{\tau}(q)$ to be decreasing, only that any decline in $\tilde{\tau}(q)$ is not faster than the corresponding increase in productivity. For example, this allows a *Political Connections Technology* where benefits from political connections decline with firm size as this puts the firm into the public eye, making any corrupt practices more difficult or where the *Political Connections Technology* is understood as a reduced-form tax evasion technology where the probability of getting caught increases with firm size. Importantly, the monotonicity assumption nests decreasing returns to scale functions that have been considered in the literature (Akcigit, Baslandze, and Lotti 2018; Garcia-Santana et al. 2020) and thus allows for testing their functional form assumptions formally.

The selection assumption is more restrictive and can be divided into two parts. First, identification requires observing a known population of productivities from which firms with access to political connections are drawn. The setup in this paper makes this particularly suitable as subsidies to connected firms are considered differential subsidies compared to non-connected firms. Hence, the TFP of non-connected firms gives their underlying productivities up to a known baseline revenue tax $\bar{\tau}$. Furthermore, we need to assume that the productivities of non-connected firms can be treated as the underlying population from which access to political connections is drawn. Again, the setup in this paper is such that the sample of non-connected firms is roughly 100 times larger than the sample of connected firms, making this population assumption a natural choice. The last component for the first part of the selection assumption is that the support of underlying productivities for connected firms is entirely contained in the support of productivities of non-connected firms. This common support assumption is similar to standard matching estimators and, as I show in the estimation part below, can be readily verified in the data.

The second part of the selection assumption puts structure on the selection rule with which firms receive access to political connections. The idea for identification is that given a selection rule with which access to connections is drawn, one can selectively sample the productivities of connected firms using the population of productivities of non-connected firms. An estimator can then simply be an average across many independently but selectively drawn productivities. The functional form restriction is required to identify the selective sampling process. To see this, note that without this assumption, one could draw arbitrary samples of non-connected firms rationalizing any productivity distribution of connected firms within the productivity support of non-connected firms. The functional form assumption allows for considerable selection. One way to see this is to think of it as a setup in which connections are formed at selective meetings where a minimal firm size is needed to access the meetings and larger, more productive firms are more likely to meet or be approached by politicians at that meeting. The pre-selection might be done directly by politicians or may capture fixed or high variable costs to rent-seeking activity that are not worthwhile for firms below a certain productivity/size threshold, capturing endogenous entry into political connections. Identification results for the parameters of the selection rule are stated in Proposition 3:

Proposition 3 (Identification of the selection of connected firms and conservative subsidy bounds)**.** *Given the assumptions in Proposition 2, the parameters of the selection rule that govern access to political connections are set-identified under the following additional restrictions:*

1. *(**conservative normalization**): The subsidy of the connected firm with the lowest possible observable TFP is zero ($\tilde{\tau}_{q=0} = 0$ for q giving the quantile of the underlying productivity distribution of connected firms).*
2. *(**rational rent-seeking**): Connected firms will never choose $\tilde{\tau} < 0$.*
3. *(**independence as minimum selection**): There is a lower bound for selection that is given by $\rho = 0$ and $\bar{z} = 0$, the case of independence between ε and z.*

*Specifically, $\bar{z}$ is point-identified under **conservative normalization** and independent of ρ. And ρ is set-identified with $\rho \in \{0, \bar{\rho}\}$, where $\bar{\rho}$ is identified from the maximum ρ for which **rational rent-seeking** still holds and another firm's $\tilde{\tau}_{q>0} = 0$. The model is rejected in case $\bar{\rho} < 0$. We can call identified subsidy distributions $\tilde{\tau}(q)$ based on the sets of parameters $\Big\{\{\bar{z}, 0\}, \{\bar{z}, \bar{\rho}\}\Big\}$ conservative bounds for actual subsidy distributions.*

This proposition makes clear how the parameters $\bar{z}$ and ρ allow for considerable selection of connected firms. $\bar{z}$ truncates the productivity distribution of non-connected firms from which access to political connections is drawn and thus shifts the entire productivity distribution of connected firms. One can also think of $\bar{z}$ as giving an entry threshold given by a fixed cost of entry into political connections. ρ allows for additional correlation between access to political connections and underlying productivity within this truncated productivity distribution. The distribution of subsidies is then identified from residual dispersion in TFP that is not explained by selective sampling-implied variation in underlying productivities.

2.3.3 Estimation of political connections

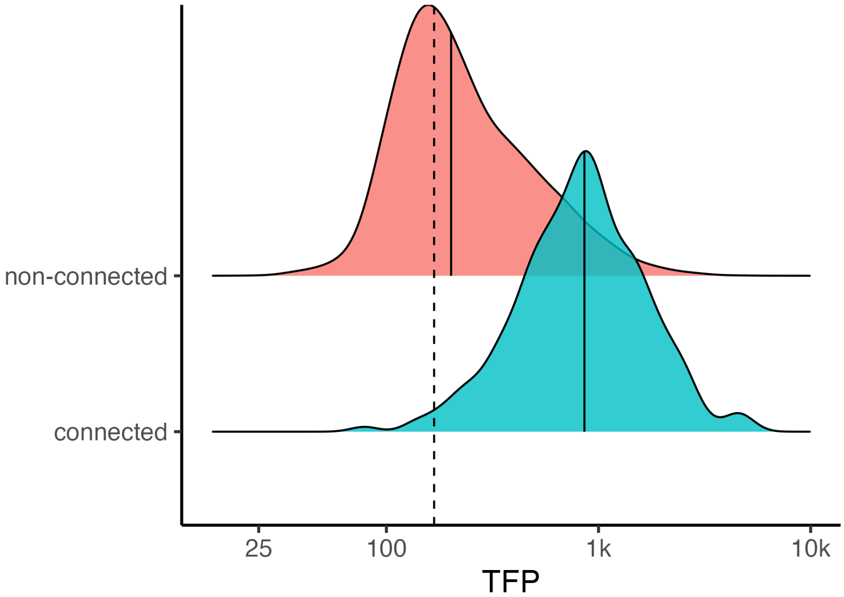
Estimation of subsidies to connected firms proceeds in two steps. The first step estimates TFP for all firms in the economy. In the second step, I use a quantile treatment effect (QTE) estimator for subsidies that builds on Propositions 2 and 3 and uses a bootstrap-based matching procedure. To estimate TFP, I follow a strictly model-consistent approach. In principle, any estimator for TFP as defined previously can be used in the first step. However, model consistency ensures consistent partial and general equilibrium counterfactuals and a cleaner identification and estimation of subsidies. In Section 5, I consider how results are affected by alternative TFP estimation that allows for wedges and further production function heterogeneity.

The model-consistent TFP estimator consists of first estimating revenue function elasticities α and β exploiting static first-order conditions of firms. These first-order conditions state that revenue spending shares on productive labor and capital equal their respective revenue elasticities. By assumption, revenue elasticities are identical across connected and non-connected firms, so it suffices to use observed revenue spending shares for non-connected firms. This has the benefit of not having to take a stand on whether reported input spending by connected firms is misreported or partly includes spending on non-productive, rent-seeking activities. Static first-order conditions of non-connected firms imply that observed revenue factor spending shares should be constant across firms. In the data, as shown among others in Hsieh and Klenow (2009), there is strong variation in revenue factor shares even within narrowly defined industries. In the baseline results, I treat the observed variation in factor shares as stemming solely from measurement error in reported labor and capital spending centred around 0. Specifically, I estimate α and β using median factor revenue shares across all non-connected firms. Given estimates for α and β, I use observed firm revenue R_{it} and the model-implied spendings on productive capital k^*_{it} and labor l^*_{it} to identify:

$$TFP_i = (1 - \tau_i)z_i = \frac{R_{it}(k^*_{it}, l^*_{it}, l^*_P, k^*_P)}{k^{*\alpha_s}_{it} l^{*\beta_s}_{it}} \tag{2.4}$$

Using model-implied spendings on productive capital and labor is crucial here. It cleans the data from measurement error in labor and capital spendings, which shuts down all the variation used in Hsieh and Klenow (2009) and abstracts from any variation in factor shares due to dynamic input choices (e.g. Asker, Collard-Wexler, and De Loecker 2014). It ensures that - conditional on α and β - all variation in TFP is estimated from variation in observed revenue. The implicit assumption here is that reported revenue is reported without measurement error. In Section 5, I consider both measurement error in reported revenue and generalize the approach to allow for both wedges and subsidies at the same time.

Figure 2.2: TFP distribution: Connected vs. non-connected firms



Notes: Distributions of firm-specific TFP for cross-section of Indonesian firms in 1997 based on Statistik Industri, the Indonesian manufacturing firm census. Connected vs. non-connected firms are identified as in Mobarak & Purbasari (2006). TFP is identified as residual from production function estimation at the 1-digit industry level (single production function across all firms) as explained in the text. The x-axis is on a log-scale. The dotted line indicates the minimum TFP of connected firms after dropping the 3 lowest TFP connected firms. Non-connected firms: N = 18,317. Connected firms: N = 241.

Figure 2.2 shows the estimated TFP distributions of connected and non-connected firms for the cross-section of Indonesian manufacturing firms in 1997. The data shows that the average connected firm has slightly less than 3.5 times higher TFP than the average non-connected firm, and there is a large overlap in the two distributions. These size differences are considerably smaller than the value-added output differences reported in Figure 2.1. Based on the baseline model, TFP for non-connected firms is exactly equal to their productivity z_i, so Figure 2.2 captures the entire productivity distribution of non-connected firms.

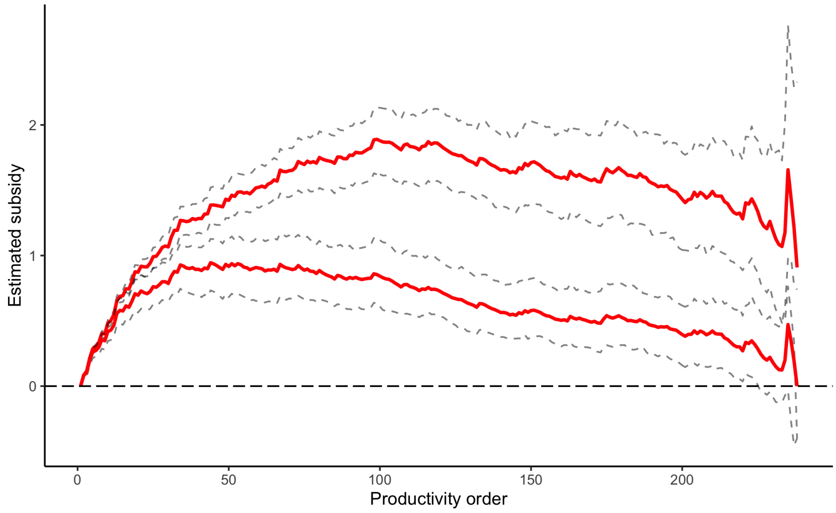
Given TFP, the second step of the estimation approach constructs a matching estimator that matches each connected firm with a comparable non-connected firm for which: $TFP_{i,NC} = \tilde{z}_{i,NC}$. For a given selection rule, Propositions 2 and 3 can be used to draw a set of connected firms from the population of non-connected firms and match them according to their ordering of productivities and TFP. In the case of sampling independent bootstrap samples, the approach matches the nth highest productivity firm in the bootstrap sample to the nth highest TFP connected firm. The productivity estimate for each connected firm is the average over all matched productivities for this specific connected firm. This non-standard, distributional matching estimator is necessary as standard matching based on observables does not work here. Standard matching approaches would require matching on observables that explain productivity but are not directly affected by political connections. In Section 5, I show how the approach can be extended to still control for further observables and that results are robust to this. Treating non-connected firms as the underlying population and selectively sampling from their productivities works well in this context as there are many more non-connected to connected firms in the data. To verify the common support assumption for productivities, we can first note that common support is fulfilled for TFP. The highest productivity of a non-connected firm is about 70% higher than the highest TFP among all connected firms. Furthermore, we know that due to **rational rent-seeking**, $\tilde{\tau}_i \geq 0$, so that $\text{TFP}_i \geq \tilde{z}_i$, which establishes that there is no connected firm with productivity higher than that of all non-connected firms in the data. For the lower bound, I assume that no connected firms have lower productivity than all other non-connected firms in the data. This assumption is very weak since connected firms are unlikely to be that unproductive, and a violation of this assumption would imply unrealistically high subsidies.

The parameters of the selection process are estimated as follows. Call the number of connected firms N_C and the sample of connected firms C_i, which is ordered by TFP. Under the assumption of **conservative normalization**, the subsidy is zero for the connected firm with the lowest possible productivity. $C_i = 1$ refers to the connected firm with the lowest productivity in the data. Setting the subsidy to zero for this connected firm gives the most conservative estimate of the lower productivity bound that is still in line with the data and the assumption of **conservative normalization** without extrapolating beyond the lowest TFP connected firm observed. It is conservative because it raises the productivity estimates for all connected firms and, in turn, lowers their subsidy estimates. Since $\bar{z}$ entirely depends on the lowest observed TFP of connected firms, this estimator is susceptible to low-TFP outliers among connected firms, which would drive up estimates of subsidies. Again, I take a conservative approach to bias my estimates against finding high subsidies by dropping the three connected firms with the lowest observed TFP. The dotted line in Figure 2.2 reports the baseline estimate for $\bar{z}$. To estimate $\bar{\rho}$, first define the truncated productivity distribution of non-connected firms $\tilde{Z}(q)_{\bar{z}}$ for any quantile q. For the case of $\rho = 0$, sampling from the truncated productivity distribution is uniform so that the subsidy distribution is given by $\tau(q) = TFP(q)/\tilde{Z}(q)_{\bar{z}} - 1$ for all uniformly spaced $(N_C - 1)$ quantiles. The estimator may also be referred to as a *quantile matching estimator*. For the lower subsidy bound, one draws bootstrap samples from the truncated productivity distribution according to the additional correlation ρ. The productivity estimate is the average productivity across bootstrap samples for each connected firm. $\bar{\rho}$ is the maximum possible correlation ρ for which another subsidy estimate than for $C_i = 1$ becomes zero.

The resulting non-parametric estimates of the *Political Connections Technology* τ_i are shown in Figure 2.3, plotted over the ordering of productivities. Both the estimated upper and the lower bound subsidy schedules follow a hump shape over productivity; the subsidy first increases and then decreases in absolute terms for highly productive firms. The shape is precisely estimated based on the 95% pointwise bootstrap confidence bands given by the grey dotted lines. The estimated shape follows from the TFP distribution of connected firms being less dispersed and less skewed than the truncated productivity distribution and is not enforced by the estimator. As shown in Section 5, the concave or hump shape also shows up when considering wedges and industry- or type-specific subsidy schedules.

The estimate for $\bar{\rho}$ is 0.93. This leads a firm with productivity similar to the largest connected firm in the sample to be more than 23x as likely to be connected as a firm close to the estimated productivity threshold. Importantly, point-estimated subsidies stay positive over the entire distribution, which is not enforced by the estimation approach except for the connected firm with the lowest TFP. We can thus use the bootstrap confidence bands as an overidentification test for the assumption that $\tilde{\tau}_i \geq 0$. The upper bound clearly passes this test except for the largest connected firm whose subsidy is imprecisely estimated. And even though the lower

Figure 2.3: Baseline subsidy estimates



Notes: Baseline non-parametric estimates of conservative subsidy bounds. Red lines give point estimates formed by average estimated productivities across bootstrap samples respectively for the lower and upper bound. Grey, dashed lines give point-wise 95% bootstrap confidence bands using 10,000 bootstrap samples. Lower bound estimated to be given by $p = 0.93$. Estimates based on assumptions explained in the text and estimated using data on cross-section of Indonesian manufacturing firms in 1997 based on Statistik Industri, the Indonesian manufacturing firm census. Connected (N = 238, dropped 3 outliers) vs. non-connected (N = 18,317) firms are identified as in Mobarak & Purbasari (2006).

bound estimate shrinks all subsidies towards zero, confidence bands in this case also clearly rule out negative subsidy estimates.

Through the lens of the model, the estimated subsidy can be directly interpreted as the total price premium paid by the government - a public mark-up. For the average subsidized connected firm, this estimated public mark-up varies from around 65% for the lower bound to about 150% for the upper bound. While it is difficult to directly compare these estimates to other estimates in the literature, Schoenherr (2019) estimates average direct cost increases for politically connected firms of around 30% for Korea. The above estimates are reasonable if indirect subsidies such as tax evasion or direct input subsidies are of a similar magnitude to direct output subsidies.[21]

2.3.4 Estimating endogenous rent-seeking activities

What can we learn from estimated subsidies about how firms invest in unobserved rent-seeking activities? In the following, I estimate a *Political Connections Technology* that flexibly captures benefits from investing in rent-seeking activities and costs and explains most variation in non-parametrically estimated subsidies. This *Political Connections Technology* that maps from rent-seeking behavior to returns to rent-seeking is central in any model of rent-seeking behavior Arayavechkit, Saffie, and Shin (2018), but has received little empirical validation to date. In this paper, estimating the *Political Connections Technology* is crucial not just to rationalize estimated subsidies, but also to quantify the costs of political connections in partial and general equilibrium where we need to know how political connections distort aggregate capi-

[21] It is important to note that while input subsidies are likely quantitatively important, they will be picked up through the output subsidy but they are not isomorphic in this setup.

tal and labor demand through rent-seeking activities.[22]

The previous literature has assumed variants of a *Political Connections Technology* where returns to rent-seeking $\tilde{\tau}_i$ are given by a decreasing returns to scale technology in rent-seeking activities p Arayavechkit, Saffie, and Shin (2018). I show in Appendix A.5 that this functional form cannot generate hump-shaped subsidy schedules as observed in the data. Hence, I extend and generalize previously used functional forms by assuming the following *Political Connections Technology* in rent-seeking p:

$$\tilde{\tau}_i(\varepsilon, z, p) = f(\varepsilon, z, p) - \text{cost}(z, p) = \varepsilon z p^{\theta_p} - c p^{\theta_c} z^{\theta_z}$$

In Appendix A.6, I provide two different micro-foundations for this functional form. One is where firms bribe and lobby politicians who need to push for regulatory changes, preferential policies and access to government contracts, and an alternative micro-foundation where firms bribe tax collectors to avoid taxes. In both cases, the first part captures benefits from connections and nests the case of decreasing returns to scale considered in the literature. θ_p captures the output elasticity with which bribes are funnelled to politicians via lobbying and obfuscatory exchanges. It also captures the degree of returns with which either politicians or tax collectors can allow for subsidies. Intuitively, this will be identified from the slope of the increasing part of estimated subsidies. ε captures the level of benefits so that not having access to political connections shuts down their benefits. Including z allows more productive firms to also be more efficient at rent-seeking activities.

The second part of the technology captures costs of political connections. In both micro-foundations, these costs capture the risk of being detected or having some benefits overturned by other politicians who oppose policies in the political process, public scrutiny or by lawsuits. The elasticity θ_c then captures the convexity or concavity of these costs and is identified from the curvature at the top of estimated τ where subsidies change from increasing to decreasing. c captures the level of these costs. Additionally, the term z^{θ_z} captures in a simple way the mechanism that costs of rent-seeking activities may be increasing in firm size, making it harder for larger firms that are in the public eye to obtain subsidies and explaining why estimated subsidy rates τ are decreasing quickly for larger firms. θ_z is identified by how fast subsidies decrease with productivity.

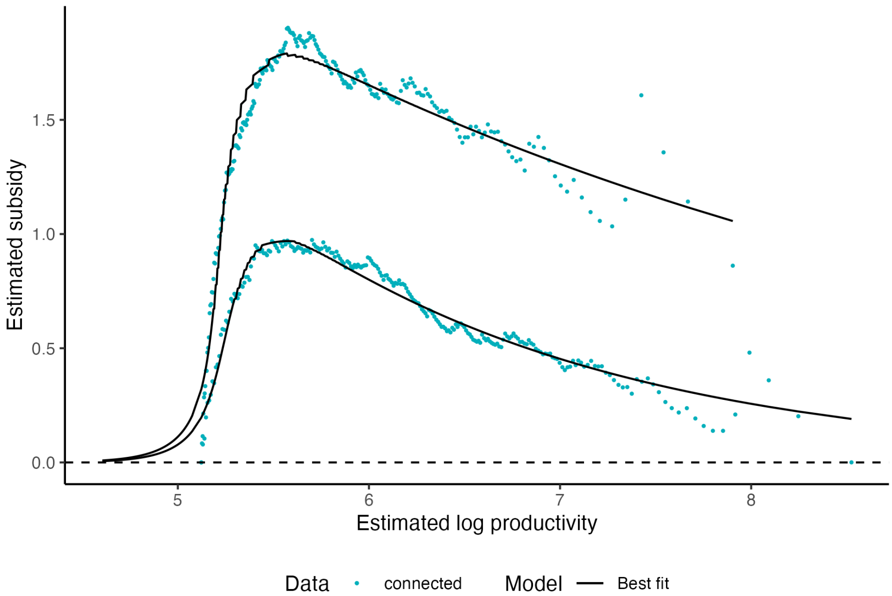
At last, I follow the previous model of rent-seeking in that connected firms employ workers and capital in a rent-seeking department to oversee all rent-seeking activities within the company. Since rent-seeking activities are not directly observed, I aggregate total rent-seeking in labor and capital using a standard Cobb-Douglas aggregator: $p \equiv k_p^{\eta} l_p^{1-\eta}$. I assume that $\eta = \frac{\alpha}{\alpha+\beta}$, which ensures that labor and capital are not differentially distorted, while ensuring a realistic relative role for capital and labor in rent-seeking. To the same effect, one could assume that rent-seeking activities p are intermediate goods bought from other firms.

Figure 2.4 shows the estimates for the lower and upper bound of the *Political Connections Technology*.[23] The economic model fits estimated subsidies almost perfectly, giving an R^2 of around

[22] Econometrically, estimating the *Political Connections Technology* also provides efficiency gains, because it can be used to reduce the variance in productivity and subsidy estimates.

[23] Parameters are estimated using non-linear least squares (NLS), minimizing the sum of squared residuals between the non-parametric subsidies and subsidies implied by the model. I use R's "L-BFGS-B" solver, which is a box constrained quasi-Newton method, to solve for optimal parameter values.

Figure 2.4: Model-implied subsidies vs. estimated subsidies



Notes: Fitting non-parametrically identified subsidies against estimates of subsidy based on functional form for the Political Connections Technology. Estimates are at the 1-digit industry level, considering a single production function across all firms. Parameters of the Political Connections Technology are chosen to minimize the sum of squared residuals between the non-parametric subsidies from the data and the implied, optimal subsidies from the model. The R^2 is 94.7 percent for the upper bound and 95.1 percent for the lower bound compared to a constant subsidy.

95% using a constant subsidy as the baseline comparison. Estimated parameters for the upper and lower bound are very similar and differ mostly in the "level" parameters. Both bounds exhibit strong decreasing returns to scale in benefits from rent-seeking activities ($\theta_p = 0.57 - 0.59$) and slightly convex costs ($\theta_c = 1.15 - 1.2$) in combination with sizable additional costs of rent-seeking activities in firm size ($\theta_z = 2.33 - 2.51$). These additional costs of size are larger for the lower bound and are important to match the faster decline in subsidies. The estimated model generates an interesting trade-off due to the dependence of rent-seeking activities on underlying productivity: Highly productive firms are also highly productive at rent-seeking activities, but they are large and visible in the public eye, making any rent-seeking activities riskier. On the other hand, less productive firms are less productive at rent-seeking activities, but they are also less prominent in the public eye, making it easier to avoid detection. Based on the model results, observed connected firms have positive subsidies because they are in the sweet spot where they are productive enough to generate subsidies in the presence of detection costs and not large enough yet to avoid too much public scrutiny. As Figure 2.4 shows, by extrapolating model-implied subsidies for lower productivities, these subsidies quickly go to zero as costs are too high compared to benefits from rent-seeking activities.

Based on the estimated model of rent-seeking, we can also look at implied spendings on rent-seeking activities. The average connected firm spends between 5-10% of total labor and capital costs on rent-seeking activities with large variation in these shares across connected firms. Connected firms with the highest subsidies spend as much as 22-27% of input spending on rent-seeking activities while spending quickly declines with firm size. The largest connected firms spend negligible shares on rent-seeking activities. How do these numbers compare to micro evidence on rent-seeking and lobbying activities by firms? Campos et al. (2021) look at judicial documents from the Odebrecht case, the anti-corruption case against a Brazilian engineering and construction conglomerate that bribed hundreds of politicians and political parties across Latin America. In this case, bribe payments alone were estimated to be around 1% of final project costs. Adding additional costs of rent-seeking going to lawyers and workers employed in rent-seeking activities, this observed magnitude is well in line with the economic model as long as we think of the Odebrecht conglomerate and its companies as being above average in size compared to other connected firms.

2.4 Quantifying the costs of political connections

This section quantifies the costs of political connections using the estimated subsidies and estimated *Political Connections Technology*. The main cost estimates are measured in output and welfare losses compared to the counterfactual economy where political connections are absent. In the last part of this section, I also quantify the benefits of public oversight to limit the role of political connections by studying counterfactual increases in auditing.[24] At last, I always separately estimate effects for the conservative upper and lower bounds (denoted LB and UB) and report bounds for all estimates throughout.

2.4.1 Baseline output and welfare losses from political connections

Table 2.2 reports the baseline estimates of the aggregate costs of political connections. The presence of political connections costs the economy between 1.0-4.7% of aggregate output and lowers aggregate wages between 1.3-5.6%. The baseline estimates come from comparing the distorted economy with political connections in 1997 to a counterfactual economy where subsidies to connected firms are entirely shut down and where the government lowers output taxes to the extent that total government revenue without subsidies stays constant. One might think of this setting as an economy where the government needs to finance a number of public goods that require a fixed amount of spending and can only do so via distortive corporate taxes. This counterfactual does not require to take a stand on how and why the government spends resources as government expenditures apart from spending on connected firms is kept constant. It is important to note that the baseline counterfactual does not abolish all distortions in the economy; it only reduces size distortions of connected firms to the benchmark of non-connected firms. I further assume that the 1997 distorted economy is in steady state and compare it to the steady-state of the counterfactual economy.[25] To solve for this counterfactual, I jointly solve for the wage that clears the labor market and the tax rate that keeps tax revenue equal to the baseline distorted economy. It turns out that abolishing subsidies to connected firms allows the government to reduce distortive output tax rates from 25% to at least 22% and up to 18%.

[24]Throughout, I use estimates of productivities and subsidies based on the estimated *Political Connections Technology*, which reduces estimation variance compared to the non-parametric estimates. The main results are almost indistinguishable when using original estimates.

[25]Transition dynamics between steady states in the static setup are not particularly interesting and relatively fast. Household savings slowly adapt to changes in capital demand. In the case where steady state capital demand rises, interest rates will first spike up and then converge back to the steady state interest rate. While labor supply is fixed, capital labor complementarity will also lead wages to slowly adapt. Relative consumption and output gains depend on household preferences, while steady-state comparisons allow to abstract from specifying them. In the case where the government productively invests tax revenue (considered further below), the stock of public capital increases only slowly, leading to slower transition dynamics and a more muted price response.

Before decomposing the aggregate costs of political connections, it makes sense to briefly discuss the baseline distorted economy as reported in Table 2.2. In the baseline distorted economy, the average connected firm is estimated to be between five to nine times larger than the average non-connected firm. This number is different from the ratio of 12 reported in Table 2.1, because we are now looking at firm gross output Y_{it} instead of net output $(1 - \tau_i)Y_{it}$ reported as value-added revenue in the data. Total output by connected firms accounts for up to 10% total output in the economy. At last, the government effectively subsidizes connected firms. Based on the upper bound estimates, the government spends about 25% of tax revenue on subsidizing connected firms.

The output and labor income costs of political connections split up into two main costs. First, is a *misallocation cost* where capital and labor is captured by highly subsidized firms instead of reallocating these resources to more productive firms in the economy. Once subsidies are eliminated, resources reallocate as connected firms downsize and prices in the economy adapt, leading more productive non-connected firms to demand more inputs. The second main cost is the *shadow cost of public funds*. There are costs to raising public funds for government expenditures captured by the distortion that output taxes bring and by the opportunity costs of public funds. In the baseline counterfactual, these costs are only captured by the distortion of the output tax. To quantify the pure *misallocation cost*, I consider a counterfactual economy where political connections are abolished, but where the *shadow cost of public funds* (the baseline distortive output tax) is kept constant. Any additional revenue gains in this counterfactual are redistributed lump-sum to households. As shown in Table 2.2, the *misallocation cost* in terms of total output makes up between zero to 45% of the costs of political connections. That is, for the lower bound, the negative effects from misallocating resources are exactly on par with the benefits from political connections.

To better understand the mechanisms that are driving the *misallocation cost*, consider first the partial equilibrium setting in which prices stay fixed (as reported in the first row of Table 2.2). The partial equilibrium counterfactual already reveals the extent of misallocation in the distorted economy as firm size differences are now only driven by differences in fundamental productivity. In partial equilibrium, abolishing political connections leaves choices of non-connected firms entirely unchanged but leads to a drastic reduction in subsidies to connected firms, leading them in turn to downsize by 40-80% for the average connected firm. In general equilibrium, the reduction in firm size by connected firms decreases demand for capital and labor, putting downward pressure on prices. Lower interest rates and lower wages incentivize all other firms in the economy to increase their capital and labor demand, leading to an increase in firm sizes for non-connected firms and pushing up prices again.[26]

As seen in Table 2.2 when reporting the relative contribution of the pure *misallocation cost*, these forces on net leave wages 3-6% lower than they were in the distorted economy with political connections. The drop in labor and capital demand from connected firms is only partly offset by non-connected firms. In principle, this could also be driven by freeing labor and capital from rent-seeking activities that can now be used in productive activities. However, according to the structural estimates, rent-seeking capital and labor comprise less than 0.2% of aggregate capital and labor and their effects on prices are correspondingly small. Households end up

[26]Note that in a stationary equilibrium, the interest rate stays unchanged because it is pinned down by household preferences: the fall in the interest rate leads households to dissave until the stationary interest rate is reached.

Table 2.2: Main results: Aggregate costs of political connections

Outcomes:	Output		Welfare		Wages		Govt Revenue	
	LB	UB	LB	UB	LB	UB	LB	UB
(Naive) PE costs	-4.11%	-4.92%	-5.4%	-5.73%	0.0%	0.0%	7.43%	26.61%
Baseline GE costs	**1.05%**	**4.67%**	**1.02%**	**4.56%**	**1.31%**	**5.59%**	**0%**	**0%**
Contribution:								
Misallocation	-2.43%	43.94%	40.61%	68.38%	-268.94%	-105.78%	12.01pp	35.9pp
Public funds	102.43%	56.06%	59.39%	31.62%	368.94%	205.78%	-12.01pp	-35.9pp
Institutions	**1.65%**	**6.76%**	**0.96%**	**5.11%**	**-1.9%**	**-1.57%**	**13.89%**	**42.17%**

Details: Costs using baseline subsidy estimates. Baseline general eq. (GE) costs are computed by comparing the observed distorted economy with a counterfactual economy where connections are shut down and distortive taxes are reduced such that govt revenue stays constant. Partial eq. results abolish subsidies to connected firms but keep prices fixed. The contribution of misallocation is quantified via a GE counterfactual where taxes stay constant and any additional tax revenue is redistributed lump-sum. Costs of worse institutions is based on an economy where govt resources are invested productively (see Section 4.2). All GE counterfactuals compare steady states. In steady state, the interest rate is pinned down by HH preferences and only the wage may change. LB and UB refer respectively to lower and upper bound estimates. Output refers to net production (without subsidies), Welfare costs are based on the percentage of consumption that households are willing to forego to keep welfare constant (and is equivalent to consumption changes here). Government revenue refers to revenue net of subsidies.

better off due to lump-sum transfers and redistributing higher firm profits. At last, given that tax revenue is saved from spending it on connected firms and output is increasing, total tax revenue from corporate taxes increases by 12-36%. About 2/3 of this increase is driven by eliminating transfers to connected firms as seen in partial equilibrium, while the remainder is driven by the general equilibrium response of output, leading all other firms to pay more taxes as they increase their output.

Given that the *misallocation cost* for the baseline estimates are up to 45% of the total output costs of political connections, the remaining 55% are explained by the *shadow cost of public funds*. These opportunity costs turn out to dominate the total costs of political connections as it would allow the government to additionally reduce distortive tax rates on all firms in the economy. According to the baseline estimates, this effect is strong enough to entirely reverse the negative wage effects when shutting down the *misallocation costs*, as all firms in the economy - especially the most productive firms - demand more labor and capital, driving up final wages by between 1.3 to 5.6%.

At last, based on the model estimates, a large part of the misallocation of resources happens across industries. That is, political connections also distort the relative size of industries, not just the allocation of resources within industries. Based on both bounds, about 10% of industries at the 4-digit level are larger due to political connections. These "connected industries" account for roughly 20% of total output and are between 14-24% larger due to political connections. However, the subsidization of connected firms within these industries comes at the detriment of the remaining 90% of industries that pay average output costs that range from 4-8%.

2.4.2 The costs of political connections from weakening institutions

The baseline estimates of the costs of political connections are based on a *shadow cost of public funds* that comes solely from distortive taxes. However, in developing countries, the costs of political connections likely also go through the quality of public goods and institutions more generally. Political connections deteriorate institutions and weaken the rule of law since subsidies crowd out resources that could be spent on alternatives. To capture this effect, I consider an alternative cost estimate based on a model where the government uses tax revenue productively to invest in productive public goods such as infrastructure, legal institutions, security and the enforcement of property rights. The subsidy and rent-seeking estimates are consistent with this reformulation; only the counterfactual changes, which now requires more assumptions.

Specifically, one needs to take a stand on how the government spends resources, how efficient it is at spending these resources on productive public goods and how public goods enter firms' production functions. I follow a standard setup for government investment (e.g. Ramey 2020). I assume that government-provided public goods enter firm TFP as: $\text{TFP}_i = (1+\tilde{\tau}_i)(1-\bar{\tau})\tilde{z}_i G^{\chi}$, where χ captures the private output elasticity of public goods. Call total tax revenue T and suppose that the government invests a constant fraction of this tax revenue productively (I^G) in public goods according to the following law of motion: $G' = (1-\delta_G)G + \kappa I^G$. Public goods G capture public capital such as roads and other infrastructure as well as any other resources the government spends that affect firm production, δ_G gives the depreciation rate of public goods and $\kappa \in (0,1]$ gives a measure of public funds that are misused (see: Pritchett 2000). I calibrate the additional parameters using either standard values or values specific to Indonesia.[27] To solve for the steady state counterfactual, I jointly solve for the wage that clears the labor market and the stationary value G^* that is consistent with steady state tax revenue T^*. I report results in the last row of Table 2.2.

Based on the quantitative results, considering the role on public goods exacerbates the aggregate costs of political connections by at least 40%. Since higher tax revenue leads to better institutions which in turn incentivizes firms to produce even more, this creates a positive feedback loop that drives up benefits from higher tax revenue. In the Indonesian data, the aggregate costs of political connections taking into account productive public goods are about 1.65-6.75% of total output, as captured by the comparison with the distorted economy with public goods. The positive feedback loop eventually leads to tax revenue and hence public good spending that is between 14-42% higher than in the distorted economy. Again, the large majority of these costs are driven by the *shadow costs of public funds*. Which numbers are more realistic - the baseline or public goods results - depends on what the Indonesian government would do with their revenue in the absence of favoring connected firms. The results in this paper indicate that investing the increased tax revenue in public goods would be far more effective than lowering tax rates, which turns out to be robust to alternative parameterizations

[27] Specifically, I assume a conservative long-run value for the output elasticity $\chi = 0.1$ taken from the literature (see Bom and Ligthart 2014). δ_G is assumed to be 3.25% following Arslanalp et al. (2010). κ is taken directly from the randomized controlled trial in Olken (2007) who studies corruption in infrastructure spending in Indonesia finding that on average about 24% of funds are lost in infrastructure projects financed by the central state, such that $\kappa = 0.76$. I assume the government spends a constant 35% of tax revenue on productive investments and returns the remainder as lump-sum transfers to households. I do not have a direct estimate of this fraction, but this number is 20% in the US (Ramey 2020). I vary this parameter and find that welfare benefits are higher, the lower this share. Hence, I take 35% as a conservative estimate.

of the benefits of public goods.[28]

2.4.3 Quantifying the benefits of public oversight

Numerous societal actors constrain the influence of politically connected firms by enforcing taxes and regulation, voicing concerns over legislature and executive orders that benefit connected firms and uncovering tax evasion and corruption. In this subsection, I use the structural model to provide evidence on the benefits of all these activities which I simply call public oversight. In the distorted baseline economy, connected firms endogenously choose their level of rent-seeking conditional on the observed level of public oversight in the Indonesian economy. To quantify the benefits of public oversight, we can now use the structural model to consider counterfactual scenarios in which we vary public oversight and study the endogenous responses of connected firms with respect to these changes in oversight. Specifically, I consider changes in public oversight by a common factor x, which, based on the structural model, simply shift the level of public oversight c in the Political Connections Technology.[29] Proportional changes in public oversight can be interpreted as proportional increases in the amount of tax audits, efforts of congressional oversight and the number of investigative reports by watchdogs such as Global Witness and Transparency International. While the costs of increasing audits is unobserved and hence a full cost-benefit analysis is beyond the scope of this paper, one can still quantify the output gains over the increase in audits x. To do so, I consider general equilibrium counterfactuals in which firms respond endogenously to the new level of public oversight and the aggregate level of tax revenue changes. To isolate the role of public oversight, I assume the government returns any additional revenue lump-sum to households and does not use it in a productive way (e.g. via lowering distortive taxes or investing extra revenue in public goods). Furthermore, note that the effectiveness of public oversight is an endogenous outcome in the model as it depends on how much connected firms invest in rent-seeking activities. More specifically, since $\theta_c > 1$ in the estimated Political Connections Technology, the marginal detection probability is decreasing in the level of audits, making any additional audit less effective, which connected firms also take into account.

To think about the benefits of public oversight, I consider the limit case of zero costs of auditing. Note that there is a trade-off even in the case of zero audit costs, because subsidizing connected firms can be beneficial up to the point where subsidies relax distortive taxes. This means that the maximum output gain over x gives an upper bound for the optimal amount of auditing. I find that while maxima differ widely across the lower and upper bound, both bounds suggest large output gains from increasing auditing. Based on the conservative lower bound estimates, maximal output gains for zero costs of audits are reached slightly below doubling auditing economy-wide and a social planner that cares about maximizing output should be willing to spend as much as 0.1% of GDP on audits. To get a rough idea of this magnitude, taking Indonesia's GDP in 1997, this fraction of GDP already amounts to roughly a 100-fold increase in the annual global budget of Transparency International in 2019.[30] For the upper bound,

[28] For example, more than doubling the depreciation rate of public goods to $\delta_G = 7.5\%$ per annum and halving the private output elasticity of public capital to $\chi = 0.05$, still gives general equilibrium output increases that are comparable to the constant tax revenue counterfactual.

[29] In Appendix A.6, I discuss in more detail the microfoundation of the Political Connections Technology as well as the exact interpretation of all parameters, including the level of public oversight c.

[30] See: https://www.transparency.org/en/the-organisation/our-operating-budget. Accessed on 12th May 2022.

output gains are as large as 2.1% of GDP, which are realized for increasing auditing threefold. Through the lens of the model, audits are effective despite not targeting heavily subsidized connected firms that pose especially large aggregate costs. Based on the model parameters, connected firms endogenously respond to a general increase in audits in such a way that audits also end up affecting the entire distribution of subsidies in a uniform way. There is an even stronger case for increasing audits in case they can be targeted at more heavily subsidized firms.

2.5 Extensions & Robustness results

In this section, I consider a variety of different robustness exercises and extensions. I start out by considering the role of variation in marginal revenue products, wedges and market power. I then turn to re-estimating non-parametric subsidies allowing for further industry- and connection-type variation, showing how the estimator can be extended to conditional matching on further observables.[31] At last, I consider how sensitive results are to measurement error and misreporting. Throughout, I relegate details to Appendix A.7. and focus mostly on the results.

2.5.1 Wedges and the costs of market power

The setup so far has abstracted from any variation in factor revenue shares. A large part of the misallocation literature following Hsieh and Klenow (2009) have used this variation to estimate wedges in static first-order conditions that pose aggregate costs for the economy. To discuss how this additional variation might affect results, Table 2.5 in Appendix A.2 reports evidence on observed labor and capital spending shares across and within industries. Connected firms have systematically lower observed factor revenue shares than non-connected firms. While factor revenue shares between the two groups of firms are only slightly lower for capital, there are large differences for labor. Differences in labor shares decline considerably when comparing shares within industries, but stay large. One more important feature of the data is that factor shares are up to 60% more dispersed for connected than for non-connected firms. Appendix A.2 provides further details and shows that these results are robust.

Through the lens of the model and in the spirit of Hsieh and Klenow (2009), systematic dispersion in factor revenue shares can be explained by reduced-form wedges that prevent firms from optimally choosing inputs. Specifically, we can define idiosyncratic labor and capital wedges based on the following distorted first-order conditions:

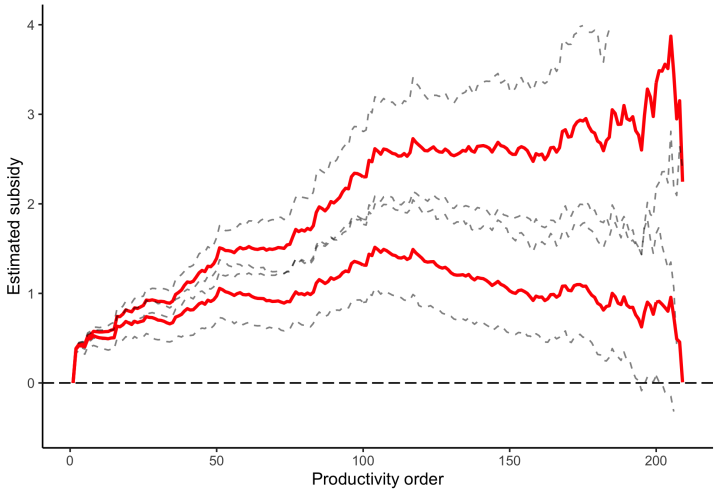
$$\alpha R_i^* = (1 + \tau_{iK}) r k_i^* \tag{2.5}$$

$$\beta R_i^* = (1 + \tau_{iL}) w l_i^* \tag{2.6}$$

where R_i^* is observed optimal firm revenue, τ_{iK} and τ_{iL} are firm-specific wedges for capital and labor choices and k_i^* and l_i^* are productive capital and labor inputs. Assume that all firms

[31] Beyond heterogeneity in industries and types, one could also control for other observables. For example, one could further match on location, ownership structure such as whether the firm belongs to a conglomerate, or even same firms who saw changes in their connections status over time. Unfortunately, I do not observe conglomerate association nor changes in connections over time. And, as I explain below, increasing the number of controls requires observing sufficiently many connected firms.

Figure 2.5: Estimated subsidies with firm-specific input cost wedges



Notes: Non-parametric estimates of conservative subsidy bounds additionally allowing for firm-specific wedges in capital and labor input costs. Red lines give point estimates formed by average estimated productivities across bootstrap samples respectively for the lower and upper bound. Grey, dashed lines give point-wise 95% bootstrap confidence bands using 10,000 bootstrap samples. Lower bound estimated to be given by $p = 0.39$, implying that the most productive connected firm was about 10 times more likely to become connected than the least productive connected firm. Estimates based on assumptions explained in the text. Connected (N = 209, after dropping same outliers as for baseline) vs. non-connected (N = 14,713) firms with observed inputs are identified as in Mobarak & Purbasari (2006).

report their productive inputs plus some potentially firm-specific fraction of their rent-seeking activities (that is: $\tilde{k}_i \in [k_i^*, k_i^* + k_{ip}^*]$ and similarly for labor). Then it follows directly that lower observed factor revenue shares translate into connected firms facing higher wedges, indicating that they face higher implicit input costs, which Hsieh and Klenow (2009) interpret as size restrictions.[32]

To quantify how wedges affect the costs of political connections, I reestimate subsidies, the *Political Connections Technology* and general equilibrium counterfactuals allowing for wedges. The estimation approach for subsidies of connected firms remains unchanged assuming that connected firms are still only selected based on productivity (and not directly on wedges). This rules out quid-pro-quo benefits where subsidies are offered conditional on how connected firms choose inputs.

Figure 2.5 shows conservative upper and lower bound estimates of subsidies. In comparison to the baseline estimates in Figure 2.3, subsidies allowing for additional idiosyncratic wedges are roughly 40% higher. This is driven by higher and more dispersed wedges for connected firms.[33] While wedges do lead to more heterogeneity, the newly estimated *Political Connections Technology* still explains about 70% of the variation in subsidies. The key economic mechanisms stay unchanged: benefits of political connections continue to exhibit decreasing returns to scale ($\theta_p \approx 0.52 - 0.56$) and costs of political connections are convex in rent-seeking activities ($\theta_c \approx 1.25 - 1.28$) as well as firm size ($\theta_z \approx 2.08 - 2.26$).

[32]The interpretation of positive wedges as size restrictions also holds for a dynamic setting: in the case where within industry variation in revenue factor shares is driven by dynamic input choices such as with time to build capital and labor or factor adjustment costs, observing lower factor shares would mean connected firms face higher or more binding adjustment costs (e.g. see Asker, Collard-Wexler, and De Loecker 2014).

[33]Note that the productivity threshold $\bar{z}$ is re-estimated so that most of the changes in the subsidy estimates are driven by the increased dispersion in wedges for connected firms rather than the level difference.

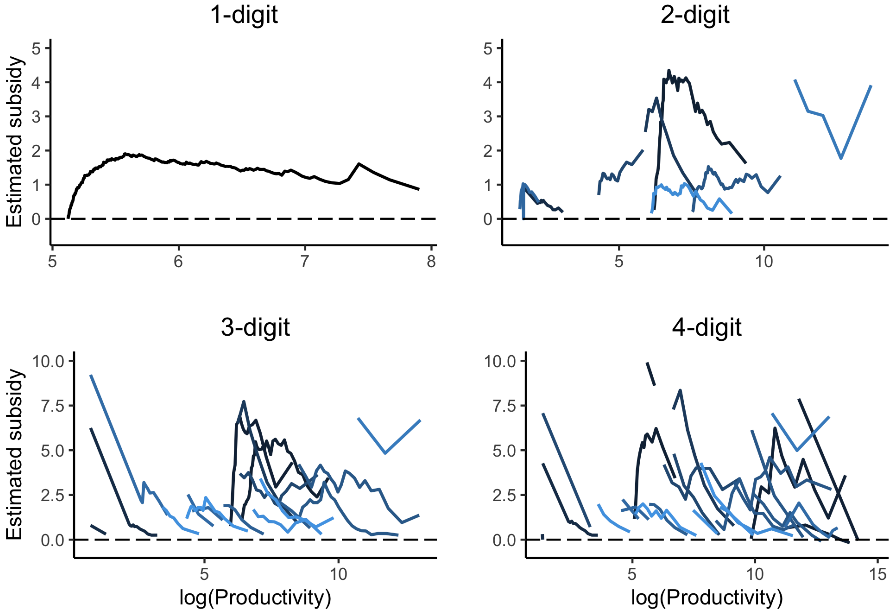
To consider the aggregate costs of political connections, one needs to take a stand on what estimated wedges capture and how these wedges may change in a counterfactual economy where political connections are abolished. Building on good evidence for Indonesia that connected firms are in less competitive industries (Hallward-Driemeier, Kochanova, and Rijkers 2021), that connected firms are much more likely to receive licenses that buy them market power (Mobarak and Purbasari 2006) and that there is a positive correlation between within-industry firm size and profit shares (shown in Appendix A.2), I interpret higher wedges of connected firms (which capture higher profit shares) as being primarily driven by market power. The benefit of the subsidy approach is that each connected firm automatically has a matched sample of comparable non-connected firms who have not benefited from political connections. One can then directly use this firm-specific matched set of comparable firms to infer counterfactual wedges. To maintain realistic variation in wedges in the counterfactual, I bootstrap 10,000 counterfactuals in which I randomly sample a single wedge for each connected firm among the set of wedges of matched firms. Abstracting from aggregate changes in market power due to abolishing political connections, connected firms in this counterfactual will only lose the market power that is associated to their connections and not the market power that they would have in either case because of their high productivity. Table 2.3 shows that aggregate costs of political connections additionally taking into account wedges and market power are at least 30% higher than the benchmark costs. Output costs are more precisely estimated and lie between 4.5-6%. Furthermore, we can quantify the contribution of the market power channel by estimating costs without re-drawing wedges. I find that market power of connected firms as measured by the counterfactual reduction in the dispersion of wedges drives between 10-35% of the total costs of political connections.

2.5.2 Industry heterogeneity

In this subsection, I consider subsidy estimates and costs of political connections under more industry heterogeneity, extending the baseline matching estimator to matching conditional on further observables. The non-parametric within-industry estimator separately draws productivities from non-connected firms within the same industry and matches firms accordingly. This introduces a trade-off as within industry matching matches firms that are more similar while at the same time reduces both the population from where productivities can be drawn and the sample with which one can match.

Subsidy estimates for four different levels of production function heterogeneity are reported in Figure 2.6 showing only upper bound estimates for expositional clarity. Within each panel, each line marks one different industry at the respective digit. Three patterns are noteworthy. First, comparing the 1-digit to the 2-digit results, we can see that the concave or even hump-shaped *Political Connections Technology* does show up in most industries (6 out of 8), indicating that it is an important feature of the data. The two 2-digit industries where this pattern is less clear have only few observations, leading to very noisy estimates. While results are slightly harder to interpret at the 3-digit and 4-digit level, we can still see hump-shape relationships between the subsidy and productivity within many industries. Secondly, the level of the subsidies increases slightly when allowing for more production function heterogeneity, reflecting slightly larger size differences within industries as reported in Table 2.1. At last, productivity estimates change considerably across the different specifications since production function elasticities now vary across industries.

Figure 2.6: Estimated subsidies with industry heterogeneity



Notes: Non-parametric identification of subsidies to connected firms as function of estimated firm-level productivities with industry heterogeneity. Panel 1-4 give estimates at 1-, 2-, 3- and 4-digit industry levels respectively using the upper bound estimator for productivities. Estimation is explained in the text. Connected (N = 238, dropped 3 outliers) vs. non-connected (N = 18,317) firms are identified as in Mobarak & Purbasari (2006).

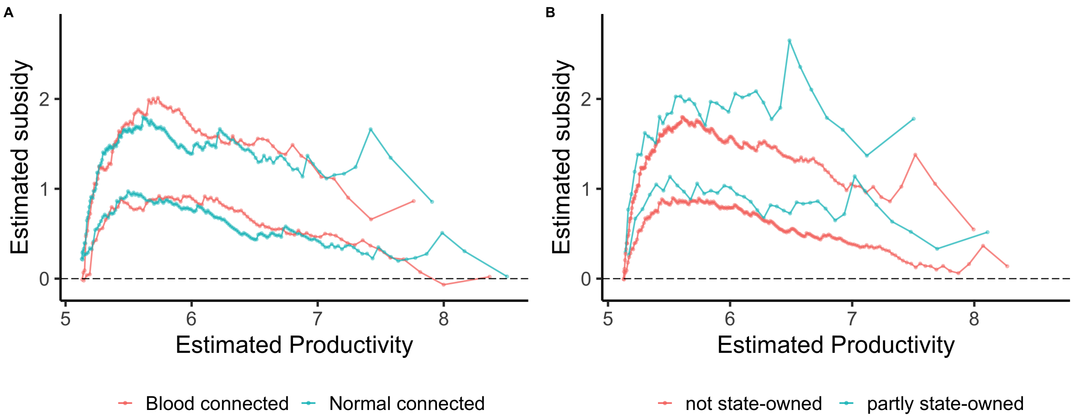
For the costs of political connections, I re-estimate the *Political Connections Technology* taking into account further heterogeneity. I find that estimated parameters at the 2-digit level maintain decreasing returns to scale in benefits from rent-seeking activities and convex costs both in rent-seeking activities and firm size. The R^2 of the noisier estimates at the 2-digit industry level still exceeds 95%. However, as for the baseline results, rent-seeking activities have only a very limited effect on aggregate effects because they only raise aggregate capital and labor by a small amount. The key effects go through extensive subsidies and firm-size distortions which misallocate productive capital and labor at marginally unproductive firms. Estimated costs of political connections turn out to be similar at the 2-digit level and much higher at the 3-digit level, which should be interpreted with care given noisier subsidy and rent-seeking activity estimates. In summary, results are broadly similar when allowing for more heterogeneity across industries.

2.5.3 Testing for further types of connected firms

Another source of heterogeneity is to consider different types of connected firms ε. For example, one might expect that some firms close to Suharto are just lucky being born into the right network, while others have worked hard and paid a high cost to obtain favors. To test for such differences, I consider two sets of type heterogeneity. For the estimation, I enforce the same productivity cutoff and the same conservative selection bounds as for the main estimates, in line with the null hypothesis of identical *Political Connections Technologies*. For each set of type heterogeneity, I then separately draw and match bootstrap samples for the differently-connected firms to obtain productivity estimates.

Panel A of Figure 2.7 reports separately estimated subsidies and productivities for “normal”

Figure 2.7: Estimated subsidies with connection-type heterogeneity



Notes: Non-parametric estimation of output subsidies and productivity allowing for different Political Connections Technologies for different types of connections. Panel A: Connection types distinguish blood connected (N = 89) and normal connected (N = 152) firms (details in Section 2). Panel B: Distinguishes connected firms that are partly state-owned (N = 39) and connected firms that are not state-owned (N = 199).

connected firms and for firms directly owned, founded and run by blood relatives of dictator Suharto. One can think of many reasons why *Political Connections Technologies* should look differently for the two sets of connected firms and why the latter set of firms should receive larger subsidies. Perhaps surprisingly, estimated *Political Connections Technologies* look almost indistinguishable, with blood connected firms receiving slightly higher subsidies. I also formally test equality of *Political Connections Technologies* by considering bootstrapped confidence bands and cannot reject equality.[34] Panel B shows separately estimated subsidies and productivities for connected firms that are at least partly state-owned and connected firms that are not. One concern with the estimates might be that state ownership changes the relationship between connected firms and politicians and thus leads to very different *Political Connections Technologies*. Again, I find no evidence for this in the data. While estimated subsidies are slightly larger for connected firms that are state-owned (in line with economic intuition), the distribution of subsidies looks very similar and as for the previous results I cannot reject equality from a statistical point-of-view. These are encouraging findings for the paper, because it alleviates concerns that unobserved type-heterogeneity or a few connected firms are biasing the results, lending credence to the baseline results. Similarity in the subsidy estimates also shows up when estimating the aggregate costs of political connections: I find largely similar, though slightly higher estimated costs as reported in Table 2.3.

2.5.4 Measurement error

The non-parametric subsidy estimation seems to rely crucially on the assumption that value-added output is correctly reported, because conditional on the production function parameters, estimated TFP only relies on variation in reported value-added output. Estimated subsidy schedules are then inferred from the relative dispersion in TFP between connected and non-connected firms. This subsection considers how sensitive subsidy estimates are in the presence

[34]Formally, I only consider point-wise overlap in confidence bands, which is not a full statistical test. However, for most points, confidence bands include point estimates, which is sufficient for rejection. Statistical power of this test is obviously limited given the few number of connected firms by type, but differences in point estimates are also economically small.

of measurement error in value-added output.

I consider four types of measurement error and bootstrap the estimation of subsidies, *Political Connections Technology* and welfare costs for both bounds. Figure 2.9 in Appendix A.7 reports average alternatively estimated subsidies across bootstrap draws for each type of measurement error. Table 2.3 reports the corresponding welfare effects. The first three types of measurement error affect all firms, connected and non-connected, and have mean zero. In Panels A and B, I consider classic, normally distributed and non-symmetric log-normally distributed measurement error. Only the second type affects the right tail of the observed output distribution and hence the magnitude of estimated subsidies. However, both forms of measurement error have little effects on the overall shape of the estimated subsidy schedules nor estimated welfare costs.

To consider a potentially more problematic case of bias, I consider measurement error that correlates directly with firm size in Panel C. Perhaps surprisingly, it turns out that this form of correlated measurement error also leaves subsidy estimates and estimated welfare costs basically untouched. The reason for this is that correlated measurement error does not affect the relative dispersion of output distributions across connected and non-connected firms. To also consider the effect of differential output distortions, I introduce measurement error that only affects connected firms in Panel D. Specifically, I assume that all connected firms systematically underreport a fixed 20% of output.[35] The results again are robust to this form of measurement error. The reason is that such selective underreporting is entirely captured by the productivity cutoff. A higher percentage of underreporting simply leads to a lower estimated cutoff, leaving the estimated levels unbiased. If anything, the upper bound welfare cost estimates are now somewhat higher. In summary, the reported estimates in this paper are robust to various forms of standard and non-standard measurement error.

2.6 Conclusion & Discussion

This paper has provided a structural approach to quantify the general equilibrium costs of political connections. Using a model where firms endogenously invest in rent-seeking activities to obtain firm-specific subsidies, I showed how to non-parametrically identify conservative bounds for these subsidies and flexibly estimate the technology with which firms invest in rent-seeking activities. Applying this methodological approach to Indonesia, I find high aggregate costs of political connections. Costs are between 1.0-4.7% of annual output and higher when accounting for market power and effects on public goods.

A number of qualifications of the results are in order. While the quantitative results are robust to further industry- and connection type-heterogeneity, wedges and different forms of measurement error, some issues are harder to assess. For example, due to data constraints, the focus of this paper has been on manufacturing plants. Political connections may play a different role in other sectors and at the firm-level. Furthermore, political connections will always remain elusive, making measurement of them difficult. This paper's measure of political connections is based on a natural experiment and arguably the most credible estimates we have. One complementary avenue for future research is to collect more direct evidence on

[35] I considered underreporting of 5%, 10% and 20% respectively, but results were all quantitatively similar and I report the results with the most measurement error.

Table 2.3: Main robustness and extension results: Aggregate costs of political connections

Outcomes:	Output		Welfare		Wages	
	LB	UB	LB	UB	LB	UB
Baseline GE costs	**1.05%**	**4.67%**	**1.02%**	**4.56%**	**1.31%**	**5.59%**
Market Power + Wedges:						
Full costs	4.57%	5.97%	4.89%	6.05%	7.65%	12.66%
Market power contrib. (%)	35.81%	10.22%	35.45%	9.56%	18.56%	8.97%
Industry heterogeneity:						
2-digit	1.09%	3.53%	0.96%	3.3%	2.99%	6.37%
3-digit	9.4%	12.23%	8.81%	12.89%	7.18%	4.91%
Type heterogeneity:						
Blood vs. normal	1.59%	4.66%	1.54%	4.54%	2.01%	5.59%
State-owned vs. not	2.03%	4.82%	1.97%	4.69%	2.58%	5.78%
Measurement Error:						
Classic	1.19%	5.07%	1.15%	4.91%	1.52%	6.15%
Non-symmetric	1.7%	5.37%	1.64%	5.24%	2.13%	6.38%
Correlated	1.17%	4.27%	1.12%	4.13%	1.47%	5.18%
Underreporting C	1.22%	5.42%	1.17%	5.24%	1.54%	6.58%

Details:

Aggregate costs of political connections under various robustness exercises and model extensions. Throughout, general eq. costs are computed by comparing the observed distorted economy with a counterfactual economy where connections are shut down and distortive taxes are reduced such that government revenue stays constant. All general equilibrium counterfactuals compare steady states. LB and UB refer respectively to lower and upper bound estimates. Output refers to net production (without subsidies), Welfare costs are based on the percentage of consumption that households are willing to forego to keep welfare constant (and is equivalent to consumption changes here). Government revenue refers to revenue net of subsidies.

rent-seeking activities and use this to validate the model-implied distribution of rent-seeking activities.

2.7 References

Abeberese, Ama Baafra, Prabhat Barnwal, Ritam Chaurey, and Priya Mukherjee. 2021. "Democracy and Firm Productivity: Evidence from Indonesia." *The Review of Economics and Statistics*, 1–30.

Akcigit, Ufuk, Salom'e Baslandze, and Francesca Lotti. forthcoming. "Connecting to Power: Political Connections, Innovation, and Firm Dynamics." *Econometrica*, forthcoming.

———. 2018. "Connecting to Power: Political Connections, Innovation, and Firm Dynamics." National Bureau of Economic Research.

Amiti, Mary, and Jozef Konings. 2007. "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia." *American Economic Review* 97 (5): 1611–38.

Arayavechkit, Tanida, Felipe Saffie, and Minchul Shin. 2018. "Capital-Based Corporate Tax Benefits: Endogenous Misallocation Through Lobbying." Working Paper.

Arslanalp, Serkan, Fabian Bornhorst, Sanjeev Gupta, and Elsa Sze. 2010. "Public Capital and Growth." Citeseer.

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

Aslund, Anders. 2019. *Russia's Crony Capitalism*. Yale University Press.

Bai, Chong-En, Chang-Tai Hsieh, and Zheng Song. 2020. "Special Deals with Chinese Characteristics." *NBER Macroeconomics Annual* 34 (1): 341–79.

Bertrand, Marianne, Francis Kramarz, Antoinette Schoar, and David Thesmar. 2018. "The Cost of Political Connections*." *Review of Finance* 22 (3): 849–76. https://doi.org/10.1093/rof/rfy008.

Bom, Pedro RD, and Jenny E. Ligthart. 2014. "What Have We Learned from Three Decades of Research on the Productivity of Public Capital?" *Journal of Economic Surveys* 28 (5): 889–916.

Brugu'es, F., J. Brugu'es, and S. Giambra. 2018. "Political Connections and Misallocation of Procurement Contracts: Evidence from Ecuador."

Campos, Nicol'as, Eduardo Engel, Ronald D. Fischer, and Alexander Galetovic. 2021. "The Ways of Corruption in Infrastructure: Lessons from the Odebrecht Case." *Journal of Economic Perspectives* 35 (2): 171–90.

Carney, Richard W., and Travers Barclay Child. 2013. "Changes to the Ownership and Control of East Asian Corporations Between 1996 and 2008: The Primacy of Politics." *Journal of Financial Economics* 107 (2): 494–513. https://doi.org/10.1016/j.jfineco.2012.08.013.

Chen, Ting, and James Kai-sing Kung. 2018. "Busting the 'Princelings': The Campaign Against Corruption in China's Primary Land Market." *The Quarterly Journal of Economics* 134 (1): 185–226.

Chernozhukov, Victor, and Christian Hansen. 2005. "An IV Model of Quantile Treatment Effects." *Econometrica* 73 (1): 245–61.

Claessens, Stijn, Simeon Djankov, and Larry HP Lang. 2000. "The Separation of Ownership and Control in East Asian Corporations." *Journal of Financial Economics* 58 (1-2): 81–112.

Costinot, Arnaud, and Andr'es Rodríguez-Clare. 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization." In *Handbook of International Economics*, 4:197–261. Elsevier.

Cremer, Helmuth, and Firouz Gahvari. 1994. "Tax Evasion, Concealment and the Optimal Linear Income Tax." *The Scandinavian Journal of Economics*, 219–39.

Diwan, Ishac, Adeel Malik, and Izak Atiyas. 2019. *Crony Capitalism in the Middle East: Business and Politics from Liberalization to the Arab Spring*. Oxford University Press.

Do, Quoc-Anh, Kieu-Trang Nguyen, and Anh N. Tran. 2017. "One Mandarin Benefits the Whole Clan: Hometown Favoritism in an Authoritarian Regime." *American Economic Journal: Applied Economics* 9 (4): 1–29. https://doi.org/10.1257/app.20130472.

Doksum, Kjell. 1974. "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case." *The Annals of Statistics*, 267–77.

Faccio, Mara. 2006. "Politically Connected Firms." *American Economic Review* 96 (1): 369–86.

Fentanes, Oscar, and Jonas Gathen. 2022. "Growth and the Plant Size Distribution over the Long-Run." *Working Paper*.

Firpo, Sergio. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *Econometrica* 75 (1): 259–76.

Fisman, Raymond. 2001. "Estimating the Value of Political Connections." *American Economic Review* 91 (4): 1095–1102. https://doi.org/10.1257/aer.91.4.1095.

Fisman, Raymond, and Yongxiang Wang. 2015. "The Mortality Cost of Political Connections." *The Review of Economic Studies* 82 (4): 1346–82.

Garcia-Santana, Manuel, Enrique Moral-Benito, Josep Pijoan-Mas, and Roberto Ramos. 2019. "Growing Like Spain: 1995-2007." *International Economic Review* forthcoming.

———. 2020. "Growing Like Spain: 1995–2007." *International Economic Review* 61 (1): 383–416.

Gonzalez, Felipe, and Mounu Prem. 2019. "Losing Your Dictator: Firms During Political Transition." *Working Paper*. https://doi.org/10.2139/ssrn.2670869.

Gonzalez, Felipe, Mounu Prem, and Francisco Urz'ua. 2018. "The Privatization Origins of Political Corporations." *Working Paper*, 46.

Guner, Nezih, Gustavo Ventura, and Yi Xu. 2008. "Macroeconomic Implications of Size-Dependent Policies." *Review of Economic Dynamics* 11 (4): 721–44.

Hadiz, Vedi R., and Richard Robison. 2013. "The Political Economy of Oligarchy and the Reorganization of Power in Indonesia." *Indonesia* 1 (96): 35–57.

Hale, Christopher D. 2001. "Indonesia's National Car Project Revisited." *Asian Survey* 41 (4): 629–45.

Hallward-Driemeier, Mary, Anna Kochanova, and Bob Rijkers. 2021. "Does Democratisation Promote Competition? Evidence from Indonesia*." *The Economic Journal* 131 (640): 3296–3321. https://doi.org/10.1093/ej/ueab023.

Haselmann, Rainer, David Schoenherr, and Vikrant Vig. 2018. "Rent Seeking in Elite Networks." *Journal of Political Economy* 126 (4): 1638–90.

Hoang, Kimberly Kay. 2018. "Risky Investments: How Local and Foreign Investors Finesse Corruption-Rife Emerging Markets." *American Sociological Review* 83 (4): 657–85.

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.

Huneeus, Federico, and In Song Kim. 2021. "The Effects of Firms' Lobbying on Resource Misallocation." *Working Paper*.

Johnson, Simon, and Todd Mitton. 2003. "Cronyism and Capital Controls: Evidence from Malaysia." *Journal of Financial Economics* 67 (2): 351–82.

Koren, Miklos, Adam Szeidl, Ferenc Szucs, and Balazs Vedres. 2015. "Political Favor Exchange in a Democracy." *Unpublished Working Paper*.

Lehmann, Erich Leo, and Howard J. D'Abrera. 1975. *Nonparametrics: Statistical Methods Based on Ranks.* Holden-day.

Martinez-Bravo, Monica, Priya Mukherjee, and Andreas Stegmann. 2017. "The Non-Democratic Roots of Elite Capture: Evidence From Soeharto Mayors in Indonesia." *Econometrica* 85 (6): 1991–2010.

Mobarak, Ahmed Mushfiq, and Denni Puspa Purbasari. 2006. "Corrupt Protection for Sale to Firms: Evidence from Indonesia." *Unpublished Working Paper, University of Colorado at Boulder*.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.

Pritchett, Lant. 2000. "The Tyranny of Concepts: CUDIE (Cumulated, Depreciated, Investment Effort) Is Not Capital." *Journal of Economic Growth* 5 (4): 361–84.

Ramey, Valerie A. 2020. "The Macroeconomic Consequences of Infrastructure Investment." National Bureau of Economic Research.

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

———. 2017. "The Causes and Costs of Misallocation." *Journal of Economic Perspectives* 31 (3): 151–74.

Rijkers, Bob, Caroline Freund, and Antonio Nucifora. 2017. "All in the Family: State Capture in Tunisia." *Journal of Development Economics* 124 (January): 41–59. https://doi.org/10.1016/j.jdeveco.2016.08.002.

Robison, Richard, and Vedi Hadiz. 2004. *Reorganising Power in Indonesia: The Politics of Oligarchy in an Age of Markets*. Routledge.

Schoenherr, David. 2019. "Political Connections and Allocative Distortions." *The Journal of Finance* 74 (2): 543–86.

Slemrod, Joel. 2001. "A General Model of the Behavioral Response to Taxation." *International Tax and Public Finance* 8 (2): 119–28.

Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Chapter 22 - Tax Avoidance, Evasion, and Administration." In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein, 3:1423–70. Handbook of Public Economics. Elsevier. https://doi.org/https://doi.org/10.1016/S1573-4420(02)80026-X.

Straub, St'ephane. 2014. "Political Firms, Public Procurement, and the Democratization Process." *Institut d'Economie Industrielle (IDEI), Working Paper* 1 (817).

Szucs, Ferenc. 2017. "Discretion and Corruption in Public Procurement." *Job Market Paper*.

2.8 Appendix

2.8.1 Further details on measuring political connections

Mobarak and Purbasari (2006) extend the work by Fisman (2001) by examining how the stock price of the universe of firms traded on the Jakarta Stock Exchange (JSX)[36] responded to adverse news about Suharto's health in various episodes between 1994 and 1997. Using daily stock price data for the 985 market trading days between 1994 and 1997, they run a set of regressions of abnormal stock returns for each firm on aggregate movements in the JSX, the average return for the industry category in which that firm belongs, movements in the exchange rate and interest rate, and an indicator variable for days when the news about Suharto's health was reported by the press. A firm is defined to be "politically connected" if the Suharto health news indicator has a negative coefficient which is significantly different from zero at the 95% confidence level. Using statistical significance as a threshold gives a firm-specific threshold that also takes into account the firm-specific variability of its stock price.[37] This identifies 29 stock listed firms as being politically connected and the authors used newspapers and other media to confirm that these firms were indeed connected.

The identities of the key personnel running these 29 politically connected firms allow Mobarak and Purbasari (2006) to identify, by proxy, other firms that are connected to Suharto, but not traded on the Jakarta Stock Exchange. The authors do this by locating all other firms that share ownership and management with those 29 firms. As Claessens, Djankov, and Lang (2000) and Carney and Child (2013) show, ownership and control is rarely separated in Southeast Asian firms including Indonesia and most firms belong to larger conglomerate structures that are owned by specific families. This allows to link stock-listed firms to a larger network of other firms of the same conglomerate, who are owned by the same family. Due to the prevalence of political connections being tied to interpersonal links between families, this allows to track connected firms beyond stock-listed firms. Specifically, Mobarak and Purbasari (2006) identify each member of the Board of Directors and Board of Commissioners of each of the 29 firms using the Indonesian Capital Market Directory 1998. They then use the publication *400 Prominent Indonesian Businessmen* to find the names of all conglomerates to which the individuals running the connected firms belong. Finally, they turn to *Conglomeration Indonesia* to identify all subsidiary firms of the 'connected' business groups and trace all other firms and conglomerates that share ownership and management. In total, this gives them 2,126 connected firms.

The implicit assumption at this point is that all relevant political connections in Indonesia go through larger conglomerates which have at least one publicly traded firm that is identified as being politically connected. Thus, this definition of political connections captures "high-level" political connections and is unlikely to capture more local connections of firms to local authorities in the bureaucracy or police. This should be kept in mind when interpreting the results

[36] The authors estimate this for 285 of the 293 firms traded on the Jakarta Stock Exchange at that time.

[37] The authors use three different definitions of firm stock returns, including the actual return, the deviation of the actual return from its average, and the abnormal return net of movements correlated with the aggregate JSX market return. They also variably define the event dates to be the day the illness occurs or the day it is reported in the press. The identities of 'politically connected' firms are roughly invariant to the particular definition of returns or event dates used. Note that using statistical significance as a filter may introduce differential bias by size. If the variability of stock prices is related to fundamentals such as firm size then statistical power will vary by size of firm and then selection will be worse for smaller firms. I have not conducted tests or simulations to assess this concern, but given that T = 985, it seems likely that power is not a relevant concern.

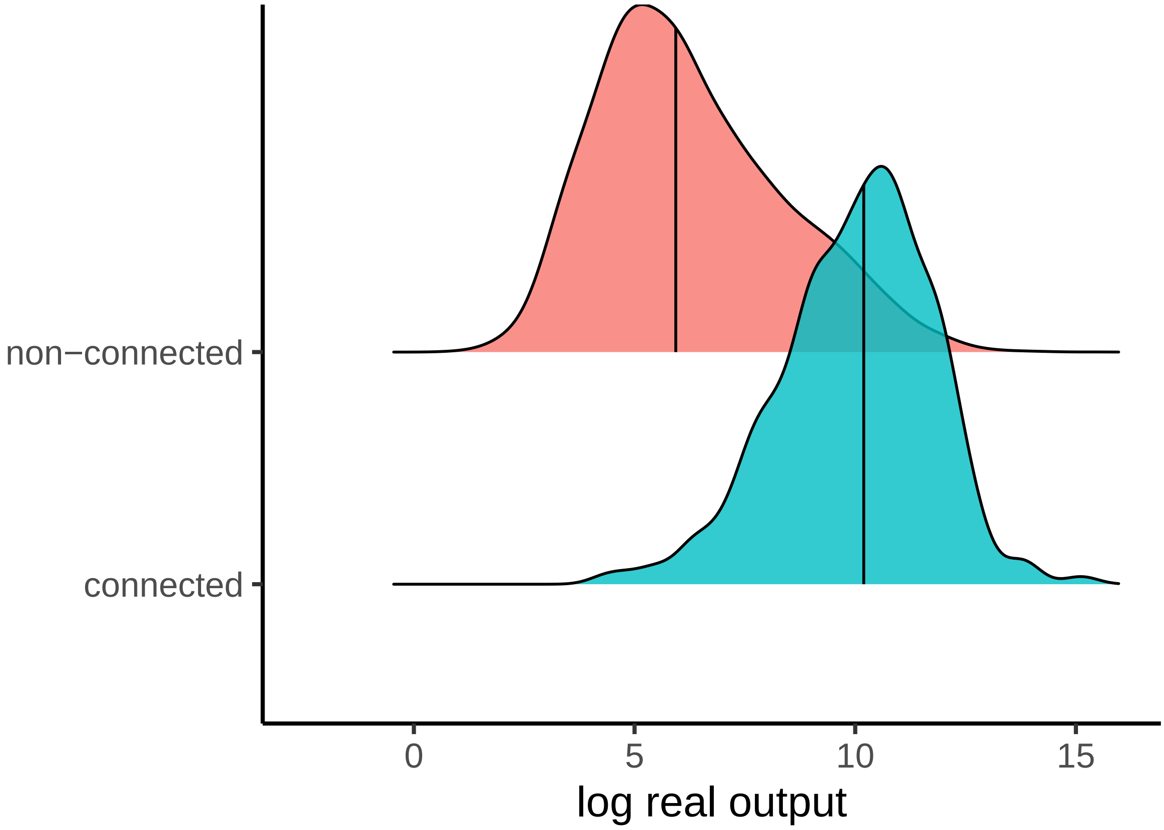
in this paper. Another key concern of using this measure of connections is that it is likely to capture only larger firms and is more likely to miss small connected firms. In the structural approach used later in the paper, results will explicitly depend on the smallest observed connected firms exactly to be robust to the idea that if all connected firms are large and successful this must not imply that connections are very beneficial, but could also be driven by the fact that we do not capture smaller and less successful connected firms in the data.

The next limitation of the data is that while the approach allows to identify a variety of connected firms, the available firm-level data to link these to is the annual manufacturing census data that captures medium- and large-sized manufacturing firms with more than 20 employees. This considerably restricts the sample: only 16 of the 29 initial stock-listed firms and 408 of the 2,126 identified connected firms are manufacturing firms, which makes up roughly 20% of firms. Based on the GGDC 10-sector database, the manufacturing sector accounted for about 34% of value-added output in 1997, which is squarely between the percentage of manufacturing firms among stock-listed firms and the percentage among all connected firms. It is unclear exactly what biases this sample selection introduces, but it may even lead to more conservative estimates of the costs of political connections given that connections are likely to play a bigger role in a number of non-manufacturing sectors such as utilities (including telecommunications and energy), mining, construction, finance and land-dependent agriculture. Of these manufacturing firms, linking them to the census is further complicated by the fact that firms are generally de-identified in the manufacturing census data. Using three broad identifying variables - province location, 5-digit industry code and (rough) number of employees - Mobarak and Purbasari (2006) can succesfully match 241 firms or 59% of connected firms to the census of manufacturing firms. Mobarak and Purbasari (2006) argue that the attrition involved in this matching step is not related to any fundamentals and should thus not differentially bias the estimates apart from underestimating the number of connected firms.[38]

In the end, this approach allows to identify 241 connected firms in the manufacturing census data. It allows to identify the snapshot of politically connected firms at the highest level for a short time period of around 1-2 years shortly before the Asian Financial crisis in 1997/8. Throughout the paper, I allow the set of connected firms to vary over time with some firms loosing their connections or seeing changes in the extent of their connections, but all results will be based on the set of connected firms in 1997 and I therefore assume that this is a representative picture of connected firms also for other years in the data. Of the 241 firms, 89 firms are identified as being owned and founded by blood connections of Suharto. 34 of these 89 firms are similarly identified as being connected by the stock market identification approach. This imperfect overlap may be due to three different problems. First, it may show that the stock market identification approach is highly imperfect in capturing all connected firms (only about 40% of connected firms are identified). This could be due to the nature of the approach only capturing firms that are linked through conglomerates that have a stock listed firm or the

[38]However, I have not been able to validate this claim and replicate this part of their analysis given that the authors could not share this part of the analysis with me. There could be a number of reasons why the matching step could introduce additional problems. For example, matching by (rough) number of employees may introduce bias against small firms as this set of firms may include more overlap in the number of employees and thus makes it less likely to find unique matches in the data. On the other hand, matching by province location may make it harder to match more successful firms in more economically active parts of the country (e.g. Java). Access to the set of all connected manufacturing firms could allow to control for potential differential misclassification in this step of the analysis.

Figure 2.8: (Gross) Output distributions: Connected vs. non-connected firms



Notes: Distributions of firm-specific real gross output (in logs and in 2010 USD) for cross-section of Indonesian firms in 1997 based on Statistik Industri, the Indonesian manufacturing firm census. Connected vs. non-connected firms are identified as in Mobarak & Purbasari (2006). Non-connected firms: N = 18,303. Connected firms: N = 241.

statistical uncertainty in the estimates, but it could also be because the approach only captures connected firms whose connections are deemed sufficiently volatile. These issues only pose a real problem for this paper if they bias the identified size distribution of connected firms, otherwise, this paper will only underestimate the costs of political connections. Second, imperfect overlap may indicate that not all blood connected firms identified in the data truly benefit from their connections. In this case, I could overestimate the costs of political connections. However, if the assumptions for the estimation of subsidies are correct, this should be picked up by the estimation approach.

2.8.2 Further empirical results

Further differences between connected and non-connected firms

In this subsection of the Appendix, I report further results on differences between connected and non-connected firms that are in part referenced in the text, but not reported. First, Figure 2.8 reports size distributions for connected and non-connected firms in 1997 using gross firm output instead of value-added output. This looks very similar to the corresponding value-added figure. In fact, the average connected firm is slightly less than 12 times as large as the average non-connected firm for both value-added and gross output measures.

Second, we can look at size differences between connected and non-connected firms within industries looking at (real) gross output figures instead of (real) value-added figures. Similar

Table 2.4: Average relative size of connected vs. non-connected firms within industry (for gross output)

		Within industry		
	unconditional	2-digit	3-digit	4-digit
Difference	11.77	12.62	11	9.44
# connected	241	241	241	241
# non-connected	18,317	18,317	18,317	18,317

Table 2.5: Median observable factor revenue shares for labor and capital for connected (C) and non-connected (NC) firms across and within industries

	labor share (va)		capital share (va)		Total share (va)	
	NC	C	NC	C	NC	C
Unconditional	0.51	0.20	0.26	0.18	0.82	0.45
Within 2-digit	0.48	0.23	0.26	0.20	0.77	0.47
Within 3-digit	0.45	0.23	0.27	0.23	0.75	0.48
Within 4-digit	0.39	0.24	0.23	0.21	0.60	0.42
Within 5-digit	0.37	0.27	0.22	0.26	0.57	0.49

Details: The table reports median factor shares of connected and non-connected firms across the different factor inputs (columns) and across industries (rows). For within-industry estimates, median factor shares in each industry are computed separately for connected and non-connected groups and are aggregated across industries using the number of connected firms within an industry as weights.

to Table 2.1 in the main text, I report differences between connected and non-connected firms in Table 2.4.

Next, Table 2.5 reports descriptive evidence on median observed labor and capital spending shares across and within industries. I compute capital and labor shares using the reported wage bill and the capital bill as a ratio over reported value-added output. For the capital bill, I use the estimated capital stock of a firm and multiply it with the effective model-based rental rate of capital. Additionally, I compute the sum of revenue shares for capital and labor. To obtain within industry estimates, I aggregate median factor shares at the industry-level across industries using the number of connected firms within an industry as weights. Results are similar for average factor shares instead.

Results are very similar: Connected firms have much lower observable labor shares, but very similar capital and materials shares. This is not an issue of selection into specific industries that have lower labor shares, but also holds within industries. Additionally, we can look at the dispersion of factor revenue shares by comparing coefficients of variation, the ratio of the standard deviation over the mean. Looking across all industries, I find that labor revenue shares are 60% more dispersed for connected than for non-connected firms. The coefficient of variation for connected firms is around 0.94 while it is around 0.57 for non-connected firms. Similar results but smaller differences in dispersion hold for capital (0.99 vs. 0.82) and total shares (0.84 vs. 0.56). These results are robust to outliers.

Correlation between firm size, market share and profit share

This subsection reports regression results for how profits correlate with market shares (defined at different industry levels). Results clearly show a positive relationship between market shares and profits, which is in line with theories where the market share is a measure of market power and is thus correlated with profits. Furthermore, the results show that connected firms seem to have even larger profit shares conditional on their market share, indicating that political connections might buy market power beyond what is expected based on firm size.

Table 2.6: Market Power Regressions: testing the relationship between profits and market share

	Profit share							
	1-digit	1-digit	2-digit	2-digit	3-digit	3-digit	4-digit	4-digit
Market Share (1-digit)	72.174*** (5.510)	60.782* (35.050)						
Market Share (2-digit)			10.172*** (0.730)	9.612*** (2.832)				
Market Share (3-digit)					4.466*** (0.261)	4.818*** (0.861)		
Market Share (4-digit)							1.890*** (0.096)	2.266*** (0.199)
Non-connected?	−0.225*** (0.033)	−0.136*** (0.042)	−0.201*** (0.033)	−0.119*** (0.040)	−0.199*** (0.033)	−0.108*** (0.040)	−0.186*** (0.033)	−0.090** (0.038)
Constant	0.359*** (0.033)		0.334*** (0.033)		0.329*** (0.033)		0.311*** (0.033)	
Industry FE?	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.016	0.203	0.017	0.206	0.023	0.213	0.029	0.222
Adjusted R^2	0.015	0.197	0.017	0.200	0.023	0.207	0.029	0.216

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered at the fixed effect level. Industry FEs are at the 4-digit level.

2.8.3 Productivity vs. demand and the relation between single good vs. differentiated inputs with CES demand

In this part of the Appendix, I show a standard result in the heterogeneous firm literature, namely that a setup with a single good produced by heterogeneous firms with DRS technology is isomorphic to a setup where firms produce differentiated goods and face CES demand. The latter setup makes it clearer that z_i in the model used in this paper can flexibly capture both productivity and demand processes.

Assume the economy is populated by a mass of identical households of total measure L who each supply labor inelastically and consume a large variety of differentiated goods according to a standard Constant Elasticity of Substitution (CES) demand system.[39] To allow for variation in demand across industries, I consider two different levels of nested preferences such that

[39] E.g. see Costinot and Rodríguez-Clare (2014), or Hsieh and Klenow (2009).

products within and across industries s can have different elasticities of substitution:

$$C = \left(\sum_s \psi_s^{\frac{1}{\sigma}} C_s^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{2.7}$$

$$C_s = \left(\int_{i \in s} \psi_i^{\frac{1}{\sigma_s}} c_{i,s}^{\frac{\sigma_s-1}{\sigma_s}} di \right)^{\frac{\sigma_s}{\sigma_s-1}} \tag{2.8}$$

where C, C_s and $c_{i,s}$ are respectively the demands for the composite consumption good, for the sector-specific composite good and for the differentiated goods. $\psi_s \geq 0$ and $\psi_i \geq 0$ are exogenous demand parameters that are sector-specific and firm-sector-specific. $\sigma \geq 1$ and $\sigma_s \geq 1$ capture the elasticities of substitution between composite goods from different sectors and between differentiated goods within a sector. Households are assumed to statically choose consumption that maximizes their utility, leading to a simple and well-known closed-form expression for product demand:

$$c_{i,s} = B_{i,s} p_{i,s}^{-\sigma_s} \tag{2.9}$$

where $B_{i,s}$ denotes a combination of the exogenous demand parameters.[40] Note that this setup assumes that households do not distinguish between goods of connected and non-connected firms within industries and in the CES setup this means that connected firms have no additional market power within industries.

On the firm side, we have individual heterogeneous firms i in industry s that differ in their firm-specific productivity $A_{i,s}$ and their political connections and that produce differentiated products $q_{i,s}$ with a standard decreasing-returns-to-scale (DRS) Cobb-Douglas production function that is industry-specific (denoted by s):

$$q_{i,s} = A_i k_i^{\tilde{\alpha}_s} l_i^{\tilde{\beta}_s} \tag{2.12}$$

where $q_{i,s}$ is firm-specific output, k & l are firm-specific capital and labor inputs, and $\tilde{\alpha}_s$ & $\tilde{\beta}_s$ are industry-specific output elasticities. Firm i statically chooses the optimal price $p_{i,s}$ given inputs k & l such that demand and supply equalize. Suppose further that firms are small so that they cannot affect the aggregate price level P nor industry-level price levels P_s and thus take product demand as given.[41] At last, political connections enter through a revenue or output subsidy τ_i such that firm-specific revenue is given by:

$$R_{i,s} = (1+\tau_i) p_{i,s} q_{i,s} = (1+\tau_i) A_i^{\frac{\sigma_s-1}{\sigma_s}} B_i^{\frac{1}{\sigma_s}} k_i^{\alpha_s} l_i^{\beta_s} \equiv (1+\tau_i) z_i k_i^{\alpha_s} l_i^{\beta_s} \tag{2.13}$$

[40] Specifically, product-specific demand is given by:

$$C_s = \psi_s C \left(\frac{P_s}{P} \right)^{-\sigma} \tag{2.10}$$

$$c_{i,s} = \psi_i C_s \left(\frac{p_{i,s}}{P_s} \right)^{-\sigma_s} \tag{2.11}$$

where P is the aggregate price index, P_s are the price indices for sectoral composite goods and $p_{i,s}$ are prices for final differentiated goods.

[41] This is a standard assumption in models of monopolistic competition, but may be violated in case where we look at large connected firms within industries that contain only few firms in total. In case it holds, the price is given by:

$$p_{i,s}^* = A_i^{-\frac{1}{\sigma_s}} B_i^{\frac{1}{\sigma_s}} \left(k_i^{\alpha_s} l_i^{\beta_s} \right)^{-\frac{1}{\sigma_s}}$$

z_i then measures a combination of demand and supply factors, which I simply call "productivity" throughout the paper. The idea is that the process z_i is seen as a highly flexible, exogenous process that is not directly affected by political connections, but benefits from political connections τ_i can directly depend on z_i and can additionally correlate due to self-selection.

2.8.4 Optimal subsidies to connected firms in the presence of distortive taxes

In this section, I formally solve for optimal subsidies to connected firms.[42] The problem of optimal subsidies is an optimal taxation problem where the government has a fixed amount of resources $\bar{T}$ it needs to levy from connected firms that are heterogeneous in productivity and tries to set firm-specific output tax rates τ_i to maximize total output for this group of firms. Note that this reduces to net subsidies instead of net taxes if $\bar{T} < 0$ and individual firms are subsidized in case $\tau_i < 0$. This encompasses arbitrary subsets of firms: e.g. the government might only be able to set some of the taxes/subsidies in an idiosyncratic way (for connected firms), while for others (non-connected) taxes could be fixed. I start with the simpler case of a partial equilibrium analysis where input prices are unaffected by the taxes. Given that the focus is on arbitrary taxes for a few firms, this is almost equivalent to the optimal taxes in general equilibrium and I deal with the general case further below.

I show that the partial equilibrium problem has a simple solution that requires setting a constant subsidy rate across connected firms. This means that more productive firms will receive higher total amounts of subsidies, but not at a higher subsidy rate. Take any subset of firms for which the government tries to maximize their output by setting idiosyncratic output tax rates. That is:

$$\max_{\{\tau_i\}_i} \sum_i z_i k_i^*(\tau_i, w, r)^\alpha l_i^*(\tau_i, w, r)^\beta + \lambda \left[\bar{T} - \sum_i \tau_i z_i k_i^*(\tau_i, w, r)^\alpha l_i^*(\tau_i, w, r)^\beta \right]$$

where $k_i^*(\tau_i, w, r)$ and $l_i^*(\tau_i, w, r)$ give optimal input choices by firms that take their idiosyncratic tax rate as given. Technically, the government optimizes over the envelope of optimal firm decisions and this is a perfect information setup where the government can set idiosyncratic taxes based on the revealed size of the firm. Given Cobb-Douglas production functions and constant input prices across firms, optimal input policies take the following well-known closed-form:

$$k_i^*(\tau_i, w, r) = [(1-\tau_i)z_i]^{\frac{1}{1-\alpha-\beta}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\alpha-\beta}} \left(\frac{\alpha}{r}\right)^{\frac{(1-\beta)}{1-\alpha-\beta}} \tag{2.14}$$

$$l_i^*(\tau_i, w, r) = [(1-\tau_i)z_i]^{\frac{1}{1-\alpha-\beta}} \left(\frac{\beta}{w}\right)^{\frac{(1-\alpha)}{1-\alpha-\beta}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha-\beta}} \tag{2.15}$$

Taking first-order conditions for any τ_i, we get the following optimal tax condition:

$$\frac{\alpha+\beta}{1-\alpha-\beta} + \lambda \left\{ (1-\tau_i^*) - \tau_i^* \frac{\alpha+\beta}{1-\alpha-\beta} \right\} = 0$$

which states that the government should equalize the marginal budget benefits from setting a higher tax rate (captured by the shadow cost of public funds λ) with the negative marginal

[42] I want to thank Matthias Meier for suggesting to do this exercise.

output effects from setting a higher tax rate. The budget benefits scale with the tax rate times the optimal output that the firm chooses based on the tax rate, while the output scales without the tax rate. The optimal tax rate can then be expressed in closed-form as a function of the shadow cost of public funds:

$$\tau_i = \frac{\frac{\alpha+\beta}{1-\alpha-\beta} + \lambda}{\frac{\lambda}{1-\alpha-\beta}}$$

Importantly, idiosyncratic productivity z_i cancels out in this expression such that optimal tax rates end up being uniform across firms and their level is determined by the need of funds.

In general equilibrium, this result changes slightly. The reason is that any tax changes now also have an indirect effect on equilibrium prices. Fortunately, this is still tractable here. Specifically, the interest rate is pinned down in steady state so that we only need to look at the effect on wages. The market clearing wage, on the other hand, can be solved for in closed-form. Given inelastic labor supply L_t, the following holds for the wage:

$$w^* = \beta \left[L_t^{-1} \sum_j^N \left[(1-\tau_j) z_j\right]^{\frac{1}{1-\alpha-\beta}} \right]^{\frac{1-\alpha-\beta}{1-\alpha}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}}$$

One can then write the general equilibrium optimal taxation problem as:

$$\max_{\{\tau_i\}_{i \in C}} \sum_i^N (1-\lambda\tau_i) z_i^{\frac{1}{1-\alpha-\beta}} (1-\tau_i)^{\frac{\alpha+\beta}{1-\alpha-\beta}} \left(\sum_j^N \left[(1-\tau_j) z_j\right]^{\frac{1}{1-\alpha-\beta}} \right)^{\frac{-\beta}{1-\alpha}} - \lambda \bar{T}$$

First-order conditions now take into account for the effect a tax change has on the entire distribution of output:

$$0 = \frac{(1-\lambda\tau_i^*)}{1-\alpha-\beta} \left[(\alpha+\beta)(1-\tau_i^*) + z_i^{\frac{1}{1-\alpha-\beta}} \frac{\beta}{1-\alpha} (1-\tau_i^*)^{\frac{\alpha+\beta}{1-\alpha-\beta}} \left(\sum_j^N \left[(1-\tau_j) z_j\right]^{\frac{1}{1-\alpha-\beta}} \right)^{-1} \right] - \lambda$$

While this does not have a closed-form form for the optimal tax rate τ_i^*, one can still say something about how the optimal tax rate is changing in productivity z_i. Using the implicit function theorem and plugging in estimated productivity values, I show that the optimal tax rate is (perhaps surprisingly) increasing in productivity. The intuition for this is that for more productive firms, charging them lower taxes increases their input demand disproportionately more, which ends up driving up input prices marginally more than other firms. This is a "small is beautiful" result that stems from the decreasing returns to scale. Importantly, these general equilibrium corrections are small because this is about a few distorted tax rates compared to the entire economy (which shows up in the formula through the inverse of the entire productivity distribution). So in the case where you start from flat taxes and relax the budgetary constraint λ (that is decrease $\bar{T}$), you optimally want to marginally lower everyones' taxes but marginally marginally more for low productive firms.

As a final remark, one could alternatively solve the optimal taxation problem in general equilibrium additionally taking into account the endogenous response of connected firms in influencing their effective tax rate. This would fix the estimated *Political Connections Technology* in this paper and solve for the optimal baseline tax rate, knowing that connected firms distort these. This is another interesting question that the setup and estimations in this paper allow to study more rigorously. I leave this for future work.

2.8.5 Estimated subsidies, hump-shape and existing Political Connections Technologies

In this section, I show that standard Political Connections Technologies considered in the literature cannot rationalize hump-shaped subsidy schedules (over productivity), which I find to be a robust feature of the data. Specifically, the literature has considered two important variants of a decreasing returns to scale Connections Technology. Variant 1 is given by: $\tilde{\tau}_i(c,p) = cp^{\theta_p}$ where c may be a constant or idiosyncratic rent-seeking efficiency and p gives rent-seeking activities. Variant 2 instead also features firm's physical productivity z: $\tilde{\tau}_i(c,z,p) = cz^{\theta_z}p^{\theta_p}$ Arayavechkit, Saffie, and Shin (2018).

To show this, I consider the following setup. Firm profits are given by:

$$\pi = (1 + \tilde{\tau}_i(c,z,p))zx^{\alpha} - w(x+p)$$

where x gives firm inputs and w gives the price of both the input and rent-seeking activity. Results do not depend on x being a single input nor on w being the same price for inputs and rent-seeking. The proof proceeds the same for both variants. I first derive first-order conditions for input and rent-seeking choices. Using these first-order conditions, I can then use the Implicit Function theorem to characterize the partial derivative $\frac{\partial p}{\partial z}$. For the Connections Technology to allow for a hump-shaped subsidy schedule, one requires that $\frac{\partial p}{\partial z} = 0$ at the peak. I show that $\frac{\partial p}{\partial z}$ can never be zero under both Connections Technologies.

Variant 1

First-order conditions are given by:

$$\frac{\partial \pi}{\partial p} = 0: \ \theta_p c p^{\theta_p - 1} z x^{\alpha} = w$$
$$\frac{\partial \pi}{\partial x} = 0: \ \alpha(1 + cp^{\theta_p}) z x^{\alpha - 1} = w$$

Solving the first equation for x and plugging it into the second equation gives the nonlinear equation F:

$$\alpha(1 + cp^{\theta_p})z\left(\frac{w}{\theta_p cz}\right)^{\frac{\alpha-1}{\alpha}} p^{\frac{(1-\theta_p)(\alpha-1)}{\alpha}} - w = 0$$

By the Implicit Function theorem,

$$\frac{\partial p^*}{\partial z} = -\frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial p}} = -\left[\frac{1}{\alpha z}\frac{(1 + cp^{\theta_p})p}{(1 + cp^{\theta_p})(\frac{\alpha + \theta_p - 1}{\alpha}) - \theta_p}\right]$$

The numerator of this term is weakly positive given that $p \geq 0$ and is non-zero at the point where estimated subsidies are maximal in the data. Furthermore, there are large parameter ranges for which model-implied subsidies under this Connections Technology are always monotonic (which is rejected in the data). For example, subsidies are strictly increasing in z if $\alpha + \theta_p < 1$ or if $\alpha + \theta_p > 1$, but $\theta_p > (1 + cp^{\theta_p})(\frac{\alpha+\theta_p-1}{\alpha})$.

Variant 2

First-order conditions are now given by:

$$\frac{\partial \pi}{\partial p} = 0 : \ \theta_p c p^{\theta_p - 1} z^{1+\theta_z} x^{\alpha} = w$$
$$\frac{\partial \pi}{\partial x} = 0 : \ \alpha(1 + cz^{\theta_z} p^{\theta_p}) z x^{\alpha-1} = w$$

Solving the first equation for x and plugging it again into the second equation gives the non-linear equation F:

$$\alpha(1 + cz^{\theta_z} p^{\theta_p}) z \left(\frac{w}{\theta_p c z^{1+\theta_z}} \right)^{\frac{\alpha-1}{\alpha}} p^{\frac{(1-\theta_p)(\alpha-1)}{\alpha}} - w = 0$$

By the Implicit Function theorem,

$$\frac{\partial p^*}{\partial z} = -\frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial p}} = -p \left[\frac{cz^{\theta_z} p^{\theta_p}(1 + \frac{1+\theta_z - \alpha\theta_z}{\alpha}) + (\frac{1+\theta_z-\alpha\theta_z}{\alpha})}{(\frac{\alpha+\theta_p-1}{\alpha})(1 + cz^{\theta_z} p^{\theta_p}) - \theta_p} \right]$$

Again, the numerator of this term is weakly positive given that $p \geq 0$ and is non-zero at the point where estimated subsidies are maximal in the data. Furthermore, there are large parameter ranges for which model-implied subsidies under this Connections Technology are always monotonic (which is rejected in the data). For example, subsidies are strictly increasing in z if $\alpha + \theta_p < 1$ or if $\alpha + \theta_p > 1$, but $\theta_p > (1 + cz^{\theta_z} p^{\theta_p})(\frac{\alpha+\theta_p-1}{\alpha})$.

2.8.6 Microfoundations of the Political Connections Technology

In the following I provide two possible microfoundations for the *Political Connections Technology* used throughout the paper that are based on two different interpretations of what political connections buy. In the first interpretation, the *Political Connections Technology* buys output subsidies, while in the second interpretation, the *Political Connections Technology* is reinterpreted as the share of taxes that connected firms pay.

The *Political Connections Technology* as an output subsidy

One interpretation of the *Political Connections Technology* is as a net output subsidy. The parametric form chosen for this technology is: $\tau_i = \varepsilon z_i p^{\theta_p} - cp^{\theta_c} z_i^{\theta_z}$. To microfound this choice, suppose the government can use part of the tax revenue to buy products from firms that are then redistributed to households. As was shown before, τ_i only captures demand beyond standard demand for a similar non-connected firm. That is, the government basically offers a contract to a connected firm saying, whatever your total demand from households, we will pay $\tau_i/(1 + \tau_i)$ percent of this demand or we subsidize households' demand by this percentage. The assumption here is that most government policies that directly or indirectly subsidize firms can be represented by this menu over τ_i instead of contracts that are fixed to quantities. That is, politicians directly bargain over subsidy rates and not absolute transfers. The micro-foundation of the parametric form of τ_i is then linked to the political process that offers subsidy rates.

Specifically, suppose that for each connected firm there exists a continuum of relevant government bills that each may promise a unit of government demand. τ_i gives at the same time the net subsidy rate obtained by a connected firm i as well as the measure of government bills that the connected firm managed to influence in its favor. Given that there are few connected firms in this economy, this model abstracts from competition for government bills across connected firms and simply assumes that all connected firms care about their own set of government bills that they can influence. There are two terms in the *Political Connections Technology*. The first term captures the amount of bills that the firm managed to influence, while the second term captures the amount of influenced bills that are overturned via audits or other public oversight. Given the continuous measure of government bills, these audits give deterministic detection rates. Let us look at each of the terms in turn.

The first part of the technology ($\varepsilon z_i p^{\theta_p}$) captures the measure of bills that the connected firm manages to influence via bribing and lobbying the politician they are connected to. There are two ways to think about this term that lead to very similar parameter interpretations. First, the politician has direct access to government bills and offers the firm a linear bribe schedule ($\tilde{\tau}_i = \text{const.} * b$), but the firm faces costs of concealment or production costs to transform rent-seeking spending p into actual bribes b so that $b = \widetilde{\text{const.}} + z_i * p^{\theta_p}$ where z_i gives the firm's productivity at concealing bribes and θ_p is now the elasticity of this concealment technology. This captures what economic sociologists call costs of obfuscating bribes as meaningful, symbolic interactions (Hoang 2018). Remember that rent-seeking spending p captures a combination of capital and labor and one can think of this as final goods (any form of bribes such as luxuries and money) or as some combination of capital and labor services. Also, the constant in the linear bribe schedule captures the politician's efficiency and one may think of this as also being potentially heterogeneous across the type of connections - an idea I explore in the paper.

Alternatively, political capital does not need to be converted ($b = p$), but the politician may face direct costs of obtaining the subsidy rates through parliamentary approval, bargaining with other politicians or filling out the paper work. For example, increasingly higher benefits to firms might require the approval of more politicians who all need to be bribed as well (in the case of $\theta_p \in (0, 1)$) or costs of bribing decline as there are increasing returns to scale in filling out paper work (in the case of $\theta_p > 1$). In these cases, θ_p captures the elasticity of costs from obtaining output subsidy rates. z_i then captures that politicians are more efficient at influencing bills if the firm is more productive (as they need to argue less). In both cases, counterfactuals have very similar interpretations. For example, one can think of doubling ε as doubling the efficiency of the politician to transform bribes into subsidies.

For the second term, suppose the politician faces risks of audits or opposition from other politicians. Remember that subsidy rates are determined by a continuum of small amendments to laws or policies. In this case, audits can overturn a fraction of subsidies. The second term ($cp^{\theta_c} z_i^{\theta_z}$) then captures the number of subsidy rates that are overruled by audits. p^{θ_c} captures the idea that benefits to politically connected firms are more likely to be contested by other politicians or the public as the number of distortionary policy and regulatory amendments increases. θ_c measures the elasticity of this opposing reaction. $z_i^{\theta_z}$ instead captures the opposition stemming not from bribes, but from extra scrutiny that larger firms in the economy receive. Importantly, c measures the level of audits in the economy.

An alternative interpretation that is not considered here is one where the *Political Connections Technology* is interpreted as a state-funded project such as a private-public partnership (PPP) or a state-owned enterprise. The idea of this interpretation is that output of connected firms is $(1+\tau_i)z_i k_i^{\alpha_s} l_i^{\beta_s}$, which can be separated into standard output $z_i k_i^{\alpha_s} l_i^{\beta_s}$ and a rent-seeking project $\tau_i z_i k_i^{\alpha_s} l_i^{\beta_s}$ that is financed entirely by the government.

The *Political Connections Technology* as a tax evasion technology

Following the redefinition of the *Political Connections Technology* as a tax evasion technology in Section 4, we can write the share of taxes that connected firms pay as: $\phi_i \equiv 1 - \tau_i\left(\frac{1-\bar{\tau}}{\bar{\tau}}\right)$ where $\bar{\tau}$ gives the official corporate tax rate. Plugging in the parametric form chosen for the *Political Connections Technology*, this can be rewritten as:

$$\phi_i = 1 - \varepsilon z_i p^{\theta_p}\left(\frac{1-\bar{\tau}}{\bar{\tau}}\right) + c p^{\theta_c} z_i^{\theta_z}\left(\frac{1-\bar{\tau}}{\bar{\tau}}\right)$$

The share of taxes that a firm pays is then determined by two terms; the first term decreases and the second term increases the share of taxes as political capital spending is increasing. Suppose the following simple setup. A tax collector is in charge of a firm's filing and has discretion over ϕ_i. The tax collector takes bribes b in the form of capital for setting a lower ϕ_i as in the previous narrative. Suppose that total taxes depend on many different rules, different documents or that it depends on a long list of entries in revenue filings to the tax administration. Suppose that the tax collector charges a bribe for reducing the tax in each document, each data entry or each part of the tax. In this case, the share of taxes paid by the firm can be expressed as a linear rule in bribes: $\phi_i = 1 - \text{const.} * b$. Now suppose that the firm needs to "produce" or "conceal" b so that $b = \text{const.} * z * p^{\theta_p}$ where θ_p is now the elasticity of this production or concealment technology and z the productivity.

For the second term, suppose the tax collector faces oversight from managers or risk of being checked up on. The tax collector conceals or calculates lower rates for each entry and managers may sporadically check up on any entry. As the number of entries becomes large, the probability of being detected equals the number of checks. Suppose for simplicity that for each check that leads to corrections, the tax collector does not face any punishment and only the tax demands are changed. The tax collector offers to reduce taxes, but does not insure the risk of corrections. Then the second term captures the number of distorted tax entries that become corrected and one can rewrite this term as const. $* f(\text{size}) * b^{\tilde{\theta}}$, where $f(\text{size})$ captures flexibly the idea that check-ups by superiors might depend on the size of the firm where larger firms are also more likely to be checked up and $\tilde{\theta}$ can be thought of as a span-of-control parameter that captures how close tax collectors are being monitored. For a high $\tilde{\theta}$, this control is high, which leaves little room for tax collectors to change tax rates for connected firms. Importantly, c captures the level of auditing.

2.8.7 Further details on extensions & robustness results

This part of the Appendix provides more details on the extensions and robustness results presented in Section 5.

Wedges and the costs of market power

Here, I provide further details on the estimation of wedges, subsidy estimation and counterfactuals with wedges. To quantify how wedges affect the costs of political connections, I reestimate subsidies, the *Political Connections Technology* and general equilibrium counterfactuals allowing for wedges. For simplicity, I assume that connected firms only report productive capital and labor inputs. Further, normalizing wedges by assuming that median reported factor shares for non-connected firms identify output elasticities, both idiosyncratic wedges and TFP can be directly estimated in the data for all firms. The estimation approach for subsidies of connected firms then remains unchanged assuming that the assumptions in Propositions 2 and 3 continue to hold. Most importantly, this means that connected firms are still only selected based on productivity (and not directly on wedges) and that wedges do not break the monotonicity of TFP, which naturally holds as long as wedges do not directly enter the *Political Connections Technology*. This rules out quid-pro-quo benefits where subsidies are offered conditional on how connected firms choose inputs.

Figure 2.5 shows conservative upper and lower bound estimates of subsidies. In comparison to the baseline estimates in Figure 2.3, estimated subsidies allowing for additional idiosyncratic wedges leads to roughly 40% higher average subsidy estimates. The reason for this directly follows from observing higher and more dispersed wedges for connected firms as they put downward pressure on size differences between connected and non-connected firms, requiring higher subsidies to explain large observed differences in size distributions.[43] While wedges do lead to more heterogeneity, non-parametric subsidy estimates can be well explained by model-implied subsidies based on optimally choosing rent-seeking activities taking into account firm's own productivity, the costs of rent-seeking activity as well as idiosyncratic wedges. Specifically, model-implied subsidies explain about 70% of the variation in subsidies for both bounds based on the R^2 with the average subsidy as the comparison. The key economic mechanisms as captured by the estimated elasticities stay unchanged: benefits of political connections continue to exhibit decreasing returns to scale ($\theta_p \approx 0.52 - 0.56$) and costs of political connections are convex in rent-seeking activities ($\theta_c \approx 1.25 - 1.28$) as well as firm size ($\theta_z \approx 2.08 - 2.26$).

Industry heterogeneity

This subsection provides further details on estimating subsidies and costs of political connections under more industry heterogeneity. The non-parametric within-industry estimator separately draws productivities from non-connected firms within the same industry and matches firms accordingly. This introduces a trade-off as within industry matching matches firms that are more similar while at the same time reduces both the population from where productivities can be drawn and the sample with which one can match. For example, in the extreme

[43] Note that the productivity threshold $\bar{z}$ is re-estimated so that most of the changes in the subsidy estimates are driven by the increased dispersion in wedges for connected firms rather than the level difference.

case of only a single connected firm within an industry, the productivity estimate will simply be the average productivity of non-connected firms above the threshold productivity $\bar{z}_s$ within this industry. Precision of non-parametric productivity estimates is driven by being able to order many firms within bootstrap samples, so the precision of the estimates declines with further industry heterogeneity. Relatedly, allowing for unrestricted industry-specific productivity thresholds $\bar{z}_s$ fixes the lowest TFP connected firm within each industry, restricting more and more subsidies as more industry heterogeneity is considered. As a solution of this bias-variance trade-off, I enforce a single quantile cutoff, meaning that the bottom x% based on TFP of non-connected firms in each industry are excluded when matching. Hence, the implied productivity threshold across industries can still vary depending on the industry-specific distributions of productivities. The cutoff x is then conservatively estimated to be the minimum productivity quantile of connected firms across all industries.[44] The lower subsidy bound is estimated similarly, enforcing a single correlation $\bar{\rho}$ across industries, which is given by the minimum $\bar{\rho}$ for which any other subsidy estimate across any industry becomes zero.[45]

For the costs of political connections, I re-estimate the *Political Connections Technology* taking into account further heterogeneity. We generally expect the parameters of the *Political Connections Technology* to vary across industries as industries differ in how closely related they are to the political system, affecting the difficulty of lobbying for preferential policies or receiving government contracts, and they differ in visibility, affecting oversight and the chance of preferential deals being detected and publicly reported. However, this variation is not summarized in a single parameter in the proposed *Political Connections Technology*. To explain estimated subsidies with further industry heterogeneity, while keeping estimation parsimonious, I keep the same functional form and estimated parameters, but allow the parameters that govern levels (ε and c) and parameters that govern elasticities to differ by a common factor across 2-digit industries. This adds two additional parameters per industry and I found this a good compromise between not overfitting, while allowing reasonable variation in the *Political Connections Technology* across industries that mimics the same hump shape pattern of subsidies and captures the same fundamental drivers of observed subsidies. Estimated parameters at the 2-digit level maintain decreasing returns to scale in benefits from rent-seeking activities and convex costs both in rent-seeking activities and firm size. The R^2 of the noisier estimates at the 2-digit industry level still exceeds 95%.

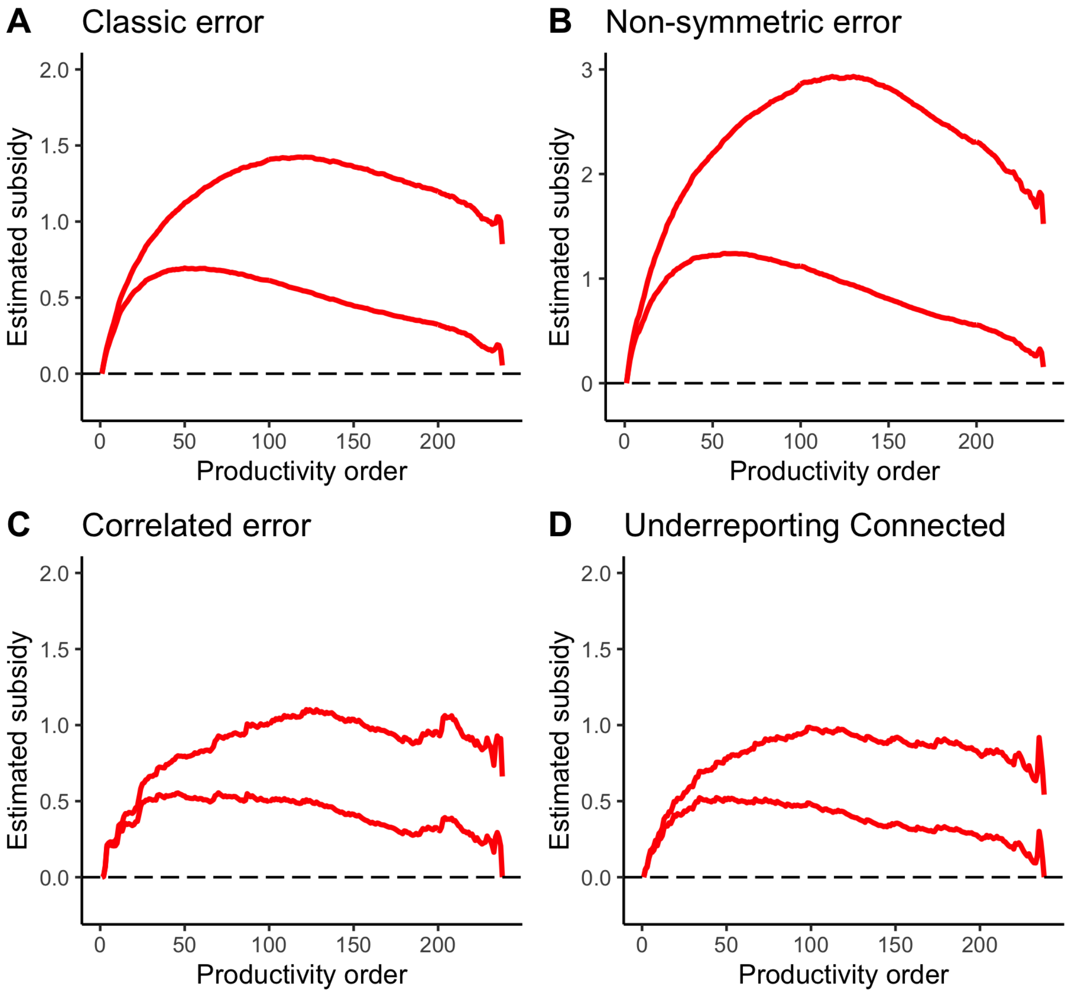
Measurement Error

I consider four types of measurement error. For each type I independently draw B type-specific realizations of measurement error for each firm. For each iteration $b \in B$, I obtain true output by purging observed output from the measurement error and then reestimate subsidies, the political connections technology and welfare costs for both bounds. I report average results for each type of measurement error. I choose B = 50. In case not otherwise stated, I

[44] This is well-defined as long as there are multiple connected firms within an industry. There is only a single connected firm for one industry at the 4-digit level, which I exclude when computing the 4-digit cutoff. The estimated cutoff is around 38.6% at the 1-digit level, 36% at the 2-digit level, 10.2% at the 3-digit level and 8.6% at the 4-digit level.

[45] This is well-defined as long as the upper bound estimates are all strictly positive. This is the case for the 2-digit and 3-digit industries. Here, I find $\bar{\rho} = 0.186$ at the 2-digit level and $\bar{\rho} = 0.157$ at the 3-digit level. At the 4-digit level, a few subsidy estimates are already negative for the upper bound. Formally, the model is rejected at the 4-digit level and given the noise at this estimation level, I abstract from it for the subsequent welfare estimates.

Figure 2.9: Estimated subsidies with different types of measurement error



Notes: Non-parametric estimation of output subsidies and productivity allowing for different forms of measurement error in reported output (value-added revenue). For Panel A - C, measurement error is mean zero and its variance is chosen such that the R2 of a regression of reported output on real output is 75 percent. Panel A: Classic normal log-additive measurement error for all firms. Panel B: Log-normal log-additive measurement error for all firms (non-symmetric). Panel C: Measurement error that positively correlates with firm size (taking reported output) for all firms. Panel D: Systematic underreporting of connected firms by 20 percent of output. Plot gives average subsidy estimates across 50 independent draws of measurement error for each type.

choose the variance of measurement error such that a regression of reported output on real output gives an $R^2 = 0.75$.

In Panels A-C, I consider measurement error that affects all firms, connected and non-connected, and that has mean zero. Panel A considers multiplicative measurement error of the form: $\tilde{y}_{it} = y_{it} * \text{error}_{it}$ where $\tilde{y}_{it}$ is reported value-added output and $log(\text{error}_{it})$ is normally distributed. Estimated subsidy schedules are almost entirely unaffected, which shows up in almost identical welfare costs. The reason is that observed log output is also close to normally distributed so that adding normally distributed errors leaves the output distribution unaffected. While not shown here, I also consider distributions with heavier or lighter tails by simply scaling normally distributed measurement error and find also no quantitatively meaningful differences in subsidy estimates.[46] To consider error that is differently distributed than

[46]Specifically, I take $log(\text{error}_{it})^{\varphi}$ with $\varphi \in \{0.5, 1.5\}$. In case of negative errors, I take $-(-x)^{\varphi}$.

log output, I consider non-symmetric measurement error by now letting error$_{it}$ be normally distributed. In this case, subsidy estimates turn out to be higher, which also translate into higher welfare costs of connections. The reason is that this measurement error led to a lower dispersion of the right tail of the observed output distributions, biasing baseline subsidy estimates downward and leading to underestimate true subsidies. Similarly, in case measurement error led to a higher dispersion of the observed output distributions on the right tail, then the baseline subsidy estimates would be overestimated. Still, based on Panel A and B, these two forms of measurement error have little effects on the overall shape of the estimated subsidy schedules nor estimated welfare costs.

To consider a potentially more problematic case of bias, I consider measurement error that correlates directly with firm size in Panel C. Specifically, I consider $log(\text{error}_{it}) = \beta_0 + \beta_1 log(\tilde{y}_{it}) + \nu_{it}$ where ν_{it} is mean zero normally distributed, β_0 is such that the overall error is mean zero and $\beta_1 > 0$. Perhaps surprisingly, it turns out that this form of correlated measurement error also leaves subsidy estimates and estimated welfare costs basically untouched. The reason for this is that this form of correlated measurement error does not affect the relative dispersion of output distributions across connected and non-connected firms. To also consider the effect of differential output distortions, I introduce measurement error that only affects connected firms in Panel D. Specifically, I assume that all connected firms systematically underreport a fixed 20% of output.[47] The results again are robust to this form of measurement error. The reason is that such selective underreporting is entirely captured by the productivity cutoff. A higher percentage of underreporting simply leads to a lower estimated cutoff, leaving the estimated levels unbiased. If anything, the upper bound welfare cost estimates are now somewhat higher.

[47] I considered underreporting of 5%, 10% and 20% respectively, but results were all quantitatively similar and I report the results with the most measurement error.

Chapter 3

Estimating government worker skills: Application to worker selection and wage setting in Indonesia

Jonas Gathen[1]

Abstract

This paper provides a new approach to estimate government worker skills that is applicable in settings where government output is unobserved and government wages are uninformative about skill differences. The approach estimates skills from wages in comparable jobs in the private sector, relates these skills to skill-related observables using Machine Learning tools and then predicts government worker skills out-of-sample. I apply the new estimation approach to rich Indonesian household-level panel data from 1988 to 2014, showing two main applications. First, I quantify that government workers are highly selected, that their skills have increased, but that their skill premium has declined over time. Second, I analyze government wage setting: the Indonesian government pays a wage premium of at least 30% conditional on skills, about 1/3 of which is driven by the large gender wage gap in Indonesia's private sector.

[1]Toulouse School of Economics; j.gathen@posteo.de. I would like to thank my advisors Stéphane Straub, Matteo Bobba, Fabrice Collard and Christian Hellwig for their continued support. For helpful feedback and comments, I also want to thank in no particular order: Olivier de Groote, Thierry Magnac, Mitch Downey, Pascal Lavergne, Guohui Jiang, Johanna Isman, Oscar Fentanes and Jens Wäckerle. At last, I am indebted to the feedback by seminar and conference participants at the IAST Political Economy workshop, the TSE Macro workshop, the EUDN development conference in Passau 2020, Stockholm University's applied micro workshop and the workshop of the World Bank's Bureaucracy Lab. All remaining errors are my own.

3.1 Introduction

How do relative skills of government workers evolve over the course of development and how does this affect changes in the quality of government services? Does the government manage to select highly skilled workers? Are government wages informative about underlying skills and does the government pay a wage premium? A key ingredient for answering these questions is a good measure of government worker skills. This paper provides a new approach to estimate such skills, showing how to estimate unobserved skills in government jobs for any worker in the economy. The paper then applies the estimation approach to rich household-level panel data from Indonesia to study (1) systematic changes in the skills of government workers and their selection and (2) government wage setting over a period of almost 30 years.

Traditional approaches in Economics measure a worker's skill – or marginal productivity of labor – either directly by observing output and making an assumption on the underlying production function (e.g. Attanasio et al. 2020; Chetty, Friedman, and Rockoff 2014) or by drawing on observed wages assuming some form of competitive wage setting process that relates wages to the marginal productivity of a worker (Polachek 2008; Meghir and Pistaferri 2011; Sanders and Taber 2012). For the skill estimation of government workers, both approaches are problematic, because government output is usually unobserved (think of the output of a bureaucracy), wage competition may not force bureaucracies to change wages and government worker wages often follow rigid wage-setting rules that disguise underlying skill differences (e.g. Biasi 2021).

This paper develops a three-step estimation approach that sidesteps these concerns. The first estimation step extracts a noisy individual-specific estimate of skills – an individual fixed effect – from wages in private sector jobs that are comparable to government jobs. The key assumption here is that there exist comparable private sector jobs for which observed wages may flexibly deviate from competitive wages but are at least a noisy function of the equilibrium price of skills, underlying worker-specific skills and experience. The second estimation step is about predicting the individual-specific noisy skill estimate using a rich set of skill-related observables. I show how Machine Learning algorithms are particularly useful here to disentangle actual skills from estimation noise and flexible shocks to wages. The key assumption here is that skills of interest are solely a (flexible) function of skill-related observables. In the third and last estimation step, one can then enforce the estimated relationship between individual-specific skills and skill-related observables to obtain an out-of-sample skill estimate for workers for whom we have not previously estimated skills. The approach thus allows to estimate skills of government workers in government jobs as well as potential skills in a government job for any worker in the economy for whom we observe these skill-related observables. The estimation approach can thus also be used in many other settings in which researchers need to estimate potential job-specific skills for workers who work in other jobs with very different skill requirements, while allowing for flexible selection into jobs based on observables and unobservables.

In the second part of the paper, I apply this estimation approach to the case study of Indonesia between 1988 to 2014 drawing on a large and high-quality representative panel that tracks individuals and households over time. The data – the Indonesian Family and Life Survey (IFLS) – is particularly suited for the estimation approach as it features (1) a large sample of government and private sector workers across many different occupations, (2) one of the longest wage and employment panels in a developing country context, and (3) an exceptional breadth of skill-related variables such as educational background, national exam scores, test results for self-administered Raven tests, cognition and memory tasks, the big-5 personality traits and elicited risk and time preferences. Drawing on the IFLS, I estimate individual-level skills in government-like jobs using private sector workers in social service jobs as the reference group given that 80% of government workers also work in social services. To do so, I first purge observed wages from a changing equilibrium skill price that I identify using the flat spot identification approach in Bowlus and Robinson (2012). I then back out individual-level fixed effects from a standard Mincerian regression on wages deflated by the skill price. In the next step, I predict the noisy individual fixed effects using skill related observables and different Machine Learning algorithms. I find that an off-the-shelf gradient boosted tree algorithm outperforms other algorithms in my context and use this estimated model to predict government worker skills for all workers in the economy.

Having estimated government worker skills, I then illustrate two main applications for which skill estimates are key. In the first application, I look at changes in government worker skills, the selection of government workers based on skills and how relative skills of government workers have evolved. I find that Indonesian government workers are strongly positively selected on skills and that their skills have also increased over time as new cohorts with higher skills entered the labor force. However, I find that the skill premium of government workers actually declined over time as skills of workers in the rest of the economy caught up. I show that a big part of this is driven by an increasing right tail of workers who would be highly skilled in government jobs but who the government does not attract for government jobs, and that changes in the hiring practices of Indonesian government workers in the process of democratization, decentralization and civil service reforms after the year 2000 have only led to small changes in how the government *de facto* selects government workers. Finally, I show evidence that uneven hiring waves across years lead to differences in cohort-specific government employment shares, which have an adverse effect on government worker selection: in years in which the government hires more, average government worker skills decline, which holds conditional on the jobs that the government opens. This is consistent with the idea that more hiring forces the government to eventually move down the distribution as the top of the distribution thins out.

In the second main application, I show how to use the skill estimates to look at government wage setting and the wage premium of government workers. First, I find evidence that Indonesian government wages are indeed less informative about underlying skills than private sector wages in comparable jobs. I find this by predicting government wages using estimated skills from comparable private sector jobs and by re-estimating skills using government wages instead. However, due to more predictable life-cycle wage progression in the public sector, government wages become more predictive of skills when accounting for experience. Second, I find that the Indonesian government pays a large wage premium of at least 30% compared to similar jobs in the private sector. Does this mean that the Indonesian government is overpaying for workers? While this could be true, I also find strong evidence for more discriminatory wage setting in the private sector. A gender wage gap in the private sector of 35% even after controlling for job, skills and experience accounts for 1/3 of the government wage premium alone.

The paper is structured as follows. Below, I discuss how the new estimation approach and empirical results relate to the existing literature. In the next section, I explain the procedure to estimate government worker skills. This approach is applied to Indonesian data in Section 3, where I also give more details on the local context. Section 4 shows two main applications for Indonesia as examples for the usefulness of the approach. Section 5 discusses extensions and the last section concludes.

Related literature

The contribution of this paper is both technical and conceptual. Technically, the paper contributes to the literature by proposing a novel estimation approach that allows to measure government worker skills for all government workers as well as any other worker in the economy. This differs from two alternative approaches that have previously been proposed in the literature. One common approach – as followed by Dal B'o, Finan, and Rossi (2013), Dal B'o et al. (2017), Besley et al. (2017), and Colonnelli, Prem, and Teso (2020) – is to use (residualized) previous private sector wages of government workers as a measure of government worker skills. Previous private sector employment can be a misleading measure of skills in cases where government workers take their first real job in the public sector or have previously worked in a very different occupation, in which case it might be a better measure of a worker's outside option. More technically, the approach in this paper – similar to Dal B'o et al. (2017) and Besley et al. (2017) – also draws on residuals from observed private sector wages but restricts to a comparable subset of government jobs and deals directly with the econometric concern that residualized wages only give a noisy estimate of individual-level skills, "regularizing" skills using skill-related observables and Machine Learning tools. Whenever rich skill-related observables are available, I thus believe that my approach provides a better way to estimate skills from wages.

The second approach followed in Best, Hjort, and Szakonyi (2017) is complementary to my approach. Best, Hjort, and Szakonyi (2017) consider a context in which individual bureaucrats' output is observable – prices paid by bureaucrats in the government procurement of goods. Similar to using wages, they then also draw on a residualized measure of output and extract an individual-bureaucrat fixed-effect. While Best, Hjort, and Szakonyi (2017) are not directly interested in the individual fixed-effects, they are interested in a variance decomposition for which noisy estimates of fixed-effects would bias their results. They thus propose a "covariance shrinkage" approach that uses bootstrapping to separate variances in the true signal and noise. The two key differences are that, first, the approach in this paper does not require to observe government output but informative wages instead, and second, that to "shrink" the noise in the estimation of individual fixed-effects this paper uses rich covariates that also allow to predict skills for any worker in the economy, while the approach in Best, Hjort, and Szakonyi (2017) uses a shrinkage approach without covariates.

Conceptually, the paper uses estimated government worker skills to contribute to the growing literature on the workings of bureaucracies, the developmental state and how the delivery of government services can be improved (Best, Hjort, and Szakonyi 2017; Decarolis et al. 2018; Rasul and Rogger 2018; Finan, Olken, and Pande 2017). The estimation approach allows for a systematic measurement of government worker skills, a key – but difficult to measure – input in the production of government services. The estimation approach thus complements a number of recent papers that have studied how the government selects government workers (Colonnelli, Prem, and Teso 2020; Dal B'o, Finan, and Rossi 2013; Jia, Kudamatsu, and Seim 2015; Bhavnani and Lee 2019; Estrada 2019), who self-selects into government jobs (Ashraf, Bandiera, and Lee 2016; Hanna and Wang 2017; Weaver 2019) and how the government remunerates workers (see: Finan, Olken, and Pande 2017). In line with Ashraf, Bandiera, and Lee (2016) and Weaver (2019), I find strong positive selection of government workers based on skills. This also means that large documented government wage premia in developing countries (e.g. see: Finan, Olken, and Pande 2017) are smaller after controlling for selection on skills; for Indonesia, I find a 30% wage premium conditional on skills, which is more than 20 percentage points lower than without controlling for skill selection.

The key novelty with respect to this literature is the broader scope of the estimation approach that allows to study all government workers over a long period of time. The approach is particularly useful in settings where government output is hard to observe, and thus especially relevant for studying higher-level bureaucracies where government output may be hard to define and measure, allowing researchers to move beyond the study of last-mile service delivery (e.g. Chaudhury et al. 2006; Banerjee, Iyer, and Somanathan 2007; Finan, Olken, and Pande 2017). Importantly, the broader scope of the approach allows me to establish at least two novel findings in the literature. First, I show that despite growing absolute skills, relative skills of government workers systematically declined in Indonesia over the past 30 years. I link this finding to the difficulty of the Indonesian government to attract the workers with the highest skills to the government. Second, I show evidence for the detrimental effect of government hiring cycles on the selection of government workers. The evidence is consistent with the idea that in years of outsized hiring, the government needs to move down the skill distribution of the applicant pool to fill all government positions.

A good sign of a new estimation approach is that it raises many interesting questions that can now be studied more rigorously: For example, what are the output or welfare costs of government hiring cycles? Or what drives the relative decline in government skills and does this go in hand with a relative decline in state capacity versus private sector capacity over the course of development? These questions are particularly well-suited for future structural work, for which the estimated government skills in this paper can function as a direct input.

3.2 Identifying government worker skills

This section outlines the three-step identification approach and its underlying assumptions. The focus is on the baseline identification setup that is also used for the subsequent empirical application, while extensions are left to Section 5.

The first step is about identifying a "noisy" signal of human capital and skills using wages in a suitable private sector job, which Proposition 4 formalizes.

Proposition 4 (Step 1: Identifying a signal of government-worker skills)**.** *Assuming there exists a subset of "government-like" jobs (potentially in the private sector) for which:*

(i) *(**wage determination process**): observed real hourly wages $W_{i,e,t}$ of individual i with experience e at time t follow:*

$$W_{i,e,t} = P_t * H_{i,e,t} * exp(\epsilon_{i,e,t}) \tag{3.1}$$

where $H_{i,e,t}$ captures human capital relevant for government jobs, P_t captures the (equilibrium) price of efficiency units of human capital and $\epsilon_{i,e,t}$ is a flexible mean-zero error term that is independently distributed of $(P_t, H_{i,e,t})$, follows a stationary process and has finite variance.

(ii) *(**experience profile**): Human capital $log(H_{i,e,t}) \equiv h_{i,e,t}$ follows a standard Mincerian experience process:*

$$h_{i,e,t} = z_i + \delta_0 * exp_{i,e,t} + \delta_1 exp^2_{i,e,t} \tag{3.2}$$

with $\delta_0 > 0$ and $\delta_1 < 0$ such that the experience profile is concave, and where $z_i = log(Z_i) = h_{i,0,t}$ denotes individual-specific human capital at labor market entry.

Then:

1. *(**flat-spot identification**): Following Bowlus and Robinson (2012), an unbiased and consistent estimator (for $N \to \infty$) for the path of the equilibrium skill price P_t (up to a level of normalization) is given by within-individual wage changes for workers in their flat-spot (FP) region:*

$$\mathbb{E}_{i\in FP}[w_{i,e,t} - w_{i,e-1,t-1}] = p_t - p_{t-1} + \mathbb{E}_{i\in FP}[(h_{i,e,t} - h_{i,e-1,t-1}) + \epsilon_{i,e,t} - \epsilon_{i,e-1,t-1}] \tag{3.3}$$

2. *(**skill signal identification**): Estimates of the experience parameters (δ_0, δ_1) are consistent and unbiased. However, estimates of individual skills z_i are only unbiased but generically inconsistent in a panel with fixed T due to the incidental parameters problem.*

The proof draws on standard results and is relegated to Appendix A.2. Proposition 4 defines the two main measures of skills used throughout this paper: (1) individual-specific human capital $h_{i,e,t}$ that incorporates experience and (2) skills at labor market entry z_i. The former is more relevant for questions related to the stock or distribution of human capital at any point in time, while the latter is more relevant for questions related to the selection of government workers and comparing government workers with different levels of experience. Proposition 4 shows how wages in a government-like job can be used to identify signals for both measures of skills. These are only noisy signals, because estimates for $h_{i,e,t}$ and z_i are not consistent and thus still include permanent and temporary components of the error $\epsilon_{i,e,t}$ that drive observed wages but are unrelated to human capital.

In the following, I discuss the two main assumptions. The assumption on the **wage determination process** follows a large human capital literature that links wages to human capital. This nests most models of the labor market, but for example rules out models of labor search and matching where informational asymmetries between workers and employees lead to a permanent disconnect between skills and wages as in Taber and Vejlin (2020).[2] The approach only assumes that wages are partly determined by underlying human capital, allowing the price of human capital to change flexibly over time and for wages to be influenced by many other (unobserved) factors as captured by the error $\epsilon_{i,e,t}$.

Allowing the price of human capital to change flexibly over time is key, as this allows for changes in the supply and demand of government-job-specific human capital in the overall economy. Such changes may be directly driven by changes in government hiring. Similarly, the approach allows for systematic changes in the price of government-like output, which

[2]It should be noted that permanent characteristics at labor market entry are generally found to be by far the main drivers of earnings inequality (e.g. Huggett, Ventura, and Yaron 2011; Keane and Wolpin 1997; Lamadon, Mogstad, and Setzler 2019; Taber and Vejlin 2020), making it a natural starting point.

could also be directly influenced by the government, and which will also show up as changes in the skill price. Importantly, the approach requires to focus on a market for government-like jobs where there is a single skill price P_t, but it does not require that the government pays that skill price or that other related markets have the same skill price.

The error $\epsilon_{i,e,t}$ in the wage-determination process allows for different factors that influence wages, such as compensating differentials, deviations from fully competitive labor markets or contractual deviations from spot-market wages. This means that firms can temporarily value individual-specific human capital in a way that differs from the market in order to attract, retain, or discourage specific individuals, or because information is imperfect. Moreover, these decisions can be correlated over time, nesting flexible time dependence of the errors. As will become clear in the second estimation step, the approach can even allow for permanent individual-specific shocks such as permanent employer-specific wage markdowns or permanent non-wage job characteristics (e.g. Lamadon, Mogstad, and Setzler 2019). However, the more stringent assumption is that errors need to be independent of human capital and the skill price; otherwise, the estimation approach will wrongfully attribute part of the error to changes in the skill price or human capital. This rules out deviations from competitive wages that are systematically correlated with experience as this would lead to inconsistent estimates of the experience profile.

The second main assumption is on the **experience profile**. Here, I take a more stringent parametric restriction, assuming a standard quadratic experience profile. The parametric assumption can easily be relaxed as I show in Section 5. The key economic restriction is that experience profiles need to exhibit a flat-spot at some level of experience in which human capital does not grow further, a condition that finds strong empirical support (e.g. Lagakos et al. 2018) and theoretical support from models of endogenous human capital accumulation that imply decreasing returns to human capital accumulation towards the end of a worker's life cycle (e.g. Magnac, Pistolesi, and Roux 2018). The flat-spot restriction is needed for the skill price identification. Following Bowlus and Robinson (2012), the idea is that the wages of workers for whom human capital is not increasing further over the life cycle reflects only changes in the underlying skill price.

Before proceeding with the second step, two further remarks are in order. First, the approach can only identify relative differences in skills across workers at a single point in time and across workers over time, but the overall level of skills is unidentified. A benefit of focusing on differences is that the setup allows for constant differences in wages, for example when labor or government-like output are priced with a constant markup or markdown. Second, the idea of wages revealing human capital implicitly assumes an underlying production function that defines the marginal product of labor for government-like output: $\frac{\partial Y^G}{\partial h_{i,e,t}}$. A statement such as "a worker is twice as skilled as another worker" thus maps differences in skills to the worker's marginal contribution to government-like output.

Having obtained a noisy estimate of skills, the second estimation step gives a consistent estimate of skills by projecting the skill signals on a rich set of skill-related observables.

Proposition 5 (Step 2: Projecting skill signals on skill-related observables)**.** *Denote the skill signals of the previous step by $\hat{z}_i$, which can be written as $\hat{z}_i = z_i + \eta_i$ with η_i being independent of z_i given the assumptions in Proposition 4. Further assuming that worker skills in the "government-like"* *job follow:*

$$z_i = f(X_i) \tag{3.4}$$

where $f()$ is any Borel measurable function and X_i are observable individual-specific characteristics, then the non-parametric regression of $\hat{z}_i = f(X_i) + \eta_i$ using observable X_i gives a consistent estimate of individual skills z_i in government-like jobs.

Proposition 5 draws on the Machine Learning literature, where the idea of flexibly projecting noisy estimates of y on explanatory variables X to obtain a consistent estimate of y is called "regularizing".[3] As in the Machine Learning literature, the goal is to remain as flexible as possible for approximating the relationship between z_i and X_i using flexible non-linear functions for $f()$, while best selecting skill-related variables X that predict skills z_i. In the labor market context, many variables are likely correlated with skills, but ex ante, it is unclear which variables will be the most predictive of skills and what the exact functional relationship looks like, let alone interaction effects between variables, making it an ideal setting for Machine Learning algorithms. The key assumption here is that skills are fully explained by observable skill-related variables, which can be restrictive in many settings. For the Indonesian application below, I use a rich set of variables related to educational background, results from self-administered cognition and intelligence tests, survey responses on risk and time preferences, spoken languages, and various literacy measures. In this setup, the key assumption would, for example, be violated if unobserved variables related to soft skills or mechanical skills are relevant for skills in government jobs, but are only insufficiently correlated with observed variables.

Again, two additional remarks are in order. First, a notable benefit of the approach is that one does not need to disentangle between which variables actually *cause* skills versus which are simply *correlated* with skills as long as the same relationship between $f(X_i)$ and skills holds for other workers in the economy. Second and related, correctly separating skills from permanent and temporary shocks to wages relies on using skill-related observables X_i that are uncorrelated with wage shocks. The following example illustrates practical difficulties with this assumption. Suppose that private sector workers in government-like jobs face a gender wage gap due to pure discrimination. Given the interest in actual skills, we do not want to include gender in X_i because the approach would then wrongly interpret the discriminatory gender wage gap as actual skill differences. However, leaving gender out of X_i will only correctly identify $f(X_i)$ if gender is uncorrelated with $f(X_i)$.

The third and final estimation step, as formalized in Corollary 5.1, exploits the mapping between skills z_i and skill-related observables X_i to obtain consistent skill estimates for any worker in the economy.

Corollary 5.1 (Step 3: Predicting skills out-of-sample)**.** *Given a consistent estimate of $f(X_i)$ over the support $S \equiv \text{supp}(X)$, the assumptions in Proposition 4 & 5, then knowledge of X is sufficient to obtain a consistent estimate of z_j for any worker j in the economy for whom X_j is observed (and for whom X_j has common support with S).*

Corollary 5.1 makes clear that the purpose of the second estimation step is not only to "regularize" noisy fixed effects. In this case, one could have also used approaches such as Ridge Regression that do not rely on covariates and the restrictive assumption that skills are solely a

[3] The proposition follows directly from the result that standard Machine Learning algorithms can approximate arbitrary non-linear functions (e.g. Hornik, Stinchcombe, and White 1989).

function of skill-related observables (e.g. see: Best, Hjort, and Szakonyi 2017). In combination with the last estimation step, the approach allows to predict skills in government-like jobs for any worker in the economy using skill-related observables. For workers in very different jobs, this allows to predict their "potential skills" in government jobs and for government workers it allows to quantify their actual skills without relying on their wages or requiring to observe government output.

At last, note that the common support assumption in Corollary 5.1 is not needed if one is willing to extrapolate on the functional form for $f(X_i)$. The approach already allows for systematic worker sorting across jobs based on individual and unobserved taste differences and sorting based on X_i. However, only in the case of perfect sorting on X_i some support of X_i will never be observed for the private sector workers in government-like jobs and hence skill predictions for any other workers with X_i outside the support will rely entirely on extrapolations of the estimated functional form $f()$. Common support on X_i can be empirically tested and is not an issue in the empirical application that I study below.

3.3 Estimating government worker skills in Indonesia

In this section I apply the identification approach to the case of Indonesia, the fourth most populous country in the world, over a period of almost 30 years from 1988 to 2014. I start by giving an overview of the context and underlying data and then go through each estimation step in turn.

3.3.1 Context & Data

Between 1988 and 2014, Indonesia moved away from a military-ruled, highly centralized authoritarian government under General Suharto (1967-1998) to a relatively consolidated, highly decentralized democracy. Incomes per capita have increased roughly 7-fold, poverty has been dramatically reduced and Indonesia has transitioned to become a middle-income country. The Suharto regime was characterized by extensive cronyism and patronage under which the bureaucracy was greatly expanded but also seen as a direct political instrument to collect votes and offer political support (Fisman 2001; Hadiz and Robison 2013; Martinez-Bravo, Mukherjee, and Stegmann 2017; Robison and Hadiz 2004). Civil Servants were obliged to become members of the political apparatus, support the party in power and hiring and promotion decisions were made to support regime stability (e.g. Kristiansen and Ramli 2006, 215; McLeod 2008). With the Asian Financial Crisis in 1997/1998 and the subsequent fall of the Suharto regime, Indonesia embarked on the *Era Reformasi*, targeting constitutional, judicial, public financial management and privatization reforms as well as an unprecedented decentralization process. In this new, decentralized system, three-fourths of the Civil Service (including teachers and health workers) are assigned to local governments in contrast to a fraction of this under the Suharto regime. Since the decentralization reforms, the central government can steer civil service hiring by setting overall quotas for the number of civil service jobs, while districts are left with a high degree of discretion as they decide on applicants and applicant requirements. Throughout the 2000s, a number of ministries started bureaucracy reform initiatives that tried to set civil service remuneration on par with the private sector and pushed for more competitive hiring and promotion practices (e.g. Horhoruw et al. 2013). Due to uneven adoption across ministries, in 2010, the government mandated public sector reform for all central and local governments, but only in 2014, a new Indonesian Civil Service law was passed.

I draw on nationally representative data from the Indonesian Family Life Survey – IFLS in short – which is particularly suited for the identification strategy in this paper. Specifically, the IFLS is a large and high-quality household- and individual-level panel dataset that collects exceptionally detailed information on individual's occupations, wages, skills and preferences. It is based on a sample of 7,224 households and 22,347 individuals tracked throughout five waves (1993, 1997-98, 2000, 2007-08, and 2014-2015), representing about 83% of the Indonesian population living in 13 of the nation's 26 provinces in 1993. Due to an intensive focus on respondent tracking, re-contact rates between any two rounds are above 90%, and 87% of the original households were contacted in all five rounds (see: Strauss, Witoelar, and Sikoki 2016). As a comparison, these re-contact rates are as high or higher than most longitudinal surveys in the United States and Europe.

Employment & wages: The IFLS data includes detailed employment data for each survey round. In addition to current employment, the survey included questions on previous employment. As in Hicks et al. (2017), this allows to create up to a 27-year annual individual employment panel from 1988 to 2014, making it one of the longest employment panel datasets available for developing countries and uniquely positioned to study life cycle wage growth (see: Lagakos and Shu 2023).[4] Employment information captures principal and secondary employment including government jobs. I focus on principal jobs as the job classification throughout this paper. For wages I use real hourly income based on total wages and total hours worked across all jobs, and by deflating nominal values using Indonesia-wide average monthly CPI-based inflation together with the specific survey months to best deal with periods of high inflation.

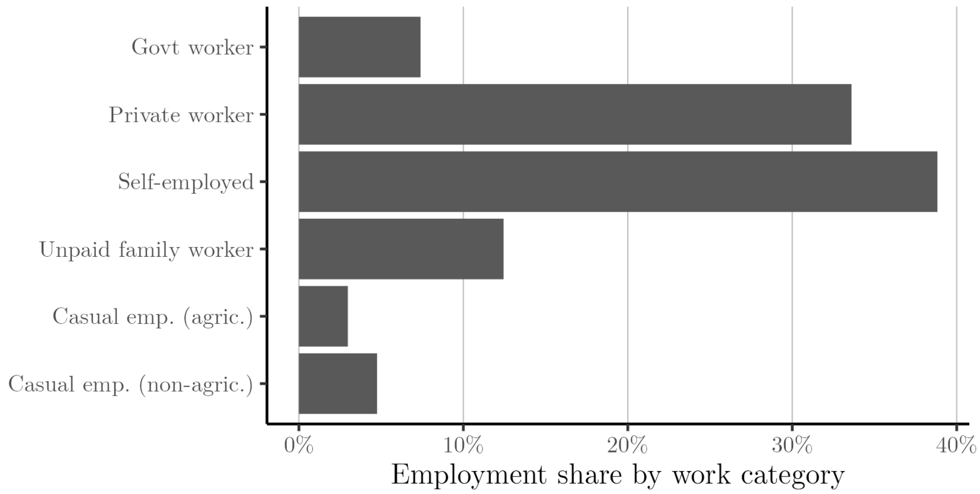
Skill variables: The IFLS data also contains an exceptional breadth of skill-related variables. I restrict myself to the 28 most important variables based on variable importance metrics in subsequent prediction tasks. These variables are dummies for the level of education, test results for self-administered Raven tests, cognition and memory tasks, the literacy of respondents, relative rankings in standard Indonesian exams, the big-5 personality traits and elicited risk and time preferences. I z-standardize all numeric variables pooling across all individuals.

Definition of government worker: In Indonesia, there are permanent civil service positions called *Pegawai Negeri Sipil (PNS)* as well as temporary civil servant positions (non-PNS). The latter are for example common in the educational sector where 60% of new teacher hires are on temporary contracts (Pierskalla and Sacks 2018). The IFLS includes both permanent and temporary government workers, defining government jobs broadly as jobs with any government office for which one receives remuneration in money or in kind. Appendix A provides further details on the hiring process of government workers in Indonesia.

Job categories, sectors and occupations: The IFLS distinguishes broad job categories (including government workers) from the job's sector at the 1-digit level and an occupational classification at the 2-digit level. To allow for sufficiently large occupational groups for government workers, I recode the occupational classification into nine different 1- to 2-digit occupations

[4] 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. Following Hicks et al. (2017), the earnings measure in this paper is the sum of all wages, profits, and benefits.

Figure 3.1: Employment shares by broad job category



Notes: Barplot of broad job categories for primary occupations over the pooled panel of employment histories 1988-2014. Source: IFLS, Pooled sample: N = 403,626. Unique individuals: n = 36,126

that captures most variation in government sector jobs. For example, I leave the 2-digit codes for "teachers" and for "government officials and executives", while I keep the 1-digit code "other service areas".

3.3.2 Descriptives

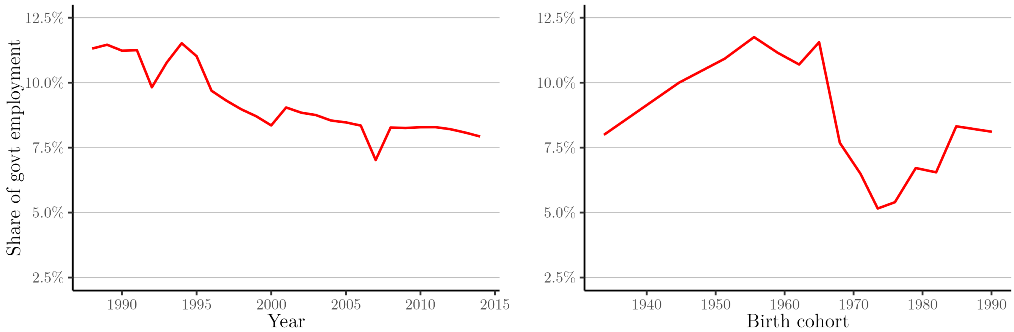
Figure 3.1 plots the share of government workers compared to the employment share of the other most common categories of work in developing countries, pooling all waves. On average, about 7% of the workforce are government workers, around 33% of workers have permanent, formal jobs in the private sector, roughly 40% of workers follow some form of self-employment and the remaining 20% of workers are casually employed or unpaid. Compared internationally, government employment in Indonesia is relatively low as it often exceeds 10% of the workforce for middle-income countries and may easily exceed 20% for high-income countries (e.g. Finan, Olken, and Pande 2017). Based on complementary government statistics, up to 50% of the government employees captured in the IFLS data should be non-permanent government workers.[5]

Figure 3.2 shows the evolution of the government workforce and hiring over time. In line with the stated government policy of reducing government employment, the left panel shows that the share of government workers has been slowly declining since 1988. The decline also shows up in the data through a large drop of hiring for birth cohorts after 1965 as reported in the right panel of Figure 3.2.

Next, we can ask what types of jobs government workers are doing in Indonesia. Figure 3.3 shows the distribution of government jobs across 10 different sectors and compares their shares to the private sector. More than 80% of government sector jobs are in social services, including education and health, with the next biggest sector being agriculture and forestry at around 6%.

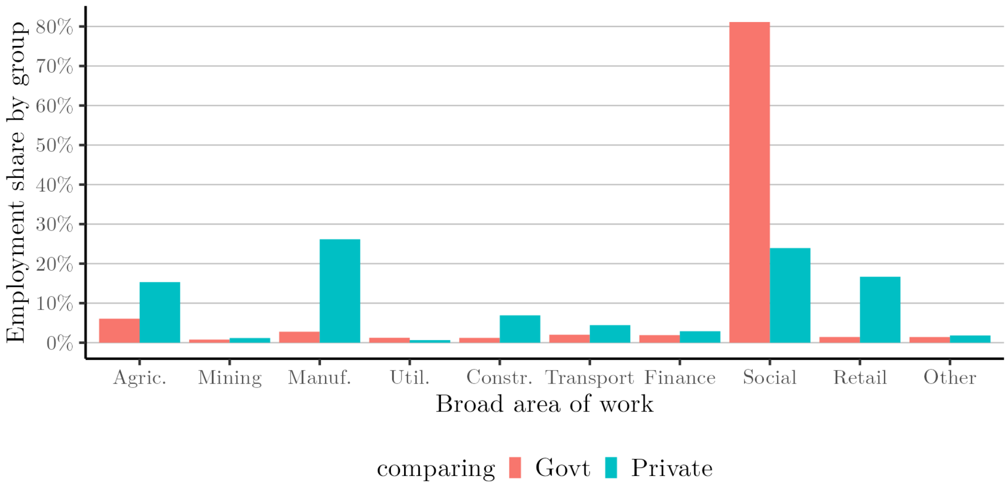
[5]Specifically, permanent government workers only make up roughly 3.5-4% of all employees in Indonesia based on a Civil Service census from the year 2015 reported in Pierskalla et al. (2020) and the total number of employees for the same year taken from Statistics Indonesia.

Figure 3.2: Evolution of government employment shares



Notes: Evolution of government employment (as primary occupation) as a share of total employment by year from 1988 to 20014 (left) and by 15 equal-sized birth cohort bins (right). Restricted to age range 25 to 58 to account for early government retirement age and tertiary education for government workers. Source: all five waves from the IFLS (1993,1997,2000,2007,2014), Pooled sample: N = 303,368. Unique individuals: n = 29,992. Unique government workers = 2,813.

Figure 3.3: Employment shares across sectors of work for government vs. private sector workers



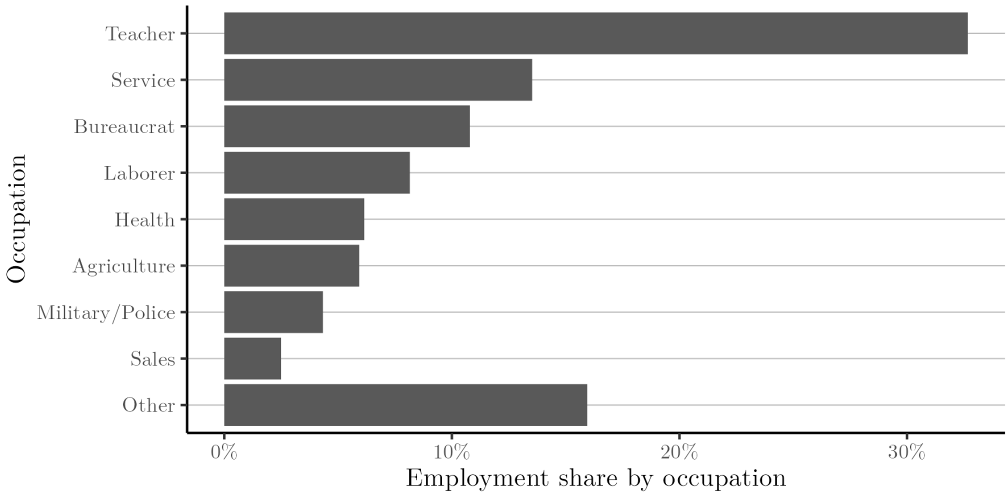
Notes: Pooled data for 1988-2014. Source: IFLS, Total observations = 165,257. Unique individuals: n = 23,377, Govt workers = 3,230. Excludes observations with missing sectoral codes.

In comparison, only about 24% of private sector jobs are in social services.

In Figure 3.4, I provide further details on which occupations government workers hold. More than 30% of government workers are teachers, more than 10% are government officials and executives, and 13.5% of government workers work in other service areas. Not surprisingly, the prevalence of the same occupations among private sector workers looks very different: while there are almost no private military/police nor bureaucrats and about twice as many government teachers than private teachers, there are 20x more private laborers and private sales workers and 12x more private agricultural workers.

Another important feature of the data on government employment are transition rates between the private and the public sector, defined as individuals who are observed switching either from a non-government job to a government job (in-mover) or the reverse (out-mover). About 10% of person-year observations are movers of which slightly more move into the government than out (around 53% vs. 47%). Among those who move, 42% are observed to move both in

Figure 3.4: Government employment shares by occupation



Notes: Barplot of occupations among government workers, 1988-2014. Source: IFLS, Pooled sample: N = 29,855. Unique govt workers = 3,238. Excludes observations with missing occupational codes. Occupation is a combination of 1- and 2-digit occupational codes reported in the IFLS data.

and out of the government.

At last, Table 3.1 reports how government workers differ from private sector workers and other workers in the economy in terms of observables, including skill variables. On average, government workers are more likely to be male, older than private sector workers, earn about 80% higher hourly wages and work slightly less than private sector workers. Importantly, government workers are on average much more skilled: they are much more likely to have received higher education, perform better on self-administered memory and word ability tasks, and they achieved higher rankings on nationally administered exams. While not reported here, one can also note that these skill measures are mostly increasing for all groups of workers over time.

3.3.3 Step 1: Estimating "noisy" skills

Following Proposition 4, to obtain "noisy" estimates of worker skills in government-like jobs, the first substep is to select a suitable government-like job for which we estimate skills. I use private sector workers in social services as the main estimation sample in this paper given the prevalence of social service jobs for government workers.[6] The choice of the estimation sample balances the need for a large enough sample and comparability with government jobs. Note that with larger sample sizes one could also separately apply the estimation approach to more disaggregated job categories. The level of aggregation that I choose alleviates the difficulty of finding exact private sector counterparts to specialized jobs such as the police. However, the underlying assumption is that even in such jobs, the general skill set needed broadly aligns with the skills needed for social service jobs in the private sector. As explained in more detail in Appendix A, civil servants in Indonesia are also not trained as specialists, but rather go through a common selection and training process that emphasizes more general skills, which is exactly the skills that the estimation approach in this paper seeks to pick up. Another cri-

[6]Specifically, I use private sector workers who have worked in a social services job before, to alleviate the issue of missing values in the sectoral category of individual's jobs and to capture individuals who transition jobs.

Table 3.1: Main observable differences: government vs. private sector vs. all other workers

Variable	Govt	Control	Private	Other
Share male	0.69	0.67	0.67	0.63
Mean age	39.32	33.09	33.80	42.12
Mean relative wage	1.81	0.99	0.90	0.93
Mean weekly hours	39.44	41.44	40.13	38.30
Share higher educ	0.52	0.16	0.09	0.03
Mean word recall	0.15	-0.07	-0.28	-0.70
Share speak Indonesian	0.41	0.39	0.30	0.18
Mean word ability	0.40	-0.04	-0.22	-0.55
Mean math IQ	0.35	0.01	-0.07	-0.16
Mean national ranking (total)	0.05	-0.03	-0.05	-0.02
Mean national ranking (language)	0.01	0.02	-0.01	-0.01
Mean national ranking (math)	0.00	0.00	-0.02	-0.01

Details: Based on pooled data and restricting to workers in working age (between 16 to 70 years old), positive work hours and non-missings in age, experience, work hours and income. N = 137,8527. Unique individuals = 24,206. Unique govt workers = 2,342. Unique control workers = 4,812. Unique private sector workers = 13,394. Unique other workers = 8,470. The control group are private sector workers who have worked in social services before. Mean relative wage gives the group-specific mean of the wage divided by the year-specific mean wage across all groups (with means trimmed at the 2.5th and 97.5th percentiles). Apart from dummy variables, all skill-related observables are z-standardized across all individuals.

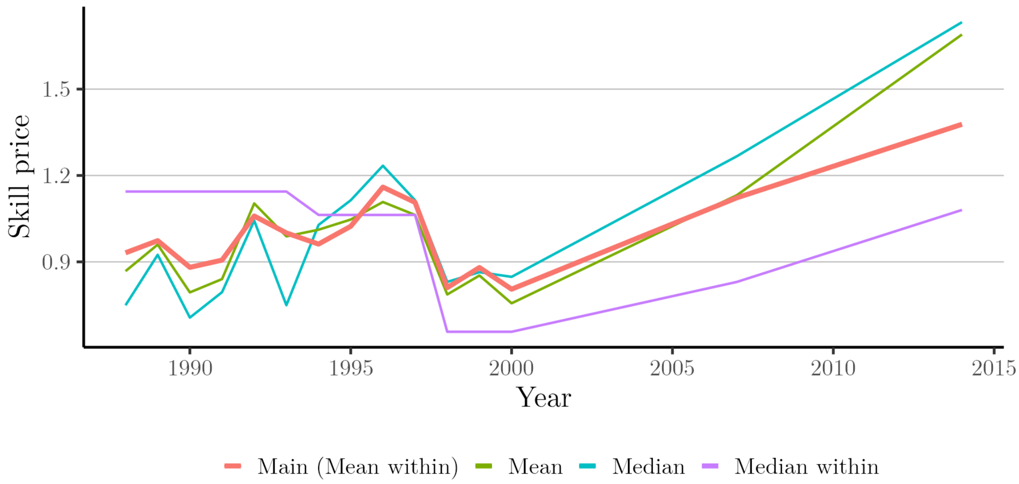
tique may be that the more multi-dimensional nature of government jobs makes performance measures difficult in the public sector (e.g. Finan, Olken, and Pande 2017). However, the same argument can be made for social service jobs in the private sector. And importantly, the skill estimation approach here specifically allows for the idea that measuring performance in a job is difficult and that wages are only noisy signals of underlying human capital and skills. In the results section, I will specifically test the informativeness of wages for underlying skills and compare the private and public sector.

After specifying the estimation sample, the next substep is to purge the observed wage from changes in the skill price. Figure 3.5 plots the estimated (equilibrium) skill price drawing on Bowlus and Robinson (2012) and Proposition 4.[7] The estimates indicate a slightly increasing price between 1988 and 1997 after which the price drops sharply with the Asian Financial Crisis, in line with a large drop in the overall demand for skilled labor in the economy and particularly in social services. After the crisis, the skill price recovers strongly, growing by more than 50% between 2000 and 2014, in line with growing demand as the economy expands and moves further into services. Overall, the increase in the skill price was much slower than the more than 7-fold increase in real GDP per capita over the entire time period, in line with a consistent growth in the supply of skilled labor. The slower increase in the skill price during the early 1990s is also in line with average years of schooling in Indonesia increasing particularly strongly in the late 1980s (Duflo 2001, 2004).

The last substep following Proposition 4 is to obtain a noisy estimate of individual-level human

[7]Skill price estimates are usually sensitive to the exact flat spot region chosen. I chose an age region from 45 to 60, which is in line with regions usually used in the literature and is consistent with the subsequently estimated wage experience profile.

Figure 3.5: Evolution of skill price



Notes: Skill price estimation following Bowlus & Robinson (2012). "Main (Mean within)" gives the baseline estimator based on Proposition 4, while "Median within" reports the same estimator using median changes instead. "Median" and "Mean" alternatively report estimates without demeaning within-individuals first. For the flat-spot region, I use the age range from 45 to 60, consistent with subsequently estimated experience profiles. For all estimates, the level of prices is normalized such that the time-series average is unity.

capital by disentangling experience effects from individual-specific skills at labor market entry. Assuming that deflated log wages follow: $\tilde{w}_{i,e,t} \equiv w_{i,e,t} - p_t = z_i + \delta_0 * exp_{i,e,t} + \delta_1 exp^2_{i,e,t}$, Table 3.6 reports estimates for the parameters of the experience profile: (δ_0, δ_1), which are identified from within-individual changes in wages and experience. Within-individual variation is key to ensure that estimated experience profiles are not biased by systematic composition changes in skills at labor market entry, such as due to less experienced cohorts entering the labor market with better education.[8] The estimated parameters give concave experience profiles with a flat-spot around 28 years of experience, in line with the assumed flat-spot region for the skill price estimation.

3.3.4 Steps 2 & 3: Predicting skills

In Step 2 of the estimation approach, following Proposition 5, the noisy individual fixed effects are flexibly projected onto a larger set of skill-related observables. I use standard off-the-shelf Machine Learning algorithms for this task. Specifically, I compare three different algorithms with a benchmark of a standard OLS regression with dummies for the educational background (primary, junior secondary, senior secondary and higher education). The three algorithms are LASSO using all variables and all their first-order interaction terms (which includes squared terms), Random Forest and Gradient-Boosted Trees. Due to the high flexibility of Machine Learning algorithms, it is important to avoid overfitting, which leads to noise in small sample estimation. To avoid overfitting, I follow standard practice and use 15-fold cross-validation and train the hyperparameters of each Machine Learning Algorithm via a simple, coarse grid search. I choose a higher than usual number of cross validation folds due to the smaller sam-

[8]Note that the IFLS only has wage information for two separate years after the year 2000 since the 2007 & 2014 waves do not ask for wages retrospectively. This means that within-worker wage experience profiles are more informed by wage profiles prior to the year 2000. However, the limitation does not preclude to estimate worker fixed effects for workers that enter the labor market after the year 2000 or even workers whose wages are only observed once.

Table 3.2: Comparing performance and similarity of different Machine Learning algorithms

Measure	Indiv.	OLS	LASSO	RF	GBM
R^2	4785	0.1097	0.1601	0.1453	0.1794
Corr(OLS,x)	4785	1.0000	0.7143	0.5007	0.5919
Corr(LASSO,x)	4785	0.7143	1.0000	0.7871	0.9321
Corr(RF,x)	4785	0.5007	0.7871	1.0000	0.8562
Corr(GBM,x)	4785	0.5919	0.9321	0.8562	1.0000

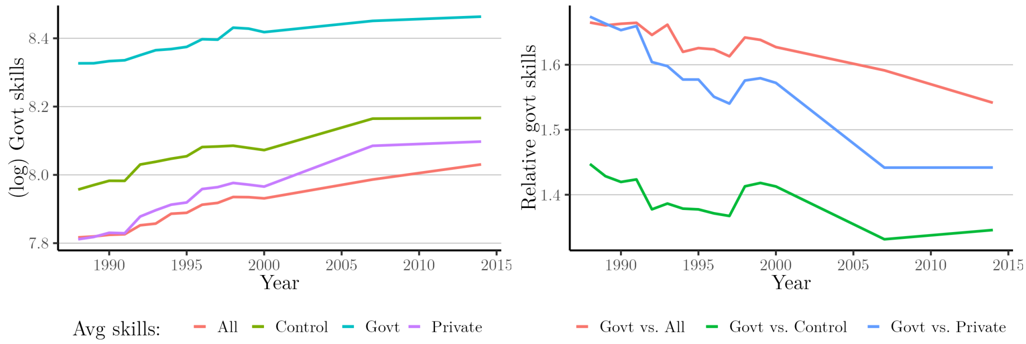
ple size, avoiding additionally separating between validation and test datasets.[9] The final individual-level dataset for "training" the skill estimation algorithm includes 4,785 unique individuals.

Table 3.2 reports the performance of each of the different algorithms in terms of its R^2. I find the GBM algorithm to perform the best with an R^2 of around 18%, while the LASSO algorithm achieves 16% and the Random Forest algorithm only 14.5%. In comparison to the simple OLS algorithm, the Machine Learning algorithms can explain about 60% more variance in the individual fixed effects, highlighting the importance of accounting for non-linearities and additional variables. Interpreting the difference between LASSO & GBM as the importance of non-linearities allows a simple variance decomposition that attributes 72% of the predictive gains compared to OLS to having additional variables and 28% to allowing for additional non-linearities.

So what information are the Machine Learning algorithms picking up and what predicts skills? Focussing on the GBM estimates as the best-performing baseline estimates, Table 3.7 in Appendix B.1 ranks variables by variable importance measure that weights the relative importance for each covariate in the prediction task taking into account nonlinearity and interactions of the different variables. We can see that the algorithm does identify the various educational background dummy variables as the most important variables for predicting skills, but among the top 15 most predictive variables there are also scores on word cognition tasks, Raven IQ-score, some of the Big-5 measures and the ability to speak the national language. Another way to see the importance of using additional covariates is to look at the correlation of predictions across different algorithms as reported in Table 3.2. Comparing the predictions of the three different Machine Learning algorithms with the OLS predictions, we can note that the correlation is at most 0.71, indicating that the OLS predictions are missing very important variation that the different Machine Learning algorithms take into account. Secondly, we can compare the different predictions of the Machine Learning algorithms among each other. For example, comparing LASSO with gradient-boosted trees, we find that their correlation is around 0.93, indicating that they both capture very similar additional variation.

[9]In larger datasets, ensemble methods combining different Machine Learning algorithms could also provide additional gains for prediction tasks and reduces model sensitivity if sufficiently regularized. Model sensitivity usually plays a big issue, but Machine Learning algorithms used on larger datasets can potentially ameliorate the issue.

Figure 3.6: Evolution government worker skills



Notes: Skill estimates show results from baseline specification with private sector social service workers (Control group in Figure) as estimation sample and GBM estimator as ML algorithm. Relative skills are in log differences. Data is pooled across all waves of the IFLS and then plotted by year.

3.4 Main empirical results

Taking the individual skill estimates based on the best-performing Machine Learning algorithm, I now show two major applications using estimates of government worker skills for any worker in the economy. I discuss each in turn.

3.4.1 Application 1: Changes in skills & selection

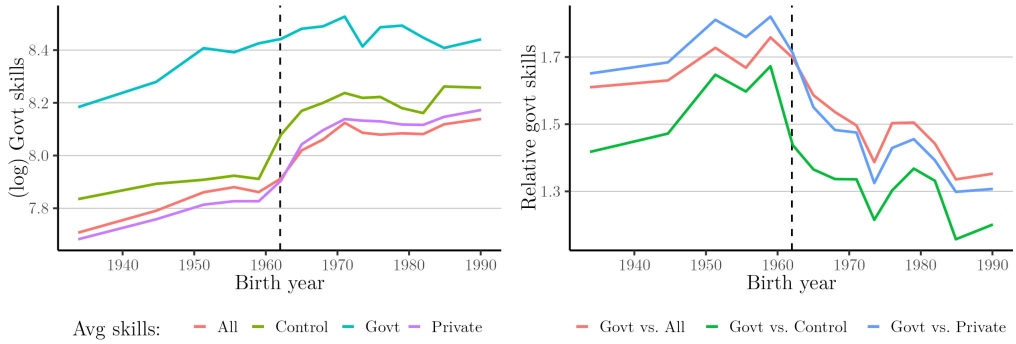
In the first application of the skill estimation approach, I quantify that government workers are highly selected, but that the quality of the government workforce and its selection changed systematically over time.

Evolution of absolute and relative skills

Figure 3.6 plots the evolution of government worker skills over time as well as relative skills of government workers compared to other workers in the economy, revealing three important facts. First, government workers are on average much more skilled: in 1988, they are roughly 45% more skilled than private sector workers in social services (the control group) and more than 65% more skilled than private sector workers, in line with skill differences reported in Table 3.1. Second, as visible on the left plot, average skills increased by roughly 30% across all workers since 1988, again, in line with general increases in education. Third, while these average increases in skills also hold separately for government workers and non-government workers, the increases were more pronounced for non-government workers such that the relative skill premium of the government has steadily declined since 1988.

What explains these changes in average skills over time? Given that the focus was on skills at labor market entry that are fixed over an individual's lifetime, changes in average and relative skills of government workers are entirely driven by changes in the composition of workers. The latter changes due to hiring of young cohorts and exiting of old cohorts as well as due to mid-career transitions into and out of the government. Figure 3.7 shows that overall skills are strongly increasing across birth cohorts, with a big increase starting around 1960, which coincides with the largest school construction program in Indonesian history (see: Duflo 2001). At the same time, the relative skill premium of government workers strongly declined across

Figure 3.7: Government worker skills across cohorts



Notes: Skill estimates show results from baseline specification with private sector social service workers (Control group in Figure) as estimation sample and GBM estimator as ML algorithm. Relative skills are in log differences. Data is pooled across all waves of the IFLS and then plotted by (binned) cohort. Cohort bins are determined by equal-sized bins in pooled data. Sample restricted to working ages of government workers (25-58). Dotted line denotes the first cohort that was treated by the INPRES school construction program studied in Duflo (2001).

birth cohorts also starting around 1960. This relative decline in government worker skills is particularly surprising given that Figure 3.2 already showed that government hiring declined in birth cohorts shortly after 1960. In the case that government jobs are in high demand, one would have expected that a decline in supply would have led to an increase in relative skills as long as the government is good at selecting workers on skills. So is the Indonesian government simply becoming worse at selecting workers? Or are government jobs becoming less attractive for highly skilled workers? Or are there maybe changes in the types of jobs that the government is offering for which the skill requirement changed systematically? The next two subsections zone in on these questions.

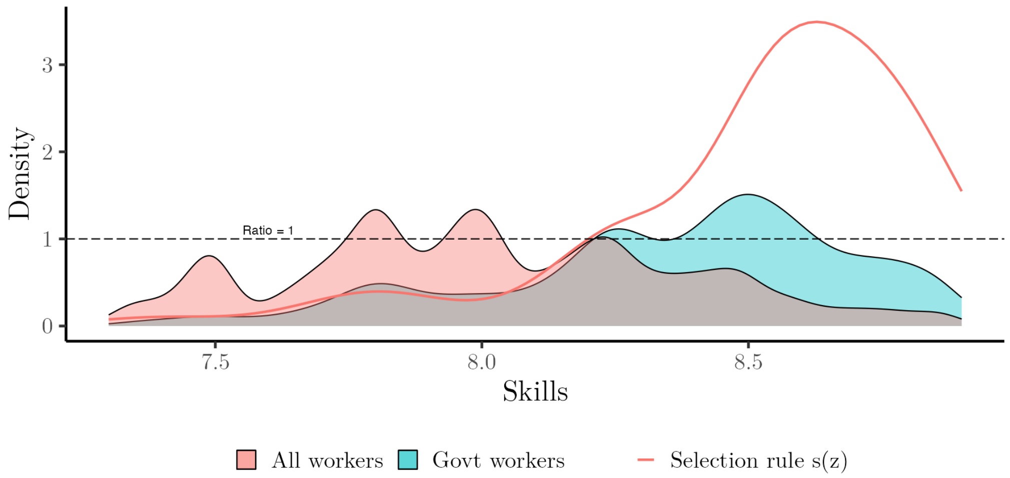
The government selection rule

Motivated by the decline in the relative skill premium of government workers, I now turn to studying the selection of government workers across the entire skill distribution. In practice, governments – including the Indonesian government – use multi-stage selection processes that include restrictions on who can apply, entry exams and a selection on a combination of observables that are revealed through CVs and information provided during the application process. At the same time, corruption, clientelism and politically-motivated hiring can influence the selection at different stages (Colonnelli and Prem 2017; Hanna and Wang 2017; Jia, Kudamatsu, and Seim 2015; Weaver 2019). This holds particularly so for Indonesia, where there is solid evidence that politics has an influence on selection into and promotion patterns within the Civil Service (Pierskalla and Sacks 2018), that there is widespread corruption among bureaucrats (Valsecchi 2016), and where civil service jobs have historically been sold in an auction-like fashion (Kristiansen and Ramli 2006). To study how "good" the government is *de facto* at selecting government workers, I define the following simple reduced-form *selection rule* $s_t(z)$:

$$s_t(z) \equiv f_t(z|(\text{selected}|\text{applied})) = \frac{f_t(z|\text{applied} \cap \text{selected})}{f_t(z|\text{applied})} \tag{3.5}$$

where $f_t(z|\text{applied})$ gives the skill distribution of the applicant pool and $f_t(z|\text{applied} \cap \text{selected})$ the skill distribution of newly selected government workers. The selection rule is above unity

Figure 3.8: Estimated government selection rule



Notes: The selection rule is based on the distribution of estimated skills for all workers and for the subset of government workers only. In both cases, the estimates are based on all unique workers, not repeated observations per worker. Estimates are based on a direct estimator of the density ratio using unconstrained Least-Squares Importance Fitting (uLSIF) taking standard values for the corresponding hyperparameters: lambda (0.2) and sigma (0.1).

at any skill level z if the state selects more people with these skills than would be expected under uniform drawing from the conditional skill distribution of applicants. A key property of a good selection rule would then, for example, be that it is increasing in skills, such that higher quality candidates are more likely to be employed.[10] An implicit assumption here is that there are more applicants than government workers to be selected, which is the empirically relevant setting for developing countries and many developed countries.[11]

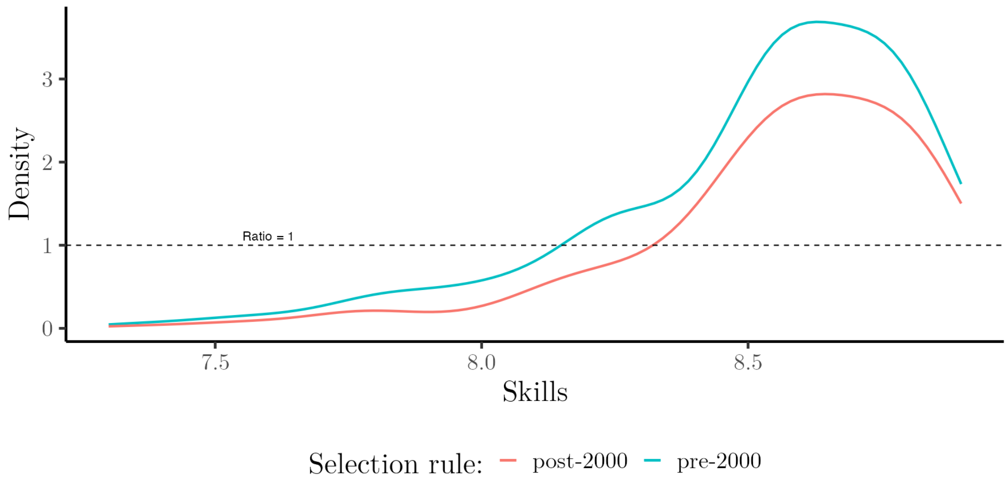
How do we measure the two conditional distributions? While the sample of selected government workers directly identifies the numerator of the selection rule, the denominator given by the applicant skill distribution turns out to be trickier. Conceptually, the applicant pool is trickier because one needs to take a stance on what the selection rule *should* capture. For example, we could interpret the selection rule narrowly by understanding applicants as individuals who have actually filled out an application form. This information may be readily available in many contexts. However, formal application requirements may already exclude highly skilled potential government workers such that a more inclusive definition of the applicant pool can be more interesting in practice. Since the IFLS data unfortunately does not measure "interest in a government job" nor whether a worker applied for a government job in the past, I look at the limit case assuming that every worker in the economy would take a government job; a useful benchmark that will be a good approximation in many developing country settings where interest in a safe government position is ubiquitous. This limit case is likely a lower bound for the selection rule because we expect low-skilled individuals to be more likely to take a government-sector job than high-skilled individuals who are more likely to have good outside options.

Figure 3.8 shows the estimated selection rule and the conditional skill density functions for

[10]This is equivalent to the two conditional densities exhibiting the Monotone Likelihood Ratio Property

[11]In 2014, there were around 2.6 million applicants for 100k open civil service positions in Indonesia (see here as of April 22, 2024).

Figure 3.9: Estimated changes in government selection rule



Notes: The two estimated selection rules are based on the distribution of estimated skills for all workers and for the subset of government workers only. For before 2000, the estimation data is restricted to workers who show up in the IFLS data in the year 2000 or before. For after 2000, the selection rule is estimated using government workers who entered the sample after 2000, and all other workers who entered the labor force after the year 2000 to focus on new entrants in the labor market. Estimates are again based on a direct estimator of the density ratio using unconstrained Least-Squares Importance Fitting (uLSIF) taking standard values for the corresponding hyperparameters: lambda (0.2) and sigma (0.1).

government workers and all workers separately.[12] Two main points are worth noting: First, the government selection rule is indeed mostly increasing. The Indonesian government is much less likely to pick low skilled workers as civil servants compared to their density in the data, which translates into a ratio below 1, and is more likely to pick high skilled workers leading to density ratios above 1.[13] Second, the selection rule drops for highly skilled workers. If indeed everyone in the economy would take a government job, the point estimates imply that the government becomes worse at selecting the most qualified workers in the economy. The more likely explanation in this case (apart from statistical uncertainty) is that we have misspecified the applicant pool and that the individuals that would be most skilled in a government sector job are actually not interested in taking government jobs.

How does the selection rule account for the decline in the skill premium of government workers? Suppose the selection rule, which the government determines through its hiring practices, stayed constant over time. In this case, a systematic rightward shift of the skill distribution in the population has ambiguous effects depending on where the distribution is. As long as the mass of applicants at the left of the selection rule's maximum moves up the curve, these workers are more likely to be selected, increasing government worker skills. However, if skills for high-skilled workers shift further, then they eventually become less likely to end up working for the government, decreasing skills.

[12]To estimate the ratio, I use a direct estimator of the density ratio using unconstrained Least-Squares Importance Fitting (uLSIF) proposed by Hido et al. (2011), which is more robust than separately estimating skill distributions for nominator and denominator and then forming the ratio of the two. Similar to density estimation, there are many different direct estimators for density ratios proposed in the literature, which I found to perform very similarly for my application.

[13]The visible spikes in the skill distributions are due to the importance of the discrete educational background and bunching of same key missing variables, which the Machine Learning algorithm simply treats as another categorical value of the respective variable.

At last, the decline in the skill premium of government workers could also be explained by changes in the selection rule itself. Between 1988 and 2014, Indonesia moved from a highly centralized autocratic government with a regime-aligned bureaucracy (Hadiz and Robison 2013; Robison and Hadiz 2004) to a democratic system with a highly decentralized bureaucracy (Blunt, Turner, and Lindroth 2012; Brinkerhoff and Wetterberg 2013). These changes were accompanied with a decline in government hiring (see Figure 3.2) as well as numerous bureaucracy reforms starting unevenly across ministries in the 2000s and a change to a computer-assisted selection that seems to have greatly reduced corruption at the selection stage (Kuipers 2021). Figure 3.9 shows estimated changes in the *de facto* selection rule when restricting to workers that started working before the year 2000 versus after 2000. To interpret changes in the selection rule, suppose that the distribution of applicants stayed the same: in this case, the estimates indeed suggest that after 2000 it became harder for low-skilled workers to work for the government, indicated by the downward shift in the selection rule. At the same time, the selection rule is also lower for high-skilled workers suggesting that the government has not become better at selecting high-skilled workers. Taken together, the relative decline in the skill premium of government workers documented in Figure 3.6 thus seem mostly driven by skill improvements in the applicant population, and only little influenced by improvements in the selection of government workers.

Relative skills & changes in government hiring

In this final subsection, I provide further evidence on a main mechanism that impacts changes in relative skills of government workers: As overall government hiring declines, the government should be able to hire relatively better workers since they have a larger pool of applicants to choose from. For Indonesia, this mechanism counteracts some of the decline in relative skills as more skilled workers end up working in the private sector.

To provide evidence for the mechanism, I consider changes in the government employment share. The idea is that there is strong year-to-year variation in how much the government is hiring, driven in part by political cycles and the discrete nature of legislation. If hiring by the government happens mostly at labor market entry, then one would expect differential exposure of birth cohorts to government hiring, as evidenced by observed differences in cohort-specific government employment shares. Conditional on filling the same job, more treated cohorts should thus show lower average skills. Table 3.3 shows regression evidence that this is indeed the case. Column 1 shows that the average skill premium of government workers is indeed declining in year-to-year changes in the cohort-specific government employment share. Column 2 shows that this effect also holds conditional on the same job, using job occupation and sector fixed effects. At last, Columns 3 & 4 show the same results additionally controlling for a time trend in skills.

3.4.2 Application 2: Government wage setting

The second major application looks at the wages of government workers; specifically at the informativeness of government wages and whether government workers are overpriced compared to the private sector.

Table 3.3: The government skill premium and changes in government hiring intensity

Dependent Variable:	Skill			
Model:	(1)	(2)	(3)	(4)
Variables				
Constant	7.866***		-1,004.170***	
	(0.001)		(73.670)	
Change govt empl share	0.465***	0.361***	0.425***	0.334***
	(0.025)	(0.046)	(0.024)	(0.045)
Govt worker	0.536***	0.202***	0.547***	0.220***
	(0.003)	(0.051)	(0.003)	(0.051)
Change govt empl share × Govt worker	-0.678***	-0.574***	-0.574***	-0.489**
	(0.065)	(0.176)	(0.064)	(0.171)
year			1.001***	0.760**
			(0.074)	(0.261)
year square			0.000***	0.000**
			(0.000)	(0.000)
Fixed-effects				
Occupation		Yes		Yes
Sector		Yes		Yes
Fit statistics				
Observations	107,759	106,854	107,759	106,854
R^2	0.19	0.35	0.22	0.37
Within R^2		0.02		0.05

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Data based on full worker panel, but restricting to workers between the age of 25 to 58 (the working age of government workers in Indonesia). Occupation and sector fixed effects reduce the sample size slightly due to missings in the occupation variable.

Are government wages uninformative?

The starting point of the estimation approach was the idea that government worker wages are potentially uninformative about underlying skills, because – as in the case of Indonesia – they follow rigid tenure schedules and allowances with little use of performance incentives or wage dispersion. For Indonesia, I find more nuanced results; government wages are far from uninformative about underlying skills, but I also find that comparable private sector wages are more informative about skills. I show this with two different exercises. First, I test whether government worker wages are indeed less informative about underlying skills than in the private sector by predicting real hourly wages for government workers and private sector workers in social services using estimated skills. Table 3.8 in the Appendix reports these results. Comparing Columns 1 & 4 shows that government worker wages are indeed less informative about underlying skills than wages in comparable private sector jobs: the R^2 is 0.114 vs. 0.157. However, this relative ranking changes once one controls for experience profiles, given a far more deterministic government wage experience schedule. For example, even when enforcing the same experience coefficients as in the comparable private sector (columns 2 & 5), life cycle human capital of government workers predicts 18.3% of the variation versus 18% in the private sector. The R^2 increases to 22.1%, roughly doubling the explanatory power compared to only job market entry skills when allowing for government-specific experience coefficients (columns 3 & 6).

The second test of the informativeness of government worker wages is to re-estimate government worker skills following all estimation stages but using government wages instead and then looking at the correlation across the two skill measures. As shown in Figure 3.10 in the Appendix, this correlation is high with an R^2 of around 74%. Assuming that baseline skills are estimated correctly, government wages are thus informative about underlying skills, but are biased. As shown in Figures 3.11 and 3.12 in the Appendix, the bias and the variance of the bias increase in skills, such that skills based on government wages are a particularly bad measure for the most skilled government workers, underestimating their skills on average and being less informative for them.

Is there a government wage premium?

The estimation approach allows for a direct measurement of the wage premium of government workers: conditional on the same level of skills, do government workers earn higher wages? The answer is: yes. Government workers earn a large wage premium of at least 30%. Specifically, Table 3.4 reports different estimates for the wage premium restricting to government workers and the control group of private sector workers in social services. Column 1 documents a large unconditional wage premium of 0.7 log points, which translates into twice as high real hourly wages. The wage premium roughly halves when controlling for worker skills (column 4), pointing to strong positive selection on skills. The wage premium also generally reduces when focussing on within-job comparisons by introducing occupation and sector fixed effects (columns 3 & 5). To avoid biasing estimated wage premia due to compositional changes over time – for example, if there are more control group workers in more recent periods – results for columns (2) to (8) include year fixed effects. At last, one may be interested in capturing the wage premium conditional on human capital that incorporates experience. Columns (6) & (7) report wage premia controlling for experience using the life-cycle skill measure $h_{i,e,t}$ that incorporates estimated human capital experience profiles. These estimates lead to the most conservative wage premia of around 0.26 log points, or a wage premium of roughly 30%. The last column considers changes in the wage premium over time and finds that the government wage premium is clearly increasing over time.

What is driving this large observed wage premium? It turns out that an important driver of the government wage premium is the absence of a gender wage gap in the public sector. Women in Indonesia's private sector face a large real hourly wage penalty compared to men in the private sector. Table 3.5 reports a raw gender wage gap of 0.432 log points, or a roughly 35% (!) wage cut for women compared to men in the private sector (column 1), which decreases to about 0.3 log points – or a wage penalty of 26% – after controlling for the same job, experience and skills (column 4). This gender wage is also stable over time (coefficient "Male x year" in column 5). The gender wage gap in the public sector, on the other hand, is much smaller, at about 1/4 of the wage gap in the private sector after controlling for the same job, experience and skills (column 4 and using the sum of the coefficients for "Male" & "Govt worker x Male"). That is, in the absence of a differential gender wage gap in the private sector, the government wage premium would be roughly 30% lower (the sum of coefficients "Govt worker" & "Govt worker x Male" in column 4 in comparison to the coefficient for "Govt worker" in Table 3 column 7).

3.5 Extensions

This section discusses main extensions of the skill estimation approach.

Table 3.4: Regression results: Government wage premium

Dependent Variable:	log(real hourly wage)							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	8.419^{***}							
	(0.007)							
Is govt worker?	0.704^{***}	0.721^{***}	0.519^{***}	0.377^{***}	0.369^{***}	0.246^{***}	0.262^{***}	0.087^{***}
	(0.012)	(0.034)	(0.042)	(0.028)	(0.032)	(0.030)	(0.032)	(0.029)
Skill				1.054^{***}	0.942^{***}			
				(0.042)	(0.049)			
Life-cycle skill						1.183^{***}	1.090^{***}	1.093^{***}
						(0.042)	(0.047)	(0.049)
Govt worker X year								0.016^{***}
								(0.002)
Fixed-effects								
Year		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation			Yes		Yes		Yes	Yes
Sector			Yes		Yes		Yes	Yes
Fit statistics								
Observations	33,583	33,583	33,362	33,583	33,362	33,583	33,362	33,362
R^2	0.09	0.13	0.19	0.24	0.26	0.27	0.29	0.29
Within R^2		0.10	0.04	0.22	0.12	0.25	0.16	0.16

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample restricted to government workers and control group (private sector workers who have worked in social services at any point in time). Standard errors are clustered at the year level (except for column 1 which assumes IID standard errors). Adding occupation and sector fixed effects reduces the sample size slightly due to missings in the occupation variable.

3.5.1 More flexible skill-experience profiles

The estimation approach for government worker skills in this paper allows for far more flexibility in skill experience profiles. A natural extension allows not only for a measure of skills at labor market entry, but also for individual-specific learning capabilities. Such a measure would then also allow to consider selection of workers on learning capabilities. A similar alternative would be to allow for multi-dimensional individual skill estimates. In both cases, the key to estimation would be to construct multiple individual-specific fixed effects and then separately regress these on skill- or learning-related observables. To see how this would look, assume that human capital $h_{i,e,t}$ evolves according to the following factor structure:

$$h_{i,e,t} = z_i * e^{g_i * \delta_e} \tag{3.6}$$

where z_i are individual time-fixed skills at labor market entry, g_i are individual-specific learning capabilities and δ_e are arbitrary experience-fixed effects. Government worker skills are defined as $z_i = h_{i,0,t}$, assuming that $\delta_0 = 0$, and the correlation between the two different fixed effects is unrestricted. The factor structure allows individual-specific skills and individual-specific wage-experience profiles, but restricts the shape of the latter to a factor structure. This can be seen as a generalization of earlier panel data methods such as the Within estimator which allows only individual-specific levels. The factor structure can match well empirically observed concave wage-experience profiles that differ in slope across individuals. Specifically, it has been well documented that wage-experience profiles are steeper for highly educated than for less educated individuals and that the variance of the slope of wage-experience profiles is increasing with observed skills such as education (Guvenen 2009; Guvenen and Smith 2014; Kahn and Lange 2014; Lagakos et al. 2018; Primiceri and Van Rens 2009). Alternatively,

Table 3.5: Regression results: Government wage premium & gender gap

Dependent Variable:	log(real hourly wage)				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Constant	8.130***				
	(0.013)				
Govt worker	1.032***	0.717***	0.510***	0.400***	0.223***
	(0.021)	(0.049)	(0.036)	(0.038)	(0.038)
Male	0.432***	0.477***	0.365***	0.303***	0.291***
	(0.016)	(0.020)	(0.018)	(0.017)	(0.025)
Govt worker × Male	-0.489***	-0.350***	-0.242***	-0.219***	-0.205***
	(0.025)	(0.026)	(0.022)	(0.023)	(0.022)
Skill			0.887***		
			(0.050)		
Life-cycle skill				1.028***	1.032***
				(0.048)	(0.050)
Govt worker X year					0.015***
					(0.002)
Male X year					0.001
					(0.001)
Fixed-effects					
Year		Yes	Yes	Yes	Yes
Occupation		Yes	Yes	Yes	Yes
Sector		Yes	Yes	Yes	Yes
Fit statistics					
Observations	33,583	33,362	33,362	33,362	33,362
R^2	0.11	0.21	0.27	0.30	0.30
Within R^2		0.07	0.14	0.17	0.17

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample restricted to government workers and control group (private sector workers who have worked in social services at any point in time). Standard errors are clustered at the year level (except for column 1 which assumes IID standard errors). Adding occupation and sector fixed effects reduces the sample size slightly due to missings in the occupation variable.

one can allow for multiple dimensions of skills by incorporating multiple factors, which are usually restricted to be orthogonal (see Ahn, Lee, and Schmidt 2013; Bai 2009). In Appendix A.5, I provide a more detailed practical guide of how to estimate multiple noisy fixed effects in this setup and discuss Monte Carlo evidence that shows its performance in realistic empirical settings.

3.5.2 Further heterogeneity in skill prices or jobs

An important limitation of the approach in this paper is that skill estimation requires the assumption that there is a single skill price in the skill estimation sample for a comparable job with the same skill requirements. While subsequent analyses conditioned on narrower within-job comparisons, the skill estimation itself relies on comparing across these jobs to value skills. If one is worried about this assumption, an alternative approach is to specify a model that maps from noisy individual fixed effects to skill-related observables, while controlling for further group fixed effects. These groups could be jobs or also geographic areas to capture different skill prices across locations. The same fixed effects should then be used in any downstream empirical analysis to ensure that results are only based on variation within fixed effects. Another alternative to allow for more price variation is to directly compute group-specific skill prices following a group-specific flat-spot identification. Again, this gives separate normalizations across groups, so that levels in skills cannot be compared across groups anymore, requiring group-specific fixed effects in later analyses.

3.6 Conclusion

This paper provided a new approach to estimate government worker skills using residualized wages in comparable jobs in the private sector, relating these to skill-related observables using Machine Learning tools and then predicting government worker skills out-of-sample. I then showed two main applications drawing on rich Indonesian household-level panel data. First, I showed evidence for the selection of government workers. Despite growing absolute skills, relative skills of government workers systematically declined in Indonesia over the past 30 years. I linked this finding to the difficulty of the Indonesian government to attract the workers with the highest skills to the government. Furthermore, I showed evidence for the detrimental effect of government hiring cycles on the selection of government workers. The evidence is consistent with the idea that in years of outsized hiring, the government needs to move down the skill distribution of the applicant pool to fill all government positions, leading to lower average skills. In the second main application, I looked at government wage setting and showed that the Indonesian government pays a wage premium of at least 30% conditional on skills, about 1/3 of which is driven by the large gender wage gap in Indonesia's private sector.

A good sign of a new estimation approach is that it raises many interesting questions – both conceptual and theoretically – that can now be studied more rigorously: For example, what are the output or welfare costs of government hiring cycles? Or what drives the relative decline in government skills and does this go in hand with a relative decline in state capacity versus private sector capacity over the course of development? All of these questions are particularly well-suited for future structural work on the functionings of bureaucracies, for which the estimated government skills in this paper can directly be used as state variables. This allows for a tractable framework to consider selection into different jobs, investments in endogenous skills and the counterfactual responses of government output. This structural work would allow to consider novel counterfactuals of practical interest that are otherwise hard to evaluate. For example, how would changes in the selection rule of government workers via changes in the application process affect private sector labor markets and investments in skills and formal education? Or eventually, how would changes in government hiring practices affect private sector versus public sector production and how should governments select workers over the course of development?

3.7 References

Ahn, Seung C., Young H. Lee, and Peter Schmidt. 2013. "Panel Data Models with Multiple Time-Varying Individual Effects." *Journal of Econometrics* 174 (1): 1–14.

Anandari, I. GAA, and Chaikal Nuryakin. 2019. "THE EFFECT OF RISK PREFERENCE ON CHOICE BETWEEN PUBLIC AND PRIVATE SECTOR EMPLOYMENT IN INDONESIA." *International Journal of Business & Society* 20.

Ashraf, Nava, Oriana Bandiera, and Scott S. Lee. 2016. "Do-Gooders and Gogetters: Selection and Performance in Public Service Delivery." *Manuscript, London School of Economics/Harvard University. URL Https://Bit. Ly/2rQgLSj*.

Attanasio, Orazio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina. 2020. "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia." *American Economic Review* 110 (1): 48–85.

Bai, Jushan. 2009. "Panel Data Models with Interactive Fixed Effects." *Econometrica* 77 (4): 1229–79.

Banerjee, Abhijit, Lakshmi Iyer, and Rohini Somanathan. 2007. "Public Action for Public Goods." *Handbook of Development Economics* 4: 3117–54.

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden." *American Economic Review* 107 (8): 2204–42.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi. 2017. "Individuals and Organizations as Sources of State Effectiveness, and Consequences for Policy." National Bureau of Economic Research.

Bhavnani, Rikhil R., and Alexander Lee. 2019. "Does Affirmative Action Worsen Bureaucratic Performance? Evidence from the Indian Administrative Service." *American Journal of Political Science*.

Biasi, Barbara. 2021. "The Labor Market for Teachers Under Different Pay Schemes." *American Economic Journal: Economic Policy* 13 (3): 63–102.

Blunt, Peter, Mark Turner, and Henrik Lindroth. 2012. "Patronage's Progress in Post-Soeharto Indonesia." *Public Administration and Development* 32 (1): 64–81.

Bowlus, Audra J., and Chris Robinson. 2012. "Human Capital Prices, Productivity, and Growth." *American Economic Review* 102 (7): 3483–3515.

Brinkerhoff, Derick W., and Anna Wetterberg. 2013. "Performance-Based Public Management Reforms: Experience and Emerging Lessons from Service Delivery Improvement in Indonesia." *International Review of Administrative Sciences* 79 (3): 433–57.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20 (1): 91–116.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the Impacts of Teachers i: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

Colonnelli, Emanuele, and Mounu Prem. 2017. "Patronage in the Allocation of Public Sector Jobs." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2942495.

Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso. 2020. "Patronage and Selection in Public Sector Organizations." *American Economic Review* 110 (10): 3071–99.

Dal B'o, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. 2017.

"Who Becomes a Politician?" *The Quarterly Journal of Economics* 132 (4): 1877–1914.

Dal B'o, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *The Quarterly Journal of Economics* 128 (3): 1169–1218.

Decarolis, Francesco, Leonardo M. Giuffrida, Elisabetta Iossa, Vincenzo Mollisi, and Giancarlo Spagnolo. 2018. "Bureaucratic Competence and Procurement Outcomes." National Bureau of Economic Research.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91 (4): 795–813.

———. 2004. "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia." *Journal of Development Economics* 74 (1): 163–97.

Estrada, Ricardo. 2019. "Rules Versus Discretion in Public Service: Teacher Hiring in Mexico." *Journal of Labor Economics* 37 (2): 545–79.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande. 2017. "The Personnel Economics of the Developing State." In *Handbook of Economic Field Experiments*, edited by Esther Duflo and Abhijit V. Banerjee, 2:467–514. Elsevier.

Fisman, Raymond. 2001. "Estimating the Value of Political Connections." *American Economic Review* 91 (4): 1095–1102.

Guvenen, Fatih. 2009. "An Empirical Investigation of Labor Income Processes." *Review of Economic Dynamics* 12 (1): 58–79. https://doi.org/10.1016/j.red.2008.06.004.

Guvenen, Fatih, and Anthony A. Smith. 2014. "Inferring Labor Income Risk and Partial Insurance from Economic Choices." *Econometrica* 82 (6): 2085–2129.

Hadiz, Vedi R., and Richard Robison. 2013. "The Political Economy of Oligarchy and the Reorganization of Power in Indonesia." *Indonesia*, no. 96: 35–57.

Hanna, Rema, and Shing-Yi Wang. 2017. "Dishonesty and Selection into Public Service: Evidence from India." *American Economic Journal: Economic Policy* 9 (3): 262–90.

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

Hido, Shohei, Yuta Tsuboi, Hisashi Kashima, Masashi Sugiyama, and Takafumi Kanamori. 2011. "Statistical Outlier Detection Using Direct Density Ratio Estimation." *Knowledge and Information Systems* 26 (2): 309–36.

Horhoruw, Maggy, Tina George Karippacheril, Wahyu Sutiyono, and Theo Thomas. 2013. "Transforming the Public Sector in Indonesia: Delivering Total Reformasi." In *Paper Presented IRSPM Conference,'Contradictions in Public Management Managing in Volatile Times', University of Rome, Tor Vergata*. Vol. 16.

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. 1989. "Multilayer Feedforward Networks Are Universal Approximators." *Neural Networks* 2 (5): 359–66.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of Lifetime Inequality." *American Economic Review* 101 (7): 2923–54.

Jia, Ruixue, Masayuki Kudamatsu, and David Seim. 2015. "POLITICAL SELECTION IN CHINA: THE COMPLEMENTARY ROLES OF CONNECTIONS AND PERFORMANCE: Political Selection in China." *Journal of the European Economic Association* 13 (4): 631–68. https://doi.org/10.1111/jeea.12124.

Kahn, Lisa B., and Fabian Lange. 2014. "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures." *The Review of Economic Studies* 81 (4): 1575–1613.

Keane, Michael P., and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3): 473–522.

Kristiansen, Stein, and Muhid Ramli. 2006. "Buying an Income: The Market for Civil Service Positions in Indonesia." *Contemporary Southeast Asia*, 207–33.

Kuipers, Nicholas. 2021. "Failing the Test." *Job Market Paper*, UC Berkeley, Political Science,.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth Across Countries." *Journal of Political Economy* 126 (2): 797–849.

Lagakos, David, and Martin Shu. 2023. "The Role of Micro Data in Understanding Structural Transformation." *Oxford Development Studies* 51 (4): 436–54.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2019. "Imperfect Competition, Compensating Differentials and Rent Sharing in the US Labor Market." National Bureau of Economic Research.

Magnac, Thierry, Nicolas Pistolesi, and S'ebastien Roux. 2018. "Post-Schooling Human Capital Investments and the Life Cycle of Earnings." *Journal of Political Economy* 126 (3): 1219–49.

Martinez-Bravo, Monica, Priya Mukherjee, and Andreas Stegmann. 2017. "The Non-Democratic Roots of Elite Capture: Evidence From Soeharto Mayors in Indonesia." *Econometrica* 85 (6): 1991–2010.

McLeod, Ross H. 2006. "Private Sector Lessons for Public Sector Reform in Indonesia." *Agenda: A Journal of Policy Analysis and Reform*, 275–88.

———. 2008. "Inadequate Budgets and Salaries as Instruments for Institutionalizing Public Sector Corruption in Indonesia." *South East Asia Research* 16 (2): 199–223.

Meghir, Costas, and Luigi Pistaferri. 2011. "Earnings, Consumption and Life Cycle Choices." In *Handbook of Labor Economics*, 4:773–854. Elsevier.

Pesaran, M. Hashem. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74 (4): 967–1012.

Pierskalla, Jan H., Adam Lauretig, Andrew Rosenberg, and Audrey Sacks. 2020. "Democratization and Representative Bureaucracy - An Analysis of Promotion Patterns in Indonesia's Civil Service, 1980-2015." *American Journal of Political Science* forthcoming.

Pierskalla, Jan H., and Audrey Sacks. 2018. "Personnel Politics: Elections, Clientelistic Competition and Teacher Hiring in Indonesia." *British Journal of Political Science*, 1–23.

Polachek, Solomon W. 2008. *Earnings over the Lifecycle: The Mincer Earnings Function and Its Applications*. Vol. 16. Now Publishers Inc.

Primiceri, Giorgio E., and Thijs Van Rens. 2009. "Heterogeneous Life-Cycle Profiles, Income Risk and Consumption Inequality." *Journal of Monetary Economics* 56 (1): 20–39.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *The Economic Journal* 128 (608): 413–46. https://doi.org/10.1111/ecoj.12418.

Robison, Richard, and Vedi Hadiz. 2004. *Reorganising Power in Indonesia: The Politics of Oligarchy in an Age of Markets*. Routledge.

Sanders, Carl, and Christopher Taber. 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Annual Review of Economics*, no. 4.

Strauss, John, Firman Witoelar, and Bondan Sikoki. 2016. *The Fifth Wave of the Indonesia Family Life Survey: Overview and Field Report*. RAND.

Taber, Christopher, and Rune Vejlin. 2020. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." *Econometrica* 88 (3): 1031–69.

Valsecchi, Michele. 2016. "Corrupt Bureaucrats: The Response of Non-Elected Officials to Electoral Accountability." *Job Market Paper*.

Weaver, Jeff. 2019. "Jobs for Sale: Corruption and Misallocation in Hiring." *Working Paper*.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.

3.8 Appendix

3.8.1 Further details on the selection and hiring process of government workers

Entering the Civil Service is remarkably well-defined in Indonesia despite large changes to the Civil Service over the time period of interest as applicants run through a centralized application process. Applicants apply to the specific position or district they are interested in, but still run through a centralized application process and due to frequent rotations and across-country stationing, are likely to end up with a position somewhere else than where they applied if they are admitted. Formal requirements of applying to the Civil Service are that individuals have to be between 18-35 years old, never been imprisoned, not be a member of a political party, be in good physical and mental health and be willing to work in any region in Indonesia. For each job opening there are then additional educational requirements, which are set by the district and which often mean an undergraduate diploma. Since 2012, all applicants have to go through a civil servant enrollment test (*CPNS*), which includes three parts: an administrative selection, a basic competence test (*Seleksi Kompetensi Dasar*) recently administered via a computer-assisted test and a specific field competence selection (*Seleksi Kompetensi Bidang*).[14]

Solely based on aggregate numbers, obtaining a Civil Service job is difficult. In 2014, prior to a 4-year public sector employment moratorium, there were more than 2.6 million applicants for 100,000 available positions, which translates into an acceptance rate of slightly below 4% (see: Anandari and Nuryakin 2019). This is similar to the 1-5% acceptance rates reported in Kristiansen and Ramli (2006) for two Indonesian regions in the early 2000s.[15]

In practice, it is unclear how well the recruitment system in place selects qualified candidates and how this changed over time. Horhoruw et al. (2013) notes that it is unclear whether reform processes since 2001 have actually led to an improvement of hiring practices beyond just a few reform-minded institutions. In Pierskalla and Sacks (2018), the authors draw on teacher censuses to show that changes in the political system after 1998 actually had negative effects on public hiring. They find that increased political competition gave local elites an incentive to use their discretionary control over state hiring to increase patronage efforts as evidenced by election-related increases in the number of contract teachers on local payrolls and increases in civil service teacher certifications. At the same time, Pierskalla et al. (2020) use data on the universe of civil servants to show that civil servants with a postgraduate education are twice as likely to be promoted after 1999 in comparison to before, indicating a combination of composition changes and more performance-related promotion patterns.

Kristiansen and Ramli (2006) draw on in-depth qualitative and quantitative evidence from interviews and focus groups with a non-representative sample of 60 civil servants in two areas of Indonesia to document that personal ties and nepotism are often named as primary reasons for hiring. Moreover, the selling of government jobs is widely practiced. Kristiansen and Ramli (2006) document that all respondents paid for their first Civil Service position and that the average reported price for these jobs is around 2.5 times the official annual initial salary of-

[14] Under the Suharto regime, the Civil Service system was organized as a military-type organisation where new recruits were not differentiated other than by level of education, which introduced generalist civil servants and abolished further specializations within the bureaucracy (McLeod 2006). Recent reforms have tried to reverse this.

[15] Note, the percentage of acceptances has increased in the author's sample, which would be in line with the model on the evolution of state capacity.

fered. This is slightly higher than the 17 months of salary reported recently in Weaver (2019).[16] There is also some evidence that the average real price for a government position has slowly increased between 1995-2004. Prices are positively correlated with the salary of the job (which in turn is mechanically tied to the education level of the civil servant) and seem to be positively correlated with the ease of rent-seeking possibilities in the specific job offered.[17] This evidence on prices is in line with a competitive auction price for government sector jobs as found in Weaver (2019). In the end, it is unclear how these unlawful hiring practices perform with respect to selecting the most qualified candidates as this depends on the correlation between quality and the ability to pay for a job or the probability of knowing someone important in the bureaucracy. Interestingly, Weaver (2019) finds that for the context they look at, this correlation is highly positive so that the selling of government sector jobs actually leads to a good selection rule in terms of quality of the new hires.

[16] The author does not share the country of study to provide additional security for their survey respondents.

[17] Among the usual rent-seeking possibilities are various forms of contract kickbacks, payment from staff in exchange for positions and hiring on projects, loan accounts structured to earn interest by the agency, provision of ghost services, inflated invoicing in collusion with contractors, procedures for tax avoidance, irregular payments for health and education services, bribes to police officers and judges, and speed money to obtain formal papers and permits (World Bank 2003; Vian 2005; Chapman 2005; Azfar 2005).

3.8.2 Proof of Proposition 4

I start with the **flat-spot identification**. Unbiasedness is given by:

$$\mathbb{E}_{i\in\text{FP}}[w_{i,e,t} - w_{i,e-1,t-1}] = p_t - p_{t-1} + \underbrace{\mathbb{E}_{i\in\text{FP}}[(h_{i,e,t} - h_{i,e-1,t-1})]}_{=0 \text{ due to flat spot assumption}} + \underbrace{\mathbb{E}_{i\in\text{FP}}[\epsilon_{i,e,t} - \epsilon_{i,e-1,t-1}]}_{=0} \quad (3.7)$$

Consistency of the estimator is given by:

$$\frac{1}{N_{FP}} \sum_{i\in\text{FP}} [w_{i,e,t} - w_{i,e-1,t-1}] \xrightarrow{N\to\infty} p_t - p_{t-1} + \underbrace{\mathbb{E}_{i\in\text{FP}}[(h_{i,e,t} - h_{i,e-1,t-1})]}_{=0 \text{ due to flat spot assumption}} + \underbrace{\mathbb{E}_{i\in\text{FP}}[\epsilon_{i,e,t} - \epsilon_{i,e-1,t-1}]}_{=0} \quad (3.8)$$

The estimator combines estimated year-to-year changes in prices, giving an unbiased and consistent estimate of the entire price path (up to a level of normalization).

Given unbiased and consistent estimates of the skill price process, one can construct an unbiased estimate of individual skills using a standard-within estimator for:

$$\widetilde{w_{i,e,t}} - \overline{\widetilde{w_{i,\cdot}}} = \delta_0 * \left(exp_{i,e,t} - \overline{exp_{i,\cdot}}\right) + \delta_1\left(exp^2_{i,e,t} - \overline{exp^2_{i,\cdot}}\right) + \left(\epsilon_{i,e,t} - \overline{\epsilon_{i,\cdot}}\right) \quad (3.9)$$

However, as is known as the incidental parameters problem, z_i cannot be consistently estimated from the above as long as $T \not\to \infty$ (e.g. see: Wooldridge 2010, Chp. 10).

Table 3.6: Regression results: Experience profile

Dependent Variable:	log wage (deflated by skill price)	
Model:	(1)	(2)
Variables		
Constant	8.193***	
	(0.015)	
experience	0.029***	0.031***
	(0.002)	(0.004)
experience square	-0.001***	-0.001***
	(0.000)	(0.000)
Fixed-effects		
Individual		Yes
Fit statistics		
Observations	20,538	20,538
R^2	0.01	0.65
Within R^2		0.01

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample are comparison workers (private sector workers in social services).

3.8.3 Additional results for skill estimation

Table 3.7: Relative variable importance

Variable	Importance
PrimaryEduc	100.000000
HigherEduc	91.345271
SeniorSecondEduc	55.312742
writeLetterIndo	46.846035
speakIndonesian	36.889641
RelativeIndonScore	20.298544
CountWordRecall	17.260944
RelativeTotalScore	15.285114
CountBackwards	14.368554
WordAbil	10.753075
Big5Open	10.375476
RavenIQ	10.116039
MathIQ	10.092988
Big5Neu	8.222708
JuniorSecondEduc	7.195992
RiskPref1	5.723784
RelativeMathScore	3.346342
CountWordRecallDelayed	3.183184
RiskPref2	1.926393
Big5Ext	1.465696

Details: Based on best-performing GBM algorithm and restricting to top 20 variables. Variable importance is based on traversing the tree and recording how much the metric (R2 here) changes every time a given variable is used for splitting. One then takes the average reductions across all base-learners for each variable and normalizes the most important variable to 100

Table 3.8: Regression results: Informativeness of government wages

Dependent Variable:	log(real hourly wage)					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	0.478**	-1.897***	-1.427***	-0.593***	-1.555***	-1.519***
	(0.211)	(0.204)	(0.204)	(0.146)	(0.149)	(0.151)
Skill	1.030***		1.166***	1.116***		1.194***
	(0.025)		(0.024)	(0.018)		(0.018)
Life-cycle skill		1.267***			1.201***	
		(0.023)			(0.018)	
experience			0.068***			0.041***
			(0.003)			(0.002)
experience square			-0.001***			-0.001***
			(0.000)			(0.000)
Fit statistics						
Observations	13,130	13,130	13,130	20,453	20,453	20,453
R^2	0.11	0.18	0.22	0.16	0.18	0.18
Adjusted R^2	0.11	0.18	0.22	0.16	0.18	0.18

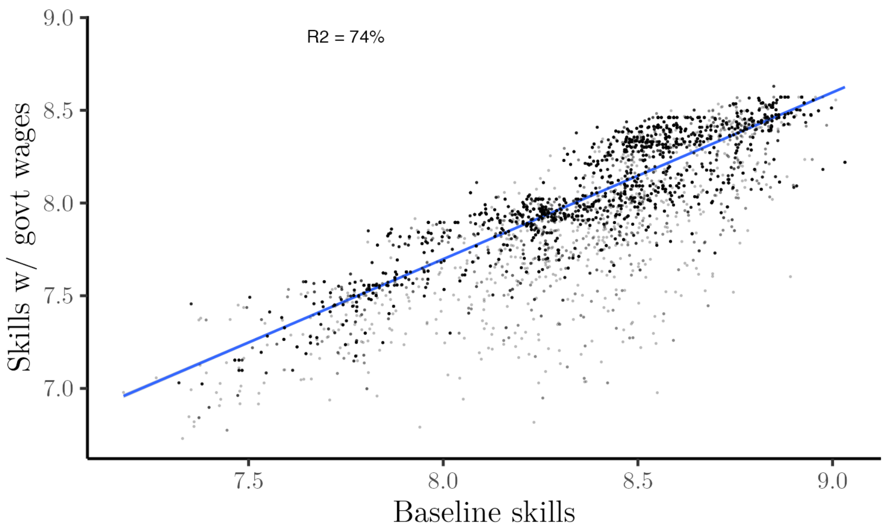
IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Columns 1-3 focus on public worker wages only, while Columns 4-6 focus on comparison workers (private sector workers in social services).

3.8.4 Additional empirical results

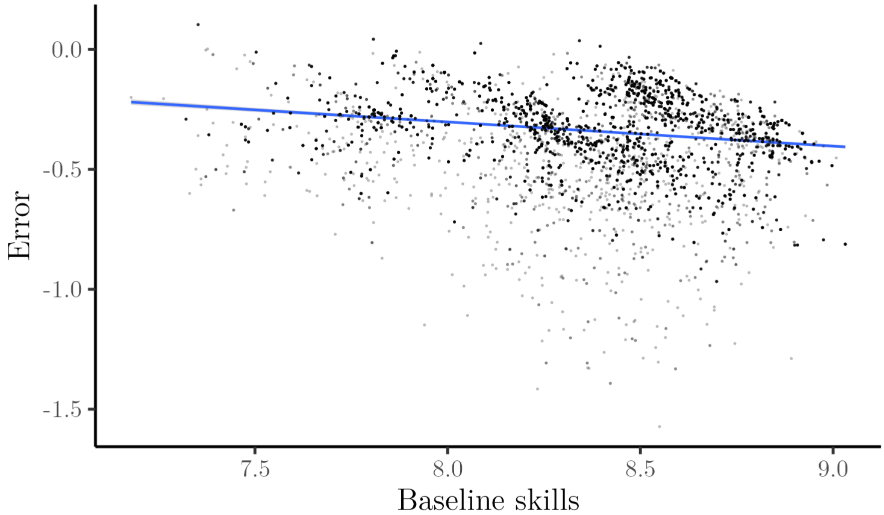
This part of the Appendix provides additional results for Section 4 that are in part referenced in the main text.

Figure 3.10: Correlation of baseline skill estimates and skill estimates based on government wages



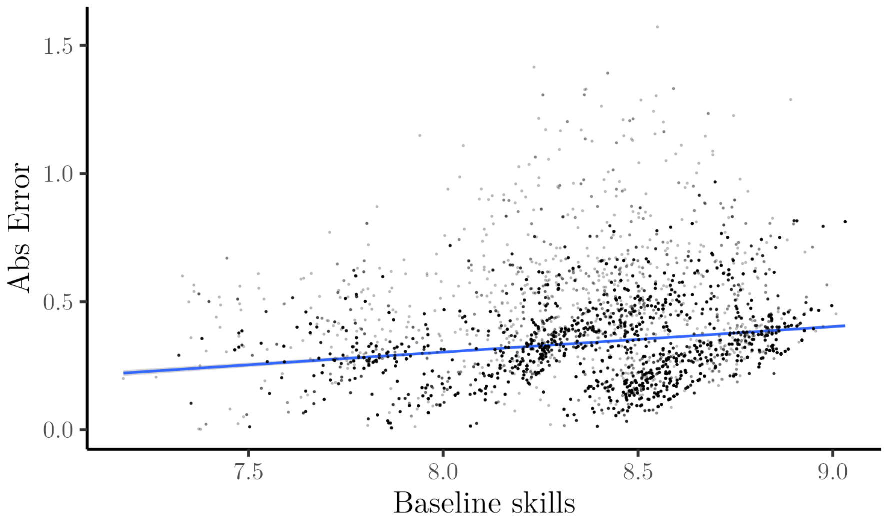
Notes: Baseline skill estimates are based on private sector social service workers as estimation sample and GBM estimator as ML algorithm (as explained in text). Skill estimates based on government wages go through the same estimation steps using all government workers for estimation sample instead (that is, the same skill price estimator and the GBM estimator).

Figure 3.11: Error of skill estimate based on government wages (compared to baseline skill estimate)



Notes: Error gives the difference between skills based on government wages and the baseline skills (in log differences, assuming multiplicative error). Baseline skill estimates are based on private sector social service workers as estimation sample and GBM estimator as ML algorithm (as explained in text). Skill estimates based on government wages go through the same estimation steps using all government workers for estimation sample instead (that is, the same skill price estimator and the GBM estimator).

Figure 3.12: Absolute error of skill estimate based on government wages (compared to baseline skill estimate)



Notes: Absolute error gives the absolute difference between skills based on government wages and the baseline skills (in log differences, assuming multiplicative error). Baseline skill estimates are based on private sector social service workers as estimation sample and GBM estimator as ML algorithm (as explained in text). Skill estimates based on government wages go through the same estimation steps using all government workers for estimation sample instead (that is, the same skill price estimator and the GBM estimator).

3.8.5 Estimation details: factor structure for experience profile

The factor structure introduced in Section 5.1 can be estimated using the two-stage estimator of Pesaran (2006).[18] Pesaran (2006) proposes to use cross-sectional averages to estimate experience factors δ_e. For this, one can first demean the series to get:

$$\tilde{\tilde{w}}_{i,e} \equiv \tilde{w}_{i,e} - \tilde{w}_{i,\bullet} = g(z_i)(\delta_e - \delta_\bullet) + (\epsilon_{i,e} - \epsilon_{i,\bullet}) \tag{3.10}$$

Using the economic structure of the problem gives $\delta_0 = 0$, so that identification of $(\delta_e - \delta_\bullet)$ also separately identifies the two terms. Using the fact that in the assumed data-generating process: $lim_{n\rightarrow\infty}\tilde{\tilde{w}}_{\bullet,e} = \overline{g(z)}(\delta_e - \delta_\bullet)$, we can write:

$$\frac{\tilde{\tilde{w}}_{\bullet,e}}{\tilde{\tilde{w}}_{\bullet,0}} \rightarrow 1 - \frac{\delta_e}{\delta_\bullet} \tag{3.11}$$

$\tilde{\tilde{w}}_{\bullet,e}$ gives as many equations as there are experience levels. However, with the previous restriction of $\delta_0 = 0$, we lose one restriction which requires to directly use: $\tilde{\tilde{w}}_{\bullet,e} \approx \overline{g(z)}(\delta_e - \delta_\bullet)$. Given that we are not directly interested in the estimates of δ nor $g(z_i)$, we can instead choose any non-zero normalization to obtain the same estimates of individual skills $log(z_i)$. For any normalization of $\overline{g(z)}$, we obtain an estimate of δ. We can then include these in the following experience-series regression for each individual to estimate $log(z_i)$ and $g(z_i)$:

$$\tilde{w}_{i,e} = log(z_i) + \widehat{\delta}_e g(z_i) + \epsilon_{i,e} \tag{3.12}$$

The individual-level skill estimates are the constant of the regression, giving the estimate $\widehat{log(z_i)}$. The derived estimate of private sector skills is generically inconsistent in a panel with fixed T due to the incidental parameters problem. This is generally true of any estimator that gives individual-specific estimates. For example, a standard Within-estimator also gives inconsistent estimates of the individual levels as long as the time dimension is not tending to infinity. Intuitively, the setup only allows us to extract a noisy signal from the even noisier wage data.

[18]Using Monte Carlo simulations, I checked that the Pesaran (2006) estimator performs much better in estimating the factor structure than alternative estimators such as a Within estimator and two different Concentrated Maximum Likelihood estimators. This performance even holds with unbalanced and potentially non-stationary panel data as observed in real-world applications. I am happy to share these results upon request. Based on the existing econometric literature, this is a novel result, because I am not aware of any studies that have looked at the performance of factor model estimators for the individual-level estimates itself (in contrast to treatment effect parameters that are estimated in the presence of individual-level effects).