The London School of Economics and Political Science

Essays on Misallocation of Production Inputs and Energy Efficiency in Indian Industries

by

Gregor Singer

A thesis submitted to the London School of Economics for the degree of Doctor of Philosophy.

London, May 2019

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others.

I declare that my thesis consists of approximately 65000 words.

Abstract

This thesis explores the causes and consequences of resource misallocation and energy efficiency in the context of firm heterogeneity in Indian manufacturing. In the first chapter, I quantify misallocation in the Indian cast iron sector. I develop a method to disentangle distortions that capture input misallocation from fundamental demand heterogeneity across firms. In the second chapter I explore why material inputs are misallocated in the cast iron sector, focusing on differences in access to suppliers through the transportation network. In the third chapter I estimate the causal relationship between industrial electricity prices and electricity productivity for a large panel of Indian plants.

Acknowledgements

I am deeply grateful for all the support and wonderful people I met on the journey that led to this thesis. First and foremost, I am indebted to my supervisors Antony Millner and Ben Groom. I owe them for their helpful guidance, constructive feedback and invaluable encouragement throughout the years. I am grateful for feedback and numerous conversations with academics at the London School of Economics and elsewhere that helped to improve this thesis. I also want to thank my great colleagues with whom I could share this journey and became friends with along the way.

I thank the Grantham Research Institute at the London School of Economics for hosting me throughout these years in its stimulating and collegial environment. I also gratefully acknowledge financial support through the PhD scholarship from the London School of Economics.

Finally, I dedicate this thesis to my parents Helga-Maria and Walter, who I feel tremendously lucky to have. I am indebted to my family and friends for their constant support, and want to thank Daphne especially for her understanding during the final year. I could not have done this without you.

Contents

List of Figures

List of Tables

Preface

This thesis explores the causes and consequences of resource misallocation and energy efficiency in the context of firm heterogeneity in Indian manufacturing. It has been widely documented that firms are heterogeneous within sectors. With firm heterogeneity, the allocation of resources between firms matters for aggregate outcomes. How efficient is the allocation of production inputs across plants? In the first chapter, I quantify misallocation in the Indian cast iron sector. I develop a method to disentangle fundamental heterogeneity between firms in terms of demand or production from input distortions that capture misallocation. In the second chapter I ask why material inputs are misallocated focusing on differences in access to suppliers through the transportation network. In the third chapter I use the heterogeneity between firms in terms of electricity prices and electricity productivity to estimate their causal relationship.

Estimating the costs of misallocation is challenging, despite its growing popularity to explain aggregate income differences around the world. We need models to define efficient allocations and to discipline data to quantify the costs of distortions that represent misallocation. When we ignore firm heterogeneity in technology and demand in these models, we infer misallocation costs that are upwards or downwards biased. The first chapter develops and estimates a structural model that addresses this problem by disentangling input misallocation distortions from fundamental heterogeneity. The model accounts for endogenous variable markups and I explore counterfactual comparisons of oligopolistic equilibria with and without misallocation. I can distinguish between welfare losses on the consumer and the producer side, as well as losses in aggregate productivities.

Using data on quantities and prices on the output and inputs side of Indian cast iron plants, I quantify the costs of input misallocation. Total welfare losses are large and equivalent to 31% of sales. The consumers bear a higher share of the misallocation incidence than the producers. In contrast, aggregate input productivities are hardly affected by misallocation. I can realign this result with the large impacts on aggregate productivity typically found in the literature. This means that while there are welfare losses, more efficient allocation in this sector would not contribute to improving sectoral material efficiency, which is receiving growing interest from an environmental policy perspective. Perhaps surprisingly, I find that welfare losses due to misallocation of input materials are 90% larger than those from misallocation of labour.

In the second chapter, I ask what the underlying drivers of these costly input material distortions are. The steel industry relies on shipping of heavy and bulky material inputs, particularly through railroads. Indian freight transportation is plagued by breakdowns due to outdated infrastructure, congestion from sharing tracks with passenger trains or state border checkpoints due to different tax systems. These issues generate indirect costs of trade such as uncertainty and delays. I account for shipping fees with observed factory gate input prices, but the misallocation distortions would capture differences in these indirect trade costs across plants. The indirect costs are likely to go up with longer routes from suppliers. Using detailed geo-located data on the Indian rail and road infrastructure, I construct a measure of supplier access and test whether differences in supplier access are systematically related to material input distortions. I find that a one standard deviation increase in supplier access reduces the costly input distortion by almost a third of its standard deviation. Placebo tests show supplier access is not significantly related to the labour input distortions, and that access to irrelevant suppliers or access to markets on the output side cannot explain the distortions on the input side.

The third chapter starts by documenting a secular increase in aggregate industrial energy efficiency and aggregate electricity productivity in India from around 2000. During the same period, industrial electricity prices almost halved. Using a large panel of Indian manufacturing plants over 16 years with information on electricity quantity and prices at the plant level, I estimate the impact of electricity prices on electricity productivity. Based on two different instruments, I recover causal estimates at the micro level that can explain these aggregate trends. While higher electricity prices reduce electricity consumption, they disproportionally decrease output, and therefore reduce electricity productivity. The causal estimates have the opposite sign of the OLS estimates. I explore mechanisms and find that higher electricity prices reduce firm size, investment, productivity and markups. This is consistent with a complementarity between electricity and modern high performance production techniques. I estimate pass-through rates and calculate that the incidence share on consumers of this large industrial electricity prices reduction was two thirds. The causal effects of industrial coal prices are of opposite sign, which has important implications for climate policy and industrial development.

Chapter I

Endogenous markups and input misallocation

I.1 Introduction

Misallocation of production factors between firms can result in large losses of aggregate income. In their seminal paper, Hsieh and Klenow (2009) estimate aggregate TFP losses of 40%-60% in Indian manufacturing. Distortions such as preferential access to credit, labour regulations that depend on firm size, or political connections have been identified, amongst others, as causes for input misallocation in this emerging literature.[1]

In this chapter, I address a problem of ignored heterogeneity across firms when calculating misallocation losses. In the growing misallocation literature following Hsieh and Klenow (2009), ignored heterogeneity across plants in terms of production or demand is conflated with misallocation distortions in a non-trivial way. As a result, inferred welfare costs from input misallocation can be upwards *or* downwards biased. In the application of this chapter, ignoring heterogeneity in markups would understate the true costs of input misallocation by up to 27%. The main contribution of this chapter is that I disentangle plant level demand and production heterogeneity from input distortions.

I focus on input misallocation between cast iron producers in India.[2] India's manufacturing sector is an interesting case for studying misallocation considering debates[3] in the literature about the contribution of reforms to resource allocation. Misallocation could have also played a role in India's slow structural transformation compared to China and other East Asian nations despite its deep economic reforms in the early 90s (Bhagwati and Panagariya, 2014).

[1]See e.g. Gopinath et al. (2017) or Midrigan and Xu (2014), Garicano et al. (2016), and Akcigit et al. (2018) respectively. See Restuccia and Rogerson (2017) or Hopenhayn (2014*a*) for recent surveys. Appendix I.A presents a brief overview of additional related literature.

[2]Cast iron is an important product in India's manufacturing sector. It has one of the highest shares of any single product in manufacturing output. India's steel sector has a 15% share in total manufacturing value added which is one of the highest in the world (UNIDO, 2016*b*).

[3]See e.g. Bollard et al. (2013); Harrison et al. (2013) vs. Nishida et al. (2014, 2015).

While the aggregate consequences of misallocation are usually expressed in aggregate TFP losses, we have not been able to say much about other margins. India has created a Resource Efficiency Panel in 2015 and sectoral material productivity is more generally receiving growing attention from an industrial competitiveness and environmental agenda (e.g. OECD, 2015; European Comission, 2013). The substantial emissions of the steel sector are in part determined by its aggregate material productivity. Most research focuses on how innovation and technology diffusion can improve sectoral material productivity, but there is little evidence on potential allocative gains. With the methodology in this chapter, I can recover the effects of misallocation on a rich set of outcomes at any level of aggregation. I can distinguish between the effects of misallocation on welfare and aggregate input productivities. I can further distinguish between the effects on consumer welfare and producer profits, which allows for an analysis of incidence of the input misallocation distortions.

It is worth to briefly conceptualise the misallocation distortions and to motivate why we need to do disentangle those from fundamental heterogeneity in the first place, before I introduce the nature of the counterfactuals. The input distortions (or "wedges") are usually measured by the plant level gaps between the marginal revenue product (MRP) of an input and the input price. Distortions can rationalise the existence of these gaps and represent additional unobserved costs for a particular plant from using that particular input. These additional cost could arise through policies like taxes and subsidies, information frictions, transaction costs, corruption, extortion, or other input constraints that vary across firms. The bundle of all these potential input distortions represent the gap between the MRP and input prices. In the absence of these distortions the MRP should be equal to the input price. In the presence of the distortions there is misallocation. Intuitively, moving a unit of an input from a low gap and low MRP plant to a plant with a larger gap and higher MRP increases aggregate output through a more efficient allocation. In the privately optimal equilibrium, plants that face larger distortions underutilise the input compared to the socially optimal outcome.

The reason why we should disentangle the distortions from fundamental heterogeneity is because the MRP and these gaps are not directly observed.[4] The distortions (or the gaps) are determined by the demand elasticities (or markups), output elasticities and the revenue share of an input.[5] Even if we were correct on average and some distortions

[4]There has been some debate on the role of data management and measurement error. See Rotemberg and White (2017) and Bils et al. (2017) respectively.

[5]To organise thoughts, take the variance of MRP which is often used as a statistic of misallocation

are over- while others underestimated, I show that mismeasurment still impacts welfare conclusions. It matters which plants face which distortions, for example whether it is the more productive plants facing the more severe distortions.[6] Any deviations of the assumed demand or output elasticities from the real ones are captured by the input distortions, while they in fact are differences in demand or production technique.[7] The literature following Hsieh and Klenow (2009), for example, assumes constant production elasticities within 4-digit sectors, and constant demand elasticities dictated by CES. There is, however, a large body of evidence that markups can vary across firms even within narrow industries.[8] Depending on the relationship between the true markups, true output elasticities and the true input distortions, the variation of input distortions across plants can increase *or* decrease when it is mismeasured. As a result, ignoring heterogeneity biases the inferred costs from misallocation upwards or downwards.

One might argue that capturing ignored heterogeneity in the inferred input distortions could be desirable, at least when it comes to heterogeneity in markups. We would capture the bundle of distortions on the input and demand side that represent deviations from a CES framework, such as excess market power. Along the lines of the theory of the second best, we only care about the joint effect of all distortions. If the firm with the additional benefits on inputs (overusing input) also has more market power (under-producing), then these distortions could offset each other in a second best world. However, when we capture ignored demand heterogeneity in input distortions, we preclude any economic evaluation and welfare analysis of the distortions. This is because a model that is based on constant demand elasticities, such as CES, cannot be used to evaluate welfare losses if we believe that we capture variable demand elasticities. We would have to change the primitives of

along the rationale described above, assuming constant input prices. Consider a simple profit maximisation problem $PQ - vX$, where either prices P depend on quantity Q or quantity on prices, and maximising with respect to input X. Without parametric functional form assumptions, we can write the first order condition in terms of the variance of the logged MRP of an input. It is equal to the variance of a combination of the inverse demand elasticity η, the output elasticity α and data on inputs and revenues:

$$\mathrm{Var}\big[\log \mathrm{MRP}\big] = \mathrm{Var}\Big[\log(1+\eta) + \log(\alpha) + \log(\frac{PQ}{X})\Big]$$

Typically, the first two terms are assumed constant (within a sector). If, in reality, demand elasticities (or markups) or output elasticities are not constant, variation in measured MRP no longer imply input distortions, but could capture demand or production heterogeneity. Vice versa, a constant MRP does not imply the absence of input distortions or misallocation. Once we account for heterogeneity, the variance in MRP can go up or down, depending on the correlation between the terms.

[6] Restuccia and Rogerson (2008) show that the correlation between firm TFP and distortion matters.

[7] This is also illustrated by the fact that the equation for input distortions in Hsieh and Klenow (2009) is the same as the equation for markups in the popular De Loecker and Warzynski (2012).

[8] See e.g. Nevo (2001), De Loecker et al. (2016) or Hottman et al. (2016).

the model as well. On the contrary, with the approach in this chapter, where I isolate the input distortions from variation in demand elasticities, we learn about whether the theory of the second best applies. In a second best world, removing only the input distortions for the counterfactual would lead to a *decrease* in welfare, a clear prediction that I can test and will reject.

The way that I disentangle misallocation from fundamental heterogeneity is through a combination of product level focus on the production side and estimating flexible endogenous demand elasticity, facilitated by using detailed quantity and price data on both outputs and inputs for a panel of plants. This is the first paper to estimate structural demand and production systems to combine it into a welfare framework to disentangle input distortions from endogenous markups.[9]

On the production side, I estimate production functions for a single product (cast iron), which is much narrower than the usual production functions that are assumed to be identical for all plants within 4- or 2-digit sectors.[10] The resulting product level output elasticities are much less likely to ignore production heterogeneity. Importantly, I am also able to address output and input price bias by estimating gross output production functions using observed quantities of outputs and inputs.[11] I recover plant level total factor productivities (TFPQ) based on the proxy method (Olley and Pakes, 1996; Wooldridge, 2009).

On the demand and output side, I observe heterogeneity across plants in the data – even for this single cast iron product category. There is significant variation in output prices, both in gross prices and prices net of sales tax, excise duty and other distributional and transport fees. This suggests a setting with differentiated products and different

[9] Some recent work incorporate separate but *exogenous* markup variation, e.g. Ho and Ruzic (2017) for different sectors, or Lenzu and Manaresi (2018), Tortarolo and Zarate (2018), Haltiwanger et al. (2018) and Eslava and Haltiwanger (2019) at the firm level. Haltiwanger et al. (2018) also show that some assumptions of the Hsieh and Klenow (2009) model do not hold using more detailed data from the US and argue that sharper estimates of distortions that are isolated from heterogeneity on the demand and production side indeed hold more informative signals. Related are also Bayer et al. (2018) and Liang (2017). Peters (2013) and Edmond et al. (2018) on the other hand focus on endogenous markups, but are not separating it from input distortions empirically. Baqaee and Farhi (2017) study misallocation in general equilibrium focusing on markups. Also related is De Loecker and Scott (2016), who estimate markups from the production and the demand side to compare them.

[10] I estimate production functions for single product plants in a single 7-digit product category. Boehm and Oberfield (2018) argue that there is significant production heterogeneity even within narrowly defined 4- or 2-digit sectors in India.

[11] When using deflated revenue, we conflate markups with physical output, stressed e.g. by Gandhi et al. (2016); Marin and Voigtländer (Forthcoming); Foster et al. (2008). When using deflated input expenditures instead of observed input quantities, we are likely to conflate quality with quantity, see Katayama et al. (2009); De Loecker (2014) and also Kugler and Verhoogen (2012).

demand conditions. I build on the Berry et al. (1995, 1999) random utility mixed model framework embedded in an oligopolistic setting, where I identify the demand parameters using production cost shifters as instruments. The estimated demand elasticities (and markups) are flexible and endogenous.[12]

Having a clean measure of input distortions is only a first step, as it does not tell us anything about the welfare costs of misallocation, which requires a counterfactual analysis. What is an appropriate counterfactual? This naturally depends on the question being asked. I construct counterfactuals, where I remove input distortions such that there are no gaps between MRP and input prices, and search for an equilibrium where all firms are best responding to each other. The estimated welfare gains can be interpreted as the gains from removing misallocation distortions that would be obtained under an oligopolistic market environment, as opposed to comparing it to a socially planned allocation for example (Behrens et al., 2018). The counterfactuals and the misallocation losses are determined by the change in the estimated distortions and by the endogenous changes in prices and quantities of all plants.

There are five further features of the counterfactual analysis worth highlighting that set this study apart from most of the input misallocation literature. First, demand elasticities, markups and marginal cost pass-through are endogenous, as they depend on the prices of all plants and demand and production fundamentals. The counterfactuals are significant changes to prices and the economy, and restricting markups to exogenous factual levels biases misallocation estimates.[13] Second, *aggregate* inputs are allowed to adjust endogenously in all equilibria. Allocative efficiency gains not only tend to increase total output, but may also affect total input use.[14] Third, I can account for observed input price differences, due to local labour markets, or input quality difference, for example.[15] Fourth, all comparative statics are at the plant level, which permits a rich analysis of outcomes at

[12]The demand elasticities are more flexible than in the Kimball (1995) model, used e.g. in Klenow and Willis (2016) or Edmond et al. (2018), where they are strictly decreasing (i.e. less elastic) in output share.

[13]Markup adjustment turn out to be important. I find markups adjust in the counterfactual for individual plants to a degree that is comparable with the original deviations of plant markups from the average markup.

[14]Recent work by Catherine et al. (2018) on collateral constraints and investment shows that the aggregate input changes are important to such an analysis. Yet, the key literature restricts aggregate inputs to be constant. I account for aggregate input changes and assume that inputs are elastically supplied given that I analyse a small industry.

[15]The size of the misallocation losses are also determined by input prices. Hsieh and Klenow (2009) need to assume constant factor prices across firms for their counterfactual analysis. Conceptually, the input distortions are separate from input prices in the literature, otherwise they would show up in input expenditures. I am able to reduce bias stemming from this source by using observed input prices. See Cheng and Morrow (2018), for example on factor price differences due to local labour markets in China.

any level of aggregation. Fifth, I can recover standard errors of all estimated distortions, welfare losses and other comparative statics, which is novel to the input misallocation literature. This is because I avoid calibration, and instead estimate the model and every parameter directly.[16] In addition, I am able to show which fundamental parameters are driving the uncertainty in estimated misallocation losses. It turns out that the estimated returns to scale are a main driver of the size of the misallocation losses, which underscores the importance of estimating it as well as providing estimates of uncertainty.

There is a trade-off between focusing on a single industry which delivers less biased misallocation costs and analysing misallocation for the entire manufacturing sector. While the focus in the literature has been on the latter, this chapter emphasises the former to maximise the signal in the distortions. Not only could the misallocation costs vary substantially across sectors, but the potential bias when ignoring heterogeneity makes it difficult to draw meaningful conclusions. Ultimately, with increasing availability of quantity (and price) data on inputs and outputs, we can get a better grip on misallocation costs for growing parts of the economy.[17]

The estimated welfare costs from labour and material input misallocation are substantial and equivalent to around 31% of the sales from the plants in the sample.[18] This is also evidence that the economy with variable markups and input distortion does not constitute a second best. The estimated counterfactual gains in compensating variation for consumers are larger than the gains in firm profits, driven by price decreases from cost pass-through. To get a sense of the bias in welfare costs from ignoring demand heterogeneity, I pretend that demand elasticities are constant, infer the wrong distortions, and use my model to calculate the bias. I find that the estimated welfare cost is between 13% to 27% lower (depending on the counterfactual) when ignoring variable markups, so we would underestimate misallocation costs in this case.

Surprisingly, aggregate input productivities are hardly affected from misallocation. Even when defining a standard Cobb Douglas aggregate production function, there are no

[16]By bootstrapping from the estimated parameters' covariance structures I can provide confidence intervals around misallocation distortions or any other outcome.

[17]There can be spillovers into other sectors. Jones (2011, 2013) studies complementarities between sectors through input-output links. Behrens et al. (2018) construct a general equilibrium model and show that distortions in one sector can impact distortions in other sectors.

[18]For the counterfactual analysis I remove misallocation distortions in input materials, labour or both, but any distortions in capital use from the factual are preserved in the counterfactual. This is because a large fraction of static capital distortions might actually be inherent adjustment costs (time-to-build) to changes in the capital stock. As Asker et al. (2014) show capital could be much more optimally allocated in a dynamic sense. David and Venkateswaran (2017) address this by explicitly modelling capital dynamically with adjustment costs.

aggregate TFP gains. This seems at first unexpected because of the large literature on TFP gains. However, I show that we can realign the results with the literature when their TFP gains are not interpreted as pure production side productivity gains, but instead as welfare gains consistent with their implicit underlying demand model. Those TFP (i.e. welfare) results are comparable with the welfare results in this chapter.[19]

The total welfare costs of misallocation of materials are around 90% larger than from misallocation of labour, and the difference is statistically significant. This is a surprising result, given that we usually think of materials as a more flexible input than labour and therefore associated with fewer distortions. While the literature often abstracts from intermediates entirely with value added production functions, distortions in materials markets appear to be important and costly, at least for the cast iron industry in India. In Chapter II of this thesis, I explore what is driving these costly material input distortions.

The rest of the chapter begins by setting up the model and estimation strategy in Section I.2. First I set up the firm problem and how they interact. This allows me to derive an expression for the input distortions which depends on production and demand parameters. I then present the production function estimation, demand estimation and welfare framework before I describe the counterfactual equilibria used for comparative statics. Section I.3 presents the data along with some descriptive statistics. Section I.4 briefly discusses the results for the production and demand estimation as well as descriptive statistics on the estimated input distortions. Section I.5 analyses misallocation losses from the counterfactual exercise before Section I.6 concludes.

I.2 Model and estimation strategy

This section sets up the model in order to identify the misallocation distortions, estimate the structural production and demand side parameters, and calculate the counterfactuals. With a slight abuse of terminology, I use the term firms and plants interchangeably. To fix ideas, we should think of single plant and single product firms, which reflects the data I use.

[19]Furthermore, note that the counterfactual gains in this chapter incorporate the full effects of misallocation including changes in aggregate inputs. When holding aggregate inputs fixed, any output gains would be necessarily attributed to aggregate TFP.

I.2.1 Firm behaviour and input misallocation

Market structure

Suppose that firms interact in a market where each firm $j \in J$ is a single product firm and is selling a differentiated product j.[20] The firms compete strategically on prices in a Bertrand-Nash fashion to maximise profits in each market (period) t separately. In the Bertrand-Nash equilibrium, all firms are individually profit maximising and best responding to each other.

Due to product differentiation, as well as heterogeneous cost structures, firms charge different prices with different markups in equilibrium. Product differentiation is consistent with the data in the application of this chapter, where prices vary across cast iron plants and are only weakly correlated with quantities. This requires differentiation, for example in terms of quality or geography. The elasticity of one firm's demand to another firms' prices varies by firm-pair. Therefore, firms with a unique high quality product can behave like a monopolist in a high quality segment. Similarly, this framework also allows for regional oligopolies where outsider firms' prices have little impact on local demand. These features will be captured by endogenous demand (cross-) elasticities as described in Section I.2.3.

Firm profit maximisation

Firms maximise profits according to:

$$\max_{P_{jt}} P_{jt}Q_{jt}(\mathbf{P_t}) - C(Q_{jt}(\mathbf{P_t}), \mathbf{c}_{jt})$$

where P_{jt} is the output price of firm j at time t, and Q_{jt} is the output quantity which depends on the vector of output prices of all firms $\mathbf{P_t}$.[21] The equilibrium output prices and quantities are not necessarily equal to realised prices and quantities. This is because we introduce an unforeseeable zero mean shock to production later, which pins down realised quantities. The equilibrium strategies of the (risk-neutral) firms do not depend on the shock, which occurs after all choices have been made. The cost function $C(.)$ depends on output quantity and the vector of firm-time specific cost parameters and shifters $\mathbf{c}_{jt}$. The

[20] In the application, we can include multi-product firms that produce this differentiated product for the demand estimation.

[21] Both output prices and input prices are at the factory gate. That is output prices are net of taxes, excise duties and shipping fees. On the input side, shipping fees are captured by the price of material inputs.

first order condition for the firm's profit maximisation problem is:

$$Q_{jt}(\mathbf{P_t}) + \frac{\partial Q_{jt}(\mathbf{P_t})}{\partial P_{jt}}\Big(P_{jt} - MC_{jt}\big(Q_{jt}(\mathbf{P_t}), \mathbf{c}_{jt}\big)\Big) = 0;$$

where I use the standard definition of the marginal costs MC_{jt} which is allowed to change with output quantity. We can rewrite this as:

$$\frac{1}{1+\eta_{jt}(\mathbf{P_t})} = \frac{P_{jt}}{MC_{jt}\big(Q_{jt}(\mathbf{P_t}), \mathbf{c}_{jt}\big)} \tag{I.1}$$

where the inverse demand elasticity is defined as $\eta_{jt} \equiv \frac{\partial P_{jt}}{\partial Q_{jt}(\mathbf{P_t})}\frac{Q_{jt}(\mathbf{P_t})}{P_{jt}}$, which is allowed to be firm and time specific and is endogenous depending on the prices of all firms. Equation (I.1) is the familiar relationship between the price elasticity of demand and the markup of prices over marginal costs used in the Learner index and reflects market power. The degree of market power falls with more elastic demand.

Input cost minimisation and input distortions τ

The input distortions enter in the cost minimisation problem. Before defining the firms problem, I highlight two assumptions. First, I follow recent literature (e.g. De Loecker et al., 2016) and frame the input cost minimisation problem as a short term cost minimisation of achieving the firm's required output Q_{jt} by choosing labour L_{jt} and materials M_{jt} conditional on installed capital K_{jt}. This avoids specifying a dynamic condition for capital optimisation, but it also precludes analysing distortions in capital inputs, as there is no capital demand condition to derive capital distortions from. As Asker et al. (2014) show, using a static condition for capital, a dynamic input, can be misleading when inferring distortions. Hence I only analyse misallocation in material and labour markets and preserve the mix of capital misallocation and adjustment costs contained in the unobserved rental rate across factual and counterfactual scenarios.[22]

Second, inputs are elastically supplied. That is, firms are assumed to be price takers on the input side, consistent with a setting where they are relatively small players in the labour market and material input market. Appendix I.B discusses this assumption. Importantly, I show that input market power (monopsony power) would be captured by the input distortions. I present evidence that input market power is not likely in this

[22]From a meta perspective, I believe we could interpret such adjustment costs as misallocation *if* they are not an inherent feature of production. That is if they are possible to change. Adjustment costs for materials, if they exist, are likely to be small and reducible, as the period of analysis are years.

setting, in favour of the assumption of elastic input supply. The firms minimise short run costs according to:

$$\min_{L_{jt},M_{jt}} r_{jt}K_{jt} + \tau^L_{jt} w_{jt} L_{jt} + \tau^M_{jt} P^M_{jt} M_{jt}$$
$$\text{s.t.} \quad F(K_{jt}, L_{jt}, M_{jt})\Omega_{jt} \geq Q_{jt}$$

where the input prices are the rental rate r_{jt}, wages w_{jt} and materials price P^M_{jt}. Variation in input prices across firms can arise through using inputs of different quality, like more expensive materials or higher skilled workers.[23] The production function $F(.)$ is assumed to have the same structure for all firms, since all firms produce the same 7-digit product, which is substantially narrower than a 4-digit sector. If a firm uses higher quality inputs, it does not produce more outputs, but higher quality outputs (with higher prices), such that the physical relationship of the weight of outputs and inputs is the same for high and low quality products. Differences in this relationship across firms are captured by firm specific Hicks-neutral total factor productivities Ω_{jt}.

Finally, the τ^L_{jt} and τ^M_{jt} are material and labour cost multipliers, that differ across firms and capture input misallocation. I follow the key literature in modelling this as a wedge that captures a range of distortions, as in Chari et al. (2007), Restuccia and Rogerson (2008) or Hsieh and Klenow (2009). Essentially, firms are assumed to behave optimally *given* their individual environment, constraints and distortions, which will enable us to infer the distortions $\boldsymbol{\tau}$. I next discuss the interpretation of these distortions.

Interpretation and identification of input distortions $\boldsymbol{\tau}$

Firms that face a higher τ^M_{jt} have higher additional costs associated with purchasing input materials, which drives wedges into the efficient allocation of inputs. What are these input distortions? The $\boldsymbol{\tau}$ can be interpreted as plant specific "taxes" or "subsidies" on the particular input. They can be actual taxes due to e.g. firm size dependent policies regarding labour taxation[24], input subsidies for a subset of goods, or land and property rights regulation that affects firms differently[25]. They can also be advantages and windfalls

[23]As described in Section I.5.1, since I measure labour L_{jt} in worked hours, I perform a robustness check where labour is measured as wage bill to capture skills, which can increase output quantity Q_{jt} not just sales price. More expensive high quality materials (per tonne) on the other hand should not increase output quantity, but rather the sales price.

[24]See Garicano et al. (2016) for a study of such policies in France.

[25]E.g. Duranton et al. (2015). In the agricultural context see Adamopoulos and Restuccia (2014*a*); Chari et al. (2017); Chen et al. (2017).

through political connections (Faccio, 2006; Akcigit et al., 2018), resulting in firms spending more than optimal amounts on certain inputs, i.e. a low τ. Other elements contained in $\boldsymbol{\tau}$ are differential overhead costs (e.g. legal and administration costs) associated with the input, market or informational frictions (David et al., 2016; Bloom et al., 2013; Allen, 2014), enforcement frictions (Boehm and Oberfield, 2018) and further plant specific barriers or advantages of using the input. Similarly, a model with constrained input access results in the same first order conditions.[26] In Chapter II, I present evidence that the material input distortions capture indirect trade costs such as delays and uncertainty from differences in access to suppliers. Anything that incentivises or constrains the use of the input away from the optimum is captured in $\boldsymbol{\tau}$.

Using the definitions of the material and labour elasticities of output, $\alpha_{jt}^M \equiv \frac{\partial Q_{jt}}{\partial M_{jt}} \frac{M_{jt}}{Q_{jt}}$ and $\alpha_{jt}^L \equiv \frac{\partial Q_{jt}}{\partial L_{jt}} \frac{L_{jt}}{Q_{jt}}$, and that the Lagrange multiplier of the minimisation problem is the marginal cost of production MC_{jt}, the cost-minimising conditions for labour and materials $X \in \{L, M\}$ are:

$$\tau_{jt}^X P_{jt}^X - MC_{jt} \alpha_{jt}^X \frac{Q_{jt}}{X_{jt}} = 0 \tag{I.2}$$

Combining both first order conditions (I.1) and (I.2) yields the expression for τ_{jt}^M and τ_{jt}^L, which are identified if the parameters on the right hand side are identified:

$$\begin{aligned} \tau_{jt}^M &= (\eta_{jt} + 1)\alpha_{jt}^M \frac{P_{jt} Q_{jt}}{P_{jt}^M M_{jt}} \\ \tau_{jt}^L &= (\eta_{jt} + 1)\alpha_{jt}^L \frac{P_{jt} Q_{jt}}{w_{jt} L_{jt}} \end{aligned} \tag{I.3}$$

In standard imperfect competition models where $\tau_{jt}^M = 1$ (and analogously $\tau_{jt}^L = 1$), the markup adjusted output elasticity is equal to the expenditure share of the input, i.e. $(\eta_{jt} + 1)\alpha_{jt}^M = \frac{P_{jt}^M M_{jt}}{P_{jt} Q_{jt}}$, or the marginal revenue product of material is equal to the input price, so $MRPM_{jt} \equiv (\eta_{jt} + 1)\alpha_{jt}^M P_{jt} Q_{jt} / M_{jt} = P_{jt}^M$. We can rationalise measured gaps between the $MRPM_{jt}$ and input price P_{jt}^M with the distortions τ_{jt}^M. The τ_{jt}^M are associated with misallocation. Intuitively, reallocating inputs from a low τ_{jt}^M firm to a high τ_{jt}^M firm increases aggregate sales with the same amount of inputs used, because of

[26] That is suppose the constrained cost minimisation is:

$$\min_{L_{jt}, M_{jt}} r_{jt} K_{jt} + w_{jt} L_{jt} + P_{jt}^M M_{jt} \qquad s.t.\ \ M_{jt} \leq \bar{M}_{jt} \qquad \textit{with multiplier } (\tau_{jt}^M - 1)$$

with the corresponding constraints for labour. See Peters (2013) for example.

the higher adjusted $MRPM_{jt}$ of the latter firm. By turning this logic around, we can infer the wedges τ_{jt}^{M} through the variation in the ratio of the estimated $MRPM_{jt}$ and observed material prices P_{jt}^{M}. Once we have recovered the τ_{jt}^{M}, we can ask how costly they are. This is the point of the counterfactual, where I remove the τ_{jt}^{M} (or τ_{jt}^{L}). Finally, Equation (I.3) demonstrates that any ignored heterogeneity in output elasticities α_{jt}^{M} or markups $1/(\eta_{jt}+1)$ would be shifted to the left hand side and be captured by the input distortions. The misattribution problem has raised doubts in the recent review by Restuccia and Rogerson (2017) amongst others. This chapter aims to disentangle this fundamental heterogeneity from input distortions.

Markup variation, misallocation and theory of the second best

I analyse misallocation from input distortions. That is, markups are allowed to vary in both the factual observed world and the counterfactual scenarios, and are determined by the estimated demand fundamentals and all prices and quantities. There is an alternative literature, that views variable markups as distortions compared to a CES world. This literature, predominately in trade, often analyses variation and changes in markups and market share reallocation in response to trade shocks.[27] One might argue that lumping the bundle of market power and input distortion together could be interesting along the lines of the theory of the second best: demand elasticities and input distortions could (partially) offset each other, such that we only care about the joint bundle. It might be tempting to conclude that the literature applying a Hsieh and Klenow (2009) approach does exactly that, as a constant markup is assumed and the inferred distortions implicitly capture the bundle of input distortions and idiosyncratic markups.[28] However, their CES demand framework is inconsistent with variable markups, so we cannot calculate welfare (or TFP) losses if we believe that the inferred distortions also capture variable markups, as it would necessarily change other parts of the model.

[27] For so-called "pro-competitive" effects of trade, see e.g. Edmond et al. (2015); Arkolakis et al. (2018). Typically the question is whether trade shrinks the variation in markups across firms due to increased output market competition. In models with CES demand, markups are constant. Models with variable demand elasticities introduce markups that vary across firms. When taking the stance that the variation in markups (i.e. demand elasticities) is not socially optimal, then we can also think of markup variation as additional misallocation of market shares. As in Dhingra and Morrow (Forthcoming), private markups are then not equal to socially optimal constant markups. See also Behrens et al. (2018) who quantify the welfare gap between equilibrium and the optimum allocation under monopolistic competition with heterogeneous sectors and firms.

[28] To see that, take the well-known contributions of both Hsieh and Klenow (2009) and De Loecker and Warzynski (2012) that infer input wedges and markups respectively from the same first order condition equation. Hsieh and Klenow (2009) assume η_{jt} to be a constant scalar, while De Loecker and Warzynski (2012) implicitly assume $\tau_{jt}^{M}=\tau_{jt}^{L}$ or simply an absent τ.

On the contrary, in this chapter, we learn whether the theory of the second best applies by isolating the input distortions. In the counterfactual that removes all input distortions but allows variation in demand elasticities, welfare gains include any effects from previously offsetting distortions. If the theory of the second best applied, welfare would go down when input distortions are eliminated as they would have previously offset variation in demand elasticities.[29] I find welfare gains from removing input distortions, thus the joint presence of input distortions and variable markups does not constitute a second best outcome.

Before I describe how I perform the counterfactual estimation and the implications for welfare, I explain how I identify and estimate the output elasticities in the next section, and the demand elasticities thereafter. The demand framework also pins down the welfare framework used for counterfactual analysis.

I.2.2 Estimating production elasticities and output shocks

We can rewrite the production function in the cost minimisation problem in logarithmic form where lower case variables indicate logarithms.

$$q_{jt} = f(k_{jt}, l_{jt}, m_{jt}) + \omega_{jt}$$

Unexpected output shock and functional form assumption

I incorporate an additional error term ϵ into the entire structural model, so that the estimation is consistent with firm behaviour and the Bertrand competition framework throughout. I provide the details in Appendix I.C.1. During or after production, once input choices have been made, an unanticipated multiplicative shock to expected firm output occurs ($\exp(\epsilon_{jt})$) and defines realised, observed output Q^r_{jt} based on anticipated equilibrium output Q_{jt}:

$$Q^r_{jt} = Q_{jt} \exp(\epsilon_{jt}) \tag{I.4}$$

For the baseline estimation and counterfactual analysis, I follow the standard in the literature and assume a Cobb-Douglas production function:

$$q^r_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \omega_{jt} + \epsilon_{jt} \tag{I.5}$$

[29] Note that there is only one output, so one demand elasticity per firm, while there are multiple inputs, so multiple input distortions. If input distortions differ by input, then a single demand elasticity cannot offset all input distortions.

The advantage of Cobb-Douglas production functions is that we can derive a simple closed form analytical solution for the conditional input demand functions which dramatically eases the search for equilibria. In Appendix I.C.3, I use a more flexible translog production function instead. The average production elasticities for this specification are reassuringly close to the Cobb-Douglas estimates.

Control function approach for identification

There are two well-known challenges with estimating production functions. They stem from unobserved productivity ω_{jt} and generate a simultaneity bias and a selection bias, as explained in more detail in Appendix I.C.1. In order to estimate the production function consistently and address these concerns, I make a set of assumptions that was first introduced by Olley and Pakes (1996), and later refined by Levinsohn and Petrin (2003), Wooldridge (2009) and Ackerberg et al. (2015), and commonly referred to as the proxy method or control function approach. The strategy is to use a control function for unobserved productivity to recover it, instead of for example, simply instrumenting for input choices.

A detailed description of my adaptation of this approach and the required assumptions are carefully explained in Appendix I.C.1. The population moment equations used for identification, where $\mathbf{\Theta}$ is the vector of all structural parameters, are:

$$E\begin{pmatrix} \epsilon_{jt}(\mathbf{\Theta}) & | & \mathbf{\Gamma}_{jt} \\ \big(\epsilon_{jt}+\zeta_{jt}\big)(\mathbf{\Theta}) & | & \mathbf{\Gamma}_{jt-1} \end{pmatrix} = 0$$

where ϵ_{jt} is the unforeseen production shock, which is uncorrelated with the current period choices and information set $\mathbf{\Gamma}_{jt}$. ζ_{jt} is the innovation in the Markov productivity process in ω_{jt} and uncorrelated to past input choices $\mathbf{\Gamma}_{jt-1}$. I use a joint estimation approach similar to Wooldridge (2009) to exploit these moment conditions.

This approach yields estimates for the structural output elasticities as well as the plant level productivity ω_{jt} and production shocks ϵ_{jt}.

An alternative for robustness checks: system GMM

To check the robustness of the estimated output elasticities from the control function approach to the invertibility condition, I also implement a dynamic panel system GMM approach following Blundell and Bond (1998, 2000). Details are reported in Appendix

I.C.2.

I.2.3 Demand structure and estimation

Dual role of the demand model

The demand model satisfies two roles. First, it allows us to estimate the elasticities of demand needed for the identification of τ_{jt}. I want to allow for flexible heterogeneous demand elasticities across (i) producers and (ii) counterfactuals by endogenising them. This avoids attributing uncaptured demand heterogeneity to the input distortions τ_{jt}.[30] As described in Section I.2.1, even though producers compete with the same product, such demand heterogeneity can arise because producers cover different geographical regions. Alternatively, product quality or brand loyalty differences introduce different price sensitivities among consumer groups or downstream firms. The second role of the demand model is to provide a structure for quantitative welfare analysis.

Heterogeneous consumers: mixed logit random utility model

The buyers of output in the application of this chapter – cast iron – are likely to be downstream firms, not consumers directly. However, to focus on the analysis of the cast iron sector, I abstract from modelling downstream sectors. Downstream firms are assumed to transform the outputs by segment into final products in a way that preserves product characteristics such that e.g. a high quality final product requires a high quality output. The downstream firms are assumed to operate with constant returns to scale and complete pass-through such that utility to consumers can be modelled as if they are buying the product directly. Schmalensee (1976), for example, shows that in competitive markets with constant returns to scale, consumer surplus can be estimated from the market of intermediate goods (i.e. the output of the cast-iron firms) instead of the final goods. What matters is that the demand elasticities are well estimated using the variation in output prices and quantities of all firms. I will therefore model consumers as if they are buying directly from the cast iron firms.[31]

Heterogeneous consumers face a discrete choice problem from which firm j to buy

[30] In the literature following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), studies have typically employed a simple CES model of demand with an assumed instead of estimated demand elasticity. This allows for some welfare analysis, but confounds heterogeneity in demand with input distortions. The literature estimating heterogeneous markups following De Loecker and Warzynski (2012) does not assume any demand model, which prevents welfare analysis.

[31] As mentioned in Section I.2.1, transport costs on the output side are not contained in the prices and implicitly captured by the estimated demand (cross-) elasticities and unobserved product characteristics.

to maximise their utility. Consumer heterogeneity in terms of price sensitivities and preferences over characteristics can be gauged by a random coefficient utility model. The seminal contribution of Berry et al. (1995), henceforth BLP, develops a random utility mixed logit approach which (i) has more realistic properties regarding demand (cross-) elasticities than either a basic logit model (independence of irrelevant alternatives) or a Kimball (1995) model (where elasticities depend only on output shares), and (ii) addresses price endogeneity. The framework is also well suited for welfare analysis. The downside is that it is not trivial to estimate this system and that algorithms for counterfactual analysis can be time-intensive to converge. Crucially, the price elasticity of demand and the markup depend on the structural parameters and distortions and prices and quantities of all firms. They are thus endogenous and vary across factual and counterfactual scenarios.

Specifying the utility function and demand elasticities

Consumers are indexed by i and choose between products j to maximise their utility:

$$U_{ijt} = (y_{it} - P^r_{jt})\theta^p_{it} + x_{jt}\theta^x_{it} + \xi_j + \xi_t + \Delta\xi_{jt} + \mu_{ijt} \equiv V_{ijt} + \mu_{ijt} \tag{I.6}$$

where y_{it} is consumer income, P^r_{jt} are realised prices (which are associated with realised quantities – these are the ones that are relevant for the consumers), x_{jt} a vector of product characteristics and a constant, ξ_j average utility from unobserved time-constant product characteristics, ξ_t average unobserved market-specific utility, and $\Delta\xi_{jt}$ the unobserved deviation from a particular product in a particular market from the unobserved averages. The unobserved ξ_j can contain the quality and the location of a product and will be absorbed by fixed effects dummies.[32] For the baseline results I only include a constant in x_{jt} as there are few time variant product characteristics (since the time invariant characteristics are absorbed in ξ_j). The parameters θ^p_{it} and θ^x_{it} are the random coefficients that determine the heterogeneity in preferences across consumers and are allowed to vary both by consumer and by market[33]. The set up in Equation (I.6) allows for heterogeneous marginal utility of income (and prices) across consumers.[34] The non-random utility can be summarised by V_{ijt}. The random utility component is μ_{ijt}, which follows an i.i.d. Type I

[32]See Nevo (2001) for a discussion of the benefits of such brand dummies. The dimensionality increases with J, and not with J^2 as in an AIDS model (Deaton and Muellbauer, 1980).

[33]We can interpret this as different consumers in different markets (periods). This precludes dynamic demand considerations.

[34]The consumer specific marginal utility of income is, however, constant with the level of income, which facilitates welfare calculation, and follows from risk-neutrality.

extreme value distribution.[35]

Appendix I.D.1 describes how the parameters in the utility function are identified and estimated. The algorithm involves an inner loop the minimises the distance between the observed market shares and the theoretically derived market shares from the utility maximisation. The outer loop addresses prices endogeneity and forms the moment conditions. The price elasticity of demand is:

$$\frac{1}{\eta_{jt}} \equiv \frac{\partial Q_{jt}}{\partial P_{jt}}\frac{P_{jt}}{Q_{jt}} = \frac{\partial(s_{jt}Y_t)}{\partial P_{jt}}\frac{P_{jt}}{s_{jt}Y_t} = \frac{\partial s_{jt}}{\partial P_{jt}}\frac{P_{jt}}{s_{jt}} = \frac{P_{jt}}{s_{jt}}\frac{1}{N}\sum_i \left(\theta^p_{it}s_{ijt}(1-s_{ijt})\right) \tag{I.7}$$

where s_{jt} is the market share of product j, Y_t the market size and s_{ijt} consumer i's expected expenditure share in product j (see Appendix I.D.1). I omit the notation with r for realised output (or market share) here, since the elasticities can be derived from any prices and quantities (so in the realised as well as in the counterfactual equilibria) conditional on the estimated parameters. Cross-elasticities can be calculated similarly and vary by firm-pair in each market.

I.2.4 Factual and counterfactual equilibria and welfare

With the estimated structural parameters, I can recover the matrix of input distortions $\boldsymbol{\tau}$ and solve for counterfactual allocations. I first discuss what the relevant counterfactual is, before I describe how I solve for it.

Misallocation costs: counterfactual distortions as weighted geometric average

What is the relevant counterfactual to evaluate the size of misallocation losses? The relevant "no misallocation" counterfactual for this chapter is the state of the economy when the distortions are removed. This would be the allocation that would occur in the same oligopolistic setting, but without input distortions. Note that this is not necessarily the optimum that a social planner might choose, which is the counterfactual in Behrens et al. (2018), for example. The counterfactual in this chapter can be interpreted as what we could achieve, if we managed to address input distortions in a market economy.

In the counterfactual, the distribution of $\boldsymbol{\tau}$ is degenerate, such that it is constant across plants. In principle, any constant $\tilde{\boldsymbol{\tau}}$ would equalise marginal revenue products of inputs across plants, adjusted for input prices (recall $\tau^M_{jt} \equiv \frac{MRPM_{jt}}{P^M_{jt}}$). A natural candidate

[35] This is a standard assumption in the literature because it facilitates inversion of market shares and exact welfare analysis. Note that the distributional assumption is not required for identification, but the instruments are key to identification (Berry and Haile, 2014).

is setting $\tilde{\boldsymbol{\tau}}$ to unity. However, with measurement error in the deflator for output and input prices, unity is no longer the appropriate counterfactual as we would artificially inflate or deflate input costs per unit across the board. Note that this would not affect production parameter estimation (quantities) or demand estimation (year fixed effects) as the measurement error is common to all firms. The measurement error multiplies prices and in- or deflates all $\boldsymbol{\tau}$ by the same proportion within each cross section.[36] Consider a simple example. Suppose that the error in the input price deflator is a change in the unit of the input price (P^M_{jt}) from dollars into cents for one period. Since the measured $MRPM_{jt} \equiv (\eta_{jt}+1)\alpha^M_{jt}P_{jt}Q_{jt}/M_{jt}$ remains unchanged, all measured τ^M_{jt} are scaled down by a factor of 100. As a result, the counterfactual $\tilde{\tau}^M_{jt}$ needs to be scaled down by a 100 as well.[37]

I take the stance that *allocative* inefficiencies between plants should be attributed to *differences* in distortions across plants alone while preserving an average of the distortions (which could be the measurement error). We can get the correct counterfactual even for $\boldsymbol{\tau}$ that are polluted with measurement error, if we assume that across the economy the true $\boldsymbol{\tau^{true}}$ are on average neither favourable nor adverse per input used (i.e. unity). The correct counterfactual is unity multiplied by the measurement error. This is equivalent to setting the counterfactual $\tilde{\boldsymbol{\tau}}$ to each period's weighted geometric average $\boldsymbol{\tau}$, where the weights are plant expenditure on that input (i.e. materials or labour).[38] The weighted *geometric* average is taken because of the nonlinear scale of $\boldsymbol{\tau}$ (the geometric average is just the exponentiated arithmetic average of $\log(\boldsymbol{\tau})$). Weights are used to account for different plant sizes such that we retrieve the average distortion per input used, not the simple average across plants.

All welfare results are robust to and qualitatively the same when using a counterfactual

[36] The Hsieh and Klenow (2009) method implicitly addresses this by defining an aggregate production function with the same structure as for the plant level and then taking ratios of plant level quantities to aggregate quantities.

[37] As another example, suppose that all true τ^M_{jt} are already unity, i.e. there is no misallocation. Again, wrong price deflators would scale the measured distortions and we would wrongly infer misallocation losses since the measured distortions are not unity but some other constant.

[38] If measurement error ϵ^{DEF}_t multiplicatively enters the input price deflator, then we work with $\tau^{M,true}_{jt}\epsilon^{DEF}_t P^M_{jt}$ in the firm's costs, where $\tau^M_{jt} = \tau^{M,true}_{jt}\epsilon^{DEF}$. This in turn means all true $\tau^{M,true}_{jt}$ are multiplicatively shifted by the same ϵ^{DEF}_t in each period. The relevant non-misallocation counterfactual is not $\tilde{\tau}^M_{jt} = 1$ but $\tilde{\tau}^{M,true}_{jt} = 1$, so $\tilde{\tau}^M_{jt} = \epsilon^{DEF}_t$. With the weighted geometric average as counterfactual we achieve this under measurement error, as long as it holds that the weighted geometric average of the true $\tau^{M,true}_{jt}$ are unity (i.e. $\exp\sum_j \ln(\tau^{M,true}_{jt}) * weight_{jt} = 1$). The counterfactual $\tilde{\boldsymbol{\tau}}$ is each period's weighted geometric average $\boldsymbol{\tau}$: $\tilde{\boldsymbol{\tau}}_{\boldsymbol{t}} = \exp\sum_j \ln(\tau^M_{jt}) * weight_{jt} = \exp\sum_j \ln(\tau^{M,true}_{jt} * \epsilon^{DEF}_t) * weight_{jt} = \exp\sum_j \ln(\tau^{M,true}_{jt}) * weight_{jt} * \epsilon^{DEF}_t = 1 * \epsilon^{DEF}_t$. When presenting statistics on distortions in the rest of the chapter I therefore use annually demeaned distortions.

of unity, but inflated due to $\tilde{\boldsymbol{\tau}}$ being above unity in most cases.[39]

Equilibria with endogenous marginal costs, markups and aggregate inputs

The counterfactuals $\tilde{\boldsymbol{\tau}}$ change the cost structure of firms, which in turn implies different best response prices and quantities in the counterfactual Bertrand Nash equilibrium conditions, along with changes in the endogenous markups. Both the factual and counterfactual equilibria are defined as the following set of equations and inequality constraints:

Definition of equilibrium: *An (internal) equilibrium satisfies profit maximisation of all plants. This consists of intersecting their best response functions, which yields the set of first order conditions (FOC), and a set of inequality constraints (SOC) for sufficiency of profit maximisation, conditional on all structural parameters and the distortions* $\boldsymbol{\tau}$*:*

$$\frac{P_{jt}}{MC_{jt}\big(Q_{jt}(\mathbf{P_t}), \mathbf{c}_{jt}(\boldsymbol{\tau})\big)} - \frac{1}{1+\eta_{jt}(\mathbf{P_t})} = 0 \qquad \text{(FOC)}$$

$$2\frac{\partial Q_{jt}}{\partial P_{jt}} + (P_{jt} - MC_{jt})\frac{\partial^2 Q_{jt}}{(\partial P_{jt})^2} - \frac{\partial MC_{jt}}{\partial Q_{jt}}\left(\frac{\partial Q_{jt}}{\partial P_{jt}}\right)^2 \leq 0 \qquad \text{(SOC)}$$

I provide detailed derivations for the terms in Appendix I.E.[40]

We are thus comparing well defined equilibria when analysing misallocation losses in an attempt to gauge the full costs of misallocation. This includes potential expansion or contraction in aggregate input use. This approach provides the advantage of explicitly endogenising the key variables (prices, quantities, input use, demand elasticities, pass-through) while preserving the estimated structural parameters (production elasticities, preferences, plant TFPQ (Ω_{jt}), etc.).[41] So far, counterfactual analyses in the input misallocation literature have assumed exogenous markups that do not change in counterfactual equilibria.

Next I briefly describe how I obtain the factual equilibrium from realised prices and quantities, and how I obtain the counterfactual.

[39] Results available upon request.

[40] Existence is proved by finding an equilibrium. Uniqueness is not proved. Even if there were multiple equilibria, we do not know which one would be reached. I could not find any numerical evidence on multiple equilibria, as a set of genetic algorithms as well as multiple starting points converged to the same equilibrium.

[41] Often, counterfactual analyses using the BLP approach do not estimate the production side, and marginal costs are simply assumed to be constant with respect to output (Berry et al., 1999; Nevo, 2000*a*; Petrin, 2002). Since I explicitly incorporate and estimate a structural model of production, I can relax this assumption and allow marginal costs to vary with output quantities according to estimated production functions.

From realised prices to (factual) equilibrium prices

First we need to recognise that we do not observe the factual equilibrium directly. Due to the unanticipated shock to production ϵ_{jt}, we observe realised prices and quantities in the data which are different to the equilibrium quantities and prices that firms expected and chose. Yet, the stage where prices and inputs are chosen (i.e. before the shock) is the relevant stage for inferring the input distortions $\boldsymbol{\tau}$ as described in Section I.2.1. Since firms are risk neutral and the shock entirely unanticipated (and mean zero), it does not influence their production input decisions, as described in Section I.2.2. I assume that firms choose the next best (realised) prices that clear their shock adjusted produced output and therefore the market.

The equilibrium quantities can be easily calculated from realised quantities, $Q_{jt} = \frac{Q^r_{jt}}{exp(\epsilon_{jt})}$ from (I.4), and equilibrium market shares are $\hat{s}_{jt} = \frac{\hat{s^r}_{jt}}{exp(\epsilon_{jt})}$ from (I.30). Given the equilibrium quantities, I search for the equilibrium prices that solve the necessary and sufficient conditions of the Bertrand Nash framework. This timing assumption harmonises the production and demand estimation with equilibrium behaviour to derive the distortions $\boldsymbol{\tau}$ from Equation (I.3). See Appendix I.E for more details on finding the factual equilibrium.

Counterfactual equilibria

Once we have obtained the factual equilibrium and $\boldsymbol{\tau}$, we can set *any* counterfactual $\tilde{\boldsymbol{\tau}}$, search for the new equilibria and perform comparative statics between the factual equilibrium and a version of the counterfactual equilibria. The three main counterfactuals that I construct either eliminate the variation in distortions in material inputs to $\tilde{\tau}^M_{jt}$, in labour to $\tilde{\tau}^L_{jt}$, or both simultaneously. The counterfactual equilibrium is pinned down by a vector of (output) prices alone, given the structural parameters. All of the comparative statics are alongside the intensive margin. Some plants can operate near or at zero output in the counterfactuals, resembling firm exit, but I do not explicitly model the extensive margin of exit and entry of new firms. For the counterfactual analysis, I use a Cobb-Douglas production function, since cost and marginal cost functions as well as conditional factor demands can be derived analytically, which makes solving for equilibrium prices more tractable. Again, Appendix I.E provides the details.

We need to assume how much capital firms choose in the counterfactual. The simplest solution is to assume that optimal installed capital follows a static optimisation condition (i.e. the same as for labour and materials). This leaves us with the unknown distribution of the rental rate for capital. While I could assume a range of values or distribution for this rental rate, I back it out from a static optimisation condition in the factual equilibrium. The median value of this inferred rental rate is 29%. As Asker et al. (2014) show, attributing all rental rate differences to misallocation could be misleading, as capital is mainly a dynamically optimised input. The rental rate contains a mix of capital distortions and capital adjustment costs. I preserve the plant specific rental rate across factual and counterfactual equilibria, i.e. I preserve the degree of capital misallocation.[42]

How noisy are misallocation losses? A parametric bootstrap.

A feature of this chapter's approach that is novel to the input misallocation literature is that I am able to derive confidence bands for any of the comparative statics.[43] I draw a set of parameter estimates $(\boldsymbol{\Theta}, \boldsymbol{\Sigma})$ from their joint asymptotic normal distribution using the estimated covariance matrices and for each draw, find the factual equilibrium (since $\boldsymbol{\epsilon}$ is different for each draw), calculate $\boldsymbol{\tau}$, find the counterfactuals and perform the comparative statics analysis.[44] This channels the information about the uncertainty in the structural parameters, such as plant productivities from the underlying Markov process, output elasticities, preference parameters and markups into the final comparative static of interest.

Calculating profit gains and average expected compensating variation

Instead of relying on the usual aggregate production functions, I calculate aggregate firm profits, compensating variation and aggregate input productivities for comparative statics. Profits as well as material and labour productivities are straightforward to aggregate, as I solve for the factual and counterfactual equilibrium prices, quantities and inputs for each plant. I use exact welfare measures for the demand side by calculating expected consumer

[42] Alternatively, I could specify a more complicated dynamic optimisation problem modelling adjustment costs such as David and Venkateswaran (2017) and use the residual variation of this as capital distortions. My approach is more conservative in terms of total misallocation losses by maintaining any capital distortions (and adjustment costs) across counterfactual equilibria.

[43] At the time of writing I am not aware of a paper that provides estimates of uncertainty around estimated input misallocation losses without calibration. The structural approach based on microdata to generate estimates of uncertainty around gains may also be useful in related counterfactual analysis studies e.g. on spatial misallocation of housing (Hsieh and Moretti, 2019) or infrastructure (Fajgelbaum and Schaal, 2017), or misallocation losses from within-country trade distortions (Costinot and Donaldson, 2016). Adao et al. (2017) for example develop a structural "mixed-CES" approach for trade models based on perfectly competitive goods and factor markets that allows for bootstrapped confidence intervals around welfare gains.

[44] I draw from the production side parameters $\boldsymbol{\Theta}$ and demand side parameters $\boldsymbol{\Sigma}$, and assuming independence between them. On the production side, ϵ_{jt} is a function of the drawn parameters $\boldsymbol{\Theta}$ and data. On the demand side, I can solve for the linear parameters $(\theta^p, \theta^x, \delta_{jt}, \boldsymbol{\xi})$ by using the draws from $\boldsymbol{\Sigma}$, the contraction mapping and the linear IV regression. I repeat the draws and analysis 330 times for all outcomes and then take the desired quantiles of the outcomes in order to get consistent confidence intervals.

compensating variation CV_{it} from moving from the factual equilibrium prices $\boldsymbol{P_t}$ to the counterfactual equilibrium prices $\boldsymbol{\tilde{P}_t}$. It is an expected welfare measure because of the random utility component μ_{ijt}. For each consumer, the CV_{it} solves:

$$\max_j U_{ijt}(y_{it} - P_{jt}, x_{jt}, \mu_{ijt}; \theta^p_{it}, \theta^x_{it}, \boldsymbol{\xi},) = \max_j U_{ijt}(y_{it} - \tilde{P}_{jt} - CV_{it}, x_{jt}, \mu_{jit}; \theta^p_{it}, \theta^x_{it}, \boldsymbol{\xi}) \quad \text{(I.8)}$$

Due to (i) additive (ii) GEV random utility disturbances, and (iii) constant marginal utility of income, I can conveniently use the Small and Rosen (1981) close form expression for CV_{it}:

$$CV_{it} = \frac{\ln(\sum_j \exp(\tilde{V}_{ijt})) - \ln(\sum_j \exp(V_{ijt}))}{-\theta^p_{it}}$$

where $\tilde{V}_{ijt}$ and V_{ijt} are the counterfactual and factual utility components defined in (I.6). We can take the average over consumers to get average expected compensating variation for each period (market) per unit. That is $CV_t = \frac{1}{N}\sum_i CV_{it}$. Multiplying this figure by the total quantity of output yields total expected compensating variation.

I.3 Data and descriptives

I.3.1 Plant and product level data

I use annual plant level panel data from the Indian Annual Survey of Industries (ASI) from 2000 to 2012. Since the Collection of Statistics Act in 1953, detailed plant level data is collected by the Ministry of Statistics and Program Implementation (MoSPI) of India, and most medium and large firms are familiar with the reporting. The mandatory nature as well as the long history of the survey makes it an arguably established and reliably data source in the developing context.[45] For the purpose of this study, the most important features of the dataset are that the output and input information is provided by product codes, both in revenue (or expenditures) and in physical quantities, which allows me to disentangle quantities from price effects. The ASI was traditionally a repeated cross-section, which researchers matched throughout the years (Bollard et al., 2013; Harrison et al., 2013). Recently it has been released in panel form, which latest research has started to use (Martin et al., 2017; Rotemberg, 2014; Allcott et al., 2016; Akcigit et al., 2016).

[45]The data is an annual census of plants with $\geq$ 100 employees (until 2004 $\geq$ 200) and a sample (around 20%) of all plants with $\geq$ 10 employees with electricity and all plants with $\geq$ 20 employees without electricity from a 4 digit sector-state strata.

A general shortcoming of the ASI data is that it covers only the formal manufacturing sector defined in its 1948 Factories Act, while a large share of manufacturing employment is in the informal sector (around 80% (Hsieh and Klenow, 2014)). However, since larger firms tend to be formal, the formal sector accounts for around two-thirds of output in manufacturing (Allcott et al., 2016). Informality is less of a concern for the cast iron industry. It is highly likely that output is even more skewed towards formal firms in this sector compared to e.g. the textiles industry in India.

I study single product cast iron plants. The aim is to compare plants that are as homogeneous as possible in their production technology to disentangle distortions from production heterogeneity within sectors. Cast iron is a 7-digit product, and to give a sense of the level of detail, the number of 2-digit sectors in the ASI manufacturing section (NIC08) are 24, the number of 4-digit sectors is 137, whereas the number of 7-digit product categories (NPCMS11) that are also manufactured in India is 5476. On average there are 40 different product categories within a 4-digit sector.

Finally, I use single product plants, because the ASI reports plant level outputs by product, but not inputs by product line. While there are ways to deal with multiproduct plants[46], they are likely to introduce further misattribution of unknown input allocations to the estimated distortions.[47]

I.3.2 Cast iron production in India

Cast iron is an iron-carbon alloy with high carbon content produced in different grades (e.g. hardness) by varying carbon, silicon and other components and processes.[48] It is used in many machines, automobile parts (such as gearboxes and cylinders), pipes and historically in construction. Cast iron is made from melting pig iron (which in turn is produced from smelting iron ore with coke and limestone in a blast furnace), coke, limestone and scrap steel, and small quantities of other metals into a desired grade and primary casting. It can be placed in the production chain between the rawer pig iron upstream

[46] See either the simpler method in De Loecker (2011) or the more advanced method De Loecker et al. (2016), for example.

[47] For the demand side, I can included products that are produced by multi-product firms to increase the coverage of the sample and improve precision of the demand elasticities.

[48] The product codes of cast iron in the ASI data are 4111102 (NPCMS11) and 71112 (ASICC). Cast iron has at least 2% carbon content (till around 3.5%-4%), while steel has less than 2% carbon content. Sometimes cast iron is loosely included in the term steelmaking. Steel on the other hand is used in construction and infrastructure, heavy machinery, white goods and tools. The advantage of cast iron over steel is a lower melting point (and costs, as well as better machinability, i.e. cutability), but tends to be brittle and have less tensile and compressive strength than steel.

and semi-finished and finished sheets, cables, pipes, blades or tins (which might be turned into tools, doorframes etc) downstream. Depending on the final use, the downstream production chain can be shorter or longer.

How significant is cast iron production in the Indian iron and steel sector? As Figure I.13 in Appendix I.F shows, a declining but considerable share of plants that produce some product in the broader classification of iron alloys of primary form (ASICC 711 or NPCMS 411) produce also cast iron (from 35% of plants in 1999 down to 20% of plants in 2009). Sales of cast iron account for a slightly more stable 25% of primary iron alloys until the financial crisis of 2008. Figure I.14 in Appendix I.F shows that around 60% of firms producing cast iron are single product firms. However, the physical output quantity produced by single product firms is typically slightly lower, as multiproduct firms tend to be larger. There is limited industry concentration amongst the cast iron plants in the sample as Figure I.15 in Appendix I.F shows. Appendix I.F provides a more detailed account of India's iron and steel sector. It also gives more detailed environmental context to this sector. The substantial carbon emissions in this sector are in part determined by aggregate material productivities. One contribution of this chapter is to study whether misallocation has an effect on these aggregate material productivities.

I.3.3 Descriptive statistics of key variables

All output and input prices as well as book values of capital (at the start of the accounting period) are deflated with industry specific and capital deflators respectively from the Office of the Economic Adviser (2019). I use output prices net of plant level subsidies, taxes or distribution costs. Labour includes workers employed through subcontractors as well as informal labourers.[49] Wages include the salaries as well as bonuses and welfare expenses. For input materials, I use the sum of the weight of input materials, and the materials price is the corresponding average price, including shipping fees. I recover the input price by dividing total expenditure on material inputs by the total weight of material inputs. Table I.1 provides some descriptive statistics on the sample of plants, after trimming the plants which are in the bottom and top 1 percentiles of either output prices, the physical output to material or labour ratio, or sales to installed capital ratio. I only keep plants that are present for at least two consecutive years in the data. The total number of plants varies by year, from around 60 to 110.

[49]In Appendix I.G.10 I measure labour as expenditure on labour instead of man-days. This provides a sensitivity analysis towards measuring skill in labour, if skill is correlated with pay. Results are qualitatively robust to this.

Figure I.1: Plant output quantities.

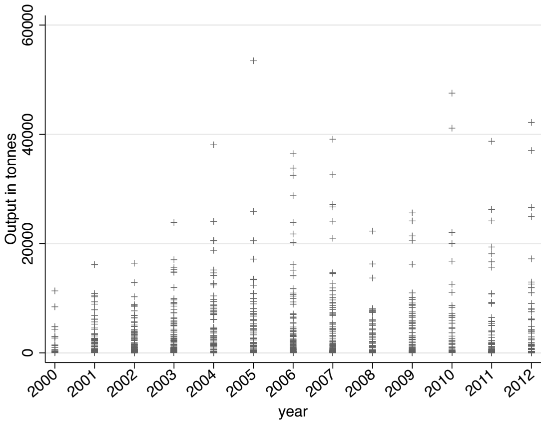


Figure I.2: Plant output prices.

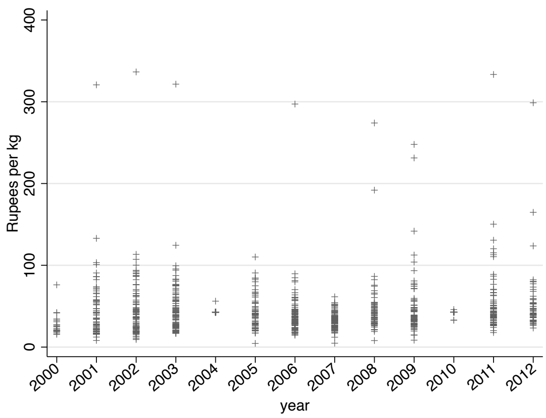


Figure I.3: Plant material input prices.

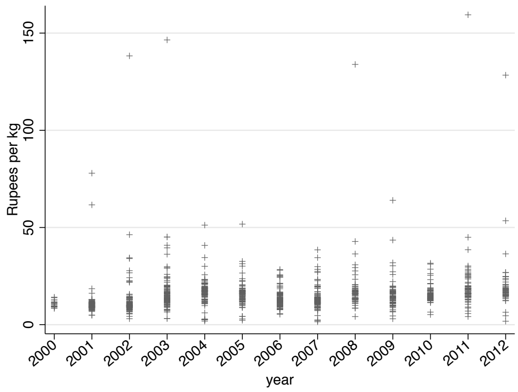


Table I.1: Descriptive statistics

	Mean	SD	p10	p90
Output quantity	3929	6345	140	10099
Output price	43.9	31.6	20.9	70.1
Materials quantity	5024	8932	154	12657
Materials price	15.7	11.4	8.8	23.0
Man-days (th.)	54.3	77.0	2.4	151
Employees	175	248	9.0	469
Daily wage (₹)	203	120	92.3	341
Capital (mil. ₹)	54.7	161	0.5	116
Observations	1001			

Notes: Quantities are in tonnes, prices in rupees per kg. All prices are deflated by 3-digit industry deflators. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

There is considerable variation in the scale in which plants operate, as shown in the plant output quantities in Figure I.1. The ratio of the 75th percentile to the 25th percentile of output quantity is 15.4. The prices for output and inputs are plotted in Figure I.2 and Figure I.3 respectively. Output prices are mostly within the range of 20 to 70 Rupees per kg with a mean of 44, which roughly matches global average steel prices over this period.[50] This suggests a setting with differentiated products. In a typical monopolistic competition framework, quantities are negatively correlated with prices in the cross section. I find no statistically significant relationship in the cross section. This suggest that there is quality differentiation. Figure I.17 in Appendix I.G.1 supports this story. It plots the output prices against the input prices and shows that they are positively correlated, consistent with quality differentiation where higher priced outputs require higher priced inputs.

[50] Figure I.2 also reveals that there is much less dispersion in output prices in 2004 and 2010 (also before any trimming), which is a feature of the underlying raw data, perhaps simply due to sample variability over the years. As I calculate annual misallocation losses, it is easy to see in Appendix I.G.8 that if anything, the misallocation losses for these two years are slightly smaller.

Average materials input prices are 16 rupees per kg and 80% are between 9 and 23 rupees per kg. Table I.1 converts labour L_{jt} in man-days into the number of employees (on average) with a mean of 175 across plant years. The wage rate is 203 rupees per day on average, which corresponds to around 4.5 USD per day. The 10th percentile at around half this figure. Despite an increase in wages from 3.3 to 5.2 USD per day over the sample period, they are still low and reflect the persistent poverty despite industrial growth described in Bhagwati and Panagariya (2014).

I.4 Results for demand, production and distortions

I.4.1 Results from demand estimation

To address the price endogeneity in the demand estimation, I use the average plant level wage w_{jt}, and the average plant level price of a tonne of material inputs P_{jt}^M as instruments for output price P_{jt}, which tend to perform well in BLP style estimations (Armstrong, 2016). A theory guided justification for the choice of these instruments, along with first stage tests, estimation results for the structural parameters and results from using alternative instruments is provided in Appendix I.G.2.

Estimated demand elasticities and decreasing markups over time

A more familiar parameter than the random coefficients in the utility equation are the demand elasticities ($\frac{1}{\eta_{jt}}$), which are determined by the estimated parameters and data as shown in Equation (I.7).[51] The estimated plant-year level demand elasticities are shown in Figure I.4, where the 90% bands are the percentiles of the distribution of the elasticities across plants within a year. We can use the equilibrium optimality condition (I.1) to express the demand elasticities as price over marginal cost markups. The median as well as the 90% bands of the cross section of markups is plotted in Figure I.5.

There is considerable variation in markups across plants within the given years. Across years, different plants are sampled, so there is natural sample variation. The years 2004 and 2010, for which I found little price variation in the raw data (Figure I.2), reassuringly also have less variation in demand elasticities and markups. The median elasticity and

[51]For the rest of the analysis, I ignore the 1% of observations where I estimate a demand elasticity larger than -1. With the standard oligopolistic models, we cannot calculate a markup or marginal costs for these observations. I do not drop these observations, but exclude them for calculating distortions τ and when comparing the factual to the counterfactuals. The 9 observations where this is the case have a median market share of 0.0004 and a maximum market share of 0.007, and their market share remains this small in all the counterfactuals.

Figure I.4: Demand elasticities

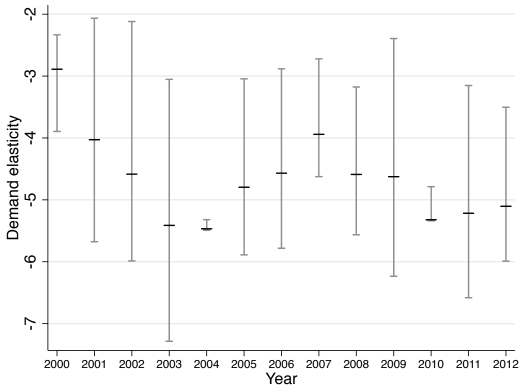
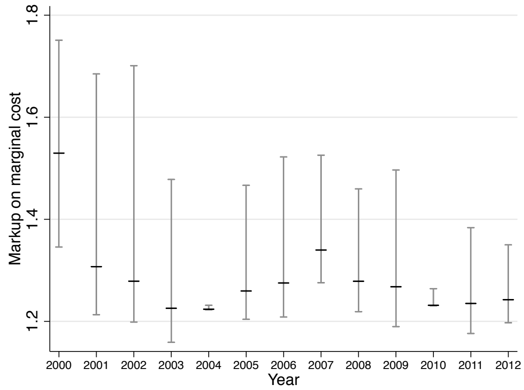


Figure I.5: Markups



Notes: The figures plot the estimated plant level demand elasticities (left) and markups (right). Plotted are the 5th, 50th and 95th percentile across plants within each year.

markup across all years are -4.82 and 1.26 respectively. De Loecker et al. (2016) find that in the Indian Prowess data during 1989–2003 the median markup for manufacturing firms is 1.34 and for basic metals firms 1.20 using the production side method of De Loecker and Warzynski (2012).

The markups decreased slightly over time. A linear regression of logged markups on years shows that markups decreased by 0.6% and 0.4% each year on average, in a pooled and a within-plant fixed effect regression respectively, significant with SE clustered at the plant. This is consistent with a story of increasing competition, particularly from large foreign low price producers form neighbouring China. I find that markups and plant total factor productivity are positively correlated. This correlation is driven by productivity pushing down marginal costs, as productivity and prices are also negatively correlated (see Appendix I.G.2). Markups and prices are negatively correlated, consistent with more elastic demand at higher price points as in Atkin and Donaldson (2015).[52]

I.4.2 Results from production estimation

The results from estimating the Cobb-Douglas production function are reported in Table I.2. Column (1) shows the baseline results with standard errors clustered at the plant level.[53] The direction of the bias in the OLS coefficients is as expected from the discussion in Appendix I.C.1. The material elasticity is upward biased from the simultaneity problem,

[52]The correlation between markups and prices, and markups and market share varies across periods, and is of opposite sign in some periods. This degree of flexibility is not possible with elasticities from Kimball (1995).

[53]The Hansen overidentification J-test for valid instruments is not rejected at the 5% level in any of the specifications. There is no standard rank test for instrument strength here as there are cross equation restrictions.

Table I.2: Estimates from a Cobb-Douglas production function

	Type of correction		Comparison to literature	
	(1)	(2)	(3)	(4)
	Simultaneity & Selectivity	None: OLS	De Loecker et al. (2016)	Collard-Wexler and De Loecker (2014)
α^K	.06***	.04***	.01	.08***
	(.02)	(.01)	(.06)	(.02)
α^L	.22***	.14***	.14	.27***
	(.05)	(.02)	(.09)	(.02)
α^M	.64***	.80***	.77	.68***
	(.05)	(.03)	(.11)	(.02)
RTS	.92***	.99***	.92	1.03***
	(.03)	(.01)		
N	443	1001	949	1498

Notes: The first two columns show the output elasticities and returns to scale with corrections for simultaneity and selectivity and without (OLS). The second two columns show results from related studies for comparison.

and the capital coefficient is (slightly) downward biased from the selectivity problem.

I perform several robustness checks regarding these estimates and the underlying invertibility condition, as discussed in Appendix I.G.3. I also use a translog production function (see Table I.7 in Appendix I.G.4) where elasticities vary by plant and year. The mean elasticities are very similar to the estimates from the Cobb-Douglas production function, with returns to scale close to unity.

Comparison to estimates from related studies

We can compare these results to other production function estimates from the relevant literature. In particular, De Loecker et al. (2016) estimate a translog production function for India for the period 1989-2003, however for the entire 2-digit basic metals sector, capturing other technologies. Collard-Wexler and De Loecker (2014) estimate a Cobb-Douglas production function for steel producers in the US between 1962-2002. The last two columns in Table I.2 compares my estimates to their estimates. De Loecker et al. (2016) estimate higher material elasticities and lower capital and labour elasticities, but the estimates of Collard-Wexler and De Loecker (2014) are remarkably close to my Cobb-Douglas and translog estimates. Arguably, the narrow technological focus of Collard-Wexler and De Loecker (2014) on steel producers in the US is more relevant for comparing elasticities than the geographic commonality but higher technological difference in De Loecker et al. (2016). Since their study in the US captures multiple decades, and not only cutting-edge technology, the production technologies are likely to be standard and similar to Indian producers in my data.

Analysis of estimated total factor productivity

Since I use output and input quantities as well as a gross output function, the control function approach estimates *physical* total factor productivity Ω_{jt} (also denoted TFPQ). Total *revenue* factor productivity TFPR is simply defined as $P_{jt} \cdot TFPQ_{jt}$. Do to the large and growing interest in TFPQ and TFPR in the literature, it is worth to briefly analyse these estimates. A more detailed analysis is in Appendix I.G.5.

The main points are as follows. First, there is evidence that more productive firms grew faster, based on comparisons of weighted and unweighted TFPQ. Second, TFPR grew by more than TFPQ, consistent with increasing prices. Together with decreasing markups, this implies that marginal costs have increased.[54] Third, the dispersion in TFPQ is smaller than in other studies, most likely due to the much narrower industry definition in this chapter. Fourth, the dispersion in TFPR is greater than the dispersion in TFPQ. This is in contrast to Hsieh and Klenow (2009). Prices and TFPQ are negatively correlated, but prices are much more dispersed than TFPQ, leading to a higher dispersion in TFPR than in TFPQ.

I.4.3 Results for the factual equilibrium and distortions

Descriptives on estimated input distortions τ: misleading SD

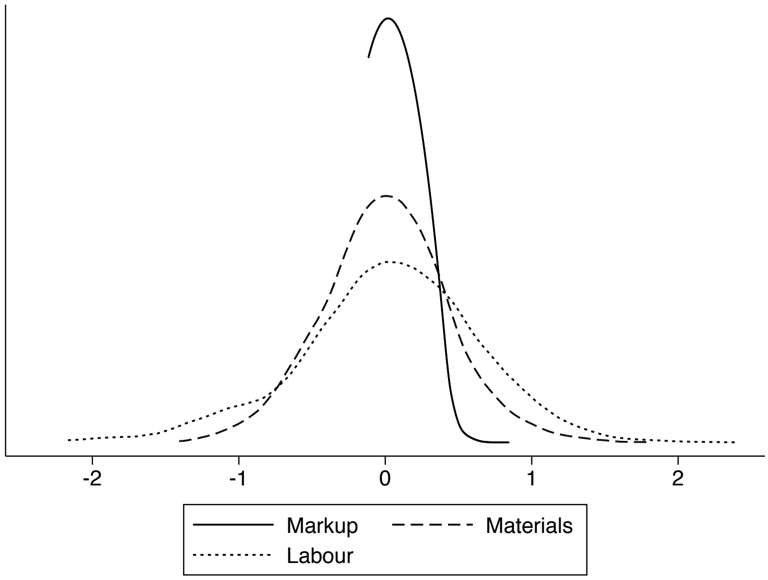
I calculate the τ_{jt}^M and τ_{jt}^L according to Equation (I.3), using the expected prices and quantities, the input expenditure, and the estimated output and demand elasticities.[55] For comparability, I also demean the $\boldsymbol{\tau}$, by dividing each by the within-year geometric weighted mean, where the weights are the input expenditures (see Section I.2.4), and take logs to transform it into a linear scale. The annual empirical density of these variables is plotted for τ_{jt}^M and τ_{jt}^L in Figure I.20 and Figure I.21 respectively for each year in Appendix I.G.7, whereas Figure I.6 pools the demeaned distortions across years.

There is pronounced dispersion in both labour and material distortions, and some

[54]Material input prices have been rising at around 3% per year, consistent with the global price increases in raw metals commodity prices (see e.g. IMF). Increasing marginal costs could also be due to changes in τ.

[55]In order to infer τ I first estimate the expected quantities Q_{jt} and prices P_{jt} as described in Section I.2.4 by using $Q_{jt} = \frac{Q^r_{jt}}{\exp(\epsilon_{jt})}$ and solving for the prices. The estimated shocks $\exp(\hat{\epsilon}_{jt})$ over the entire sample have a mean of 1.017 and the 90% range of estimates is [0.55,1.45]: on the extremes of this interval the plants have an estimated shock that decreased and increased output by 45% and 45% respectively. The log of the expected prices are plotted against the log of realised prices (i.e. after production shock ϵ) across all years in Figure I.19 in Appendix I.G.6 and shows that they are similar. The ratio of realised to expected prices ($\frac{P^r_{jt}}{P_{jt}}$) has mean 1.01 and the 90% range of estimates is [0.91,1.15], much tighter than quantity ratio, which is consistent with a convex elastic downward-sloping demand curve.

Figure I.6: Dispersion in markups, τ_{jt}^{M} and τ_{jt}^{L} across all years



Notes: Plotted are the kernel densities of the logged markup, $\ln(\tau_{jt}^{M})$ and $\ln(\tau_{jt}^{L})$ divided by the respective weighted means, where the weights are plant materials and labour expenditure. Used kernel is epanechnikov with bandwidth 0.2.

annual densities are multi-peaked. There is no clear trend in the degree of dispersion over time, but some years appear to exhibit less dispersion than others, in part influenced by sampling variability over time. A range of [-0.5,0.5] on the axis corresponds to a τ of [0.6,1.65]. Such values for τ imply that the plant faces a distortion "as if" it had to pay only 60% of the input price or pay a 65% tax on the input respectively.[56]

The standard deviations of the distortions, or alternatively the standard deviation of the marginal revenue product of an input, is often used as a statistic for misallocation (e.g. Hsieh and Klenow, 2009; Asker et al., 2014).[57] This chapter shows that, while being popular, this statistic can be misleading. The standard deviation of $\log(\tau_{jt}^{L})$ is 0.60 compared to 0.40 for $\log(\tau_{jt}^{M})$. The equality of variances is rejected in the robust Levene (1960) and Brown and Forsythe (1974) tests.[58] This might suggest that misallocation of labour is

[56] Recall also that the dispersion in τ is separate and in addition to any dispersion in plant input prices and wages disparities. Input prices are likely to reflect quality. The dispersion in material input prices is slightly smaller than the dispersion in $\boldsymbol{\tau}$, and the dispersion in wages is larger.

[57] Recall $\tau_{jt}^{X} = MRPX_{jt}/P_{jt}^{X}$. Often, the variation in TFPR is used as a summary of the $MRPX$ of all inputs.

[58] I can only conduct the test for the unweighted densities, since statistical significance is non-trivial to expand to weighted samples here. Figure I.22 in Appendix I.G.7, which plots the standard deviations corresponding to the plotted densities, shows that for some years, the standard deviation is greater in τ_{jt}^{M} than in τ_{jt}^{L}, but insignificantly so. For most years, the standard deviation is statistically significantly larger for τ_{jt}^{L}, where the statistical significance is obtained for the unweighted samples. Insignificant are the differences in the years 2001, 2002, 2006 and 2009. The hypothesis of equal distributions in the Kolmogorov–Smirnov test is strongly rejected.

more costly than misallocation of materials in this industry. However, as the next section shows, the opposite is the case in terms of welfare losses. Importantly, this implies that the pure variation in $\boldsymbol{\tau}$ is not a sufficient statistic to rank welfare losses, at least not across inputs. The size of the welfare losses depend on which plants face the distortions and how it affects other plants through the market structure, which cannot be captured in the variation of $\boldsymbol{\tau}$ or the marginal revenue products alone.

Statistical significance of estimated distortions

Since I can also derive confidence intervals for all individual τ_{jt}, I can test whether differences in τ_{jt} across plants are also statistically significant. I use two groups of plants, those with a τ_{jt} smaller than the 30th and those with a τ_{jt} larger than the 70th percentile. No plant is categorised in the opposite group in any of the bootstrapped versions of the τ_{jt}. Furthermore, for around 90% of all plants, the positive or negative logged demeaned distortion is significantly different from zero at the 10% level.

Adjusting distortions for markups and correlation between distortions

Figure I.6 also depicts the density of the logged demeaned markup. There is no mass below zero since this corresponds to a markup of less than zero percent. The dispersion of markups is significantly smaller, and the tests for equal variances and distributions are strongly rejected. We can also compare the distribution of $\ln(\tau^M_{jt})$ and $\ln(\tau^L_{jt})$ with a version for each with constant markups (i.e. a "naive" version), as shown in Figure I.23 and I.24 in Appendix I.G.7. While the "naive" and correct distortions appear similar, there are significant differences in the inferred welfare losses as discussed in Section I.5.4.[59]

Finally, it is interesting to ask whether the inferred distortions are correlated. Figure I.25 in Appendix I.G.7 plots the distortions against each other and the correlation is nearly zero and insignificant. Ex-ante we might expect that a firm that is constrained in one input is likely to be constrained in other inputs as well. While it is the case for some firms,

[59]If they looked very different, then we would likely obtain welfare *losses* from correcting the (wrong) distortions. Pooled across years, the standard deviation drops from 0.40 to 0.37 for $\ln(\tau^M_{jt})$ in the naive version, but stays roughly constant at 0.60 for τ^L_{jt}. The decrease in the variation in the former is because the naive $\ln(\tau^M_{jt})$ is negatively correlated with the estimated markup (-0.33***). This translates into a positive correlation with the inverse of the markup, so there is variation added to the naive $\ln(\tau^M_{jt})$ (See Equation (I.3)), increasing the standard deviation to 0.40 in the correct $\ln(\tau^M_{jt})$. Since the correlation between the naive $\ln(\tau^L_{jt})$ and the markup is low (0.05), there is hardly any change in the standard deviation. Despite similar pooled standard deviations, the consequences of mismeasured distortions in terms of welfare are more substantial as shown in Section I.5.4. For the same degree of dispersion, individual distortions can still be severely mismeasured. It matters which plants face which distortions for the size of the welfare losses.

the opposite is the case for other firms, being constrained in one input and having an advantage in the other input. One reason for not observing a stronger correlation between distortions could be that firms that are severely disadvantaged in both inputs are likely to be not competitive and exit or do not enter the market. The fact that the input distortions are uncorrelated also provides evidence against potential concerns that there is a distortion on the output side instead of the input side, as that would predict a perfect correlation between the input distortions.

I.5 Results from the counterfactual analysis

I begin this section by analysing the welfare consequences of input misallocation for consumers and producers. Then I discuss the effects on aggregate input productivities, before I compare the welfare losses to a version with constant markups. I end this section by showing the effects of misallocation on the size distribution of plants.

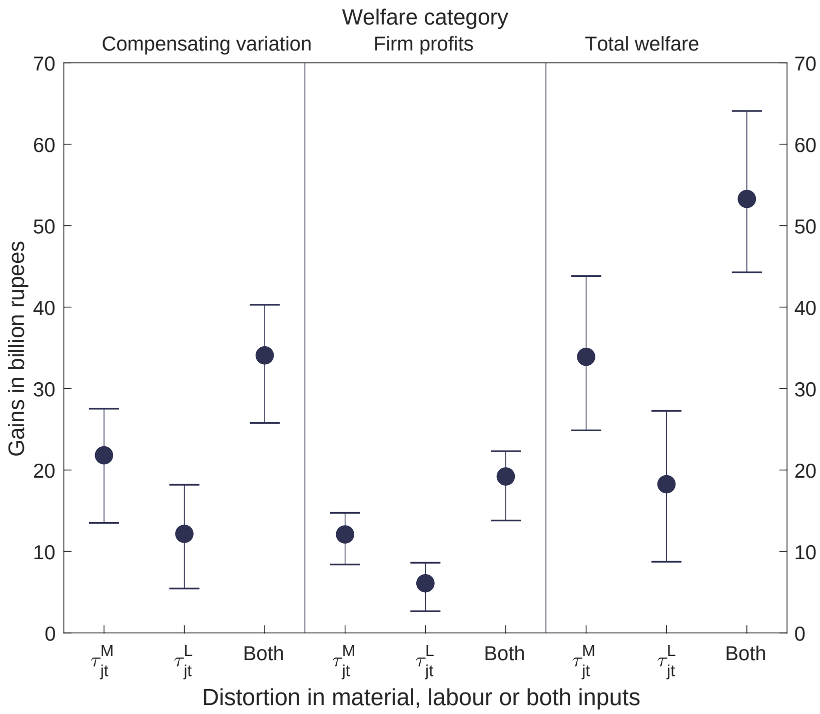
I.5.1 Welfare and profits

The gains from removing misallocation distortions are shown in Figure I.7. The first panel shows expected total compensating variation, the second aggregate profits and the third the total welfare gains, all in billions of rupees. For each panel, the gains from removing material or labour misallocation distortions individually are also reported.[60] The bootstrapped 95% confidence intervals are shown around the point estimates and derived as described in Section I.2.4. These figures are the total across all sample years. Table I.9 in Appendix I.G.8 shows the gains for each year individually. There is some variation in the gains across years, but without census data we cannot reliably interpret these as changes in misallocation.[61]

[60]Removing both input distortions can be slightly larger or smaller than the sum of gains, because the full interdependencies are taken into account through the model. The sum of the gains from removing either labour or material distortions is close to removing them jointly. This suggests that removing distortions from one input does not affect the gains from removing distortions in the other input.

[61]Almost all of the compensating variation, profits and total welfare estimates are positive and significant with few exceptions. The variation in the point estimates across years arise from three confounded factors, however. First, the sample size is different for each year, as some firms enter and some firms drop out of the sample, responsible for scale effects in aggregate compensating variation and profits. Second, due to the unbalanced sample, the composition of firms changes, and in some periods more dispersed τ firms may be sampled than in others, leading to higher estimated misallocation losses. Third, the actual degree of misallocation could have improved or worsened throughout the years.

Figure I.7: Welfare gains from removing misallocation distortions



Notes: Plotted are the respective welfare gains from eliminating material distortions ($\tilde{\tau}^M_{jt}$), labour distortions ($\tilde{\tau}^L_{jt}$) or both, summing across all 13 years of the sample. Bootstrapped 95% confidence intervals are shown around the point estimates.

Higher consumer than producer incidence

Who gains from removing misallocation – producers or consumers? The incidence of the changes in distortions depends on the pass-through rate of marginal costs to prices, the demand elasticities, market power and interdependencies between all plants. These are in turn determined endogenously by the estimated fundamental demand and production parameters. For all three counterfactuals, the compensating variation is around double the size of the profit gains, and the bootstrapped difference statistically significant.[62]

While individual plants experience increases or decreases in their marginal costs, depending on their initial level of $\boldsymbol{\tau}$, there is a more efficient allocation of inputs across plants, which benefits consumers through (average) price reductions. Indeed, the drop in average price per sold output quantity is 6%, 7% and 14% for the three counterfactuals respectively. While there are winners and losers on the firm side, the winners win more than the losers lose. Average profits are increasing despite average price declines, driven by

[62] I calculate the compensating variation only arising from the output produced by plants in the sample, and do not extrapolate to the whole market size, as for example in Nevo (2000*a*), in order to have comparable numbers for the consumer and producer gains.

the allocative efficiency gains.[63]

Higher misallocation in materials than labour inputs

A surprising result is that the welfare losses from misallocation of input materials are higher than those from misallocation of labour. As the right panel in Figure I.7 shows, the point estimate is 89% higher at almost 34 billion rupees vs. 18 billion rupees. The difference is statistically significant with a bootstrapped p-value of 0.03**.[64] This comes at a surprise for a prior that labour is a less flexible input with higher potential for misallocation, particularly in the Indian context.

There is hardly any evidence in the literature comparing misallocation of materials and labour directly.[65] But perhaps, differences in access to materials plays a bigger role than labour market distortions in this industry, as materials are an important production input with a high estimated output elasticity. For example, the political connectedness of firms (e.g. Faccio, 2006; Akcigit et al., 2018) could be more relevant for distorted spending on material inputs than on labour inputs. In Chapter II, I show that differences in geographic access to suppliers can partially explain these costly input material distortions.

The result that material misallocation costs are substantial is important for policy making for two reasons. First, targeting the allocative barriers in the material market could be easier, both politically and practically. As described in Hasan and Jandoc (2014) or Dougherty (2009) many Indian states have stricter labour firing laws for firms above a certain employee threshold. This could be a source of variation in τ_{jt}^{L}, but removing such polices could be politically challenging, due to valid concerns of employee protection. From a practical point of view, improving workforce mobility is challenging (Bryan et al., 2014) despite widespread regional structural mismatches in the labour market (Bryan and Morten, Forthcoming; Munshi and Rosenzweig, 2016).

Second, the (unmodelled) short term costs associated with reallocation in the materials market are likely to be much smaller than in the labour market, simply due to the fact that goods are reallocated and not people. Reallocation of labour involves hiring and

[63] Section I.G.14 shows that the variation in markups also decreases in the counterfactual for misallocation in materials.

[64] Comparing the difference in all bootstrapped runs is the appropriate way to test the difference to account for interdependencies, rather than comparing the individual confidence intervals.

[65] Hsieh and Klenow (2009) conclude that the misallocation in capital markets is higher than in labour markets in their study based on value added production functions. Dias et al. (2016) use the Hsieh and Klenow (2009) model for gross output with material inputs, and also find that capital misallocation is higher than labour misallocation. Slightly altering their model also sheds light on comparing labour and material misallocation in a Hsieh and Klenow (2009) style model. I analyse this in section I.G.15.

firing, and even if the new equilibrium features higher employment, there are undoubtedly transitional costs for the labourers, which may be large (e.g. Walker, 2013). Of course, addressing the variation in τ_{jt}^{M} also reallocates market share among firms, which necessarily also involves labour reallocation. But intuitively as well as empirically[66], the degree of layoffs (and hiring) is larger when removing labour markets distortions.

There are possible concerns that the losses from material misallocation are larger than for labour. First, in terms of external validity, other industries, particularly those that primarily rely on labour, are likely to have higher misallocation losses from labour. Second, I measure labour inputs as the number of total man-days, which does not account for the impact of skill on output in the production function. Misallocation of talent can play a role (Hsieh et al., Forthcoming). I construct a robustness test that accounts for skills by measuring labour as the total wage bill instead of man-days. If skills are paid a premium, the wage bill captures skills as well. As Appendix I.G.10 shows, the gap between losses from material and labour misallocation increases, if anything.[67] Third, wage disparities (which are slightly larger than material input price disparities) may be interpreted as misallocation themselves, and reducing those could offer additional gains. I keep plant level material input prices as well as wages constant throughout the counterfactuals, because I assume that firms are price takers on the input side. While this is an arguably realistic assumption for materials, given that there are much fewer and larger upstream firms (see Appendix I.B), it might not be the case for labour.

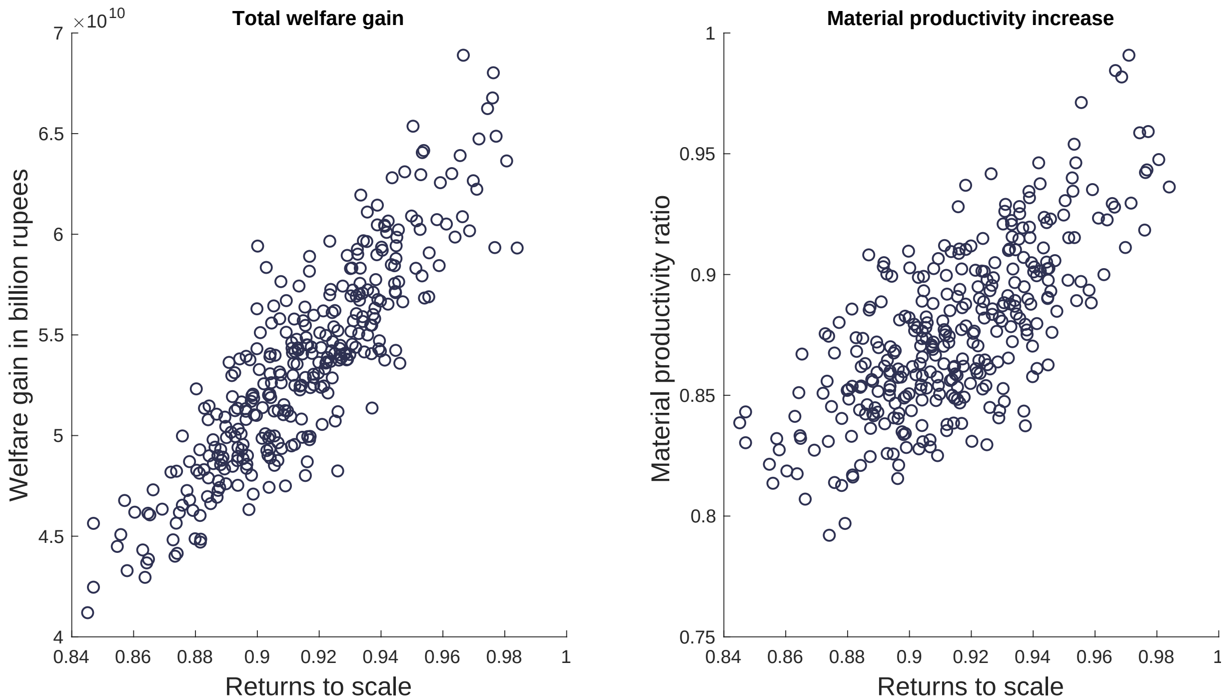
Accounting for changes in tax revenues

The interpretation of the misallocation distortions $\boldsymbol{\tau}$ is broad and could pick up any form of barriers and implicit costs. For the extreme case, that all of the $\boldsymbol{\tau}$ are only plant level input tax differences, I also calculate the implied impact on government tax income. I add the difference in tax income between the counterfactual and factual equilibrium to total welfare, as reported in Appendix I.G.11. Across the whole period, including government tax income changes in the calculation slightly increases the welfare gains from removing misallocation distortions.

[66] The counterfactual $\tilde{\tau}_{jt}^{L}$ involves more layoffs and more hiring than the counterfactual $\tilde{\tau}_{jt}^{M}$, despite the lower welfare gains.

[67] Note that I do not construct an analogous robustness check for materials. This is because output is measured in quantity. Higher material quality (i.e. expenditure) is likely to increase the quality and the price of output, not necessarily its quantity. Higher skilled labourers, on the other hand, more likely increase the quantity and the price of output.

Figure I.8: Correlation between returns to scale, welfare and material productivity gains



Notes: Plotted are the respective outcomes from the bootstrapped runs, where the returns to scale is a function of the bootstrapped underlying structural parameters.

Correlation of distortions and TFPQ

As Restuccia and Rogerson (2008) show, the correlation between distortions τ_{jt} and plant level productivity ($TFPQ = \Omega_{jt}$) matters significantly for welfare losses.[68] In the counterfactual, previously high τ plants (constrained) tend to grow while low τ plants shrink.[69] If the constrained high τ_{jt} plants are also the more productive plants (high Ω_{jt}) the aggregate welfare effects are larger. In Appendix I.G.12, I show that the estimated correlations between τ_{jt} and Ω_{jt} are low, and that a higher correlation would have implied even higher welfare effects.

Insights from bootstraps: returns to scale are important

An advantage of the estimated structural model in this chapter is that we can examine the sensitivity of the welfare gains with respect to specific underlying parameters that vary across bootstraps. Figure I.8 (left panel) plots the total welfare gains in each bootstrapped run against the returns to scale, which are a function of the underlying bootstrapped

[68]Of course plant size also matters, so we should think of it as correlation weighted by size to be precise, see also Hopenhayn (2014*b*) for a theoretical exploration.

[69]Whether firms shrink or grow also depends on the (estimated) interdependencies between firms.

Table I.3: Total welfare gains and statistics from factual equilibrium across bootstraps

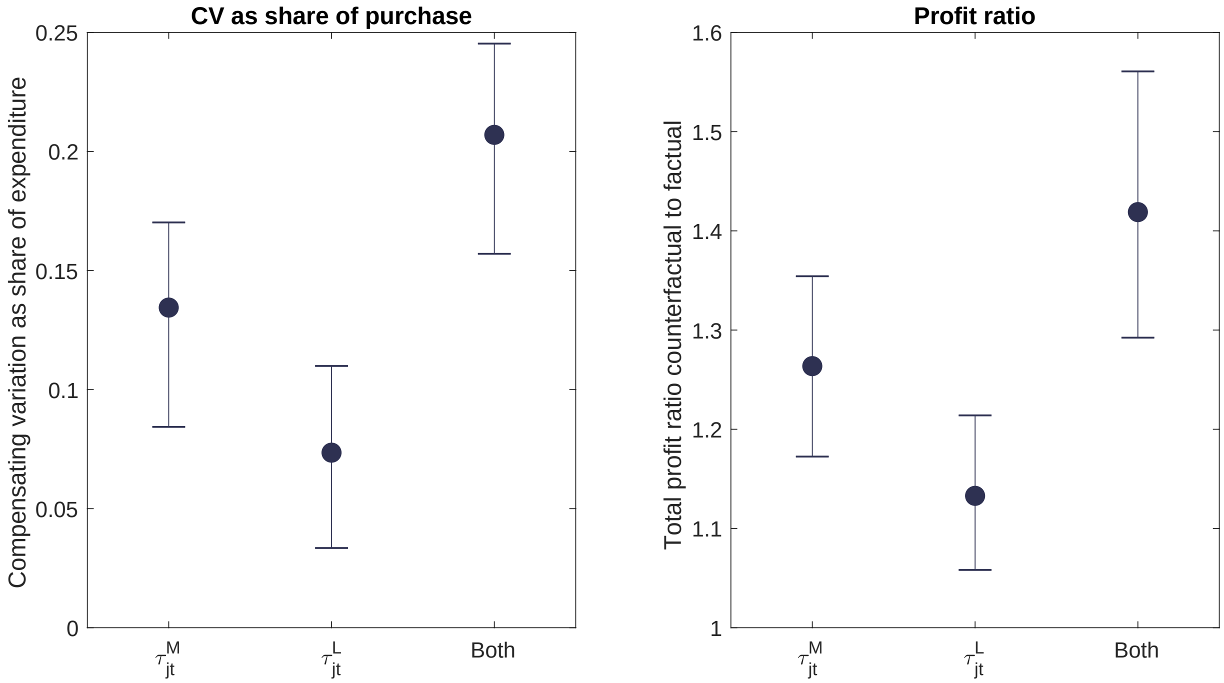
	(1)	(2)
Returns to scale	0.90*** [31.4]	1.04*** [19.8]
SD(Ω_{jt})		0.21*** [4.1]
SD($\frac{1}{\eta_{jt}}$)		0.19*** [7.5]
Corr(Ω_{jt}, τ^M_{jt})		0.15*** [3.2]
Corr(Ω_{jt}, τ^L_{jt})		0.18*** [4.0]
Corr($\frac{1}{1+\eta_{jt}}$, τ^M_{jt})		-0.17*** [-4.8]
Corr($\frac{1}{1+\eta_{jt}}$, τ^L_{jt})		-0.08*** [-2.7]
Corr(τ^M_{jt}, τ^L_{jt})		0.02 [0.9]
R^2	0.80	0.88

Notes: The table shows the estimates from an OLS regressions of total welfare gains (from both material and labour distortions) on statistics in the factual equilibrium. There are 330 bootstrapped runs, and each run is equivalent to one observation for this regression. Coefficients are standardized and t-statistics in square brackets based on robust standard errors.

structural production side parameters. The returns to scale are a significant driver of the size of the estimated welfare losses from misallocation. The 95% confidence interval means that returns to scale of 0.97 vs 0.86 are associated with around 40% higher welfare losses of 65 vs. 45 billion rupees. Table I.3 shows a regression of the total welfare gains on the bootstrapped returns to scale. The R^2 is high – almost 80% of the welfare variations across bootstraps can be explained by the variation in estimated returns to scale alone. The second column includes additional statistics from the factual equilibrium. As expected, the variation in plant level productivity Ω_{jt} as well as its correlation with plant level distortions are positively and significantly associated with welfare losses from misallocation, as discussed in the previous paragraph.

The size of the standardized coefficient on the returns to scale is much larger than of any of the other coefficients. Hopenhayn (2014*b*) shows that the original Hsieh and Klenow (2009) model is in theory highly sensitive to their constant returns to scale assumption. Indeed, when I use their model with data on the entire manufacturing sector or just my sample, and assume returns to scale of 0.92, all gains in their model are eliminated and some even turn negative, as Section I.5.3 shows. This corroborates the theoretical insights from Hopenhayn (2014*b*) empirically. There are two important messages emerging from this. First, the model and estimations of misallocation losses presented in this chapter – even though sensitive to the returns to scale – are much less sensitive than coarser approaches following Hsieh and Klenow (2009). Second, the uncertainty in the welfare gains is primarily driven by the uncertainty around the returns to scale. It is therefore

Figure I.9: Interpretation of the size of welfare gains



Notes: Plotted are the respective welfare gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both, summing across all 13 years of the sample. The left panel expresses the compensating variation per unit purchased as a share of the unit price (i.e. as share of expenditure on the products in the sample). Bootstrapped 95% confidence intervals are shown around the point estimates.

critical to accurately estimate the returns to scale for different sectors to have reasonably unbiased welfare estimates, and if possible, construct confidence intervals around them.

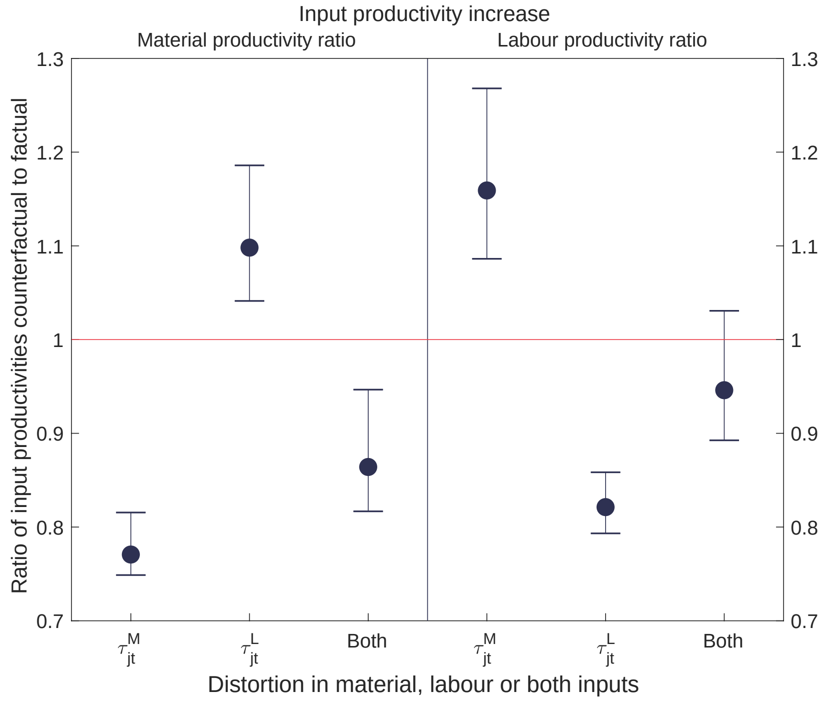
Large size of misallocation losses

Finally, it is left to discuss whether the reported welfare gains of 34, 18 and 53 billion rupees in the respective counterfactuals for $\tilde{\tau}_{jt}^{M}$, $\tilde{\tau}_{jt}^{L}$ and both, are large. Since I only cover the cast iron industry, it is most appropriate to set the gains into perspective of the size of this industry. Total sales are 171.5 billion rupees, so the total welfare gains are 20%, 11% and 31% of total sales of the plants in the sample.

I plot the relative levels of total compensating variation and profits in Figure I.9 (for annual figures see Table I.8 in Appendix I.G.8). I express the expected average compensating variation *per unit purchased* as share of the (factual) weighted average unit price.[70] I express the profit ratio as total profits in the counterfactual over total profits in the factual equilibrium. Removing misallocation in materials and labour increases consumer welfare equivalent to a price drop of 21%, while profits grow by 42%. Considering

[70]This is a useful statistic since each simulated consumer purchases one unit (see Equation (I.8).

Figure I.10: Aggregate input productivity gains from removing misallocation distortions



Notes: Plotted are the respective input productivity ratios (counterfactual to factual equilibria) from eliminating material distortions ($\tilde{\tau}^M_{jt}$), labour distortions ($\tilde{\tau}^L_{jt}$) or both, pooling across all 13 years of the sample. Bootstrapped 95% confidence intervals are shown around the point estimates.

that these comparative statics do not entail any technological innovation or diffusion, nor change the inherent factor price differences between plants and regions, the welfare gains from reallocation in material input and labour markets are sizeable. While general equilibrium effects across sectors are beyond the scope of this chapter, we know that under complementarity between sectors, misallocation in one sector can have large indirect effects on the economy through an input output structure (Jones, 2011, 2013).

I.5.2 Input productivities

Figure I.10 reports the changes in aggregate physical material productivity (physical output/physical materials input) and aggregate physical labour productivity (physical output/worker) expressed as a ratio of counterfactual to factual input productivity. The figures are pooled across all years, i.e. taking the sum of the quantities across years before calculating the ratios.[71]

[71] Table I.10 in Appendix I.G.9 reports the annual ratios as well as pure aggregate physical output changes. Table I.11 in Appendix I.G.9 reports the outcomes in terms of aggregate revenue productivities. The changes in physical input productivities are the more relevant metric to discuss as we do not need to deal with deflation (see Appendix I.G.9).

Distortions increase aggregate input productivity of same input

The first result is that material productivity *decreases* when material distortions are removed, and *increases* if labour distortions are removed. The same holds analogously for labour productivity. This might be initially surprising but has an intuitive explanation. When the distortions in the material market are lifted, plants with previous constraints use relatively more materials and plants with previous preferential access use relatively fewer materials. But due to the allocative efficiency improvements, the former plants expand their material use more than the latter plants reduce their use. On aggregate, the removal of frictions amounts to higher aggregate incentives to use that input relative to other inputs. This means that improving material misallocation can actually decrease aggregate material productivity through increased incentives to use that input. This result appears much like Jevon's paradox (Jevons, 1865), where increases in energy efficiency increase aggregate energy intensity due to a rebound effect, because energy is cheaper to use.

No misallocation losses in aggregate input productivities

When both input distortions are removed, aggregate input productivities slightly decrease (in the case of labour productivity insignificantly). This is because both aggregate outputs and inputs grow in the counterfactual.[72] The aggregate analysis masks high heterogeneity in changes of input productivities across plants, which are discussed in Appendix I.G.13. As for the welfare gains, the returns to scale in production matter significantly for the size of the gains in input productivities as seen in the right panel of Figure I.8. The results suggests that improvements in sectoral material efficiency require innovation and technology adoption, at least in this sector, as there is a limited role of improving allocative input distortions. Yet, with the introduction of new technologies, reallocation can still play an important part in the dynamics of the industry, as documented by Collard-Wexler and De Loecker (2014) for the US steel industry.

Emissions per welfare dollar higher from misallocation

One way we could include the environmental externalities of production (see Appendix I.F) into the welfare calculation is to compare the increase in emissions with the increase

[72]One immediate concern is that the distortions in capital have been preserved across factual and counterfactuals. I reran the entire analysis where I also remove any differences in the inferred rental rates (i.e."τ^K") in the same way as the labour and materials distortions, and calculate counterfactuals. The conclusions hardly change: the point estimates for the ratio in physical material productivity and labour productivity are slightly below one. Full results including welfare analysis with capital distortions available on request.

in welfare in a back of the envelope calculation. We know that materials use increases in the counterfactuals, as well as profits and consumer welfare. Since we do not have a baseline of consumer welfare, but only the compensating variation, we could just add it, in an admittedly simplistic way, to the profit increase. If we take the percentage increase of emissions as the percentage increase of material inputs,[73] and compare it to the welfare increase, we find that welfare grows 39% more than emissions. Therefore, while pure quantity input intensities are slightly increasing, the emission intensity of welfare is decreasing in the counterfactual. This is driven by the comparatively larger gains in welfare than in emissions from removing misallocation distortions.

I.5.3 Comparison to aggregate TFP results in the literature

The aggregate input productivity results may be somewhat surprising. Perhaps most prominently, Hsieh and Klenow (2009) found that aggregate TFP would increase by 40-60% in India if it equalised its marginal products to US levels. If sectoral TFP increased by such an extent, we would also expect the input productivities to increase in a similar fashion (with a near constant returns to scale). I can define an analogous standard aggregate Cobb Douglas production function to examine the impact on "aggregate TFP". Using my estimated elasticities from the plant level, I find that aggregate TFP is roughly the same in the factual and counterfactual equilibria. Why is there this apparent difference in this chapter to some of the key literature? There are three explanations that can realign these results.

The first is of interpretive nature. The TFP results of Hsieh and Klenow (2009) can be regarded as welfare results, as they are equivalent to gains in the utility of a representative consumer with CES demand as in Dixit and Stiglitz (1977) or Melitz (2003). Appendix I.G.15 shows this explicitly. When their results are interpreted as gains in utility rather than gains in physical productivity, they are consistent with the welfare gains in this chapter.

Second, the analysis of Hsieh and Klenow (2009) assumes that aggregate inputs do not change when removing distortions. They construct a statistic that summarises the loss from misallocation into a TFP measure, keeping aggregate inputs constant. When I use the plant level estimates of TFPQ (Ω) to calculate their efficient benchmark TFP_s^*

[73]Of course there might be non-linear relationships, both in the mapping from inputs to emissions as well as from emissions to welfare

[74], I obtain gains in TFP_s^* that are roughly 100-200% larger than the baseline aggregate Cobb-Douglas TFP. This is similar to their 127% TFP gain estimate for India in the 90s. However, this requires aggregate inputs to remain constant.[75] When distortions are removed, there is little reason why aggregate input demand should remain constant in the counterfactual equilibrium.[76] In this chapter, I can distinguish between TFP and welfare while comparing equilibria where the optimising behaviour of all plants, and potential aggregate input growth is taken into account. The results suggest that it is not aggregate physical TFP that increases, but welfare, which is a subtle but important distinction.

Third, there is a set of methodological and empirical differences to keep in mind when comparing the divergent TFP findings. For example, I observe plant specific input prices, use gross output instead of value added production functions, estimate all elasticities and returns to scale, allow for endogenous plant varying markups, and don't restrict the correlation between TFPQ and input distortions. Appendix I.G.15 provides more details and replicates the Hsieh and Klenow (2009) model and shows how the TFP results respond to changes to the above mentioned elements.[77] The returns to scale, in particular, affect TFP gains in their model substantially and can even turn them negative.

I.5.4 Markup changes and ignoring markups

In this chapter, I have accounted for markups that are endogenous and variable. This section first briefly examines how much the markups change endogenously. I then use a measure of distortions that ignores variation in markups to calculate counterfactual gains from mismeasured "naive" distortions.

[74] $TFP_s^* = \left[\sum_i TFPQ_{si}^{\sigma-1}\right]^{\frac{1}{\sigma-1}}$, the value depends on σ, here between 3 and 6.

[75] This is best seen when considering their equation for TFP losses, where the inputs only drop out if they remain constant across the inefficient and efficient equilibria (see Appendix I.G.15 for more details):

$$\frac{Y_s}{Y_s^*} = \frac{\left(TFP_s K_s^{\alpha_s} X_s^{\beta_s} L_s^{1-\alpha_s-\beta_s}\right)}{\left(TFP_s^* K_s^{\alpha_s} X_s^{\beta_s} L_s^{1-\alpha_s-\beta_s}\right)} = \frac{TFP_s^{\theta_s}}{TFP_s^*}$$

[76] When attempting to find a counterfactual equilibrium where I constrain aggregate inputs to factual levels, a set of algorithms with a range of starting points fail to converge to a point where firm first order conditions would be satisfied. This suggests that – at least in our case – there is no counterfactual equilibrium with the same level of aggregate inputs.

[77] I cannot fully nest their approach in my approach due to the substantial difference in both demand models. Ho and Ruzic (2017) nest a model allowing for non-constant returns to scale, and the difference in inferred misallocation losses are substantial for the US. In recent work, Haltiwanger et al. (2018) reject some of the key assumptions of the Hsieh and Klenow (2009) model using more detailed data. With my data, I can reject the same tests of the validity of the necessary assumptions in Hsieh and Klenow (2009).

Table I.4: Bias from constant markups

	Compensating Variation			Profits			Total welfare		
	$\tilde{\tau}_{jt}^M$	$\tilde{\tau}_{jt}^L$	*both*	$\tilde{\tau}_{jt}^M$	$\tilde{\tau}_{jt}^L$	*both*	$\tilde{\tau}_{jt}^M$	$\tilde{\tau}_{jt}^L$	*both*
Welfare losses baseline (bil. Rs.)	21.8	12.1	33.9	12.2	6.1	18.3	34.1	19.2	53.3
Bias with constant markups	-23%	-30%	-14%	-18%	-22%	-12%	-21%	-27%	-13%

Notes: The first row shows the baseline welfare losses from misallocation in inputs, where $\tilde{\tau}_{jt}^M$ refers to the counterfactual with removed material distortions, $\tilde{\tau}_{jt}^L$ with removed labour distortions and *both* to both removed. The bias in welfare losses is calculated from a counterfactual, where "naive" distortions are removed using this chapter's model. Naive distortions are inferred from ignoring the variation in markups.

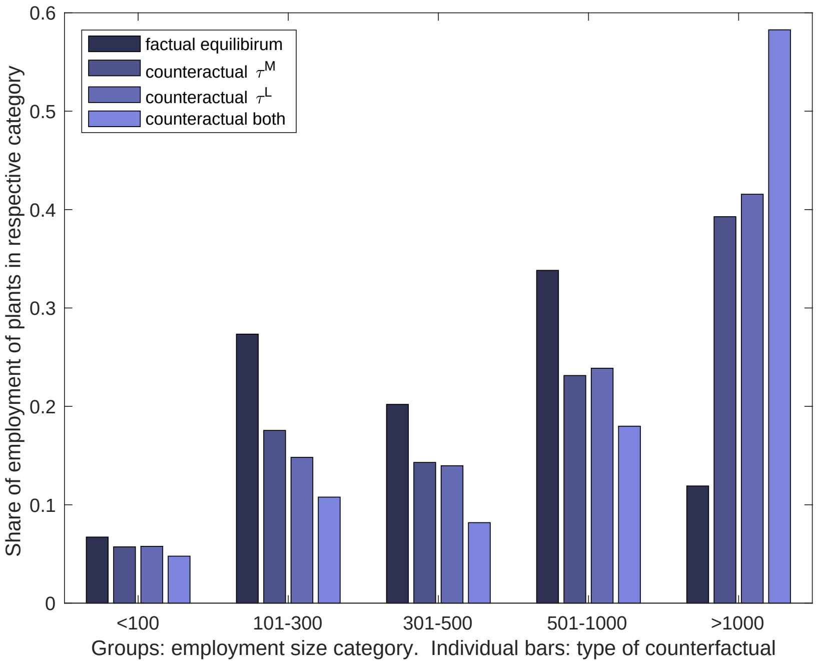
Markup variation across counterfactuals

On average, the endogenous markup changes by 5% for each plant between the factual and counterfactual equilibrium.[78] This is large compared to the 7% average deviation of markups from the average markup in the factual equilibrium. This suggest that accounting for the endogeneity of markups is important. With exogenous markups, we would be over- or undercounting input misallocation losses.

Mismeasuring distortions with constant markups generates bias

I obtain "naive" $\boldsymbol{\tau}$ by setting the demand elasticities η_{jt} (and therefore markups) constant when calculating this version of $\boldsymbol{\tau}$ from Equation (I.3).[79] I solve for the new counterfactuals, where I remove these naive $\boldsymbol{\tau}$ instead. The markups are still allowed to adjust endogenously in the counterfactual. That is, I mismeasure the $\boldsymbol{\tau}$ but still use the same model with the same primitives that allows for variable markups in counterfactual equilibria. As Table I.4 shows, the inferred welfare costs from misallocation are substantially *lower* than in the baseline version.[80] Mismeasuring distortions can considerably bias welfare conclusions. Disentangling distortions from fundamental heterogeneity can thus be important to avoid detecting allocative inefficiencies where there are in fact none, or failing to detect them where they in fact exist.

Figure I.11: The effect of misallocation distortions on the size distribution of plants



Notes: The vertical axis is the share in total employment by the plants belonging to a category. The categories from left to right are the size category in terms of plant employment: <100, 101-300, 301-500, 501-1000 and >1000. The left bar in each grouping is the factual equilibrium, the second bar the counterfactual with removed material distortions $\tilde{\tau}_{jt}^{M}$, the third bar the counterfactual with removed labour distortions $\tilde{\tau}_{jt}^{L}$ and the fourth bar the counterfactual where both distortions are removed.

I.5.5 The effect of misallocation on the size distribution of plants

Before I conclude this chapter, I briefly discuss the effects of the input distortions on the size distribution of firms. Hasan and Jandoc (2014) show that on average, 84% of India's manufacturing workers are in small firms, but 50% of China's workers are in large firms. They hypothesise that this is one of the proximate reasons for lower growth in India, as small firms tend to be less productive.[81] Figure I.11 shows that input misallocation is part of the reason for skewed firm distributions. In the counterfactuals, a much larger share of workers is in large firms. This can be viewed as complementary microlevel evidence to Bento and Restuccia (2017) who document a positive relationship between average

[78] Appendix I.G.14 reports further statistics.

[79] They are displayed in Figures I.23 and I.24 and briefly described in Section I.4.3. Despite the resemblance of the standard deviation of the naive and the baseline distortions, the welfare bias is still considerable, as it matters which distortions are measured for which plants.

[80] Ho and Ruzic (2017) present evidence that the Hsieh and Klenow (2009) model understates misallocation losses for the US manufacturing sector when markups are industry specific instead of common to all industries. Their markups are constant within industries.

[81] Kothari (2014) argues that developing countries typically have a thick left tail in the firm size distribution, in part driven by lower demand for high quality products predominately produced in larger firms, with some evidence for India. See more recent discussions on firm size distributions in the developing context e.g. in Cirera et al. (2018).

productivity and average firm size across countries, and show that distortions keep average firm size small.[82]

I.6 Conclusion

This chapter develops an approach that disentangles input misallocation distortions from fundamental heterogeneity in demand and markups across plants. I can distinguish between effects on producers and consumers, and effects on aggregate productivities. This provides a nuanced picture for the Indian cast iron industry. Removing input distortions in one input decreases the aggregate input productivity of the same input while improving the aggregate input productivity of the other input. This is in part driven by substitution to the input that is becoming more efficiently allocated. I find no evidence that removing misallocation distortions in both inputs would lead to improvements in input productivities. Since aggregate input productivities are determined by aggregate outputs and inputs, I allow aggregate inputs to adjust in counterfactual equilibria, which is in contrast to previous studies in this literature. This result is relevant for policies aimed at improving aggregate material efficiency. At least for the Indian cast iron industry, the results suggest that there are no allocative gains, and that within-firm innovations and technology diffusion are a more promising way to improve aggregate material efficiency.

I find that input misallocation significantly affects the size distribution of plants, keeping plants artificially smaller. There are also significant welfare losses from misallocation. The welfare losses are higher for consumers than for producers, driven by the price effects of input misallocation. The welfare losses from misallocation of materials are larger than those from labour. Even though I ignore any direct welfare costs on the employee side, this is a surprising result. While misallocation in material input markets has received little attention in the literature, the results suggest that these distortions could play a bigger role in explaining differences in performance of materials dependent sectors across countries.

The chapter provides methodological novelties to the literature. By combining production and demand into a full structural model I disentangle endogenous markups from input distortions. In the case of the Indian cast iron industry, ignoring variation in markups

[82] In line with Hsieh and Klenow (2014), I find that larger plants are more negatively affected from the input distortions. In terms of age, older plants are more adversely affected from material input distortions but benefit from labour distortions. The results on distortions and size distribution are also consistent with the finding of Martin et al. (2017). They show that dismantling small scale reservations by removing restrictions on firm size (India's SSI policy) led to output growth driven by the expansion of previously size constrained firms. See also Alfaro and Chari (2014) who analyse the firm size distribution in India following the end of the licence Raj.

biases the estimated welfare costs from misallocation downwards. I can provide confidence intervals around welfare cost and any other outcome, as I estimate all parameters in the model. This, for the first time in this literature, provides measures of uncertainty around aggregate misallocation losses. The developed approach can be applied to products and countries where quantity and price data on outputs and inputs is available. While there are disadvantages such as higher data requirements and computationally more intensive procedures, the benefits are detailed insights that admit a rich set of outcomes, which hopefully can be useful both for tailoring further research on misallocation as well as informing policy. Shortcomings are that I focus on misallocation in a static sense without dynamic considerations, ignore general equilibrium implications, and focus on the intensive margin ignoring firm entry. These provide interesting avenues for future research.

Appendix to Chapter I

I.A Further underlying and related literature

This chapter is related to four different types of literature. The theoretical and empirical misallocation literature in economic growth and development, the production and demand estimation literature from industrial organisation, a more policy oriented environmental material efficiency literature and the literature on Indian economic development of manufacturing industries.

In their surveys, Restuccia and Rogerson (2013, 2017) categorise the recent input misallocation literature into indirect and direct approaches.[83] Indirect approaches use wedges that capture a bundle of distortions, and typically aim to answer the question of how severe misallocation is (using data) or could be (using simulation). Direct approaches typically model or evaluate a particular distortion, often borrowing constraints, to analyse a particular cause of misallocation. This chapter relates more to the former literature.

Two of the most influential papers in this literature have been the theoretical analysis by Restuccia and Rogerson (2008), and Hsieh and Klenow (2009) using micro data. Restuccia and Rogerson (2017) argue that the literature attributes a significant role to misallocation from a variety of sources. Much of the focus in the literature is on misallocation in a static sense, but a few papers, such as Peters (2013); Da-Rocha et al. (2017) also emphasise the dynamic consequences of such static misallocation. For more detailed surveys of this literature, the reader is referred to the dedicated surveys by Syverson (2011); Hopenhayn (2014*a*); Restuccia and Rogerson (2013, 2017). There is a literature that decompose aggregate productivity changes into within-firm and across-firm components based on Olley and Pakes (1996) or Petrin and Levinsohn (2012), for example.[84] The nuanced difference in this literature is that these approaches only identify realised reallocation gains over the years, but not the level of misallocation compared to an optimum.

[83] One of the earliest articles on resource misallocation dates back to the study of monopoly power in the US by Harberger (1954). The more recent literature is based on new trade theory models with an emphasis on heterogeneous firms. Most papers analyse the manufacturing sector, some the agricultural sector (e.g. Adamopoulos and Restuccia, 2014*b*; Restuccia and Santaeulalia-Llopis, 2017; Adamopoulos et al., 2017), and few the service/retail sector (e.g. Vries, 2014).

[84] See also decompositions by Baily et al. (1992), Foster et al. (2001), Griliches and Regev (1995) or Baqaee and Farhi (2017).

Second, this chapter is related to an empirical industrial organisation literature to identify parameters while distinguishing different margins of heterogeneity. On the side of estimating production functions, this chapter follows the control function approach.[85] The literature on this topic dates back to Marschak and Andrews (1944) and is summarised in Griliches and Mairesse (1999) or Eberhardt et al. (2010). On the demand side, I implement a discrete choice random utility mixed model approach of Berry et al. (1995), which is based on the characteristics of products in order to address the representative consumer restriction and dimensionality problem of more traditional demand systems (e.g. AIDS by Deaton and Muellbauer (1980)), and allows for more realistic cross elasticities than more basic random utility logit models. For a survey, see e.g. Ackerberg et al. (2007).[86] Combining these two approaches on the production and demand side are novel to the misallocation literature.[87] This chapter is also related to price cost markup estimation. While there is literature to estimate markups from the demand side[88] and the production side[89], this chapter combines both.[90] I take the estimated markups from the demand system and use the identifying equation from the production side to disentangle the markups from input distortions that drive input misallocation.

This chapter is also relevant for the literature on material efficiency. In the policy sphere, there has been growing attention for sustainable material use due to environmental and economic considerations, see e.g. OECD (2015); European Comission (2013) or the creation of the dedicated Indian Resource Panel in late 2015. While emphasis is often on within-firm innovation, there is little evidence on whether across-firm misallocation could complement efforts in this respect. Baptist and Hepburn (2013) offer descriptive evidence that higher intermediate input intensity is correlated with lower TFP, which could point towards misallocation. This chapter is a first rigorous analysis of the impact of

[85]This approach has been introduced by Olley and Pakes (1996) and further developed by Levinsohn and Petrin (2003); Wooldridge (2009); Ackerberg et al. (2015). See also De Loecker et al. (2016) for a recent implementation with some innovations and Gandhi et al. (2016) or Forlani et al. (2016) for some criticism and alternative suggestions.

[86]Adao et al. (2017) for example apply a Berry et al. (1995) inspired demand system to a gravity trade model to depart from CES.

[87]Hsieh and Klenow (2009) for example assume values for the parameters on a demand side CES model, and take production side parameters from US ratios of aggregate data.

[88]See e.g. Stone (1954),Deaton and Muellbauer (1980), Goldberg (1995) and Berry et al. (1999).

[89]See e.g. Hall (1986, 1988), Roeger (1995) or De Loecker and Warzynski (2012)

[90]Forlani et al. (2016) unravel productivity and markup variation, but not input distortions. Pozzi and Schivardi (2016) or De Loecker and Scott (2016) also estimate supply and demand parameters, but no input distortions. In the bounding exercise of David and Venkateswaran (2017) to separate capital distortions from markups, the upper bounds ignore materials input distortions, as they are assumed to be absent for estimation.

misallocation on aggregate material resource efficiency.

Fourth, this chapter contributes to the literature on misallocation in manufacturing sectors in India. Bollard et al. (2013) and Harrison et al. (2013) find little *changes* in misallocation in India's manufacturing sector over time. This is a counterintuitive result as many economists and policy makers thought that the Indian reforms would impact allocative efficiency substantially. However, Nishida et al. (2014) and Nishida et al. (2015) show that these previous approach may be misleading and the method based on Petrin and Levinsohn (2012) yields opposite results. All these studies also use the Indian Annual Survey of Industries data.

I.B Input distortions and monopsony power

If inputs are not elastically supplied, i.e. plants are not price takers and have some monopsony power, the cost minimisation problem changes to:

$$\begin{aligned}\min_{L_{jt},M_{jt}} r_{jt}K_{jt} + \tau_{jt}^{L_{adj}} w_{jt}(L_{jt})L_{jt} + \tau_{jt}^{M_{adj}} P_{jt}^{M}(M_{jt})M_{jt} \\ \text{s.t.} \quad F(K_{jt}, L_{jt}, M_{jt})\Omega_{jt} \geq Q_{jt}\end{aligned}$$

where $\tau^{M_{adj}}$ is the new input distortion adjusted for monopsony power, and the input prices are some functions of the input quantities. The first order condition with respect to materials (and the analogue can be derived for labour) is:

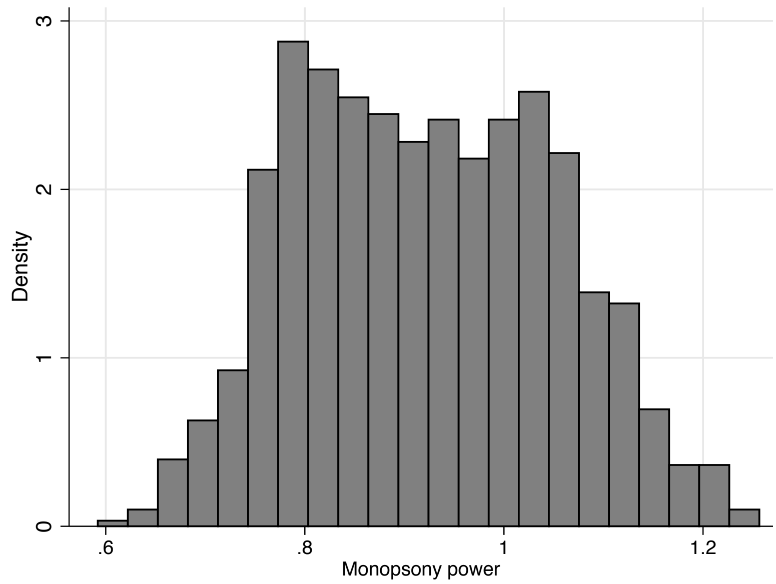
$$\underbrace{\tau_{jt}^{M_{adj}}(\psi_{jt}+1)}_{\equiv \tau_{jt}^{M}} = (\eta_{jt}+1)\alpha_{jt}^{M}\frac{P_{jt}Q_{jt}}{P_{jt}^{M}M_{jt}} \tag{I.9}$$

where $\psi_{jt} \equiv \frac{\partial P_{jt}^{M} M_{jt}}{\partial M_{jt} P_{jt}^{M}}$ is the inverse input price elasticity of input demand. Note that if we ignore $\tau_{jt}^{M_{adj}}$, and use that $MRPM \equiv (\eta_{jt}+1)\alpha_{jt}^{M}\frac{P_{jt}Q_{jt}}{M_{jt}}$, we can write $(\psi_{jt}+1) = \frac{MRPM}{P_{jt}^{M}}$. That is $(\psi_{jt}+1)$ is the ability to pay an input a lower price than its marginal revenue product, a common definition of market power on the input side, or monopsony power. The measured input distortion τ_{jt}^{M} captures input market power as well as other input distortions $\tau^{M_{adj}}$.[91]

How likely is it that the measured input distortion τ_{jt}^{M} represent monopsony power in this case? First, if plants that are close to each other and operate in the same input market have similar monopsony power, and τ_{jt}^{M} is primarily driven by monopsony power, then the

[91] See also Morlacco (2019), who studies monopsony power based on similar first order conditions.

Figure I.12: Histogram of estimated input market power ($\psi_{jt} + 1$)



Notes: The figure plots the histogram of monopsony power ($\psi_{jt} + 1$). A ratio larger than one suggest that the input price is below the marginal revenue product for that input.

τ_{jt}^M should vary more across states then within states. Using a variance decomposition as in Davis et al. (2013), I find that on average, the variance between states is 31% and 39% of the total variance for the logged τ_{jt}^L and τ_{jt}^M respectively, so most of the variance is within states. Second, for material inputs, monopsony power is likely higher if there are many upstream suppliers and few cast iron plants using that particular input. One of the main inputs is pig iron. We can compare the number of plants using pig iron as an input and the number of plants producing pig iron in the raw data. The number of plants producing pig iron is between 24 and 60 throughout the years. The number of plants using pig iron is between 280 and 800 throughout the years. It is unlikely that the comparatively large number of smaller plants can exert input market power over the few large plants. Third if monopsony power is plant specific and applies to both inputs, labour and materials, then the τ_{jt}^L and τ_{jt}^M should be correlated. Figure I.25 in Appendix I.G.7 shows that they are not correlated, however.

Fourth, I construct two, admittedly heuristic, plant specific proxies for monopsony power. If relatively larger plants can exert more market power on the input side as well, a larger market share of a plant in a given state can proxy for monopsony power. The correlations between the market share within a state and the material or labour distortions are small (<0.02) and statistically insignificant. The second proxy for monopsony power is based on a direct estimate of $\frac{\partial P_{jt}^M M_{jt}}{\partial M_{jt} P_{jt}^M}$. I recover those heuristically by regressing logged input prices on a second order polynomial in input quantities controlling for district and time fixed effects:

$$\log(P^M_{jt}) = f(\log(M_{jt})) + \boldsymbol{\lambda_d} + \boldsymbol{\kappa_t} + \upsilon_{jt} \tag{I.10}$$

Based on the coefficients, I can compute ψ_{jt}. I plot the histogram of input market power ($\psi_{jt}+1$) in Figure I.12. The majority of plants has a negative elasticity ψ_{jt} and therefore an input market power ($\psi_{jt}+1$) below one. A negative ψ_{jt} means that input prices decrease for larger quantities, which can be related to quantity discounts instead of input market power. The correlation between $\log(\psi_{jt}+1)$ and $\log(\tau^L_{jt})$ and $\log(\tau^M_{jt})$ is (-0.12 and -0.19) respectively, which is inconsistent with distortions capturing monopsony power.

I.C Estimation details for the production side

I.C.1 Control function approach to production side estimation

The production side identification problem

The fundamental problem is that we do no observe total factor productivity ω_{jt}, which is likely correlated with inputs and causes endogeneity problems.[92] Furthermore, we need to pin down plant total factor productivities for the counterfactual analysis.

Unexpected output shock

In order to avoid adding an ad-hoc error term just for the estimation, I incorporate an additional error term ϵ into the entire structural model, so that it is consistent with firm behaviour and the Nash competition framework throughout, as detailed below. Splitting up the combined error term into a so-called transmitted (to inputs) component ω_{jt} and untransmitted component ϵ_{jt} is common in the productivity literature (Griliches and Mairesse, 1999). The way we can interpret this in the context of the conduct model in this chapter is that the equilibrium prices and output are treated as expected prices and output. Firms maximise profits by choosing the expected prices in line with Bertrand-Nash competition. They base their production input decision on achieving the desired expected

[92]In traditional production function estimation total factor productivity ω_{jt} has often been treated as regression error term. This has been recognised as problematic for a long time (Marschak and Andrews, 1944), as it is very likely correlated with the input choices. Researchers often resorted to using an index number approach, essentially retrieving the output elasticities from the mean or median of the first order condition (I.3). If we assume that on the mean or median, the associated τ and $(\eta+1)$ are unity, then we could use this approach, at least for constant production elasticities. In the estimation strategy I use, the mean or median τ and $(\eta+1)$ vary by year and are not always close to unity, which would bias the index number estimates.

output by minimising costs. During or after production, an unanticipated multiplicative shock to expected firm output occurs ($\exp(\epsilon_{jt})$) and defines realised, observed output Q^r_{jt}:

$$Q^r_{jt} = Q_{jt} \exp(\epsilon_{jt}) \tag{I.11}$$

I assume that the input decisions have been made by the time this shock materialises and that this shock is entirely unpredictable by the firm. It could likewise also be interpreted as measurement error in the output variable. The firm productivity ω_{jt} on the other hand fully enters into the decision of the input variables. The realised, observed plant output in logs is therefore:

$$q^r_{jt} = f(k_{jt}, l_{jt}, m_{jt}) + \omega_{jt} + \epsilon_{jt} \tag{I.12}$$

This is the basic equation I want to estimate but in order to implement it I need to make further identifying and functional assumptions.

Functional form assumption

For the baseline estimation and counterfactual analysis, I follow the key literature and assume a Cobb-Douglas production function:

$$q^r_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \omega_{jt} + \epsilon_{jt} \tag{I.13}$$

With Cobb-Douglas, as opposed to a translog production function for example, we can derive a closed form analytical solution for the conditional input demand functions which dramatically eases the search for equilibria. However, in Appendix I.C.3, I make a second order Taylor approximation around the unknown production function which results in a translog specification and can be viewed as a generalised approximation to a CES production function. The production estimates from this more flexible translog approximation are reported in the results as well, and are on average reassuringly close to the Cobb-Douglas estimates.[93]

[93] Essentially the Cobb-Douglas is a simplification of the translog specification by enforcing some parameter restrictions.

Simultaneity and selection biases

The main identification challenge is that we are unlikely to get consistent estimates when running OLS on Equation (I.13) since we do not observe productivity ω_{jt}. While the shock ϵ_{jt} is assumed to be unexpected and unknown to the firm, and therefore uncorrelated with input choices, the productivity ω_{jt} is known to the firm, and highly likely to influence input choices. If a firm experiences a positive productivity shock, it is likely to use more variable inputs, creating a positive bias in the coefficient. This problem is commonly referred to in the productivity literature as simultaneity or transmission bias (Marschak and Andrews, 1944; Griliches and Mairesse, 1999).

A second identification issue comes from sample selection. Similar to Heckman's selection problem, we only observe firms that are in production. Firms' survival is positively correlated with productivity. But firms' decisions to exit are negatively correlated with installed capital, conditional on unobserved productivity. As Ericson and Pakes (1995) argue, capital serves as buffer to shocks. Therefore, surviving firms have an expected productivity that is decreasing in installed capital. This creates a downward bias in the capital coefficients in an OLS regression omitting productivity. Moreover, since I only use single-product firms for estimating the production function, there is an additional self-selection problem of single-product firms turning multi-product which is positively related to productivity. Conditional on productivity, firms with higher installed capital (or labour) are more likely to introduce a second product, as in the model of Mayer et al. (2014). Again, "surviving" single product firms (i.e. those not turning multiproduct) have an expected productivity that is decreasing in installed capital (and perhaps labour).

Note that in the selection problem the bias arises in the more persistent variable (capital), becoming more severe with more dynamics, whereas in the simultaneity problem, the bias arises in the flexible inputs (material), becoming more severe with higher flexibility.[94] Both can cause inconsistent and biased estimates of all coefficients, so we should address both.

Addressing simultaneity and selectivity

I assume that material inputs are a function of several observed variables and a scalar unobserved variable, productivity:

$$m_{jt} = m(k_{jt}, l_{jt}, \mathbf{z}_{jt}, \omega_{jt}) \tag{I.14}$$

[94]This is one reason why value added production functions have traditionally often been used in the literature, to avoid highly flexible materials as input in the estimation.

where $\mathbf{z}_{jt}$ are additional variables which I discuss below. Additional to this scalar unobservable assumption, it is assumed that this function is (conditionally) monotonically increasing in productivity ω_{jt}.[95] Therefore, the material demand function can be inverted for productivity, which we can later flexibly include in the estimating equation:

$$\omega_{jt} = h(k_{jt}, l_{jt}, \mathbf{z}_{jt}, m_{jt}) \tag{I.15}$$

where $h(.)$ is an unknown function, which we can approximate with polynomials or semi-parametrically. The choice of variables in $\mathbf{z}_{jt}$ could be important, but was omitted in the pioneering applications (Olley and Pakes, 1996; Levinsohn and Petrin, 2003). Gandhi et al. (2016) argue that having no additional variables in $\mathbf{z}_{jt}$ leaves the production function non-parametrically non-identified.[96] The second reason why $\mathbf{z}_{jt}$ may be important is that we need to be comfortable with the assumption that productivity is the only unobservable driving material demand, which is more likely if we control for factors such as input prices. For robustness checks, I include in $\mathbf{z}_{ijt}$:

$$\mathbf{z}_{jt} = (p^M_{jt}, IMP_{jt}, \eta_{jt}, p_{jt}, s_{jt}, \mathbf{G}_t)$$

that is log material input prices p^m_{jt}, import status of material goods IMP_{jt}, the inverse demand elasticity η_{jt}, log output prices p_{jt}, market share s_{jt} and firm location dummies $\mathbf{G}_t$.[97]. The last three variables have also been used in the proxy equation in De Loecker et al. (2016). However, they have neither observed input prices nor a measure of the demand elasticity. As Forlani et al. (2016) argue, variations in demand elasticity or market power are likely to drive material demand and thus are important to include. Note that the framework of De Loecker and Warzynski (2012) retrieves markups after production side estimation and can therefore not included markups in the estimation. The advantage of the approach in this chapter is that we can recover demand elasticities from independent demand side estimations.

A central concern is that the misallocation wedges τ_{jt} are likely to influence material demand. Since they are identified from the input prices, demand elasticity, output prices and market shares (Equation I.3)), it is sensible to include them in $\mathbf{z}_{jt}$. The monotonicity

[95] Only relatively mild conditions are necessary that the marginal product of materials is increasing in ω_{jt} (Levinsohn and Petrin, 2003) . This is easier to prove in the case where investment acts as the proxy variable (Pakes, 1996).

[96] As otherwise only the shock in productivity from Equation I.17 identifies it, which is unobserved and later assumed to be orthogonal to all input choice lags.

[97] As indicated in the results, I use a full $\mathbf{z}_{jt}$ for robustness checks, but not for the baseline results

assumption needs to hold only conditional on $\mathbf{z}_{jt}$. Figure I.18 plots the relationship between material inputs and productivity and shows a monotonic relationship. For concerns that the τ_{jt} still violate the scalar unobservable assumption, I furthermore implement a Blundell-Bond system-GMM estimator for the production function as a robustness check, as detailed further below.

In practice, I address the self-selection problem of firms from being single-product into exit and into multi-product firms by following and augmenting the strategies of Olley and Pakes (1996) and De Loecker et al. (2016). Essentially, I estimate the probability of being in the sample $Prob_{jt}$ with a discrete model, and the predicted probability from this estimation will be included in the final estimation.[98]

[98]In more detail: In Ericson and Pakes (1995), productivity follows a Markov process and the exit decision depends on a threshold value of productivity $\underline{\omega_{jt}}$. A draw below this threshold value makes it more profitable to sell the firm since its sell-off value is higher than discounted net profits based on the current productivity draw. However, the productivity threshold also depends on installed capital, and is decreasing in it since discounted profits are higher for higher capitalised firms. The sell-off value is assumed to increase less in capital than discounted profits increase in capital. In short, the firm exits if $\omega_{jt} < \underline{\omega_{jt}}(k_{jt})$. In the model of Mayer et al. (2014), the number of products is increasing (as a step function) in productivity draws. The multiproduct threshold productivity $\overline{\omega_{jt}}$ is again decreasing in capital, since bigger firms are more likely to be able to set up new product lines. In short, the firm becomes multi-product if $\omega_{jt} > \overline{\omega_{jt}}(k_{jt})$. Putting these elements together, the conditional probability of being in the single-product sample, indicated by $sp_{jt} = 1$ is:

$$\begin{aligned}
&Pr\big[sp_{jt} = 1 \mid \underline{\omega_{jt}}(k_{jt}), \overline{\omega_{jt}}(k_{jt}), \omega_{jt-1}\big] \\
=&Pr\Big[\underline{\omega_{jt}}(k_{jt}) < \omega_{jt} < \overline{\omega_{jt}}(k_{jt}) \mid \underline{\omega_{jt}}(k_{jt}), \overline{\omega_{jt}}(k_{jt}, l_{jt}), \omega_{jt-1}\Big] \\
=&\tilde{g}\Big(\tilde{d}\big(\underline{\omega_{jt}}(k_{jt}), \overline{\omega_{jt}}(k_{jt})\big), \omega_{jt-1}\Big) \\
=&g\Big(k_{jt-1}, i_{jt-1}, l_{jt-1}, \mathbf{z}_{jt-1}, m_{jt-1}\Big) \equiv Prob_{jt} \qquad \text{(I.16)}
\end{aligned}$$

where i_{jt-1} is investment. Conditionally on knowing the thresholds and previous period productivity, the probability that current productivity lies within the thresholds can be written as an unknown function of these elements. The reason why I use the notation of a function $\tilde{d}$ to summarise both thresholds will become apparent below. Since capital is a function of previous period capital and previous period investment, and productivity a function of given variables from the invertibility condition, we can write the survival probability as an unknown function of these previous period variables. I estimate this probability with a discrete model, and the predicted probability $\hat{Prob}_{jt}$ will be included in the final estimation. In this estimation I include whether the plant belongs to the census or the sampled sector, as this is additional critical information whether the plant is contained in the sample. I do not empirically restrict the threshold productivities to be decreasing in its arguments, but estimate the function $g(.)$ flexibly.

Markov productivity process

A common convenient[99] assumption in the production function and productivity literature, including the proxy approach, is that productivity follows a first order Markov process (see influential early papers of Hopenhayn (1992); Hopenhayn and Rogerson (1993)):

$$\begin{aligned}
\omega_{jt} &= E\big(\omega_{jt} \mid \omega_{jt-1}, IMP_{jt-1}, sp_{jt} = 1\big) + \zeta_{jt} \\
&= \Psi\Bigg(\omega_{jt-1}, IMP_{jt-1}, \tilde{d}\big(\underline{\omega_{jt}}, \overline{\omega_{jt}}\big)\Bigg) + \zeta_{jt} \\
&= \Psi\Bigg(h(k_{jt-1}, l_{jt-1}, \mathbf{z}_{jt-1}, m_{jt-1}), IMP_{jt-1}, g^{-1}(\omega_{jt-1}, Prob_{jt})\Bigg) + \zeta_{jt} \\
&= \tilde{\Psi}\Bigg(h(k_{jt-1}, l_{jt-1}, \mathbf{z}_{jt-1}, m_{jt-1}), IMP_{jt-1}, Prob_{jt}\Bigg) + \zeta_{jt} \qquad \text{(I.17)}
\end{aligned}$$

The productivity process is not completely exogenous[100], but is allowed to depend on the firm's import status, because of potential international technology spillovers and depends on the firm being in the sample and single product ($sp_{jt} = 1$) which ultimately depends on its survival probability from (I.16). Therefore the productivity is an unknown function of the elements in the last equation and the shock to productivity ζ_{jt}.[101]

The Markov assumption implies that the CDF of ω_{jt} is a decreasing function of ω_{jt-1}, i.e. that high ω_{jt-1} firms stochastically dominate low ω_{jt-1} firms. By construction:

$$E[\zeta_{jt} \mid \omega_{jt-1}, IMP_{jt-1}, \tilde{d}\big(\underline{\omega_{jt}}, \overline{\omega_{jt}}\big), sp_{jt} = 1] = 0 \qquad \text{(I.18)}$$

Note that the Markov assumption implies that ζ_{jt} is not only uncorrelated with all lagged variables in the function $h(.)$, but through the capital accumulation equation also with current capital k_{jt}, if current capital is only a function of previous period capital and previous period investment and depreciation, which have all been realised before the

[99]It is convenient, as with higher order Markov processes, we need a longer history of the data and effectively lose observations.

[100]It can also be allowed to depend additionally on R&D expenditures as in Doraszelski and Jaumandreu (2013), but I have no data on this.

[101]The second equality states that the conditional expectation is a function of its conditioning variables including the function of the productivity thresholds since they define the range for $sp_{jt} = 1$ if $\underline{\omega_{jt}}(k_{jt}) < \omega_{jt} < \overline{\omega_{jt}}(k_{jt}, l_{jt})$, where ω_{jt-1} enters the function $\Psi(.)$ twice. For the third equality, I use the control function for productivity, and inverting the unknown function $\tilde{g}(.)$ from Equation (I.16) to write $\tilde{d}\big(\underline{\omega_{jt}}(k_{jt}), \overline{\omega_{jt}}(k_{jt}, l_{jt})\big) = \tilde{g}^{-1}(\omega_{jt-1}, Prob_{jt})$. I assume that this inversion exists. A sufficient condition would be that there is indeed a function $\tilde{d}(.)$ in which $\tilde{g}(.)$ is monotonous, despite $\tilde{g}(.)$ being increasing in $\overline{\omega_{jt}}$ and decreasing $\underline{\omega_{jt}}$ individually, so e.g. the gap $\tilde{d}(.) = \overline{\omega_{jt}} - \underline{\omega_{jt}}$. For more discussion on the assumptions on an inversion involving one threshold, see Olley and Pakes (1996). The last equation shows that I address the selection problem in the productivity process by conditioning on the probability or propensity of being in the sample, i.e. between the thresholds.

productivity shock ζ_{jt} is incurred, a common assumption in the literature. Similarly, if labour hiring and firing takes enough adjustment time, ζ_{jt} could also be uncorrelated with current labour inputs. I check my results for either assumption on labour timing. These orthogonality assumptions in the productivity Markov process are crucial for the identification of the production function parameters.

Estimated equations and moments

Following Wooldridge (2009), we can write down two equations for the production function, where we substitute in for ω_{it} from Equation (I.15) and from Equation (I.17):[102]

$$q^r_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + h(k_{jt}, l_{jt}, \mathbf{z}_{jt}, m_{jt}) + \epsilon_{jt} \tag{I.19}$$

$$q^r_{jt} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \\ + \tilde{\Psi}\Bigg(h(k_{jt-1}, l_{jt-1}, \mathbf{z}_{jt-1}, m_{jt-1}), IMP_{jt-1}, Prob_{jt}\Bigg) + \zeta_{jt} + \epsilon_{jt} \tag{I.20}$$

For estimation we need to specify the unknown functions $h(.)$ and $\tilde{\Psi}(.)$. For $h(.)$ I use a third order polynomial with all interactions in its arguments.[103] For the Markov process in productivity I use an AR(1) process, so $\tilde{\Psi}(.)$ becomes a linear function.[104]

For consistent estimates, we need to specify the instrument matrix for each of the two equations, which requires assumptions on timing. For the first Equation (I.19), the shock to production ϵ_{jt} is unexpected and incurred during or after production and therefore not linked to current (or past) firm input choices. We can use the full set of current and past variables as instruments for themselves, which I denote as the information set $\mathbf{\Gamma}_{jt}$.

However, for the second Equation (I.20), the joint error term contains ζ_{jt}, which is part of ω_{jt}, which the firm is assumed to know before the beginning of production. So clearly, this is correlated with current input choices. Since ζ_{jt} is the non-anticipated innovation in the Markov productivity process, it is not correlated with past input choices, however. It depends on the assumption of the flexibility of inputs, which current input choices are problematic from an econometric point of view. If we believe that current capital is set in the last period (last period investment, depreciation and capital stock), then current period capital is not correlated with ζ_{it}. I follow the literature in assuming this. For labour,

[102]See Appendix I.C.3 for the translog version.

[103]As in De Loecker et al. (2016) for example. I also check the results with higher-order polynomials. One can alternatively use non-parametric methods as in Levinsohn and Petrin (2003), at the expense of a much more complicated estimation procedure.

[104]Similar to the assumption in Forlani et al. (2016). I also check the results' robustness with higher order polynomials.

it depends how flexibly hiring and firing takes place. Most likely it is partially dynamic, so there is not complete digression on the size of the labour force each period (multi-year contracts). I allow current labour choices to be correlated with current productivity, but also check robustness with a version where it is fully dynamic, i.e. determined in the previous period. I denote the set of instruments for the second Equation (I.20) as the information set $\boldsymbol{\Gamma}_{jt-1}$, where k_{jt} is contained since it is determined in the previous period, and depending on the labour assumption, l_{jt} is contained or only l_{jt-1}.

By rearranging the equations we can formulate the set of population moment equations, where the errors are a function of all parameters $\boldsymbol{\Theta}$:

$$E\begin{pmatrix} \epsilon_{jt}(\boldsymbol{\Theta}) & | & \boldsymbol{\Gamma}_{jt} \\ \big(\epsilon_{jt}+\zeta_{jt}\big)(\boldsymbol{\Theta}) & | & \boldsymbol{\Gamma}_{jt-1} \end{pmatrix} = 0$$

We can form the analogous stacked sample moments and write the criterion function $\tilde{\boldsymbol{Q}}(\boldsymbol{\Theta})$ to be minimised:

$$\text{Define: } \boldsymbol{r}_{jt}(\boldsymbol{\Theta}) \equiv \begin{pmatrix} \epsilon_{jt}(\boldsymbol{\Theta}) \\ \big(\epsilon_{jt}+\zeta_{jt}\big)(\boldsymbol{\Theta}) \end{pmatrix} \text{ and } \tilde{\boldsymbol{\Gamma}}_{jt} \equiv \begin{pmatrix} \boldsymbol{\Gamma}_{jt} & 0 \\ 0 & \boldsymbol{\Gamma}_{jt-1} \end{pmatrix}$$

$$\text{Set of sample moment conditions:} \quad \frac{1}{JT}\sum_j\sum_t[\tilde{\boldsymbol{\Gamma}}'_{jt}\hat{\boldsymbol{r}}_{jt}(\boldsymbol{\Theta})] = \boldsymbol{0}$$

$$\hat{\boldsymbol{\Theta}} = \min_{\boldsymbol{\Theta}} \tilde{\boldsymbol{Q}}(\boldsymbol{\Theta}) = \frac{1}{JT}\left[\sum_j\sum_t[\tilde{\boldsymbol{\Gamma}}'_{jt}\hat{\boldsymbol{r}}_{jt}(\boldsymbol{\Theta})]'\boldsymbol{W}\sum_j\sum_t[\tilde{\boldsymbol{\Gamma}}'_{jt}\hat{\boldsymbol{r}}_{jt}(\boldsymbol{\Theta})]\right]$$

where the weighting matrix $\boldsymbol{W}$ is clustered on plants accounting for non-identically distributed and autocorrelated errors.

Production elasticities

Having estimated the vector of parameters $\hat{\boldsymbol{\Theta}}$, I simply use the residual of the first stage Equation (I.19) to get the estimate $\hat{\epsilon}_{it}$. I recover productivities $\hat{\omega}_{jt}$ by subtracting the production function with plugged in estimates $\hat{\boldsymbol{\Theta}}$ from the predicted values $\hat{q}_{jt}$ in the first stage equation. The estimate for the production elasticity of inputs α_{jt} is simply the corresponding coefficient, for example for materials:[105]

$$\hat{\alpha}^M_{jt} = \hat{\beta}_m$$

[105] For the translog it varies by plant: $\hat{\alpha}^M_{jt} = \hat{\beta}_m + \hat{\beta}_{lm}l_{jt} + \hat{\beta}_{km}k_{jt} + \hat{\beta}_{mm}m_{ijt}$, see Appendix I.C.3.

I.C.2 An alternative for production function estimation: dynamic panel system GMM

As an alternative to the proxy approach, I also implement a quasi-differenced dynamic panel system GMM approach (Blundell and Bond, 2000). This serves as a robustness check, as the material demand equation depends on $\boldsymbol{\tau}$, which we do not observe. Everything else equal, firms with higher $\boldsymbol{\tau}$ demand less materials. I include the factors that drive $\boldsymbol{\tau}$ in $\boldsymbol{z}_{jt}$ (such as input prices or demand elasticity), but the system GMM approach serves as a test whether this is enough for the scalar unobservable and invertibility condition required for the proxy approach. Shenoy (2015) uses a dynamic panel method in his analysis of input misallocation for Thai rice farmers due to a similar concern. He also develops a test for the scalar unobservable assumption, and argues that with input constraints, the dynamic panel approach tends to perform better in his setting (Shenoy, 2016).

I maintain the first order Markov assumption for the productivity process, which I further specify into an AR(1) process. But I allow for a firm specific time-invariant component of productivity ν_j:

$$\omega_{jt} = \rho\omega_{jt-1} + \nu_j + \zeta_{jt}$$

Analogous to Equation I.17, we could condition the productivity process on sample selection and import status, or R&D expenditures. Quasi-first differencing the log production function (I.12) by ρ eliminates unobserved ω_{jt}:

$$q^r_{jt} = -\rho q^r_{jt-1} + f(k_{jt}, l_{jt}, m_{jt}) - \rho f(k_{jt-1}, l_{jt-1}, m_{jt-1}) + \nu_j + \zeta_{jt} + \epsilon_{jt} - \rho\epsilon_{jt-1}$$

By first differencing this equation, I can implement an Arellano and Bond (1991) estimator by using lagged independent variables as instrument. I use a more efficient Blundell and Bond (2000) system GMM estimator, where I use lagged differenced independent variables as instruments for the level equation in addition.[106] This estimator is implemented for both the Cobb-Douglas and the translog versions.

[106]The appropriate instruments need to take into account that q^r_{jt-1} is correlated with ϵ_{jt-1}, and l_{jt-1} and m_{jt-1} and k_{jt} with ζ_{jt-1}.

I.C.3 Estimation with translog production function

I also provide the production function estimates for a more flexible translog specification. To repeat, the realised, observed output in logs by firm is:

$$q_{jt}^{r} = f(k_{jt}, l_{jt}, m_{jt}) + \omega_{jt} + \epsilon_{jt} \tag{I.21}$$

Since we don't know the functional form of the production function we can form a second order Taylor approximation[107] with approximation error ν_{jt} around the point $\mathbf{X} = \mathbf{1}$ (so $\mathbf{x} = \mathbf{0}$):

$$\begin{aligned} q_{jt}^{r} = f(\mathbf{0}) &+ \frac{\partial f}{\partial k_{jt}}_{|\mathbf{x}=\mathbf{0}} k_{jt} + \frac{\partial f}{\partial l_{jt}}_{|\mathbf{x}=\mathbf{0}} l_{jt} + \frac{\partial f}{\partial m_{jt}}_{|\mathbf{x}=\mathbf{0}} m_{jt} \\ &+ \frac{1}{2}\frac{\partial f^2}{(\partial k_{jt})^2}_{|\mathbf{x}=\mathbf{0}} k_{jt}^2 + \frac{1}{2}\frac{\partial f^2}{(\partial l_{jt})^2}_{|\mathbf{x}=\mathbf{0}} l_{jt}^2 + \frac{1}{2}\frac{\partial f^2}{(\partial m_{jt})^2}_{|\mathbf{x}=\mathbf{0}} m_{jt}^2 \\ &+ \frac{\partial f^2}{\partial k_{jt}\partial l_{jt}}_{|\mathbf{x}=\mathbf{0}} k_{jt}l_{jt} + \frac{\partial f^2}{\partial k_{jt}\partial m_{jt}}_{|\mathbf{x}=\mathbf{0}} k_{jt}m_{jt} + \frac{\partial f^2}{\partial l_{jt}\partial m_{jt}}_{|\mathbf{x}=\mathbf{0}} l_{jt}m_{jt} \\ &+ \omega_{jt} + \epsilon_{jt} + \nu_{jt} \end{aligned}$$

Since the derivatives evaluated at $\mathbf{x} = \mathbf{0}$ are constant across time and firms within the same product category, we can interpret them as the regression coefficients,[108] which yields a second order translog[109] production function:

$$q_{jt}^{r} = \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \beta_{kk} k_{jt}^2 + \beta_{ll} l_{jt}^2 + \beta_{mm} m_{jt}^2 \tag{I.22}$$

$$+ \beta_{kl} k_{jt} l_{jt} + \beta_{km} k_{ijt} m_{jt} + \beta_{lm} l_{jt} m_{jt} + \omega_{jt} + \epsilon_{jt} \tag{I.23}$$

In this equation, the additional error term ν_{ijt} which comes from approximation is assumed to be zero and omitted. We cannot identify this term (which can in principle be "transmitted" to inputs) separately to the shock ϵ_{jt}, so assume it is zero and the approximation is perfect. This is of course a silent assumption in all of the production estimation literature. Compared to Cobb-Douglas or CES functional form assumptions, the translog specification is more flexible allowing for significant amount of curvature, and is thus less likely to suffer from functional form assumption or approximation bias.

[107] And Young's theorem of equal cross-partials.

[108] The factor of a half is incorporated in the coefficient for the quadratic terms.

[109] The transcendental logarithmic function was introduced by Christensen et al. (1971, 1973); Berndt and Christensen (1973). See also Griliches and Ringstad (1971) who propose a similar generalisation of the approximation for estimating CES functions by Kmenta (1967).

Analogous to the main text, we can write down two equations for the production function, where I substitute in for ω_{it} from Equation I.15 and from Equation I.17:

$$\begin{aligned} q^r_{jt} &= \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \beta_{kk} k^2_{jt} + \beta_{ll} l^2_{jt} + \beta_{mm} m^2_{jt} \\ &\quad + \beta_{kl} k_{jt} l_{jt} + \beta_{km} k_{jt} m_{jt} + \beta_{lm} l_{jt} m_{jt} + h(k_{jt}, l_{jt}, \mathbf{z}_{jt}, m_{jt}) + \epsilon_{jt} \qquad \text{(I.24)} \\ q^r_{jt} &= \beta_0 + \beta_k k_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \beta_{kk} k^2_{jt} + \beta_{ll} l^2_{jt} + \beta_{mm} m^2_{jt} \\ &\quad + \beta_{kl} k_{jt} l_{jt} + \beta_{km} k_{jt} m_{jt} + \beta_{lm} l_{jt} m_{jt} \\ &\quad + \tilde{\Psi}\Big(h(k_{jt-1}, l_{jt-1}, \mathbf{z}_{jt-1}, m_{jt-1}), IMP_{jt-1}, Prob_{jt} \Big) + \zeta_{jt} + \epsilon_{jt} \qquad \text{(I.25)} \end{aligned}$$

The estimate for the production elasticity of material inputs is, for example:

$$\hat{\alpha}^M_{jt} = \hat{\beta}_m + \hat{\beta}_{lm} l_{jt} + \hat{\beta}_{km} k_{jt} + \hat{\beta}_{mm} m_{ijt}$$

For the Cobb-Douglas version, we apply the parameter restrictions that $\beta_{lm} = \beta_{km} = \beta_{mm} = 0$, so that $\hat{\alpha}^M_{jt} = \hat{\beta}_m$.

I.D Estimation details for the demand side

I.D.1 Estimation of the demand model

To repeat for convenience, consumers are indexed by i and need to decide to buy from a firm j to maximise their utility from using product j:

$$U_{ijt} = (y_{it} - P^r_{jt})\theta^p_{it} + x_{jt}\theta^x_{it} + \xi_j + \xi_t + \Delta\xi_{jt} + \mu_{ijt} \equiv V_{ijt} + \mu_{ijt} \qquad \text{(I.26)}$$

where y_{it} is consumer income, P^r_{jt} realised prices (which are associated with realised quantities – these are the ones that are relevant for the consumers), x_{jt} a vector of product characteristics and a constant, ξ_j average utility from unobserved time-constant product characteristics, ξ_t average unobserved market-specific utility, and $\Delta\xi_{jt}$ the unobserved deviations from a particular product in a particular market from the unobserved averages. The unobserved ξ_j can contain the quality and the location of a product and ξ_j and ξ_t will be absorbed by fixed effects dummies. For the baseline results I only include a constant in x_{jt} as there are few time variant product characteristics (since the time invariant characteristics are absorbed in ξ_j). The non-random utility can be summarised by V_{ijt}. The random utility component is μ_{ijt}, which follows an i.i.d. Type I extreme value distribution.

We can further specify the random parameters into a mean and variance component:

$$\begin{pmatrix} \theta^p_{it} \\ \theta^x_{it} \end{pmatrix} = \begin{pmatrix} \theta^p \\ \theta^x \end{pmatrix} + \begin{pmatrix} \sigma^p & 0 \\ 0 & \sigma^x \end{pmatrix} \begin{pmatrix} \nu^p_{it} \\ \nu^x_{it} \end{pmatrix}, \quad \boldsymbol{\nu}_{it} \sim P(\boldsymbol{\nu})$$

where ν_{it} are draws from a multivariate normal distribution. Therefore the consumer heterogeneity has three dimensions, the random utility shock μ_{ijt} as well as the two ν_{it} draws. I estimate the means $\begin{pmatrix} \theta^p \\ \theta^x \end{pmatrix}$ and the variances $\Sigma \equiv \begin{pmatrix} \sigma^p & 0 \\ 0 & \sigma^x \end{pmatrix}$ of the random coefficients. We can rewrite the utility function with a mean ($\equiv \delta_{jt}$) and individual consumer part, which simplifies the estimation algorithm:

$$U_{ijt} = \theta^p_{it} y_{it} \underbrace{-\theta^p P^r_{jt} + \theta^x x_{jt} + \xi_j + \xi_t + \Delta\xi_{jt}}_{\equiv \delta_{jt}} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt} + \mu_{ijt}$$

Derived theoretical market shares

Consumer i purchases from firm j if it yields the highest utility, compared to the products from all other firms or the outside option $j = 0$. The outside good also serves to normalise the utility by setting the mean and individual components in the outside good utility to zero.[110] Define as set A_{jt} the set of consumers which strictly prefer product j.[111] The integral over the consumers that belong to this set is the theoretical (realised) market share s^r_{jt} of firm j in period t:

$$\begin{aligned} s^r_{jt} &= \int_{A_{jt}} dP(\boldsymbol{\nu}, \boldsymbol{\mu}) = \int_{A_{jt}} dP(\boldsymbol{\mu} \mid \boldsymbol{\nu}) dP(\boldsymbol{\nu}) = \int_{A_{jt}} dP(\boldsymbol{\mu}) dP(\boldsymbol{\nu}) \\ &= \int_{A_{jt}} \frac{\exp(\theta^p_{it} y_{it}) \exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})}{\exp(\theta^p_{it} y_{it}) \left[\exp(\delta_{0t} - \sigma^p \nu^p_{it} P^r_{0t} + \sigma^x \nu^x_{it} x_{0t}) + \sum_{j=1} \exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})\right]} dP(\boldsymbol{\nu}) \\ &= \int_{A_{jt}} \frac{\exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})}{1 + \sum_{j=1} \exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})} dP(\boldsymbol{\nu}) \end{aligned} \tag{I.27}$$

where the third equality in the first row follows from assuming that the random coefficient and the random utility shocks are independent. The fourth equality uses the Type I extreme value distributional assumption about the random utility shocks, and the fifth equality uses that I normalise the components of the utility of the outside good ($j = 0$) to zero.

[110] So everything except the terms $\theta^p_{it} y_{it}$ and μ_{i0t}, see e.g. Nevo (2000*b*) for more details.

[111] So $A_{jt} = \{(\boldsymbol{\nu}_{it}, \boldsymbol{\mu}_{it}) \mid U_{ijt} > U_{ilt} \forall l = 0, 1..., J\}$

Minimising the distance between theoretical and observed market shares

The theoretically predicted market shares can be used to find parameter values that match them to empirically observed realised market shares $\hat{s^r}_{jt}$. The problem is that the parameters enter in a nonlinear fashion into the market shares, which is a difficult minimization problem and more importantly, cannot address price endogeneity concerns in the usual linear way. The main contribution of BLP and Berry (1994) is to show how we can estimate the parameters while taking price endogeneity into account in a linear fashion.[112] For logit models without random coefficients, this is easily achieved by an analytic relationship between market share ratios and mean utility δ_{jt} in (I.27). Berry (1994) and BLP solve the integral in (I.27) by simulating consumers and using a contraction mapping. The following sketches the procedure and algorithm. For a more detailed account, see BLP, Berry (1994) or Nevo (2000*b*).

The algorithm operates on an inner and an outer loop. The inner loop first solves for δ_{jt}, and then linearly estimates the mean coefficients (θ^p, θ^x) and $\boldsymbol{\xi}$. The outer loop solves for Σ. The inner loop finds a δ_{jt} for a given Σ that sets the observed $(\hat{s^r}_{jt})$ and the theoretical market shares (s^r_{jt}) equal: $\min_{\delta_{jt}} ||s^r_{jt} - \hat{s^r}_{jt}||$. The theoretical market shares are calculated via numerical integration by simulation by drawing a number of consumers N (a consumer is defined by $\boldsymbol{\nu}_{it}$ after integrating the random utility component μ_{ijt} out):

$$s^r_{jt} \approx \frac{1}{N} \sum_i \frac{\exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})}{1 + \sum_{j=1} \exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})} \tag{I.28}$$

For the baseline I simulate $N = 2000$ consumers, but also check robustness of the point estimates with $N = 10000$ consumers. Berry (1994) proves that there exists a unique δ_{jt} which matches the theoretical and empirical market shares under mild regularity conditions. Based on this, BLP employ a contraction mapping (nested fixed point algorithm) where for each step h, the new δ^{h+1}_{jt} is found conditional on Σ by:

$$\delta^{h+1}_{jt} = \delta^h_{jt} + ln(\hat{s^r}_{jt}) - ln(s^r_{jt})$$

[112]Estimating demand systems has been the focus of a large literature over decades. While allowing flexible substitution patterns between products, the drawbacks of the popular classic Almost Ideal Demand System (Deaton and Muellbauer, 1980) are the number of parameters required to be estimated and the requirement of a representative consumer. Therefore, parts of the relevant IO literature have moved from a product space approach towards a characteristic space approach, which I employ in this chapter as well. The BLP model has been further extended (e.g. Nevo (2001)) and used in a variety of contexts, often for merger analysis (e.g. Nevo (2000*a*)), but also for welfare consequence due to e.g. trade policy changes (Berry et al., 1999) or the introduction of a particular product (minivan) (Petrin, 2002).

which is iterated until the change in δ_{jt}, so $ln(\hat{s^r}_{jt}) - ln(s^r_{jt})$, is below a tolerance level.[113]

Identifying the linear preference parameters

Thereafter, I obtain the linear parameters (θ^p, θ^x) through a linear IV regression of δ_{jt} from its definition on:

$$\delta_{jt} = -\theta^p P^r_{jt} + \theta^x x_{jt} + \xi_j + \xi_t + \Delta\xi_{jt} \tag{I.29}$$

where I instrument the endogenous price P^r_{jt} with plant cost shifters and use the appropriate product and time dummies. The price endogeneity arises from correlation with the unobserved taste shocks $\Delta\xi_{jt}$, which might allow changing prices without consequences for quantities sold, for example.

Moment conditions in outer loop

From the IV regression I also calculate $\Delta\xi_{jt}$ with which I form the objective GMM function to be minimised to obtain a solution for Σ:

$$\hat{\Sigma} = \arg\min_{\Sigma} \boldsymbol{\Delta\xi}' Z W Z' \boldsymbol{\Delta\xi}$$

where Z is the instrument matrix and $W = (Z'Z)^{-1}$ is a weighting matrix. The underlying moment conditions are that the unobserved deviation in mean utility $\boldsymbol{\Delta\xi}$ are orthogonal to the instrument matrix. I will further discuss the choice of the instrument matrix when I present the results in Section I.4.1. The outer loop searches over the parameter space of the nonlinear parameters in Σ, and for each iteration, the inner loop and linear IV regression are performed. This procedure solves for all structural demand side parameters. The estimation performs better with analytical Jacobians, which are provided, along with further estimation details and the analytic robust standard errors of the estimates in Appendix I.D.2.

[113] I use a tolerance level of on average 10^{-13} with a maximum tolerance of 10^{-12} for an individual jt. Davis and Schiraldi (2014) provide a faster convergence to the unique vector of fixed points via a Newton-Rhapson algorithm. An alternative to the inner loop contraction mapping is to use a MPEC (mathematical program with equilibrium constraints) algorithm that takes the market shares as constraints to the GMM objective function, see Dubé et al. (2012).

Outside good and observed market shares

The estimation uses data on market shares (of sold quantities), rather than quantities themselves. Since the size of the market also includes the outside good, we need to quantify the outside good. BLP, which analyse the car market, for example, take as total market the population that *can* buy a vehicle. Here, I take as market size Y_t the total amount of the particular product sold by Indian firms, both by the firms in the sample and outside the sample by using the plant specific sampling multiplier in the data. Therefore the plant level quantity sold is:

$$Q^r_{jt} = \hat{s}^r_{jt} Y_t \tag{I.30}$$

and $\sum_{j\geq 1} Q^r_{jt} < 1$ due to the outside good. An increase in the production of an in-sample firm would therefore not increase Y_t, but $\sum_j Q^r_{jt}$.

Price elasticities of demand

The price elasticity of demand is:

$$\frac{1}{\eta_{jt}} \equiv \frac{\partial Q_{jt}}{\partial P_{jt}} \frac{P_{jt}}{Q_{jt}} = \frac{\partial (s_{jt} Y_t)}{\partial P_{jt}} \frac{P_{jt}}{s_{jt} Y_t} = \frac{\partial s_{jt}}{\partial P_{jt}} \frac{P_{jt}}{s_{jt}} = \frac{P_{jt}}{s_{jt}} \frac{1}{N} \sum_i (\theta^p_{it} s_{ijt}(1 - s_{ijt})) \tag{I.31}$$

where $s_{ijt} \equiv \frac{\exp(\delta_{jt} - \sigma^p \nu^p_{it} P_{jt} + \sigma^x \nu^x_{it} x_{jt})}{1 + \sum_{j=1} \exp(\delta_{jt} - \sigma^p \nu^p_{it} P_{jt} + \sigma^x \nu^x_{it} x_{jt})}$. I omit the notation with r for realised output (or market share) here, since the elasticities can be derived from any prices and quantities in any equilibrium conditional on the estimated parameters. Cross-elasticities can be calculated similarly and vary by firm-pair in each market.

I.D.2 Demand side algorithm details, Jacobian and standard errors

For the outer loop that searches over Σ, both an interior-point (see e.g. Byrd et al. (2000)) as well as sequential quadratic programming (SQP) algorithm perform well (see e.g. Nocedal and Wright (2006)). As starting values for Σ, I use a vector of zeros, but checked robustness with various positive and negative starting values. I supply a starting value for δ_{jt} from a model without random coefficients where it has an analytical solution: $\delta^{start}_{jt} = \log(\hat{s^r}_{jt}) - \log((\text{outside market share})_t)$.

The Jacobian of the objective function can be solved for analytically and can be supplied to the optimisation algorithm for speed improvements. The derivative of the objective function (let us call it $f()$ for simplicity) with respect to the two elements of Σ on the diagonal ($\boldsymbol{\sigma}$) is, with a slight abuse of notation:

$$\frac{\partial f(\boldsymbol{\Delta\xi}(\boldsymbol{\delta}(\boldsymbol{s^r}(\boldsymbol{\sigma}))))}{\partial \boldsymbol{\sigma}} = \frac{\partial f(.)}{\partial \boldsymbol{\Delta\xi}} \frac{\partial \boldsymbol{\Delta\xi}}{\partial \boldsymbol{\delta}} \frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{s^r}} \frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\sigma}} = 2ZWZ'\boldsymbol{\Delta\xi} \left(\frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\delta}}\right)^{-1} \frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\sigma}}$$

The second last component of the Jacobian $\left(\frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\delta}}\right)^{-1}$ is a square matrix with a size equal to the number of observations (so products times markets or periods). The elements of the matrix can be calculated by taking the derivative of Equation (I.28), which I repeat for convenience:

$$s^r_{jt} \approx \frac{1}{N} \sum_i \frac{\exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})}{1 + \sum_{j=1} \exp(\delta_{jt} - \sigma^p \nu^p_{it} P^r_{jt} + \sigma^x \nu^x_{it} x_{jt})} \equiv \frac{1}{N} \sum_i s^r_{ijt}$$

$$\begin{aligned} \frac{\partial s^r_{jt}}{\partial \delta_{jt}} &= \frac{1}{N} \sum_i s^r_{ijt}(1 - s^r_{ijt}) \quad \forall j \\ \frac{\partial s^r_{jt}}{\partial \delta_{mt}} &= -\frac{1}{N} \sum_i s^r_{ijt} s^r_{imt} \quad \forall m \neq j \end{aligned}$$

The last component of the Jacobian $\frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\sigma}}$ is again obtained by taking the derivative of the market share equation with respect to each of the k diagonal elements σ^k of Σ (with associated draw ν^k_{it}), so:

$$\begin{aligned} \frac{\partial s^r_{jt}}{\partial \sigma^p} &= \frac{1}{N} \sum_i \nu^p_{it} s^r_{ijt} (\sum_j P^r_{jt} s^r_{ijt} - P^r_{jt}) \\ \frac{\partial s^r_{jt}}{\partial \sigma^x} &= \frac{1}{N} \sum_i \nu^x_{it} s^r_{ijt} (x_{jt} - \sum_j x_{jt} s^r_{ijt}) \end{aligned}$$

The standard errors of $\hat{\Sigma}$ are obtained by taking the square root of the diagonal terms of its covariance matrix. The covariance matrix of the GMM estimate $\hat{\boldsymbol{\sigma}}$ is:

$$V\hat{COV}(\hat{\boldsymbol{\sigma}}) = N\,(G'ZWZ'G)^{-1} \left(G'ZW\hat{V}WZ'G\right) (G'ZWZ'G)^{-1}$$

where G is the gradient of the moment conditions, for which we can use part of the Jacobian of the objective function above:

$$G \equiv \left(\frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\delta}}\right)^{-1} \frac{\partial \boldsymbol{s^r}}{\partial \boldsymbol{\sigma}}$$

and W is the 2SLS weighting matrix and $\hat{V}$ a consistent heteroskedasticity robust estimator of the moment conditions:

$$W = (Z'Z)^{-1}$$
$$\hat{V} = \frac{1}{N}\sum_{jt}\Delta\xi_{jt}z'_{jt}z_{jt}$$

The linear structural parameters $\{\theta^p, \theta^x\}$ depend on the non-linear structural parameters Σ and are solved for via the inner loop in the algorithm. In order to obtain a covariance matrix of the linear parameters I bootstrap from the estimated $V\hat{COV}(\hat{\boldsymbol{\sigma}})$ and solve the inner loop for each draw with an associated $\{\theta^p, \theta^x\}$. I recover the standard errors of the linear parameters from the resulting sampling distribution of $\{\theta^p, \theta^x\}$.

I.E Details for estimating equilibria

Both the factual and the counterfactual equilibria are determined by a vector of prices, since the market shares (and output quantities) are a function of prices and the structural parameters. The input quantities for the counterfactual can be derived from the cost minimisation conditions once the equilibrium prices (and quantities) are found. Note that the contraction mapping in the inner loop was only needed to identify the structural demand parameters. For the equilibria I only search over prices taking the structural demand and production parameters as given. The strategy is to (1) first find the factual equilibrium and associated equilibrium prices (i.e. before shock ϵ introduces noise into realised observed quantities and prices), then (2) calculate the implied factual $\boldsymbol{\tau}$, and (3) use a counterfactual $\tilde{\boldsymbol{\tau}}$ to obtain prices that solve for the counterfactual equilibrium.

The equilibrium conditions are:

$$\frac{P_{jt}}{MC_{jt}\big(Q_{jt}(\mathbf{P_t}), \mathbf{c}_{jt}(\boldsymbol{\tau})\big)} - \frac{1}{1+\eta_{jt}(\mathbf{P_t})} = 0 \qquad \text{(FOC)}$$

$$2\frac{\partial Q_{jt}}{\partial P_{jt}} + (P_{jt} - MC_{jt})\frac{\partial^2 Q_{jt}}{(\partial P_{jt})^2} - \frac{\partial MC_{jt}}{\partial Q_{jt}}\left(\frac{\partial Q_{jt}}{\partial P_{jt}}\right)^2 \leq 0 \qquad \text{(SOC)}$$

For both, the factual and counterfactual equilibria, I use the Hessian (SOC) as a constraint in the optimisation to ensure profit maximisation (and not minimisation). We can rewrite the Hessian with markets shares instead of quantities using Equation (I.30): $Q_{jt} = s_{jt}Y_t$. Note that I do not use superscript r in this section since I am now using equilibrium, not realised, quantities and prices. The Hessian is:

$$H_{jt} = Y_t \left(2\frac{\partial s_{jt}}{\partial P_{jt}} + (P_{jt} - MC_{jt})\frac{\partial^2 s_{jt}}{(\partial P_{jt})^2} - Y_t \frac{\partial MC_{jt}}{\partial s_{jt}} \left(\frac{\partial s_{jt}}{\partial P_{jt}} \right)^2 \right)$$

where

$$\begin{aligned}
\frac{\partial s_{jt}}{\partial P_{jt}} &= \frac{1}{N} \sum_i \left(\theta^p_{it} s_{ijt}(1 - s_{ijt})\right) \\
\frac{\partial^2 s_{jt}}{\partial (P_{jt})^2} &= \frac{1}{N} \sum_i \left((\theta^p_{it})^2 s_{ijt}(1 - s_{ijt})(1 - s_{ijt} - s_{ijt})\right) \\
\text{where } s_{ijt} &\equiv \frac{\exp(\delta_{jt} - \sigma^p \nu^p_{it} P_{jt} + \sigma^x \nu^x_{it} x_{jt})}{1 + \sum_{j=1} \exp(\delta_{jt} - \sigma^p \nu^p_{it} P_{jt} + \sigma^x \nu^x_{it} x_{jt})} \\
MC_{jt} &= \frac{P_{jt}}{1 + \eta_{jt}(\mathbf{P_t})}
\end{aligned}$$

When using a Cobb-Douglas production function, the marginal cost function[114] has an analytical closed form and is a function of output Q_{jt}, output elasticities and distortions

[114]The cost function is:

$$\begin{aligned}
C_{jt} = \left(\frac{Q_{jt}}{\Omega_{jt}}\right)^{\frac{1}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} &\Bigg[\left(\frac{r_{jt}\alpha^M_{jt}}{\alpha^K_{jt}}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{w_{jt}\tau^L_{jt}\alpha^M_{jt}}{\alpha^L_{jt}}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(P^M_{jt}\tau^M_{jt}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \\
&+ \left(\frac{r_{jt}\alpha^L_{jt}}{\alpha^K_{jt}}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{\alpha^L_{jt}P^M_{jt}\tau^M_{jt}}{\alpha^M_{jt}}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(w_{jt}\tau^L_{jt}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \\
&+ \left(\frac{\alpha^K_{jt}w_{jt}\tau^L_{jt}}{\alpha^L_{jt}}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{\alpha^K_{jt}P^M_{jt}\tau^M_{jt}}{\alpha^M_{jt}}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(r_{jt}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \Bigg]
\end{aligned}$$

τ^L_{jt}, τ^M_{jt}:

$$MC_{jt} = \frac{(s_{jt}Y_t)^{-1}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}} \left(\frac{s_{jt}Y_t}{\Omega_{jt}}\right)^{\frac{1}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \Bigg[\left(\frac{r_{jt}\alpha^M_{jt}}{\alpha^K_{jt}}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{w_{jt}\tau^L_{jt}\alpha^M_{jt}}{\alpha^L_{jt}}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} (P^M_{jt}\tau^M_{jt})^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}}$$
$$+ \left(\frac{r_{jt}\alpha^L_{jt}}{\alpha^K_{jt}}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{\alpha^L_{jt}P^M_{jt}\tau^M_{jt}}{\alpha^M_{jt}}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} (w_{jt}\tau^L_{jt})^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}}$$
$$+ \left(\frac{\alpha^K_{jt}w_{jt}\tau^L_{jt}}{\alpha^L_{jt}}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{\alpha^K_{jt}P^M_{jt}\tau^M_{jt}}{\alpha^M_{jt}}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} (r_{jt})^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \Bigg]$$

Hence the remaining component of the Hessian is:

$$\frac{\partial MC_{jt}}{\partial s_{jt}} = \left(\frac{1-(\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt})}{(\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt})^2}\right) (s_{jt}Y_t)^{\frac{1}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}-2} \Omega_{jt}^{-\frac{1}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}}$$
$$\Bigg[\left(\frac{r_{jt}\alpha^M_{jt}}{\alpha^K_{jt}}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{w_{jt}\tau^L_{jt}\alpha^M_{jt}}{\alpha^L_{jt}}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} (P^M_{jt}\tau^M_{jt})^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}}$$
$$+ \left(\frac{r_{jt}\alpha^L_{jt}}{\alpha^K_{jt}}\right)^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{\alpha^L_{jt}P^M_{jt}\tau^M_{jt}}{\alpha^M_{jt}}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} (w_{jt}\tau^L_{jt})^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}}$$
$$+ \left(\frac{\alpha^K_{jt}w_{jt}\tau^L_{jt}}{\alpha^L_{jt}}\right)^{\frac{\alpha^L_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \left(\frac{\alpha^K_{jt}P^M_{jt}\tau^M_{jt}}{\alpha^M_{jt}}\right)^{\frac{\alpha^M_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} (r_{jt})^{\frac{\alpha^K_{jt}}{\alpha^K_{jt}+\alpha^L_{jt}+\alpha^M_{jt}}} \Bigg]$$

For the factual equilibrium, I search over prices that minimise the summed squared distances between the markets shares that are a function of these prices $\boldsymbol{s}$ and the equilibrium market shares $\hat{\boldsymbol{s}}$ that we know from Equation (I.30), $\hat{s}_{jt} = \frac{\hat{s^r}_{jt}}{exp(\epsilon_{jt})}$:

$$\boldsymbol{P_t}^{fact} = \arg\min_{\boldsymbol{P_t}} -\hat{\boldsymbol{s}}'\boldsymbol{s}(\boldsymbol{P_t})$$
$$s.t.\ \ H_{jt}(\boldsymbol{P_t}) \leq 0 \ \ \forall \ \ jt$$

Using realised prices as starting values for the search typically is most efficient, but I also check the robustness with alternative starting values. I use the first order equilibrium conditions above to infer the $\boldsymbol{\tau}$. Once we set the counterfactual $\tilde{\boldsymbol{\tau}}$, we can search for the counterfactual equilibrium prices. For the counterfactuals, I use the first order equilibrium conditions as the objective function. I minimise the squared distances between the variable marginal costs inferred from the prices and demand elasticity, and the variable marginal costs from the derivative of the cost function:

$$MC_{jt}^{D} = \frac{P_{jt}}{1+\eta_{jt}(\mathbf{P_t})} =$$

$$MC_{jt}^{C} = \frac{1}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}\left(\frac{s_{jt}Y_t}{\Omega_{jt}}\right)^{\frac{1}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}-1}$$

$$\left[\left(\frac{r_{jt}\alpha_{jt}^{M}}{\alpha_{jt}^{K}}\right)^{\frac{\alpha_{jt}^{K}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\left(\frac{w_{jt}\tau_{jt}^{L}\alpha_{jt}^{M}}{\alpha_{jt}^{L}}\right)^{\frac{\alpha_{jt}^{L}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\left(P_{jt}^{M}\tau_{jt}^{M}\right)^{\frac{\alpha_{jt}^{M}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\right.$$

$$+\left(\frac{r_{jt}\alpha_{jt}^{L}}{\alpha_{jt}^{K}}\right)^{\frac{\alpha_{jt}^{K}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\left(\frac{\alpha_{jt}^{L}P_{jt}^{M}\tau_{jt}^{M}}{\alpha_{jt}^{M}}\right)^{\frac{\alpha_{jt}^{M}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\left(w_{jt}\tau_{jt}^{L}\right)^{\frac{\alpha_{jt}^{L}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}$$

$$\left.+\left(\frac{\alpha_{jt}^{K}w_{jt}\tau_{jt}^{L}}{\alpha_{jt}^{L}}\right)^{\frac{\alpha_{jt}^{L}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\left(\frac{\alpha_{jt}^{K}P_{jt}^{M}\tau_{jt}^{M}}{\alpha_{jt}^{M}}\right)^{\frac{\alpha_{jt}^{M}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\left(r_{jt}\right)^{\frac{\alpha_{jt}^{K}}{\alpha_{jt}^{K}+\alpha_{jt}^{L}+\alpha_{jt}^{M}}}\right]$$

Therefore:

$$\boldsymbol{P_t}^{counterfact} = \arg\min_{\boldsymbol{P_t}} -\boldsymbol{MC^C}(\boldsymbol{\tau})'\boldsymbol{MC^D}(\boldsymbol{P_t})$$
$$s.t.\ \ H_{jt}(\boldsymbol{P_t}) \leq 0 \ \ \forall \ \ jt$$

Using the factual equilibrium prices as starting values for the counterfactual prices is typically most efficient, but I checked a range of alternative starting values.

All plant-years where I estimate an elasticity larger than -1 in the original demand estimation are ignored in terms of finding any equilibrium and are ignored for the comparative statics as well.[115] Furthermore I constrain the elasticity to be smaller than -1 in the factual and counterfactual estimation such that the relationship between prices and marginal cost markups is defined. This also improves stability in finding the equilibria since the algorithm does not move over the discontinuity. I also supply a lower bound of zero on prices to the algorithm.

For both, the factual and the counterfactual equilibria, I solve separately for each market (i.e. time period) since they are independent. This reduces the dimension over which prices are searched and speeds up the algorithm considerably. I use a sequential quadratic programming (SQP) algorithm (Nocedal and Wright, 2006) which almost always performs faster than an interior point algorithms for these purposes.

[115] As noted in the main text, there are only 9 observations with a median market share of 0.0004.

Figure I.13: Cast iron producing plants as share of all iron alloy producing plants

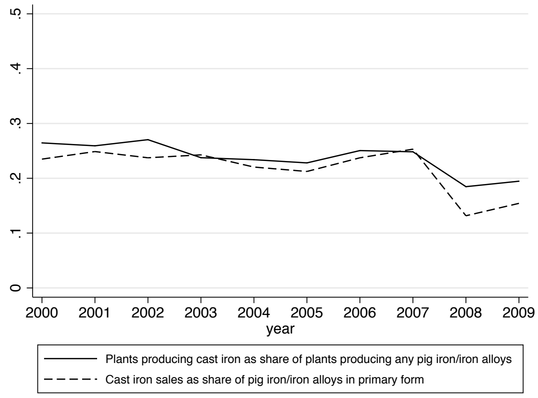
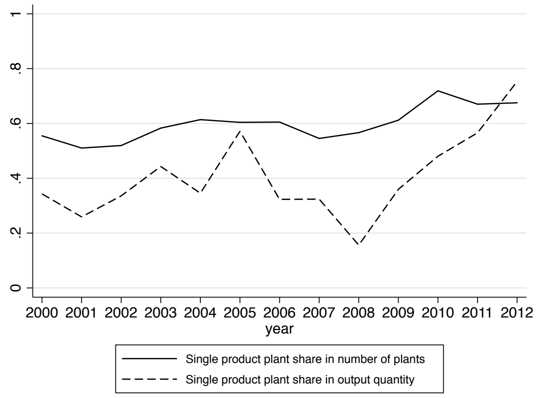


Figure I.14: Share of single product plants in cast iron manufacturing



While neither the existence nor the uniqueness of equilibria is proven analytically, the algorithm always finds an equilibrium, including for the bootstrapped equilibria where I have different simulated structural parameters for each draw. This at least proves existence for this sample. I perform some checks on the global nature (and uniqueness) of the minimum by using a range of different starting values which converge to the same equilibrium.

I.F Details on the Indian iron and steel industry

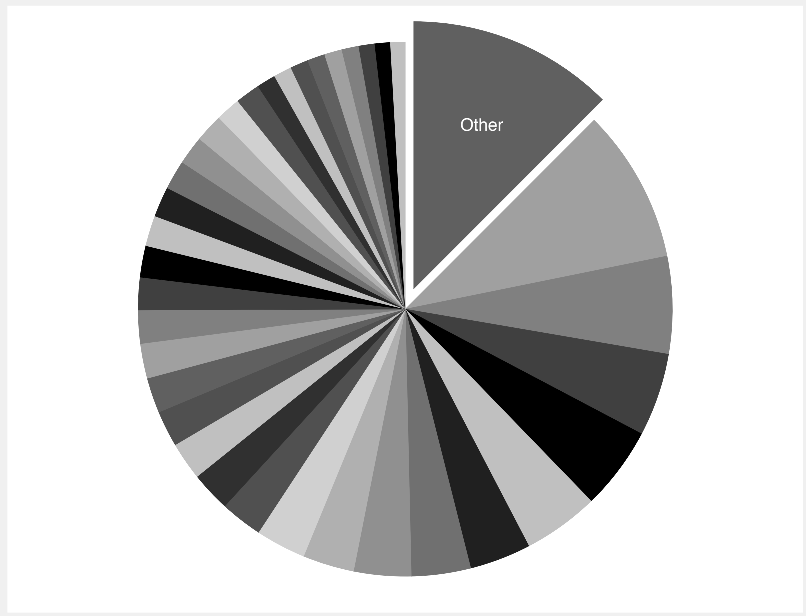
Iron and steel industry in India

The iron and steel sector in India is an interesting sector in itself. In 2015 India has been the third largest producer both of total crude steel and pig iron after China and Japan (WSA, 2016*b*), up from ninth place for both in 2000 (WSA, 2010). The share of basic iron and steel (2710 in ISIC3) in value added of total manufacturing was 15% in 2007 (latest year available in UNIDO (2016*b*) Indstat). After Bahrain, which have a small total manufacturing sector, this is the largest value added share any basic iron and steel sector has in their national manufacturing value added in the world (see Figure I.16). Figure I.13, I.14 and I.15 show more descriptive statistics on the cast iron sample described in the main text.

Raw materials in the iron and steel industry and carbon emissions

In terms of raw materials used in the iron and steel sector, India was the fourth largest producer of iron ore after China, Australia and Brazil (WSA, 2016*b*) in 2015, and the

Figure I.15: Industry concentration: 35 biggest players in 2004

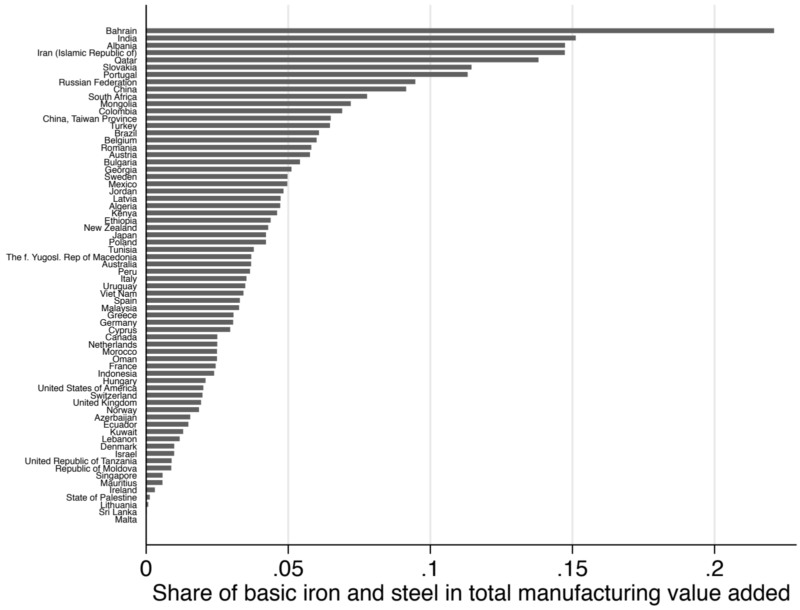


Notes: The market shares of the 35 largest plants and "other" plants are shown in 2004. The calculation is based on the single product firms in the final sample.

third largest coal producer (EIA, 2015). Even after the raw material mining stage, the production chain and process in the iron and steel industry has substantial environmental significance due to its heavy use of coal. The main raw materials iron ore and coal as well as alloying elements such as nickel or chromium are relatively abundant resources. Coal has typically the dual role of providing the heating for smelting and melting in the production chain, but is also directly required to adjust the carbon content of the products. Carbon is often burnt out in the melting process and needs to be re-added accordingly. Carbon emissions come therefore from the heat generation as well as the process of production directly. For the upstream production of pig iron, coke is the main reducing agent to turn iron ore into pig iron saturated with carbon in the smelting process.

Globally, in 2013 around 15% of total coal consumption is accounted for by the iron and steel industry (World Coal Association, 2014), more than twice the entire consumption of the EU (BP, 2016). India accounts for around 9% of global coal consumption, primarily through its electricity generating sector which is heavily reliant on coal (BP, 2016). Therefore even with substitution to a different low-carbon fuel for electricity generation and heating, coal is likely to remain a necessary ingredient in iron and steel manufacturing, with accompanying process emissions. According to the UNFCCC (2016), the iron and steel sector accounts for more than a third of all process emissions in Annex I countries. Per tonne of produced steel in India around 3.1 to 3.8 tonne of $CO2$ are emitted (IPCC, 2007). With India's steel production at 89 million tonnes in 2015 (WSA, 2016*b*), this

Figure I.16: Share of basic iron and steel in total manufacturing value added



Notes: Calculations based on data from UNIDO (2016*b*).

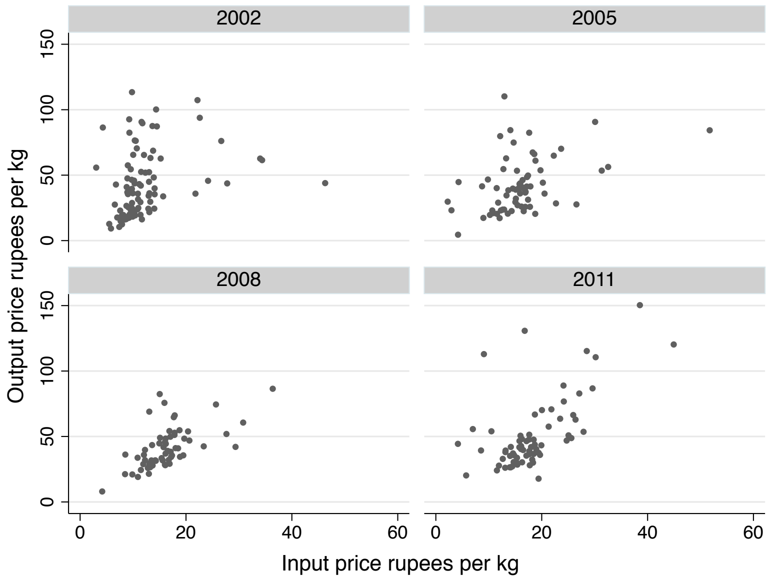
implies around 328 million tonnes of total CO2 emissions of the Indian iron and steel sector, around 82% of UK's or 13% of India's total emissions in 2015 (EC, 2016).

Reduction of process emissions can be achieved through process and product innovation. In terms of process innovation, pulverised coal injection techniques can save around 30% of coal (WSA, 2016*a*), and emissions can be reduced through ex-post carbon capture and storage. In terms of product innovation, it is often cited that 75% of steel types have been introduced in the past 20 years. The Eiffel Tower would only require a third of today's material and old automobiles would only require two thirds of today's steel (WSA, 2017). However, if there are barriers to reallocation from less efficient to more efficient firms, then removing barriers and reallocating inputs could in principle decrease aggregate process emissions, without any plant level process or product innovations, which this chapter can shed light on.

I.G Further results and robustness checks

I.G.1 Output and input prices

Figure I.17: Output and input prices for selected years



Notes: The figure plots deflated output and input prices of the cast iron plants in rupees per kg. Input prices are recovered by dividing total expenditure on material inputs by the total weight of material inputs.

I.G.2 Detailed results from demand estimation

Instrument choice and first stage

In order to consistently estimate the mixed logit demand model we need instruments for the endogenous price. Taste shocks captured in $\Delta\xi_{jt}$, for example, are likely to be positively correlated to prices in the mean utility regression (I.29). If higher utility can be derived from a product, producers are likely to be able to raise prices without compromising on sales. Note that including product and year fixed effects $\boldsymbol{\xi}$ already goes a long way by accounting for time or firm invariant taste characteristics as well as other sources of endogeneity bias that do not vary across these dimensions.

There are several candidates for instruments which have been used in demand estimation. The profit maximisation condition for plants is a useful guide for instrument choice around which the literature can be structured. Rewriting condition (I.1) yields $P_{jt} = \frac{1+\eta_{jt}(\mathbf{P_t})}{MC_{jt}\left(Q_{jt}(\mathbf{P_t}),\mathbf{c}_{jt}\right)}$. Much of the literature relies on an internal instrument that drives the numerator $1 + \eta_{jt}(\mathbf{P_t})$.[116] However, since I have endogenous marginal costs that depend

[116] Hausman et al. (1994) estimate a nested logit demand system for beer and use other cities' beer prices

on observed marginal cost shifters, I can use these external instruments for prices. Shocks to input prices are assumed to be uncorrelated with shocks to taste $\Delta\xi_{jt}$, conditional on average product quality which is controlled for by product fixed effects. Armstrong (2016) shows that internal BLP-style instruments, in particular in the Bertrand Nash structure, tend to perform poorly in small samples and also lose identifying power asymptotically. He recommends cost shifters which are consistent over a broad range of cases.

I use the average plant level wages w_{jt}, and the average plant level prices of a tonne of material inputs P^M_{jt} as instruments for output prices P_{jt}. At the solution of the model, the first stage Kleibergen-Paap F statistic is 21.81, rejecting a weak instrument hypothesis. The Hansen overidentification Chi-Square J statistic is smaller than 0.001, with a p-value close to unity so the hypothesis of valid instruments can not be rejected. The point estimate of the IV regression of mean utility on price is -17.15*** as shown in Table I.5. A plain OLS regression in (I.29) yields an estimate of -14.60*** , which is significantly positively biased as expected.

An alternative instrument was proposed by Foster et al. (2008) who estimate productivity (TFPQ) and a demand function. They propose to use TFPQ (Ω_{jt}) or the innovation in the productivity process (ζ_{jt}) as instruments for prices. In my data, the correlation between Ω_{jt} and P_{jt} is -0.18***, and between ζ_{jt} and P_{jt} -0.09*. This makes intuitive sense, considering that higher productivity leads to higher output quantity which is associated with lower prices. However, Forlani et al. (2016) find in their joint estimation of demand and productivity that shocks in consumer taste and shocks to productivity are negatively correlated, which would make both TFPQ (Ω_{jt}) and the innovation in it (ζ_{jt}) unsuitable instruments.[117] Including TFPQ (Ω_{jt}) as an additional instrument for estimating Equation (I.29) yields a point estimate of -17.20*** for the instrumented price, very similar to my baseline result of -17.15*** in Table I.5. The first stage F statistic is 16.05, and Hansen's J test of valid instruments cannot be rejected with a p-value of 0.70.

---

to instrument for a city's beer price. BLP themselves use other products' characteristics values of the same firm and the same product's characteristics values of other firms (via a competition channel). Both do not rely on additional data, but require stronger assumptions, as discussed therein and in Nevo (2001). Reynaert and Verboven (2014) suggest an improved set of BLP instruments based on Chamberlain (1987) optimal instruments (see also more recent work on this in Gandhi and Houde (2016)). The rank condition of these types of instruments can be rationalised by noting that they affect the demand elasticities, upon which the price choice depends in the firm's optimisation condition.

[117]I also find that productivity shocks ζ_{jt} and unobserved average utility shocks $\Delta\xi_{jt}$ are negatively correlated (-0.04), but insignificantly. I avoid using ζ_{jt} as instrument also for practical reasons. I would need to estimate the production side to generate the instrument for the demand side. But when estimating the demand side first, I can use demand side estimates as instruments for the invertibility condition in the production estimation.

Table I.5: Estimates of demand parameters

	Point estimate	SE
θ^p	-17.15***	0.40 (bootstrapped)
σ^p	-5.83***	0.13 (robust)
σ^x	0.02	1.19 (robust)
$\boldsymbol{\xi}$	Yes	.

Notes: The table shows the estimates for the structural parameters on the demand side. The number of observations is 989. The standard errors of the linear parameter θ^p depends on the non-linear parameters. Details for calculation of the standard errors are in Appendix I.D.2.

Estimated demand parameters

I omit the estimates for the dummies (θ^x, ξ_j, ξ_t) in Table I.5. The estimate for the mean price coefficient θ^p (-17.15***) is negative and highly statistically significant, as is the standard deviation of the mean price coefficient σ^x (-5.83***).[118] This means that there is significant variation in the random coefficient on price. On the other hand, the variance on the constant (σ^x) is small and insignificant, so presents little evidence of a random intercept. See Appendix I.D.2 for details on the calculation of the standard errors.

I.G.3 Further results and robustness checks for the production side

Table I.6 shows the baseline estimates in Column (1) and the OLS results in Column (2). Since I am only using observations that are part of consecutive spells of data because of the assumed timing structure of the model, I have fewer observations than for the plain OLS result in Column (2). Observations that belong to a plant that has consecutive spells but are in years without consecutive spells are not used for the estimation. The OLS results are robust to only using the same consecutive-spells sample of Column (1) as well. Importantly, since there are no lags in the first equation of this GMM system, I can calculate ϵ_{jt} and Ω_{jt} also for the non-consecutive observations (and can use them for the counterfactual exercise).

Column (3) controls for simultaneity only, which reduces the material elasticity compared to Column (2) in the Cobb-Douglas specification. In the translog version the mean increases, but with a much higher standard deviation.

I perform several robustness checks for the crucial scalar unobservable and invertibility condition in the control function approach. Column (4) includes a variable in the invertibility condition that is aimed to capture variation in the unobserved τ. I construct a variable

[118]This translates into a variance of 33.9. The negative sign of the standard deviation is irrelevant as the square, the variance, is always positive. If I take positive starting values for σ^p instead of zero, I get an equal size "positive" standard deviation as solution.

Table I.6: Estimates from a Cobb-Douglas production function

	Type of correction				
	(1) Simultaneity & Selectivity	(2) None: OLS	(3) Simultaneity only	(4) Sim. & Selec. w. augmented z	(5) BB system-GMM
α^K	.06***	.04***	.10***	.05**	.08
	(.02)	(.01)	(.02)	(.02)	(.09)
α^L	.22***	.14***	.19***	.20***	.06
	(.05)	(.02)	(.05)	(.05)	(.10)
α^M	.64***	.80***	.68***	.65***	.73***
	(.05)	(.03)	(.06)	(.05)	(.07)
RTS	.92***	.99***	.97***	.90***	.86***
	(.03)	(.01)	(.02)	(.03)	(.06)
N	443	1001	512	443	512

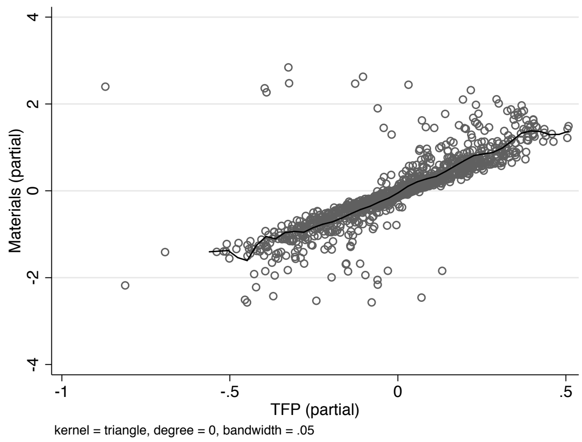
Notes: The columns show the output elasticities and returns to scale for different type of corrections for simultaneity and selectivity. Column (4) includes an additional variable in the material demand equation $(\hat{\eta}_{jt}+1)\frac{P^r_{jt}Q^r_{jt}}{P^M_{jt}M_{jt}} = \frac{\tau^M_{jt}}{\alpha^M_{jt}}\frac{P^r_{jt}Q^r_{jt}}{P_{jt}Q_{jt}}$. Column (5) is based on a Blundell-Bond system GMM estimator. Clustered standard errors at the plant level are in parentheses.

$(\hat{\eta}_{jt}+1)\frac{P^r_{jt}Q^r_{jt}}{P^M_{jt}M_{jt}}$ which is similar to the definition of τ and include it in $\mathbf{z}_{jt}$.[119] Empirically, the estimates are very close to the main specification in Column (1) suggesting no violation of the invertibility condition (I.15). In Column (5), I use the Blundell-Bond system GMM estimator as a further check for the invertibility. Compared to OLS it also reduces the bias in the OLS material elasticity in the same direction as my main specification, particularly for Cobb-Douglas, but yields less precise estimates. We can also inspect the monotonicity required for the invertibility condition, similar to Levinsohn and Petrin (2003). Figure I.18 plots material use against productivity, where I partialled out a polynomial of the other input variables, and fitted a local kernel with a tight bandwidth. For most of the density, the smoothed mean is indeed monotonically increasing in productivity (as expected), with a couple of outliers.

Finally, the autoregressive parameter estimate in the productivity process is 0.87*** for Cobb-Douglas and 0.81** for translog, very similar to the annual persistence parameter of 0.8 in Foster et al. (2008). The coefficient for the predicted probability of being in the sample is positive and significant for predicting productivity as expected and in line with the corrected selection bias in the capital elasticity.

[119] It is the same apart from using realised rather than expected revenue, and omitting the output elasticity, which is constant across plants in the Cobb Douglas case.

Figure I.18: Monotonicity of material demand in productivity



Notes: Plotted are the residuals from a regression of log materials m_{it} on a third order polynomial in labour and capital, against the residuals of a regression of log productivity $\hat{\omega}_{jt}$ on the same polynomial.

Table I.7: Estimates from a translog production function

	Type of correction					
	(1) Simultaneity & Selectivity	(2) None: OLS	(3) Simultaneity only	(4) Sim. & Selec. w. augmented	(5) BB system-GMM	(6) Sim. & Selec. w. investment
α^K	.07	.05	.08	.07	.09	.05
	(.14)	(.05)	(.10)	(.11)	(.07)	(.14)
α^L	.28	.13	.08	.28	.04	.14
	(.15)	(.05)	(.21)	(.14)	(.15)	(.34)
α^M	.60	.81	.82	.59	.79	.76
	(.22)	(.13)	(.34)	(.19)	(.17)	(.39)
RTS	.95	.98	.97	.94	.92	.96
	(.08)	(.06)	(.15)	(.06)	(.05)	(.16)
N	443	1001	512	443	511	410

Notes: The columns show the output elasticities and returns to scale for different types of corrections for simultaneity and selectivity. Column (4) includes an additional variable in the material demand equation $(\hat{\eta}_{jt}+1)\frac{P^r_{jt}Q^r_{jt}}{P^M_{jt}M_{jt}} = \frac{\tau^M_{jt}}{\alpha^M_{jt}}\frac{P^r_{jt}Q^r_{jt}}{P_{jt}Q_{jt}}$. Column (5) is based on a Blundell-Bond system GMM estimator. Column (6) is based on an investment function instead of a material demand function for the invertibility condition, following the original Olley and Pakes (1996). Standard deviations across the entire sample are in parentheses.

I.G.4 Translog estimates

The estimates from a translog production function are presented in Table I.7. The translog elasticities vary by plant and year, but the mean elasticities are very similar to the estimates from the Cobb-Douglas production function, with returns to scale close to one. For the translog elasticity, the standard deviation across all plant-years are reported in parentheses. In Column (6) I additionally use an investment function instead of a material demand function for the invertibility condition, following the original Olley and Pakes (1996). These estimates are also less precise because of the lumpiness and zeros in investment data.

I.G.5 Analysis of estimated plant total factor productivities

How has productivity evolved and how dispersed is it? I recover *physical* total factor productivity Ω_{jt} (also denoted TFPQ). Total *revenue* factor productivity TFPR is simply defined as $P_{jt} \cdot TFPQ_{jt}$.[120]

I find that TFPQ has not significantly changed over time within firms. A linear within-plants fixed effects regression of logged TFPQ on years yields small and insignificant results. The same goes for a pooled (across plants) regression suggesting that average TFPQ stagnated. This also suggests that entry of productive and exit of unproductive firm did not play a major role either. However, when I weight the pooled regression by output quantity, TFPQ has increased by 1% per year, significant with SE clustered at the plant. This comes from changes in the weights, and likely due to more productive firms growing faster compared to less productive firms, or from larger firms becoming more productive than smaller firms. When we interpret the weighted TFPQ as a form of aggregate TFPQ, the results suggest that despite stagnating average TFPQ, aggregate TFPQ seems to have slightly increased over the sample period.

TFPR increased over the years (using deflated prices), highly significant with SE clustered at the plant. On average, each year the TFPR increases by 2%, both in the pooled and the within estimation. This rise in inflation adjusted prices, together with the result on decreasing markups (Section I.4.1) suggests increasing marginal costs. Indeed, marginal costs have been rising by a little over 2% per year for the pooled and within specification. This is mainly due to an increase in input prices. Material input prices have been rising at around 3% per year, consistent with the global prices increases in raw metals commodity prices (see e.g. IMF). Increased marginal costs could also have been driven by changes in τ.

The estimated *dispersion* in TFPQ is smaller than in some other studies in the literature.[121] The ratio in TFPQ of the 90th percentile plant to the 10th percentile plant is 1.83 in this sample, much smaller than the ratio reported for India in Hsieh and Klenow (2009), which is over 20. This is likely due to three aspects. First, Hsieh and Klenow (2009) don't observe prices and quantities and cannot estimate TFPQ directly. Second, they use a value added instead of gross output production function. Third, I look at a much narrower industry. My ratio is more in line with the ready-mix concrete producers in

[120] The reported numbers are based on my baseline CD specification of Column (1) that I also use for the counterfactual analysis.

[121] The dispersion of productivity (both TFPQ and TFPR) across plants is of interest in itself as it has become an important feature and subject of analysis in various disciplines, as reviewed in Syverson (2011).

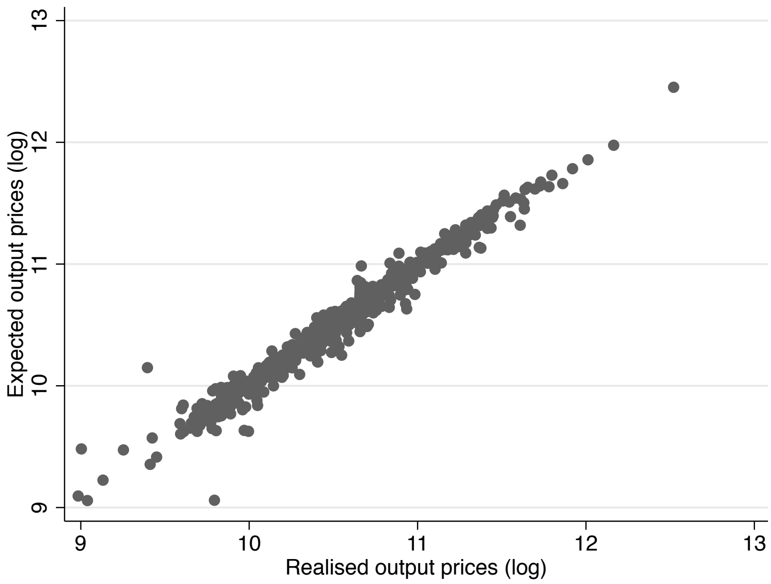
the US of around 1.91 reported in Syverson (2004*a*). Finally, the higher dispersion in these studies could also arise, from lumping Ω_{jt} and ϵ_{jt} together, which I disentangle. Indeed the 90th to 10th percentile ratio of the comparable $(\Omega_{jt}\exp(\epsilon_{jt}))$ in my data is 2.5.

The dispersion in TFPQ is smaller than the dispersion in TFPR. The 90th to 10th TFPR ratio is 3.3, larger than the average ratios of 1.92 in the US within 4-digit sectors as reported in Syverson (2004*b*), but smaller than the ratio of 5 reported for India in Hsieh and Klenow (2009). Interestingly, I estimate a lower dispersion for TFPQ then for TFPR, which is the opposite for Hsieh and Klenow (2009) and for Foster et al. (2008). I also find a robust and significant negative correlation between TFPQ and prices in the data, rationalised by a standard downward-sloping demand curve. However, because the dispersion in prices is much larger than the dispersion in TFPQ, combining both leads to a dispersion in TFPR that is smaller than that of prices but larger than that of TFPQ.

I.G.6 Realised and expected prices

Figure I.19 plots the realised (after shock ϵ) prices against the expected prices.

Figure I.19: Expected and realised prices



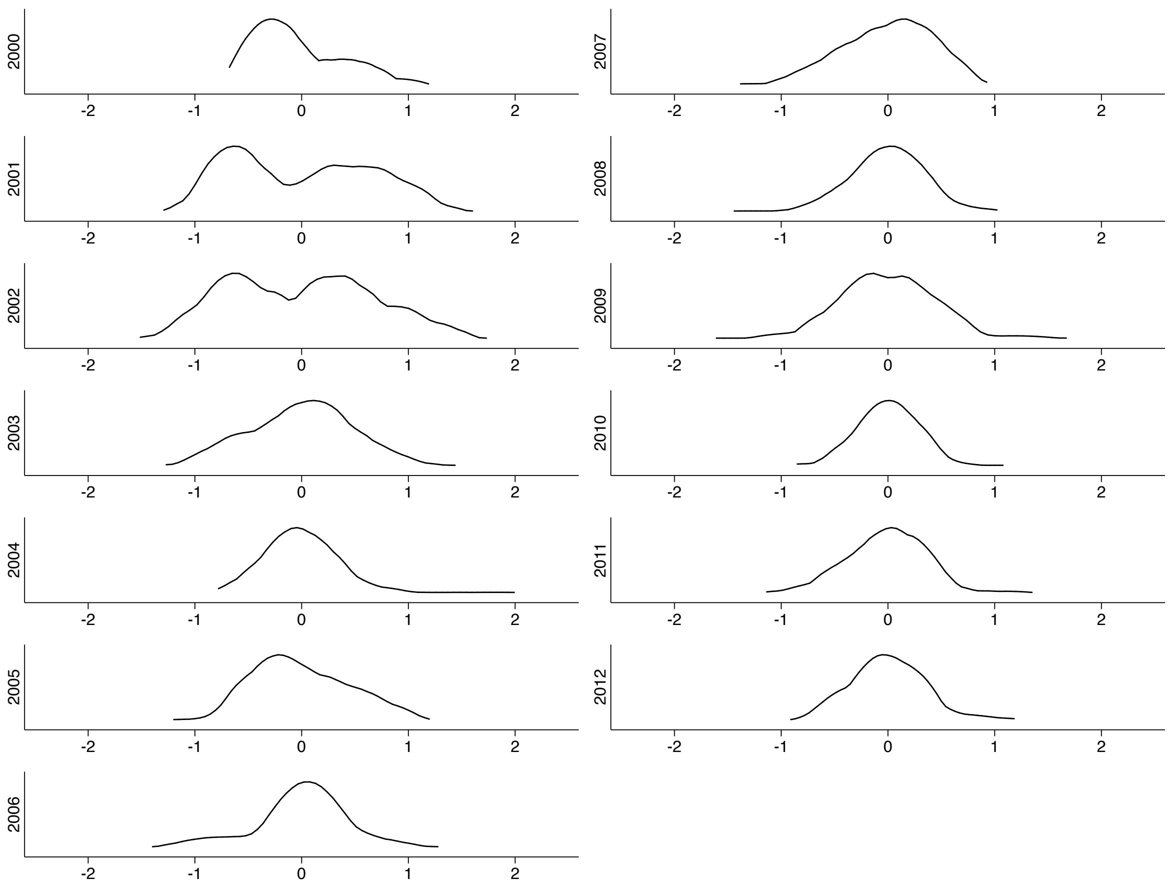
Notes: Plotted is the the realised observed prices $\log(P^r_{jt})$ against the equilibrium prices $\log(P_{jt})$.

I.G.7 Additional descriptive figures on distortions

Figures I.20 and I.21 plot the distortions by year. Figure I.22 plots the annual standard deviations. We can also compare the distribution of $\ln(\tau^M_{jt})$ and $\ln(\tau^L_{jt})$ in their baseline versions with variable markups with a version for each with constant markups as shown in

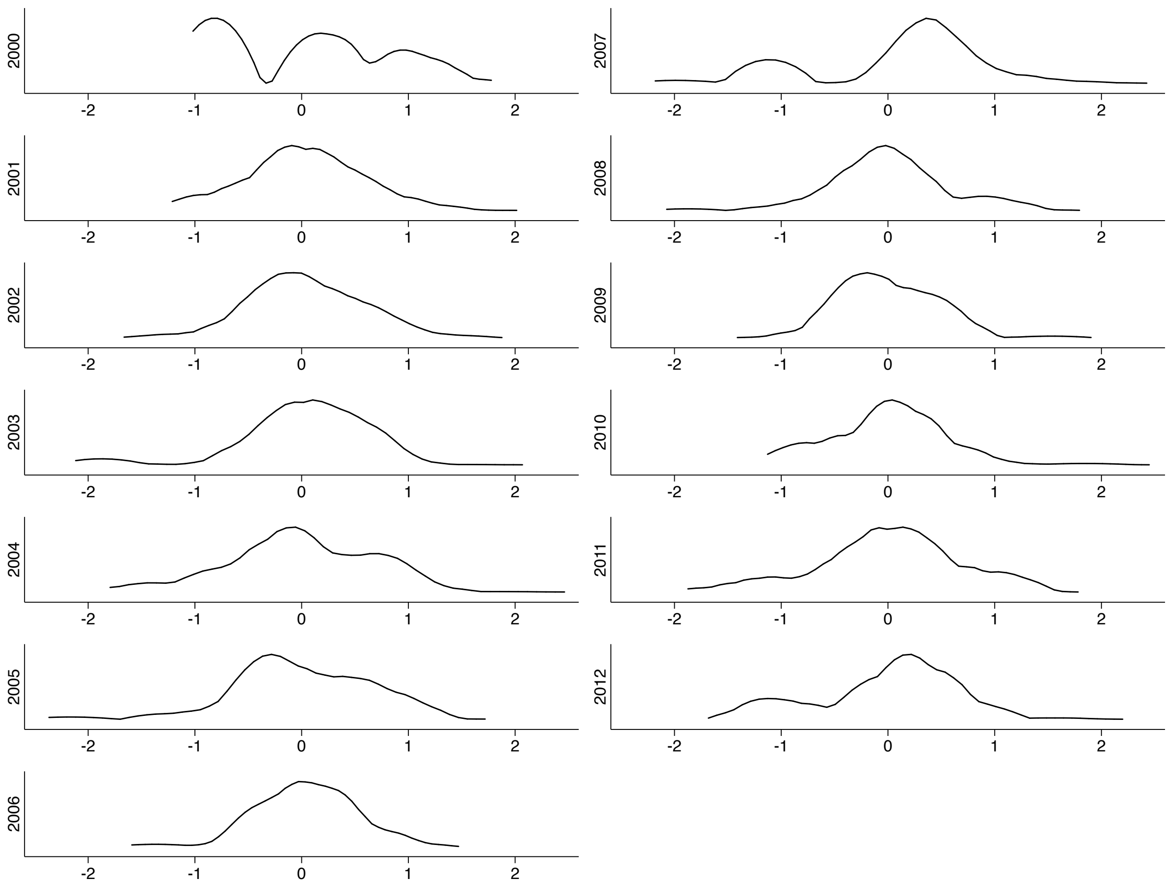
Figure I.23 and I.24. Figure I.25 plots $\ln(\tau_{jt}^M)$ against $\ln(\tau_{jt}^L)$.

Figure I.20: Dispersion in τ_{jt}^M by year



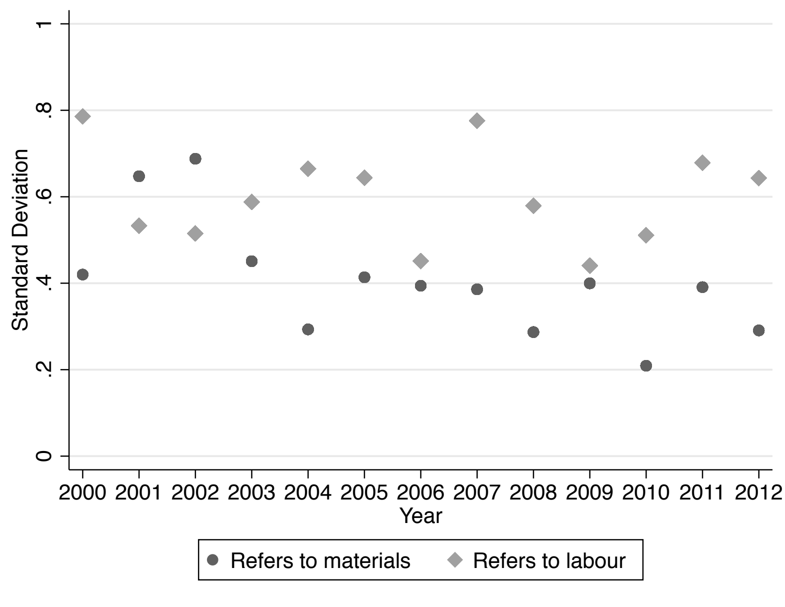
Notes: Plotted is the kernel density of $\ln(\tau_{jt}^M)$ divided by the weighted geometric mean of τ_{jt}^M, where the weights are plant material expenditure. Used kernel is epanechnikov with bandwidth 0.2.

Figure I.21: Dispersion in τ_{jt}^L by year



Notes: Plotted is the kernel density of $\ln(\tau_{jt}^L)$ divided by the weighted geometric mean of τ_{jt}^L, where the weights are plant labour expenditure. Used kernel is epanechnikov with bandwidth 0.2.

Figure I.22: Standard deviation fo τ_{jt}^M and τ_{jt}^L by year



Notes: Plotted are the standard deviations of the demeaned and weighted $\ln(\tau_{jt}^M)$ and $\ln(\tau_{jt}^L)$.

Figure I.23: Markup correction for τ_{jt}^M

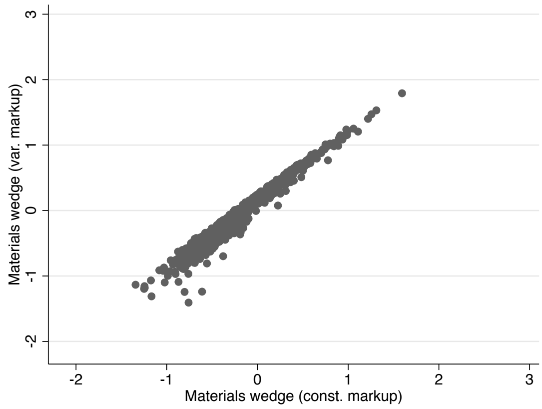
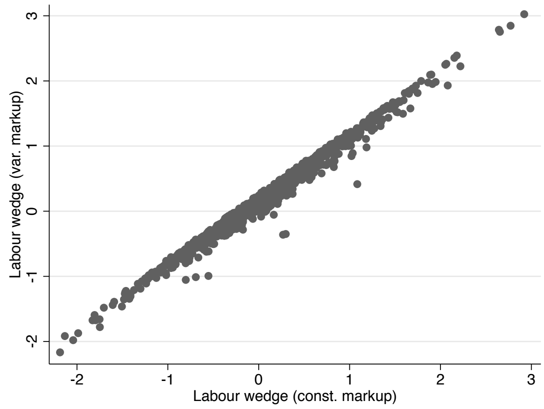
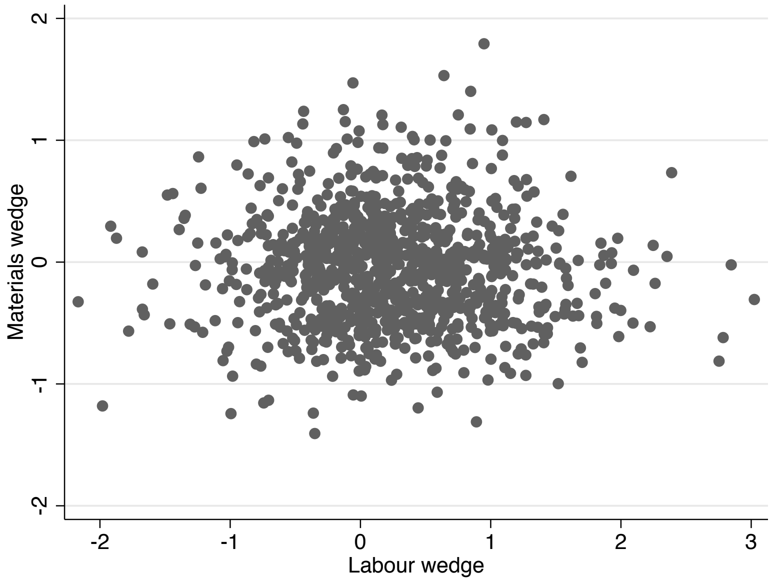


Figure I.24: Markup correction for τ_{jt}^L



Notes: Plotted are the pooled demeaned $\ln(\tau_{jt}^M)$ and $\ln(\tau_{jt}^L)$. The vertical axis corresponds to the input distortions corrected for markups and the horizontal axis corresponds to a "naive" version where an average markup is used to calculate input distortions instead.

Figure I.25: Correlation between τ_{jt}^M and τ_{jt}^L



Notes: Plotted are $\ln(\tau_{jt}^M)$ and $\ln(\tau_{jt}^L)$ divided by the respective weighted means, where the weights are plant materials and labour expenditure. Pooled across all years.

Table I.9: Welfare gains in billion rupees

	Compensating Variation			Profits			Total welfare		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	0.15	0.14	0.27	0.1	0.08	0.18	0.25	0.21	0.45
	[0.11,0.21]	[0.07,0.2]	[0.23,0.34]	[0.07,0.13]	[0.03,0.1]	[0.11,0.2]	[0.19,0.33]	[0.1,0.3]	[0.37,0.53]
2001	1.57	0.39	2.12	1	0.21	1.32	2.57	0.6	3.43
	[1.13,2.25]	[0.2,0.62]	[1.66,2.76]	[0.72,1.19]	[0.09,0.3]	[0.98,1.43]	[1.94,3.37]	[0.29,0.91]	[2.78,4.15]
2002	3.27	0.56	4.08	2.04	0.32	2.48	5.31	0.88	6.56
	[2.36,4.2]	[0.28,0.92]	[3.22,4.81]	[1.45,2.41]	[0.15,0.48]	[1.79,2.77]	[3.92,6.51]	[0.43,1.4]	[5.33,7.45]
2003	2.49	0.89	3.29	1.1	0.39	1.41	3.59	1.28	4.7
	[1.76,3.57]	[0.45,1.41]	[2.57,4.26]	[0.72,1.38]	[0.18,0.54]	[1.02,1.64]	[2.5,4.9]	[0.64,1.92]	[3.83,5.77]
2004	2.49	1.66	3.68	0.86	0.49	1.22	3.35	2.15	4.9
	[1.93,3.02]	[0.89,2.34]	[3.22,4.02]	[0.48,1.12]	[0.22,0.66]	[0.68,1.38]	[2.53,4.1]	[1.1,2.97]	[4.06,5.34]
2005	2.07	1.4	4.07	1.17	0.8	2.5	3.23	2.2	6.57
	[1.53,2.87]	[0.66,2.4]	[3.08,5.34]	[0.81,1.42]	[0.35,1.29]	[1.71,3.29]	[2.42,4.24]	[0.99,3.73]	[5,8.59]
2006	2.1	0.97	3.25	1.13	0.45	1.72	3.23	1.42	4.97
	[1.52,3.03]	[0.48,1.51]	[2.55,4.28]	[0.78,1.4]	[0.2,0.63]	[1.22,2.09]	[2.43,4.32]	[0.69,2.13]	[3.97,6.4]
2007	1.3	1.47	2.6	0.79	0.82	1.6	2.1	2.28	4.2
	[1.04,1.77]	[0.81,2.18]	[2.13,3.34]	[0.59,0.93]	[0.35,1.12]	[1.04,1.83]	[1.68,2.64]	[1.14,3.24]	[3.4,5.02]
2008	0.07	0.6	0.59	0.14	0.37	0.44	0.21	0.97	1.03
	[-0.1,0.33]	[0.28,0.98]	[0.3,0.95]	[0.04,0.24]	[0.16,0.55]	[0.26,0.56]	[-0.06,0.57]	[0.44,1.51]	[0.6,1.49]
2009	2.43	0.81	3.06	1.4	0.41	1.7	3.83	1.22	4.76
	[1.66,3.81]	[0.38,1.31]	[2.33,4.27]	[0.98,1.81]	[0.18,0.61]	[1.27,2]	[2.68,5.62]	[0.56,1.89]	[3.76,6.22]
2010	0.51	0.87	1.11	0.19	0.34	0.44	0.7	1.22	1.55
	[0.39,0.68]	[0.44,1.35]	[0.79,1.44]	[0.13,0.23]	[0.15,0.5]	[0.26,0.54]	[0.54,0.91]	[0.59,1.84]	[1.07,1.98]
2011	1.8	1.18	3.52	1.55	0.76	2.98	3.35	1.95	6.5
	[1.18,2.95]	[0.61,2.01]	[2.84,5.45]	[0.96,2.23]	[0.35,1.07]	[2.14,4.86]	[2.18,5.08]	[0.96,2.98]	[5.06,10.16]
2012	1.54	1.22	2.44	0.63	0.66	1.22	2.17	1.88	3.66
	[1,2.3]	[0.6,1.99]	[2,3.18]	[0.38,1.01]	[0.28,1.03]	[0.86,1.54]	[1.41,3.29]	[0.88,2.96]	[2.99,4.74]
Total	21.81	12.16	34.08	12.09	6.09	19.21	33.9	18.26	53.28
	[16.08,30.12]	[6.13,18.87]	[27.86,42.38]	[8.4,14.74]	[2.66,8.61]	[13.8,22.31]	[24.87,43.82]	[8.74,27.27]	[44.27,64.09]
Per year	1.68	0.94	2.62	0.93	0.47	1.48	2.61	1.4	4.1
	[1.24,2.32]	[0.47,1.45]	[2.14,3.26]	[0.65,1.13]	[0.2,0.66]	[1.06,1.72]	[1.91,3.37]	[0.67,2.1]	[3.41,4.93]

Notes: The table shows the respective welfare gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both in billion rupees. The last three columns sum the consumer side compensating variation and the profits for total welfare gains. The last two rows report the total across all years and the implied average per year. Bootstrapped 95% confidence intervals in brackets (see Section I.2.4).

I.G.8 Tables with annual welfare gains

Table I.8: Compensating variation as share of consumer expenditure and profit growth

	Compensating Variation			Profit growth		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	0.13	0.11	0.23	1.22	1.16	1.38
	[0.09,0.17]	[0.06,0.17]	[0.19,0.29]	[1.14,1.29]	[1.06,1.24]	[1.24,1.46]
2001	0.24	0.06	0.32	1.51	1.11	1.66
	[0.17,0.34]	[0.03,0.09]	[0.25,0.42]	[1.35,1.63]	[1.05,1.17]	[1.48,1.82]
2002	0.31	0.05	0.39	1.71	1.11	1.86
	[0.22,0.4]	[0.03,0.09]	[0.31,0.46]	[1.47,1.87]	[1.05,1.18]	[1.6,2.03]
2003	0.16	0.06	0.21	1.29	1.1	1.37
	[0.11,0.23]	[0.03,0.09]	[0.16,0.27]	[1.18,1.39]	[1.04,1.16]	[1.25,1.47]
2004	0.15	0.1	0.22	1.2	1.11	1.28
	[0.11,0.18]	[0.05,0.14]	[0.19,0.24]	[1.1,1.3]	[1.05,1.17]	[1.15,1.38]
2005	0.13	0.09	0.26	1.28	1.19	1.61
	[0.1,0.19]	[0.04,0.16]	[0.2,0.35]	[1.19,1.37]	[1.08,1.36]	[1.38,1.94]
2006	0.1	0.05	0.16	1.21	1.08	1.32
	[0.08,0.15]	[0.02,0.08]	[0.13,0.21]	[1.14,1.28]	[1.04,1.13]	[1.21,1.46]
2007	0.09	0.1	0.18	1.18	1.19	1.36
	[0.07,0.12]	[0.06,0.15]	[0.15,0.23]	[1.13,1.23]	[1.08,1.28]	[1.24,1.47]
2008	0.01	0.06	0.06	1.05	1.14	1.17
	[-0.01,0.04]	[0.03,0.11]	[0.03,0.1]	[1.02,1.1]	[1.06,1.23]	[1.1,1.24]
2009	0.15	0.05	0.19	1.33	1.1	1.4
	[0.1,0.24]	[0.02,0.08]	[0.15,0.27]	[1.22,1.5]	[1.04,1.16]	[1.29,1.56]
2010	0.05	0.08	0.1	1.07	1.13	1.16
	[0.04,0.06]	[0.04,0.13]	[0.08,0.14]	[1.04,1.1]	[1.06,1.21]	[1.09,1.22]
2011	0.1	0.06	0.19	1.34	1.17	1.65
	[0.06,0.16]	[0.03,0.11]	[0.15,0.29]	[1.2,1.52]	[1.08,1.24]	[1.44,2.2]
2012	0.1	0.08	0.15	1.15	1.15	1.29
	[0.06,0.14]	[0.04,0.12]	[0.12,0.2]	[1.09,1.26]	[1.06,1.26]	[1.19,1.43]
Total	0.13	0.07	0.21	1.26	1.13	1.42
	[0.1,0.18]	[0.04,0.11]	[0.17,0.26]	[1.17,1.35]	[1.06,1.21]	[1.29,1.56]

Notes: The table shows the respective welfare gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both. The first three columns express the compensating variation per unit purchased as a share of the unit price (i.e. as share of expenditure on the products in the sample). The profit ratio is total profits in the counterfactual divided by total profits in the factual equilibrium. Bootstrapped 95% confidence intervals in brackets (see Section I.2.4).

I.G.9 Tables with annual input productivity gains

Table I.10 and I.11 report the ratio in the physical and revenue productivities respectively, i.e. the ratio between the input productivity in the counterfactual and the factual equilibria.[122]

[122]The first three columns in both tables show that output increases more than revenues (except in one case in 2008), as the average price decreases which contributes to the consumer welfare gains. Therefore the ratios of the counterfactual and factual productivities is lower in the revenue productivity outcomes. Since we have decreasing prices across the counterfactuals, we would need to correct for this and inflate the revenue productivity accordingly. But since I can measure output in weight, the physical productivity is a directly suited metric for deflated value per unit.

Table I.10: Physical output and productivity ratios

	Output ratio			Material productivity			Labour productivity		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	1.27	1.26	1.52	0.89	1.19	1.03	0.9	0.55	0.55
	[1.16,1.35]	[1.11,1.37]	[1.32,1.63]	[0.87,0.93]	[1.09,1.31]	[0.97,1.13]	[0.87,0.96]	[0.52,0.58]	[0.53,0.59]
2001	1.61	1.15	1.87	0.75	1.1	0.81	1.1	0.8	0.87
	[1.37,1.83]	[1.06,1.24]	[1.56,2.1]	[0.72,0.79]	[1.04,1.17]	[0.78,0.88]	[1.02,1.24]	[0.79,0.83]	[0.83,0.95]
2002	1.98	1.16	2.27	0.65	1.09	0.74	1.55	0.85	1.18
	[1.59,2.2]	[1.07,1.26]	[1.82,2.42]	[0.59,0.75]	[1.04,1.18]	[0.65,0.86]	[1.32,1.86]	[0.81,0.89]	[1.05,1.35]
2003	1.38	1.15	1.52	0.89	1.1	0.98	1.04	0.88	0.98
	[1.21,1.53]	[1.06,1.23]	[1.34,1.65]	[0.85,0.95]	[1.05,1.18]	[0.93,1.07]	[0.99,1.15]	[0.85,0.93]	[0.93,1.09]
2004	1.36	1.21	1.53	0.65	1.12	0.77	1.42	0.81	1.08
	[1.22,1.44]	[1.1,1.3]	[1.32,1.58]	[0.59,0.76]	[1.05,1.25]	[0.68,0.92]	[1.24,1.72]	[0.76,0.86]	[0.97,1.23]
2005	1.36	1.24	1.79	0.82	1.12	0.95	1.31	0.81	0.95
	[1.23,1.49]	[1.1,1.42]	[1.49,2.05]	[0.8,0.85]	[1.04,1.26]	[0.88,1.1]	[1.18,1.53]	[0.8,0.84]	[0.9,1.08]
2006	1.26	1.12	1.42	0.87	1.05	0.92	1.12	0.88	0.95
	[1.16,1.36]	[1.05,1.19]	[1.27,1.56]	[0.86,0.9]	[1.02,1.1]	[0.9,0.98]	[1.05,1.23]	[0.86,0.91]	[0.9,1.02]
2007	1.26	1.28	1.56	0.92	1.15	1.05	1.1	0.62	0.69
	[1.18,1.33]	[1.12,1.42]	[1.35,1.71]	[0.89,0.95]	[1.07,1.25]	[0.99,1.14]	[1.06,1.17]	[0.59,0.65]	[0.66,0.74]
2008	1.04	1.21	1.21	0.93	1.1	1.01	0.98	0.92	0.93
	[0.98,1.11]	[1.08,1.34]	[1.1,1.32]	[0.93,0.95]	[1.04,1.2]	[0.97,1.09]	[0.96,1]	[0.87,0.98]	[0.9,0.98]
2009	1.38	1.14	1.51	0.74	1.06	0.8	1.05	0.95	0.99
	[1.22,1.6]	[1.06,1.22]	[1.33,1.72]	[0.73,0.78]	[1.02,1.12]	[0.77,0.87]	[1.02,1.17]	[0.91,1.01]	[0.95,1.1]
2010	1.12	1.21	1.28	0.95	1.06	1.02	1.01	0.83	0.86
	[1.08,1.16]	[1.1,1.33]	[1.17,1.37]	[0.94,0.97]	[1.02,1.13]	[0.98,1.08]	[1,1.04]	[0.8,0.89]	[0.84,0.9]
2011	1.35	1.22	1.72	0.71	1.09	0.78	1.35	0.86	1.17
	[1.2,1.53]	[1.1,1.34]	[1.49,2.12]	[0.66,0.78]	[1.03,1.17]	[0.71,0.89]	[1.21,1.64]	[0.83,0.91]	[1.04,1.38]
2012	1.3	1.25	1.53	0.71	1.11	0.79	1.08	0.79	0.92
	[1.17,1.49]	[1.1,1.41]	[1.36,1.73]	[0.68,0.82]	[1.04,1.21]	[0.74,0.9]	[1.01,1.22]	[0.77,0.83]	[0.85,1.07]
Total	1.35	1.2	1.58	0.77	1.1	0.86	1.16	0.82	0.95
	[1.22,1.47]	[1.08,1.3]	[1.4,1.73]	[0.75,0.82]	[1.04,1.19]	[0.82,0.95]	[1.09,1.27]	[0.79,0.86]	[0.89,1.03]

Notes: The table shows the respective gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both. Outcome variables are the ratio of the counterfactual to the factual. Bootstrapped 95% confidence intervals in brackets (see Section I.2.4).

Table I.11: Revenue and revenue productivity ratios

	Revenue ratio			Material productivity			Labour productivity		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	1.28	1.13	1.39	0.9	1.06	0.95	0.91	0.49	0.51
	[1.19,1.33]	[1.04,1.16]	[1.23,1.45]	[0.89,0.92]	[1.03,1.12]	[0.93,0.99]	[0.89,0.94]	[0.47,0.51]	[0.49,0.55]
2001	1.69	1.11	1.81	0.78	1.05	0.78	1.15	0.77	0.84
	[1.49,1.83]	[1.05,1.16]	[1.56,1.92]	[0.78,0.8]	[1.03,1.1]	[0.78,0.81]	[1.1,1.23]	[0.76,0.78]	[0.83,0.87]
2002	1.7	1.11	1.73	0.56	1.05	0.56	1.32	0.82	0.9
	[1.46,1.75]	[1.05,1.16]	[1.46,1.76]	[0.54,0.59]	[1.02,1.1]	[0.54,0.63]	[1.24,1.47]	[0.79,0.84]	[0.87,0.99]
2003	1.34	1.1	1.4	0.86	1.06	0.9	1.01	0.85	0.9
	[1.21,1.4]	[1.04,1.14]	[1.26,1.45]	[0.81,0.88]	[1.03,1.1]	[0.87,0.94]	[0.98,1.07]	[0.83,0.87]	[0.88,0.95]
2004	1.02	1.06	1.03	0.49	0.99	0.51	1.07	0.71	0.73
	[0.93,1.04]	[1.01,1.08]	[0.91,1.06]	[0.42,0.6]	[0.96,1.04]	[0.44,0.64]	[1.02,1.17]	[0.7,0.72]	[0.7,0.78]
2005	1.32	1.2	1.55	0.79	1.08	0.83	1.27	0.78	0.83
	[1.21,1.4]	[1.09,1.32]	[1.34,1.63]	[0.77,0.81]	[1.04,1.18]	[0.8,0.88]	[1.15,1.43]	[0.77,0.8]	[0.81,0.87]
2006	1.24	1.07	1.31	0.86	1	0.85	1.11	0.84	0.88
	[1.16,1.3]	[1.03,1.1]	[1.2,1.36]	[0.84,0.88]	[1,1.02]	[0.84,0.88]	[1.06,1.18]	[0.83,0.85]	[0.86,0.91]
2007	1.15	1.15	1.28	0.84	1.03	0.86	1	0.55	0.56
	[1.08,1.17]	[1.06,1.21]	[1.13,1.32]	[0.8,0.86]	[1.01,1.06]	[0.82,0.89]	[0.99,1.03]	[0.54,0.57]	[0.55,0.59]
2008	1.07	1.12	1.17	0.96	1.02	0.98	1.01	0.85	0.9
	[1.03,1.11]	[1.05,1.18]	[1.1,1.23]	[0.94,0.99]	[1.01,1.05]	[0.94,1.03]	[0.99,1.03]	[0.84,0.87]	[0.89,0.93]
2009	1.41	1.07	1.44	0.76	0.99	0.77	1.08	0.89	0.95
	[1.26,1.55]	[1.03,1.09]	[1.29,1.57]	[0.73,0.8]	[0.98,1.01]	[0.73,0.81]	[1.05,1.13]	[0.87,0.91]	[0.91,1.03]
2010	1.05	1.09	1.12	0.9	0.95	0.89	0.95	0.75	0.75
	[1.02,1.06]	[1.03,1.12]	[1.04,1.14]	[0.87,0.92]	[0.93,0.98]	[0.86,0.92]	[0.93,0.97]	[0.73,0.77]	[0.73,0.78]
2011	1.34	1.15	1.52	0.7	1.03	0.69	1.34	0.81	1.03
	[1.23,1.44]	[1.07,1.23]	[1.36,1.66]	[0.63,0.75]	[1.01,1.07]	[0.61,0.78]	[1.22,1.54]	[0.79,0.84]	[0.94,1.13]
2012	1.13	1.14	1.24	0.62	1.02	0.64	0.94	0.73	0.75
	[1.07,1.17]	[1.06,1.21]	[1.13,1.3]	[0.58,0.69]	[1.01,1.05]	[0.58,0.74]	[0.88,1.04]	[0.69,0.76]	[0.71,0.8]
Total	1.27	1.11	1.36	0.72	1.02	0.74	1.09	0.77	0.81
	[1.17,1.32]	[1.05,1.16]	[1.22,1.42]	[0.7,0.76]	[1.01,1.06]	[0.71,0.79]	[1.05,1.14]	[0.76,0.77]	[0.8,0.85]

Notes: The table shows the respective gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both. Outcome variables are the ratio of the counterfactual to the factual. Bootstrapped 95% confidence intervals in brackets (see Section I.2.4).

I.G.10 Measurement of labour input

Instead of using man-days as labour input L_{jt} we could also use the wage bill as labour input , so $L_{jt}^{alt}=L_{jt} * w_{jt}$. If higher skill correlates with higher salaries, this alternative measurement of labour input accounts for difference in quality of labour across plants. Rerunning the entire analysis with L_{jt}^{alt} yields the results reported in Table I.12, I.13 and I.14. I omit confidence intervals for simplicity, but almost all point estimates are within the range of the confidence intervals of the baseline version.

Table I.12: Welfare gains in billion rupees using L_{jt}^{alt}

	Comp. Variation			Profits			Total welfare		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	0.15	0.08	0.22	0.12	0.04	0.16	0.28	0.12	0.38
2001	1.79	0.20	2.09	1.32	0.12	1.51	3.11	0.31	3.60
2002	3.53	0.25	3.94	2.49	0.17	2.74	6.01	0.42	6.68
2003	2.65	0.47	3.04	1.36	0.23	1.52	4.01	0.70	4.56
2004	2.95	0.93	3.63	0.91	0.28	1.08	3.86	1.21	4.71
2005	2.39	0.62	3.29	1.56	0.39	2.16	3.96	1.01	5.45
2006	2.55	0.47	3.09	1.56	0.24	1.87	4.11	0.71	4.96
2007	1.57	0.66	2.06	1.01	0.38	1.34	2.58	1.04	3.39
2008	0.25	0.25	0.70	0.28	0.18	0.70	0.53	0.43	1.40
2009	2.83	0.42	3.18	2.00	0.23	2.19	4.83	0.65	5.37
2010	0.61	0.41	0.87	0.25	0.17	0.37	0.86	0.58	1.24
2011	2.44	0.63	3.48	2.20	0.42	3.21	4.64	1.05	6.69
2012	1.99	0.58	2.38	0.93	0.33	1.24	2.93	0.91	3.62
Total	25.70	5.96	31.97	16.00	3.18	20.09	41.70	9.14	52.07
Per year	1.98	0.46	2.46	1.23	0.24	1.55	3.21	0.70	4.01

Notes: The table shows the respective welfare gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both in billion rupees. The last three columns sum the consumer side compensating variation and the profits for total welfare gains. The last two rows report the total across all years and the implied average per year.

Table I.13: Compensating variation as share of consumer expenditure and profit growth using L_{jt}^{alt}

	Comp. Variation			Profit growth		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	0.13	0.06	0.19	1.25	1.09	1.34
2001	0.27	0.03	0.32	1.60	1.05	1.69
2002	0.34	0.02	0.38	1.77	1.05	1.85
2003	0.17	0.03	0.19	1.31	1.05	1.34
2004	0.17	0.05	0.21	1.18	1.05	1.21
2005	0.16	0.04	0.22	1.34	1.08	1.47
2006	0.13	0.02	0.15	1.25	1.04	1.31
2007	0.11	0.05	0.14	1.21	1.08	1.27
2008	0.03	0.03	0.08	1.10	1.06	1.25
2009	0.18	0.03	0.21	1.44	1.05	1.48
2010	0.06	0.04	0.08	1.08	1.06	1.12
2011	0.13	0.03	0.19	1.43	1.08	1.62
2012	0.13	0.04	0.15	1.20	1.07	1.26
Total	0.16	0.04	0.19	1.31	1.06	1.39

Notes: The table shows the respective welfare gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both. The first three columns express the compensating variation per unit purchased as a share of the unit price (i.e. as share of expenditure on the products in the sample). The profit ratio is total profits in the counterfactual divided by total profits in the factual equilibrium.

Table I.14: Physical output and productivity ratios using L_{jt}^{alt}

	Output ratio			Material productivity			Labour productivity		
	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
2000	1.29	1.15	1.43	0.90	1.09	0.96	0.79	0.78	0.74
2001	1.74	1.08	1.88	0.78	1.04	0.81	1.05	0.85	0.83
2002	2.12	1.07	2.27	0.67	1.03	0.71	1.45	0.86	1.13
2003	1.40	1.07	1.46	0.89	1.04	0.93	0.93	0.84	0.86
2004	1.40	1.11	1.49	0.61	1.04	0.66	1.63	0.88	1.32
2005	1.45	1.10	1.64	0.84	1.03	0.89	1.20	0.80	0.88
2006	1.31	1.05	1.39	0.90	1.01	0.92	1.04	0.89	0.91
2007	1.31	1.11	1.42	0.93	1.04	0.96	1.08	0.83	0.93
2008	1.11	1.09	1.31	0.94	1.03	0.94	0.96	0.88	0.90
2009	1.51	1.07	1.59	0.76	1.02	0.78	0.97	0.89	0.88
2010	1.14	1.10	1.22	0.96	1.01	0.97	1.01	0.92	0.98
2011	1.50	1.12	1.74	0.71	1.04	0.74	1.23	0.82	0.98
2012	1.41	1.11	1.53	0.71	1.04	0.74	1.12	0.86	1.07
Total	1.43	1.09	1.55	0.77	1.03	0.81	1.11	0.86	0.95

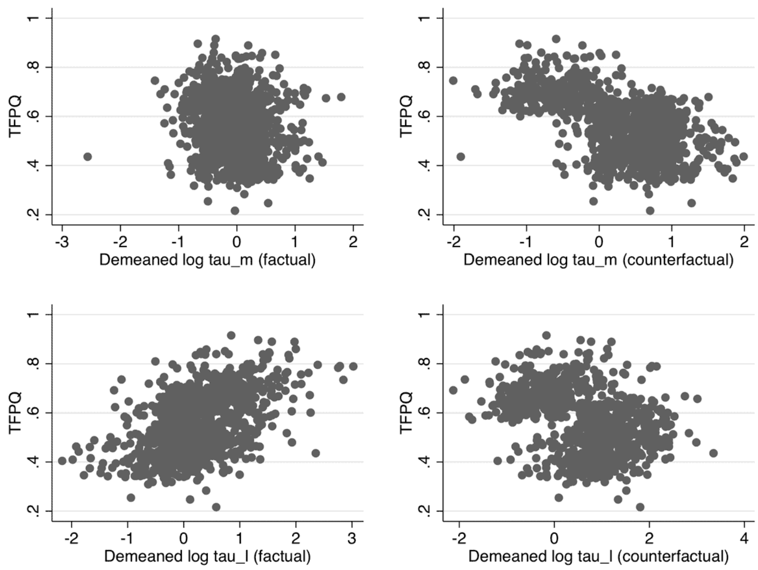
Notes: The table shows the respective gains from eliminating material distortions ($\tilde{\tau}_{jt}^{M}$), labour distortions ($\tilde{\tau}_{jt}^{L}$) or both. Outcome variables are the ratio of the counterfactual to the factual.

I.G.11 Wedges as tax income

When the $\boldsymbol{\tau}$ are interpreted as taxes, removing differences in $\boldsymbol{\tau}$ also affects government revenues. In Table I.15, I add the tax revenue changes to the welfare calculations. The difference to the baseline version is statistically significant when using bootstrapped differences.

I.G.12 A counterfactual of more negatively correlated TFPQ and τ

Figure I.26: Correlation between TFPQ and τ



Notes: Plotted are the plant TFPQ (Ω_{jt}) and the annually demeaned $\ln(\tau_{jt}^{M})$ or $\ln(\tau_{jt}^{L})$. The left panels correspond to the factual equilibrium. The right panel corresponds to an alternative counterfactual, where the $\boldsymbol{\tau}$ are reduced for all above average productivity plants and increased for all below average productivity plants until the correlation is -0.5.

Figure I.26 shows the correlations between plant level TFPQ ($\boldsymbol{\Omega}$) and $\boldsymbol{\tau}$. In the factual equilibrium (the two left panels), there is no significant correlation between Ω and τ^M and a slight positive correlation with τ^L. I construct an alternative counterfactual where the $\boldsymbol{\tau}$ are set such that their weighted geometric average is preserved, but the $\boldsymbol{\tau}$ are reduced for all above average productivity plants and increased for all below average productivity plants. This does not remove misallocation. The two right panels depict this counterfactual and the resulting correlation between $\boldsymbol{\Omega}$ and $\boldsymbol{\tau}$. There are substantial welfare gains from moving from the left to the right panels. The size of the welfare gains depends on how strongly the (artificial) correlation differs in the counterfactual compared to the factual. In this case (correlation to both around -0.5) the welfare gains are roughly half of the welfare gains of the baseline results where the $\boldsymbol{\tau}$ are removed instead.

Table I.15: Welfare gains in billion rupees with tax income adjustments

	Total welfare			Taxes on materials			Taxes on labour			Total welfare with taxes		
	$\tilde{\tau}^M_{jt}$	$\tilde{\tau}^L_{jt}$	*both*	$\tilde{\tau}^M_{jt}$	$\tilde{\tau}^L_{jt}$	*both*	$\tilde{\tau}^M_{jt}$	$\tilde{\tau}^L_{jt}$	*both*	$\tilde{\tau}^M_{jt}$	$\tilde{\tau}^L_{jt}$	*both*
2000	0.25	0.21	0.45	-0.08	0.01	-0.08	0.05	0.02	0.07	0.23	0.24	0.44
	[0.19,0.33]	[0.1,0.3]	[0.37,0.53]	[-0.09,-0.05]	[0.01,0.03]	[-0.1,-0.05]	[0.03,0.07]	[0,0.03]	[0.03,0.09]	[0.15,0.32]	[0.12,0.35]	[0.33,0.54]
2001	2.57	0.6	3.43	0.05	0.19	0.13	0.85	0.12	0.97	3.48	0.91	4.53
	[1.94,3.37]	[0.29,0.91]	[2.78,4.15]	[-0.38,0.5]	[0.1,0.29]	[-0.35,0.55]	[0.45,1.18]	[0.03,0.24]	[0.47,1.33]	[2.29,4.68]	[0.42,1.42]	[3.27,5.76]
2002	5.31	0.88	6.56	0.27	0.48	0.26	1.26	0.2	1.24	6.85	1.56	8.05
	[3.92,6.51]	[0.43,1.4]	[5.33,7.45]	[-0.4,0.86]	[0.28,0.75]	[-0.38,0.83]	[0.61,1.65]	[0.05,0.4]	[0.54,1.58]	[4.68,8.14]	[0.77,2.51]	[6.14,8.99]
2003	3.59	1.28	4.7	-0.14	0.21	-0.06	1.01	0.28	1.15	4.46	1.76	5.8
	[2.5,4.9]	[0.64,1.92]	[3.83,5.77]	[-0.66,0.34]	[0.06,0.35]	[-0.63,0.42]	[0.43,1.4]	[0.07,0.54]	[0.53,1.62]	[2.73,6.15]	[0.84,2.75]	[4.3,7.46]
2004	3.35	2.15	4.9	-0.43	0.03	-0.43	-0.11	0.13	-0.18	2.81	2.31	4.29
	[2.53,4.1]	[1.1,2.97]	[4.06,5.34]	[-0.69,-0.35]	[-0.13,0.15]	[-0.69,-0.02]	[-0.47,0]	[-0.04,0.19]	[-0.55,0]	[1.9,3.35]	[1.22,3.09]	[3.22,4.75]
2005	3.23	2.2	6.57	-0.37	0.72	-0.22	0.9	0.55	1.42	3.77	3.47	7.77
	[2.42,4.24]	[0.99,3.73]	[5,8.59]	[-0.76,0.09]	[0.33,1.31]	[-0.9,0.31]	[0.49,1.25]	[0.13,1.12]	[0.62,1.91]	[2.41,5.1]	[1.52,6.02]	[5.43,10.17]
2006	3.23	1.42	4.97	-0.1	0.1	0.02	0.88	0.22	1.09	4.01	1.73	6.08
	[2.43,4.32]	[0.69,2.13]	[3.97,6.4]	[-0.56,0.33]	[-0.01,0.21]	[-0.53,0.45]	[0.46,1.25]	[0.05,0.42]	[0.49,1.53]	[2.66,5.56]	[0.79,2.69]	[4.47,7.97]
2007	2.1	2.28	4.2	-0.44	-0.11	-0.43	0.3	0.3	0.54	1.96	2.47	4.31
	[1.68,2.64]	[1.14,3.24]	[3.4,5.02]	[-0.81,-0.3]	[-0.3,0.01]	[-0.82,-0.24]	[0.11,0.41]	[0.06,0.54]	[0.13,0.8]	[1.41,2.53]	[1.23,3.54]	[3.21,5.32]
2008	0.21	0.97	1.03	0.13	0	0.19	0.13	0.18	0.29	0.47	1.16	1.51
	[-0.06,0.57]	[0.44,1.51]	[0.6,1.49]	[0,0.27]	[-0.11,0.08]	[-0.01,0.39]	[0.06,0.2]	[0.04,0.34]	[0.11,0.48]	[0.17,0.89]	[0.51,1.83]	[1.01,2.13]
2009	3.83	1.22	4.76	0.21	0.1	0.25	1.22	0.15	1.28	5.26	1.47	6.3
	[2.68,5.62]	[0.56,1.89]	[3.76,6.22]	[-0.43,0.9]	[0,0.17]	[-0.41,0.89]	[0.63,1.8]	[0.03,0.27]	[0.63,1.89]	[3.18,7.93]	[0.67,2.29]	[4.34,8.79]
2010	0.7	1.22	1.55	-0.06	-0.01	0.03	0.08	0.14	0.19	0.72	1.35	1.76
	[0.54,0.91]	[0.59,1.84]	[1.07,1.98]	[-0.11,0.01]	[-0.14,0.06]	[-0.09,0.19]	[0.01,0.1]	[0.02,0.24]	[0.03,0.27]	[0.5,0.93]	[0.68,1.99]	[1.2,2.27]
2011	3.35	1.95	6.5	0.68	0.91	1.39	1.16	0.51	1.61	5.18	3.37	9.51
	[2.18,5.08]	[0.96,2.98]	[5.06,10.16]	[-0.24,1.58]	[0.46,1.46]	[-0.07,2.53]	[0.62,1.66]	[0.13,1.01]	[0.73,2.28]	[2.99,7.52]	[1.59,5.31]	[6.99,12.81]
2012	2.17	1.88	3.66	-0.03	0.14	0.17	0.35	0.4	0.63	2.49	2.42	4.47
	[1.41,3.29]	[0.88,2.96]	[2.99,4.74]	[-0.22,0.18]	[-0.07,0.31]	[-0.16,0.35]	[0.12,0.49]	[0.1,0.77]	[0.2,1.03]	[1.49,3.77]	[1.1,3.87]	[3.27,5.81]
Total	33.9	18.26	53.28	-0.3	2.78	1.2	8.09	3.19	10.31	41.68	24.23	64.8
	[24.87,43.82]	[8.74,27.27]	[44.27,64.09]	[-4.74,3.59]	[1.1,4.64]	[-4.28,5.8]	[3.89,11.07]	[0.72,5.84]	[4.36,14.29]	[26.9,55.02]	[11.37,37.39]	[48.41,81.18]
Per year	2.61	1.4	4.1	-0.02	0.21	0.09	0.62	0.25	0.79	3.21	1.86	4.98
	[1.91,3.37]	[0.67,2.1]	[3.41,4.93]	[-0.36,0.28]	[0.08,0.36]	[-0.33,0.45]	[0.3,0.85]	[0.06,0.45]	[0.34,1.1]	[2.07,4.23]	[0.87,2.88]	[3.72,6.24]

Notes: The table shows the respective welfare gains from eliminating material distortions ($\tilde{\tau}^M_{jt}$), labour distortions ($\tilde{\tau}^L_{jt}$) or both in billion rupees. Bootstrapped 95% confidence intervals in brackets (see Section I.2.4).

Table I.16: Determinants of plant level changes in input productivities

Dependent variable:	Δ plant mat. prod. (log)			Δ plant lab. prod. (log)		
Counterfactual:	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*	$\tilde{\tau}_{jt}^{M}$	$\tilde{\tau}_{jt}^{L}$	*both*
Dep. var. (exp) 10th - 90th percentiles:	[.8, 1.4]	[.9, 1.2]	[.8, 1.5]	[.8, 1.4]	[.4, 1.7]	[.4, 1.8]
τ_{jt}^{M} (log demeaned)	-1.03***	-0.01**	-0.89***	0.96***	-0.00**	0.36***
	(0.00)	(0.04)	(0.00)	(0.00)	(0.04)	(0.00)
τ_{jt}^{L} (log demeaned)	0.00	1.00***	0.48***	0.00	-1.00***	-0.91***
	(0.33)	(0.00)	(0.00)	(0.32)	(0.00)	(0.00)
TFPQ (log)	-0.01**	-0.00	-0.01**	-0.01**	-0.00	-0.00**
	(0.04)	(0.33)	(0.02)	(0.04)	(0.33)	(0.02)
Markup	-0.07***	-0.01**	-0.05***	-0.06***	-0.00**	-0.02***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
N	979	979	979	979	979	979
R^2	0.99	0.99	0.99	0.99	1.00	1.00

Notes: Coefficients are standardized. p-values in parentheses are based on clustered standard errors at the plant level. Dependent variables are the log of the ratio of the input productivities in the counterfactual to the factual (i.e. the difference Δ in logs). Only observations with an initial demand elasticity < -1 are included.

I.G.13 Heterogeneity in plant input productivity changes

Table I.16 reports the 10th and 90th percentile of the input productivity ratios (between counterfactual and factual) in the table header. There is substantial heterogeneity in the comparative statics across plants. For the third column, for example, where both distortions are removed, the 10th percentile is a decrease in material productivity of 20% while the 90th percentile is an increase of 50%.

I run regressions where the dependent variable is the change in the log input productivities between the counterfactual and the factual equilibrium at the level of plants. I regress this on the two $\boldsymbol{\tau}$, plant productivity TFPQ (Ω) and initial markups. The table reports standardized coefficients. The R^2 is extremely high, and the variation in input productivity growth across plants can be well explained by the initial input distortions.

Removing distortions in one input has large effects on the changes in plant level input productivity of the same input. This is intuitive. Plants with a high τ^M face costs using materials, and when these costs are reduced in the counterfactual, the plant has incentives to use the previously constrained materials relatively more intensively, thereby decreasing the ratio of output to materials. This also provides intuition of why aggregate input productivities do not increase when distortions in both markets are removed. The plants that grow (i.e. previously constrained through high τ) also use the input relatively more intensively and addressing both distortions compensates each distortion's effects on the input productivities.

Table I.17: Change in markup variation

	Factual	Counterfactual $\tilde{\tau}_{jt}^{M}$	Counterfactual $\tilde{\tau}_{jt}^{L}$	Counterfactual $\tilde{\tau}_{jt}^{M}$ and $\tilde{\tau}_{jt}^{L}$
Median	1.26	1.25	1.27	1.26
95th/5th	1.30	1.20	1.32	1.23
90th/10th	1.21	1.14	1.22	1.16
75th/25th	1.09	1.07	1.10	1.07

Notes: The table shows the median and ratios of percentiles of the markups in different equilibria, pooled across years.

I.G.14 Changes in markup dispersion

Table I.17 reports the median and the ratios of markups of the the 95th to th 5th percentile, the 90th to the 10th percentile and the 75th to the 25th percentile across all years. The first column reports the ratios for the factual equilibrium and the other columns for the corresponding counterfactual equilibria. The median markup is similar, but the variation in markups is lower in the counterfactuals except for the one where only labour misallocation is removed ($\tilde{\tau}_{jt}^{L}$).

I.G.15 Modifying and discussing Hsieh and Klenow (2009)

The way Hsieh and Klenow (2009) calculate losses from misallocation is by calculating the output gap $\frac{Y}{Y^*}$:

$$\frac{Y}{Y^*} = \prod_{s=1}^{S} \frac{\left(TFP_s K_s^{\alpha_s} X_s^{\beta_s} L_s^{1-\alpha_s-\beta_s}\right)^{\theta_s}}{\left(TFP_s^* K_s^{\alpha_s} X_s^{\beta_s} L_s^{1-\alpha_s-\beta_s}\right)^{\theta_s}} = \prod_{s=1}^{S} \frac{TFP_s^{\theta_s}}{TFP_s^{*\theta_s}} = \prod_{s=1}^{S} \left[\sum_i \left(\frac{A_{si}}{\overline{TFP_s^*}} \frac{\overline{TFPR_s}}{TFPR_{si}} \right)^{\sigma-1} \right]^{\frac{\theta_s}{\sigma-1}} \tag{I.32}$$

where I follow the notation in their paper, and adding X_{si} as material input for firm i in sector s with corresponding output elasticity β_s. The other components are:

$$\frac{A_{si}}{\overline{TFP_s^*}} = \frac{(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}} / K_{si}^{\alpha_s} (vX_{si})^{\beta_s} (wL_{si})^{1-\alpha_s-\beta_s}}{\left[\sum_i [(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}} / K_{si}^{\alpha_s} (vX_{si})^{\beta_s} (wL_{si})^{1-\alpha_s-\beta_s}]^{\sigma-1} \right]^{\frac{1}{\sigma-1}}} \tag{I.33}$$

$$\frac{TFPR_{si}}{\overline{TFPR_s}} = \frac{P_{si}Y_{si}}{P_sY_s} \frac{(\sum_i K_{si})^{\alpha_s} (\sum_i vX_{si})^{\beta_s} (\sum_i wL_{si})^{1-\alpha_s-\beta_s}}{K_{si}^{\alpha_s} (vX_{si})^{\beta_s} (wL_{si})^{1-\alpha_s-\beta_s}} \tag{I.34}$$

As in their paper, I calibrate the demand elasticity σ to 3, $\theta_s = \frac{P_sY_s}{PY}$ because of the perfect competition assumption in the final good sector, and I take the output elasticities from

Table I.18: Output gains from replication and extension of Hsieh and Klenow (2009)

	2 factor value added model			3 factor gross output model		
	CRS	RS=0.95	RS=0.92	CRS	RS=0.95	RS=0.92
Manufacturing	95 [93,98]	37 [36,40]	12 [10,15]	41 [39,42]	2 [1,3]	-16 [-16,-15]
Basic metals	146 [127,175]	71 [64,87]	37 [30,53]	54 [51,59]	10 [9,12]	-9 [-11,-7]
Basic metals with estimated elasticities	-	-	-	32 [29,34]	-3 [-5,-1]	-19 [-20,-17]

Notes: Calculated is the average gain from the four periods 2000-2003, in percent. The square brackets contain the minimum and maximum of the four years. The columns have different returns to scale assumptions, and the elasticities are scaled to fit them accordingly. The elasticities corresponding to the first two rows are from the NBER-CES database. The last row uses the estimated elasticities from this chapter, and they are scaled to fit the returns to scale accordingly. The last column of this row just uses the estimated returns to scale (0.92).

the CES-NBER database for each sector.[123]

The first point to note is that aggregate sectoral output Y_s can be written in two ways in their paper:

$$Y_s = TFP_s^* K_s^{\alpha_s} X_s^{\beta_s} L_s^{1-\alpha_s-\beta_s} \tag{I.35}$$

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{I.36}$$

where the first equation is the aggregate production function, and the second is the utility function of the representative consumer.[124] The output gap in Equation (I.32) can not only be interpreted as pure production side TFP gap, but equally as gap in utility. The latter interpretation is in line with the results of this chapter.

Second, I replicate their analysis and adjust two assumptions. First I either use their 2 factor value added production model,[125] or add materials as a third factor for gross output production functions. The second adjustment are the assumed returns to scale, where I use their constant returns to scale, 0.95, or my estimated 0.92 returns to scale. The results are reported in Table I.18. I replicate the analysis for the entire manufacturing sector (first row), or for the basic metals sector only (second and third row). In addition, I use the

[123] Trimming of outliers is analogously done to their paper by pooling sectors in a year and trimming top and bottom 1% of both ratios, $\frac{A_{si}}{TFP_s^*}$ and $\frac{TFPR_{si}}{TFPR_s}$, and then recalculating all sector level variables.

[124] The demand function is derived through a cost minimisation of a representative purchaser in industry s, who minimises the expenditure of all varieties in industry s ($\sum_i P_{si}Y_{si}$) subject to $Y_s \geq \bar{Y}_s$. That is subject to some minimum level of CES output or "utility" where $Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$, which describe the preferences over varieties with an elasticity of substitution σ.

[125] That is dropping X_{si} from the analysis and using value added instead of revenues for $P_{si}Y_{si}$.

estimated input elasticities instead of the CES-NBER elasticities in the third row.[126]

In their paper they report output gains between 100 and 128 percent for the years 1987 and 1994 respectively. I use the years 2000 to 2003 and calculate gains of 95 percent, close to their estimates from earlier years. When I reduce the assumed returns to scale to 0.95 or 0.92, the hypothetical gains fall dramatically to only a tenth of the gains under CRS.[127] Moving to a gross output model shrinks the gains as well.[128] With 0.92 returns to scale, some of the gains are actually negative – India would be better off with misallocation than without. Part of the problem why these strange results arise in their model lies with the definition of the counterfactual $\tilde{\boldsymbol{\tau}}$. In this chapter, I use a geometric average instead of unity, which addresses measurement error in the distortions that is constant across plants, for example in the output elasticities (that are scaled by the returns to scale), as discussed in Section I.2.4.

[126]When I use my sample of cast iron plants only and apply my estimated elasticities to their model, I obtain a positive correlation between the implied logged TFPQ from their model, and my estimated logged TFPQ. The R^2 is 0.27, so there is a substantial difference between the productivities.

[127]For a theoretical analysis of the sensitivity of these types of models to the returns to scale assumption see Hopenhayn (2014*b*).

[128]Gandhi et al. (2017) explore the theoretical and empirical relationships between value added and gross output production functions. See also Dias et al. (2016) which find larger gains in a Hsieh and Klenow (2009) 2 factor value added type model (their Table A1) than in a gross output 3 factor model (their Table 4).

Chapter II

Input misallocation and geographical supplier access

II.1 Introduction

A growing literature documents losses from misallocation of production inputs. While the vast majority of studies focuses on capital misallocation or labour misallocation, for example from financial frictions or hiring and firing laws, we know much less about the degree and the determinants of misallocation of input materials. Chapter I shows that the welfare losses from misallocation of material inputs are substantial for the cast iron industry in India, and are even higher than the losses from misallocation of labour.

This chapter asks: What is causing these costly input material distortions and misallocation losses? I present evidence that access to suppliers is a significant driver of these input distortions in the Indian cast iron industry. This is the first study to show that precisely estimated input distortions that lead to misallocation capture differences in supplier access.

The steel industry relies on extensive shipping of heavy and bulky material inputs. Geography and transport infrastructure, particularly railroads, are thus important features in the production process. Any issues with sourcing are likely to play a role in the documented misallocation from material input distortions. I begin this chapter by analysing the issues in Indian freight transport. Sourcing inputs through the transportation network in India is characterised by frequent delays and uncertainty. Infrastructure is outdated, leading to breakdowns, freight trains share congested tracks with passenger trains, and there are state border checkpoints for tax purposes that delay shipments, to name a few. With longer sourcing routes, these issues are likely to become more severe.

Shipping fees also increase with longer sourcing routes. The advantage of this analysis is that I observe plant specific input prices that are measured at the factory gate. They explicitly contain input shipping fees and the model accounts for differences in these observed input prices. Any differences in estimated input distortions are therefore net of input shipping fees and the measure of supplier access is not simply picking up those. In the trade literature, shipping fees are often called direct trade costs. There is a large body of evidence that suggest that direct trade costs cannot account for the implied total costs of trade (Anderson and Van Wincoop, 2004). There are indirect costs of trade, such as search costs, contracting costs or costs associated with delay and uncertainty. The freight transport issues in India, such as uncertainty, are examples of such indirect costs of trade.[1]

While the estimated input distortions are net of shipping fees, they would certainly capture differences in indirect trade costs. And if indirect trade costs increase with poorer geographical access to suppliers, the estimated material input distortions should decrease with better supplier access. This chapter tests this hypothesis. I estimate input distortions for the cast iron industry as described in Chapter I. The advantage of this methodology is that distortions are cleaned from differences in fundamentals across plant, such as demand conditions and market power, productivity, and production techniques.

I construct a measure of supplier access by combining the cost to reach suppliers with the size of supplier industries in the around 540 districts of India. The cost to reach suppliers is based on the fastest path between any bilateral pairs of districts. I collect geo-located data on the entire rail and road infrastructure in India. Using information on the types of rails and roads (e.g. motorways vs. tertiary roads), I construct a weighted network graph to compute the fastest path between district pairs.

Over the sample period from 2000 to 2012, there was hardly any variation in the railway infrastructure, which is the main mode of transport for this industry. Therefore, the infrastructure component is time invariant in the measure of supplier access. Even if there was time-variation, placement of railway infrastructure, as well as the location choice of plants is non-random.[2] To address these issues, I condition the analysis on district fixed effects, and therefore use the variation in supplier access that is driven by the expansion and contraction of supplier industries in distant districts over time.

A one standard deviation increase in supplier access is associated with almost a third of a standard deviation decrease in the input material distortion. The estimate is robust to using lagged supplier access, addressing potential reverse causality concerns. To provide

[1] While storage can smooth over uncertainties, it is costly and therefore also part of indirect costs of trade.

[2] Endogenous infrastructure placement is a central challenge in the literature on the effects of infrastructure investments. There are some strategies to address this. Faber (2014), for example, constructs a hypothetical infrastructure network based on construction costs and a minimum spanning tree. Banerjee et al. (2012) use the areas on the straight line between start and end point of a transportation link.

more evidence for the causality of the relationship, I present three types of placebo test. The first shows that supplier access has no significant relationship with the estimated labour distortion. The second shows that a measure of access to irrelevant "supplier" industries, e.g. textiles, rubber or food, is not related to the material input distortion. The third placebo test shows that access to markets on the output side cannot explain the distortion on the input side. It is thus only the access to relevant input suppliers that is associated with the material input distortions that drive misallocation losses.

The policy implications of this analysis are nuanced. Differences in shipping fees mirror the geographic reality that shipping goods across space is costly. While theoretically beneficial for manufacturing, it is also costly to build infrastructure to equalise shipping fees for all plants. The findings of this chapter are conditional on the observed differences in input shipping fees, however. The distortions pick up the indirect trade costs, and those are in turn driven by supplier access. Addressing reliability, delay and shipping uncertainty, for example, could reduce the impact of supplier access on indirect trade costs, and could thus reduce misallocation losses.

This chapter contributes to a rapidly growing literature on the determinants of misallocation in manufacturing industries. The focus in the literature has been on financial frictions[3], labour market distortions[4], trade barriers[5] or land misallocation and property rights[6]. Other sources for misallocation include R&D subsidies (Acemoglu et al., 2018), informational frictions (David et al., 2016; Allen, 2014), entry barriers and innovation (Peters, 2013), differential state taxes (Fajgelbaum et al., 2018), or misspecification and

[3]See for example Buera et al. (2011). These also include limited enforcement (Amaral and Quintin, 2010), imperfect corporate control (Caselli and Gennaioli, 2013), limits on issuable debt (Midrigan and Xu, 2014), borrowing constraints (Moll, 2014; Gopinath et al., 2017; Wang, 2017), financial subsidies (Buera et al., 2013) and other various sources of capital misallocation (David and Venkateswaran, 2017; Gorodnichenko et al., 2018). There is ample micro-evidence on capital misallocation in the developing context, summarised by Banerjee and Duflo (2005).

[4]An early paper in this literature is Hopenhayn and Rogerson (1993) on the the ease of job reallocation. Lagos (2006), Guner et al. (2008) and Garicano et al. (2016) study the impacts of versions of size-dependent labour laws on misallocation and welfare. Hsieh et al. (Forthcoming) study the misallocation of talent based on racial and gender discrimination. Munshi and Rosenzweig (2016), Hsieh and Moretti (2019) and Bryan and Morten (Forthcoming) present evidence on spatial misallocation of labour due to lack of local insurance in India, migration constraints in Indonesia, and housing constraints in the US respectively.

[5]There is a trade literature on misallocation or the "competitive" effects of trade openness. Higher output market competition can affect markups and reallocate market shares to increase allocative efficiency (Epifani and Gancia, 2011; Barseghyan and DiCecio, 2011; Edmond et al., 2015; Galle, 2016; Arkolakis et al., 2018, e.g.). A smaller literature also analyses the effect of trade openness on input markets (e.g. Amiti and Konings, 2007; Goldberg et al., 2010; De Loecker et al., 2016).

[6]Duranton et al. (2015) analyse the degree of misallocation comparing Indian districts using an Olley and Pakes (1996) decomposition. They find that misallocation of land and buildings is large, partially explained by the repeal of the Urban Land Act. For studies in the agricultural context, see Adamopoulos and Restuccia (2014*a*), Chari et al. (2017) or Chen et al. (2017).

measurement error[7]. Atalay (2014) documents dispersion of material input prices and its possible reasons in two narrow US industries. Conceptually, the input distortions in the misallocation literature are different from input price differences. An advantage of this chapter and Chapter I is that I observe material input prices, and can disentangle misallocation estimates from those.

Most of the literature on measuring misallocation abstracts from intermediate inputs entirely by using value added production functions.[8] There is even less work on potential determinants of misallocation of material inputs. As an exception, Boehm and Oberfield (2018) study misallocation of intermediate inputs in India due to court congestion that generates a hold up problem in their model. In current work in progress, Hornbeck and Rotemberg (2019) examine the impact of railroad expansion on input reallocation in historic US manufacturing using a growth decomposition.[9] This chapter aims to address the literature gap in misallocation of input materials.

This chapter is also related to the literature studying the effect of transport networks on productivity and welfare.[10] Perhaps most closely related are papers that examine the effects of changes in market access.[11] Donaldson and Hornbeck (2016) analyse the contribution of historical railroads on agricultural productivity in the US. Alder (2017) compares the road construction projects in India with a counterfactual following a Chinese infrastructure model. Allen and Atkin (2016) estimate the impact of road expansion in India on the volatility of grain prices and farmer production choices, and Huang and Xiong (2018) analyse Chinese road expansion. As in this chapter, they use geo-located transport infrastructure data to compute intra-national trade costs and a measure of market access.

Two aspects set this chapter apart from this literature. First, these studies use access to output markets where the entire economic activity of other regions are taken into account.

[7] Asker et al. (2014) show that adjustment costs can account for capital "misallocation". Bils et al. (2017), Rotemberg and White (2017) and Haltiwanger et al. (2018) show that measurement error, data cleaning procedures and misspecification respectively affect misallocation estimates. The estimates used in this chapter are based on the carefully specified and estimated model in Chapter I which aims to minimise misattribution of fundamental heterogeneity to misallocation distortions.

[8] Exceptions are Jones (2013) or Dias et al. (2016), for example.

[9] The citation of Hornbeck and Rotemberg (2019) refers to a presentation at LSE and conversations with one of the authors. At the time of submission, there was no working paper version available.

[10] There is a longer literature that analyses the effects of transport infrastructure starting with Fogel (1964), which as received growing attention more recently. Allen and Arkolakis (2014), Allen and Arkolakis (2019) and Fajgelbaum and Schaal (2017) develop frameworks for estimating the impact of transport infrastructure investment on welfare in spatial equilibrium, where the latter two also account for traffic congestion. See Redding and Rossi-Hansberg (2017) for a review.

[11] Redding and Venables (2004) provide a theoretical foundation of both market access on the output side and supplier access on the input side. Redding (2010) and references therein provide a summary of the earlier literature on market access.

In contrast, I am using a measure of supplier access on the input side. Moreover, I only take access to *relevant* potential suppliers into account. Naturally, to study distortions on intermediate inputs, the access to these suppliers instead of market access is the object of interest. Second, while these studies rely on variation in infrastructure construction for changes in market access, I use growth in distant supplier industries for variation in supplier access.

A related literature estimates the intra-national costs of trade from price differences instead of using fastest path algorithms (Fackler and Goodwin, 2001; Anderson and Van Wincoop, 2004). Donaldson (2018) infers trade costs from price differences of single origin goods and finds that railroads decreased trade costs in colonial India. Atkin and Donaldson (2015) use price differentials at the barcode level for Ethiopia and Nigeria. Importantly, they adjust for markups which would otherwise distort trade costs estimates when using prices. Asturias et al. (2018) examine the impact of road construction on allocative efficiency in India using price gaps. The misallocation in their model comes from dispersion of markups. In contrast, this chapter accounts for variable endogenous markups as fundamental differences in demand, and misallocation stems from distortions on the input side. I provide some evidence that the findings are robust to controlling for proxies for monopsony power as well.

The remainder of the chapter starts with a discussion of the relevant issues in Indian freight transportation in Section II.2, which are likely to be captured as indirect trade costs in the estimated input distortions. Section II.3 describes the plant level data as well as the construction of a weighted network graph from geo-located rail and road data. Section II.4 first describes how I recover input distortions. I then construct a measure of supplier access and discuss the identification strategy and main specification. The results along with placebo tests and discussions are presented in Section III.4 before Section III.5 concludes.

II.2 Freight transport issues in India

India's freight transport infrastructure has often been criticised by industry and policy makers alike. Its poor state has been identified as a key constraint for the efficient running and expansion of heavy industry and steel in particular (NCAER, 2015). Inadequate road quality and severe congestions result in high and uncertain transit times, with average truck speeds at around a third of those in developed countries (NTDPC, 2014).

While the share of freight traffic on rail is at around 30% overall (NTDPC, 2014), it is more important for the steel industry with a rail share of around 70% (EY, 2014). There are issues with rail shipping that mirror the problems with road shipping. Passenger trains and freight trains share the same tracks, leading to congestion. Further delays are frequent due to outdated infrastructure operating above capacity limits, breakdowns, different rail gauges requiring different wagons, and numerous state border checkpoints for tax purposes which can take days or weeks to clear (EY, 2013; NTDPC, 2014; EY, 2014; NCAER, 2015). Freight trains are only travelling at an average speed of 25km/h (Appendix II.A.1) and the Government of India is investing heavily to increase the speed and reliability with current rail infrastructure projects (NTDPC, 2014). Van Leemput (2016) estimates that India faces higher internal than international trade barriers.

Steel plants require to ship a large amount of heavy and bulky inputs.[12] The above described issues suggest that there are two types of costs to shipping. One type are the shipping fees, the direct trade costs. The other types are indirect trade costs, such as delay and uncertainty. In fact, we know from the trade literature, that these "indirect" costs of trade are large (Anderson and Van Wincoop, 2004). They can also include search costs and contract enforcement costs (Startz, 2018). Using a specific railway line in India, Firth (2017) presents evidence that the *variance* of shipping time causes the bulk of costs to firms and constrains their operation.

Crucially, the shipping fees are explicitly included in the (factory gate) input prices and are thus observed. The indirect trade costs, however, are not accounted for. Differences in indirect trade costs would therefore be captured by the estimated input distortion. Plants that need to source from further away are more likely to experience any of these issues in freight transportation and likely have higher indirect trade costs and distortions. The aim of this chapter is to test this hypothesis.

I use the Enterprise Survey in India from World Bank (2005*a*) for some motivating evidence. The survey asks firms whether transportation is an obstacle for their growth. Around a third of the firms answer that transportation is an obstacle. I regress the logged demeaned input distortion τ^M_{jt}, which I will derive below, on the district average of the responses in Table II.1. Plants that claim that they face transportation obstacles also have higher estimated material input distortions τ^M_{jt}.

[12]One tonne of steel requires the transportation of more than four tonnes of input materials (NCAER, 2015). Chapter I provides a more detailed description of the cast iron industry in India.

Table II.1: Transport obstacles as distortions: Some evidence using World Bank (2005*a*)

	(1)
Transport obstacle?	0.24**
	(0.10)
N	27
R^2	0.19

Notes: The regression is at the district level in 2005. The dependent variable is τ_{jt}^{M} demeaned by the weighted geometric mean and in logs. The independent variable is the district average of the survey question whether transportation is an obstacle to firm growth from World Bank (2005*a*). Year is 2005, and only respondents from the metals or minerals industry are used. Robust standard errors are in parentheses.

II.3 Plant data and infrastructure network

There are three main data sources that I use for this chapter. The first one is the ASI plant data, the second is geo-located data for administrative boundaries, and the third is the universe of geo-located roads and railways.

II.3.1 Plant data

The input distortions are recovered from plant level panel data from the Annual Survey of Industries (ASI) from 2000 to 2012. I focus on single product plants that produce a cast iron. Cast iron is an important product of the basic steel sector in India and is a sector that requires substantial transportation. The narrow focus delivers a clean estimate of input distortions, as discussed in Chapter I. Crucially, I use information on quantity and prices on the output and the input side. I end up with 926 observations for the final analysis. For a much more detailed account of the Indian cast iron sector, the plant level data, the rationale of focusing on single products and descriptive statistics I refer the reader to Chapter I.

II.3.2 Geographic data

The geo-located data of administrative boundaries is from Database of Global Administrative Areas (GADM) (2016). The data of transport infrastructure is from OSMF (2016).

Matching plant data with the location of firms

There are no exact geographical identifiers of plants in the ASI due to confidentiality. By matching the panel and the cross-sectional versions of the ASI, I obtain the districts of plants which are only contained in the latter version. The district centroids act as plant locations for the rest of the chapter.[13] I matched the ASI districts to the 594 geo-coded district data via fuzzy string matching within states, with extensive manual matching and checks until all districts were matched.

Information on transport infrastructure

I use data on railroads and roads. Transportation via inland waterways is negligible, as it is severely underdeveloped in India (NTDPC, 2014).[14] The share of imported materials in total materials is around 2% in quantity and value terms for this sample. I therefore ignore international sourcing. The transport network contains information about the type of each edge, for example broad vs. narrow gauge rails, or motorways vs. secondary roads. This information is used to assess the speed for each edge of the network.

There is, however, no temporal information on the opening of railroad tracks and roads. Figure II.4 in Appendix II.A.1 shows that the route kilometres of railways during the sample period for 2000-2012 only increased by 3.8%, almost all of it in the last 4 years. Furthermore, the average speed of goods trains was nearly constant (see Figure II.4), due to little investment in upgrading of existing infrastructure. I therefore treat the transport infrastructure as constant over time using a snapshot from the end of the sample period. For roads, the picture is slightly different, and there has been an increase in total road length (Ministry of Road Transport and Highways, 2016). Since the steel industry relies predominately on rails (see Section II.2), I ignore this temporal variation. I provide a robustness check using only railways, with very similar results.

Construction of geographical network

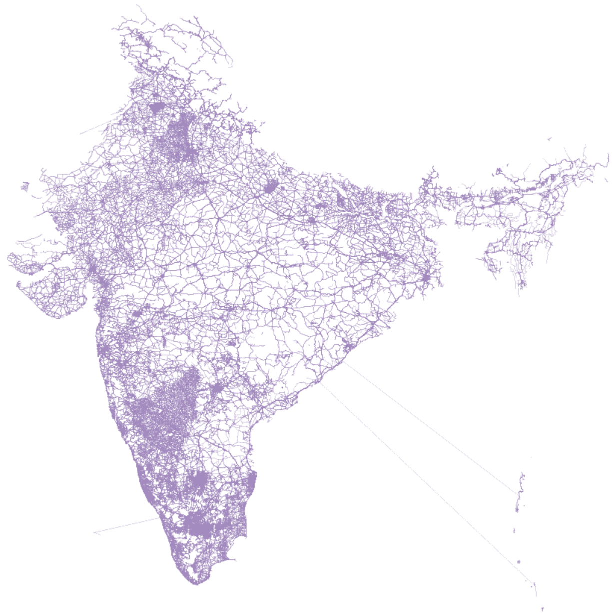
In order to run network analysis algorithms to calculate access to suppliers, we need a weighted undirected network graph with connected plant locations (district centroids). I prepare the infrastructure data by keeping only segments that can be used for shipping, i.e. deleting abandoned rail tracks, rural bridleways etc. I then perform a series of network preparation and cleaning tasks, for example, to make sure that road intersections contain nodes and that relevant nodes are snapped to each other. I connect the plant locations (district centroids) with a straight line to the nearest point in the network.[15] The full

[13] I also repeatedly drew random points in each district as plant location and estimate the effect for each draw. The average effect is close to the reported estimates.

[14] The share of transportation via waterways in India is less than 1% of tonne-km (Raghuram, 2004), at least an order of magnitude lower than in Bangladesh, the US, China or Germany (Rangaraj and Raghuram, 2007; NCAER, 2015)

[15] I am assigning a low speed to travelling on this edge, see Table II.2.

Figure II.1: Indian rail and road transport network



Notes: The map shows the undirected network graph of 1.6 mil. edges and 1.2 mil. nodes, based on Indian rail and road infrastructure.

network contains around 1.6 million edges and 1.2 million nodes, which can still be handled well with a standard computer and optimised network algorithms. Figure II.1 shows the entire network graph of roads and railways that are used for the analysis.

Speed assumptions and edge weights

The weights for the network edges are determined by how fast goods can be shipped on a particular piece of infrastructure. This depends on length and speed. The length can be calculated, but we need assumptions for speed. I exploit the information on the type of infrastructure and assign them into speed classes. For example, travelling on railroads and motorways is faster than on tertiary roads. Table II.2 shows the speed assumptions for different edge types, which are based on reported figures from the literature.[16] Only the relative speed values matter for the way that supplier access will be constructed. Figure II.2 shows the prevalence of speed classes in terms of route kilometres.

[16]The assumptions for average speed of goods trains are supported by information from the Ministry of Railways (see Figure II.4). Average truck speeds on roads in India are typically a third of the counterparts in more developed countries (NTDPC, 2014). EY (2013) estimates average truck speeds at around 20 km per hour. Baum-Snow et al. (2018) assume 25 km per hour for Chinese roads. Allen and Atkin (2016) use a value of 20 miles per hour for non-highway roads for India. Alder (2017) uses speeds of 35 km per hour based on a survey by the World Bank (2005*b*) which argues that truck speeds are typically less then 40 km per hour in India. The (short) direct connection from the district centroid (plant location) to the closest point in the network is assigned a low speed of 5 km per hour. I compared the calculated fastest path times between a few district pairs to the duration using Google Maps, and the results are reassuringly close.

Figure II.2: Share of speed classes in network

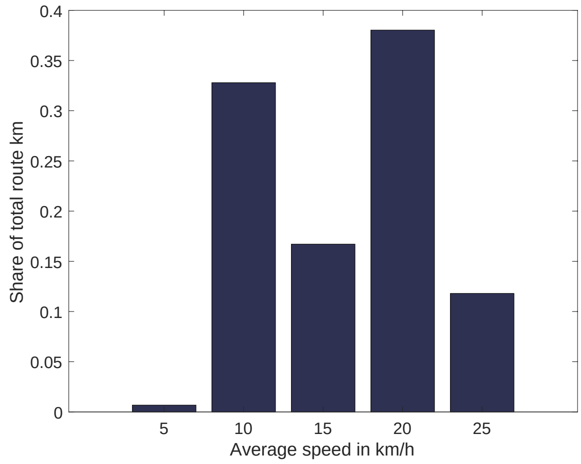


Table II.2: Average speed by edge types

Rail edge type	*Road edge type*	*Speed in km/h*
Broad gauge rail	Motorway, motorway link	25
Narrow gauge rail, light rail	Primary, primary link, trunk, trunk link	20
	Secondary, secondary link	15
Funicular, yard, platform, station, freight station, turntable	Tertiary, tertiary link, tertiary unclassified, road, minor	10
	Connection of plant location to network	5

Notes: The right table shows the speed assumptions for different types of infrastructure. The left figure shows the shares of the speed classes in total route km of the network.

II.4 Recovering distortions and empirical strategy

The question this chapter asks is whether differences in access to suppliers can explain the material input distortions faced by cast iron plants. I first show how I estimate the input distortions. Then I construct a measure of supplier access. Finally I will provide the empirical strategy and specification.

II.4.1 Measuring input distortions

The input distortions τ_{jt}^M and τ_{jt}^L vary by plant j and time t for material inputs M_{jt} and labour L_{jt} respectively. I use the model and follow the estimation in Chapter I, and only briefly sketch the set-up here.[17] Firms are profit maximising and interact with differentiated products in a Nash-Bertrand setting. They minimise short run costs:

$$\begin{aligned} \min_{L_{jt}, M_{jt}} \; & r_{jt}K_{jt} + \tau_{jt}^L w_{jt} L_{jt} + \tau_{jt}^M P_{jt}^M M_{jt} \\ \text{s.t.} \quad & F(K_{jt}, L_{jt}, M_{jt})\Omega_{jt} \geq Q_{jt} \end{aligned}$$

As in Chapter I, the rental rate r_{jt} is not observed and contains the true rental rate along with any capital distortions and adjustment costs. Plant specific wages w_{jt} and input prices P_{jt}^M are observed and capture local labour market differences and input shipping

[17]Alternatively, see Singer (2018).

fees.[18] The distortions are modelled as implicit input taxes, which capture a variety of disadvantages that a particular plant faces by using that particular input. This can include, for example, actual taxes, corruption and legal costs, constrained access or information costs. Importantly, differences in indirect trade costs such as uncertainty, delay or search costs would also be captured by the material input distortions. The aim of this chapter is to provide evidence on what these material input distortions capture. Combining the first order condition of the cost minimisation problem with the Bertrand profit maximising conditions yields the following equations for the input distortions:

$$\tau_{jt}^{X} = (\eta_{jt} + 1)\alpha_{jt}^{X} \frac{P_{jt} Q_{jt}}{P_{jt}^{X} X_{jt}} \quad \forall X \in (M, L) \tag{II.1}$$

which depends on the plant specific inverse demand elasticity η_{jt}, on the input specific output elasticity α_{jt}^{X}, and the revenue share of the input expenditure in equilibrium $\frac{P_{jt} Q_{jt}}{P_{jt}^{X} X_{jt}}$. The required demand and production side parameters are estimated as in Chapter I.

It is worth to briefly discuss the assumption that inputs are elastically supplied, i.e. that plants are price takers on the input side.[19] As Appendix II.B.1 shows, differences in monopsony power would be captured by the distortions. If access to suppliers is related to monopsony power, then this matters for the interpretation of my findings, particularly whether the relationship between distortions and supplier access represents indirect trade costs or monopsony power. My results are robust to including year and district or plant fixed effects which eliminates monopsony power that is constant along these dimensions.[20] Nevertheless, I control for two constructed proxies for plant specific monopsony power, and construct an input distortion net of monopsony power in robustness checks in Appendix II.B.1. The results are similar which suggest that the distortions capture indirect trade costs rather than monopsony power.

By focusing on a single product, the elasticities and the recovered distortions are relatively well measured. They are disentangled from fundamental heterogeneity across plants within sectors, both in terms of production technique and demand conditions. This puts us in a position to explain a distortion with a signal to noise ratio that is much higher than usually found in the literature.

[18] For example, see Cheng and Morrow (2018) for local labour markets in China, and Meunier et al. (2016) for geographical differences in input shipping costs for US cement plants.

[19] The input suppliers are allowed to have arbitrary market power as long as single cast iron plants cannot affect input prices in different ways.

[20] That is, cast iron plants are allowed to have monopsony power ($dP_{jt}^{M}/dM_{jt} \neq 0$), as long as they share a common materials price elasticity of material input consumption ($d\log P_{t}^{M}/d\log M_{t}$) or if the plant specific elasticity is fixed over time ($d\log P_{j}^{M}/d\log M_{j}$).

II.4.2 Measuring supplier access

A first measure for access to suppliers is the district rail penetration, which is simply the total rail km divided by the district size in square km. Adamopoulos (2011) uses a similar measure on a country level for international comparisons, and I use this measure for some additional results. Since it only varies at the district level, we cannot condition on district fixed effects and it is thus likely endogenous. It also does not capture access to suppliers through the transportation network outside of the district.

To obtain a measure of a plant's potential to reach suppliers, I construct a measure of supplier access SA_{dt} for each district d in period t. Redding and Venables (2004) provide a theoretical foundation of both market access on the output side and supplier access on the input side. This is the access to *potential* suppliers, not necessarily the actual supplier choices made by plants.[21] It is similar to the measure of market access in Donaldson and Hornbeck (2016), but on the input side and only considering relevant input suppliers (as opposed to population size):

$$SA_{dt} = -\sum_{h} T_{dh} N_{ht} \tag{II.2}$$

where T_{dh} are the costs in district d to source from district h and a function of the fastest path through the transport network. N_{ht} is the share of the relevant suppliers in district h in the country-wide value of the supplier industry at time t.[22] The relevant supplier industries are mainly pig iron and coking coal, derived from the detailed information of input use of the cast iron plants in the data. I also use a version of SA_{dt} for robustness checks where I exclude the same district such that $h \neq d$. For T_{dh} I follow the literature and use a function that is concave in the fastest path FP_{dh} from h to d, which also captures mobilisation costs:

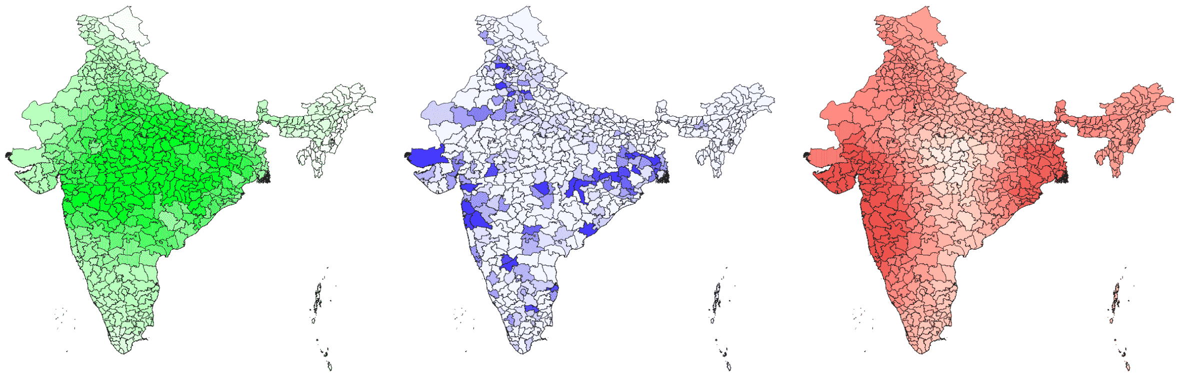
$$T_{dh} = 1 + FP_{dh}^{0.8} \tag{II.3}$$

where the value 0.8 as well as the structure of the function is in line with recent studies

[21] I do not observe plant to plant trade. The measure should therefore be interpreted as access to potential suppliers. This is standard in the literature, dating back to the measure of "market potential" in Harris (1954).

[22] For each district I multiply the plant level values with their sampling multiplier, where the sampling multiplier constitutes how many plants are represented by each plant, to recover a measure based on the universe of plants. Note that when using logged SA_{dt}, it does not matter whether we take N_{ht} as the share or the absolute size of supplier industries, as it will be absorbed by the year fixed effects.

Figure II.3: Average supplier access, supplier presence and change in supplier access



Notes: The left map shows the average supplier access of districts. The middle map shows the average size of supplier industry. The right map shows the average of the absolute deviation of the supplier access from its average within districts. Darker shading mean higher values.

relating travel time to costs.[23] I also use a linear version $T_{dh} = FP_{dh}$ for robustness checks. I calculate the fastest route FP_{dh} using network algorithms. Since we have an undirected graph with positive weights, I can use Dijkstra's 1959 algorithm using the distance divided by the speed as edge weights.[24] The resulting histogram of bilateral shipping times is plotted in Figure II.5 in Appendix II.A.2. The median shipping time is 62 hours, and manual inspection yields shipping times between district pairs that are close to estimates using Google Maps for the same district pairs.

As mentioned in the introduction, this chapter stands out from the literature by combining (i) directly measured costs T_{dh} with (ii) only including *relevant* potential supplier industries that I recover from the input product codes. Based on the calculated fastest path FP_{dh}, Figure II.3 plots the average SA_{dt} over the sample period for all districts in India in the left map. The middle map plots the location of the supplier industries (average value over time). The right map plots the average of the absolute deviation of the SA_{dt} to its within district average, i.e. a measure of how much it changed over time. This is a summary of the variation that I use for identification to which I turn next.

[23] See e.g. Baum-Snow et al. (2018), Alder (2017), Roberts et al. (2012). Also Au and Henderson (2006) use a concave relationship between distance and shipping costs. I also performed some robustness checks by varying parameters (γ_1, γ_2) in $T_{dh} = 1 + \gamma_1 FP_{dh}^{\gamma_2}$, with similar results.

[24] With my network of 1.6 million edges and 1.2 million nodes, it takes only around a minute to calculate the fastest path matrix with an optimised algorithm. The maximum running time for Dijkstra's algorithm with a Fibonacci heap (Fredman and Tarjan, 1987) for one source node to all other nodes is $O(|E| + |V| \log |V|)$ where $|E|$ is the number of edges and $|V|$ the number of nodes. Other studies that used Dijkstra's 1959 algorithm to analyse economic outcomes are e.g. Faber (2014) or Donaldson and Hornbeck (2016). If an exact vector based network is not available but only rasterized data then a fast marching algorithm can be used (Allen and Arkolakis, 2014; Faber, 2014; Allen and Atkin, 2016; Alder, 2017).

II.4.3 Identification and estimation

There are two sources of variation in supplier access SA_{dt}, the time-invariant network component T_{dh} and the time-variant geography of shares of the relevant potential suppliers N_{ht}. There are two endogenous location decisions. One is the location decision of plants, which likely depends on the transport infrastructure. The other is the location decision of infrastructure, which is placed strategically, not randomly. Both give rise to omitted variables and reverse causality concerns. Despite these concerns, others have used changes in infrastructure (e.g Donaldson and Hornbeck, 2016)[25], but I cannot use this variation even if I wanted to, as there were hardly any changes in railway infrastructure for this sample period in India (see Appendix II.A.1).

My strategy is to use variation in the growth or decline of (distant) supplier industries as the identifying source for variation in SA_{dt} by using district fixed effects. For example, if the supplier industry grows in a distant district A, and district A is better connected to district B than to C, then the supplier access for plants in district B improves compared to those in C. The main specification is:

$$\log(\tau_{jt}^M) = \phi SA_{dt} + \boldsymbol{X}_{jt}\boldsymbol{\chi} + \boldsymbol{\lambda_d} + \boldsymbol{\kappa_t} + \iota_{jt} \tag{II.4}$$

where τ_{jt}^M is the material input distortion estimated from the structural model and SA_{dt} the supplier access for district d at time t.[26] $\boldsymbol{X}_{jt}$ is a vector of control variables, such as plant age or legal form, $\boldsymbol{\lambda_d}$ are district fixed effects, $\boldsymbol{\kappa_t}$ year fixed effects and ι_{jt} an error.

If the growth in supplier industries in other districts is uncorrelated to the shocks to distortions ι_{jt}, the effect is identified. There may be further reverse causality concerns warranted if current plant distortions affect supplier industries in other districts, or endogeneity issues when plants and supplier industries are hit by correlated shocks. To at least partially address these concerns, I use lagged supplier access SA_{dt-1} and obtain similar estimates. I also include plant fixed effects in reported robustness checks, which along with further robustness checks are reported in Section II.5.2.

As argued above, the estimates of τ_{jt}^M taken from Chapter I are relatively clean. However, they are still estimated. Since they are used as a dependent variable there is no classical measurement attenuation bias problem. The advantage of the method in Chapter I is that I can recover a distribution of τ_{jt}^M for every single plant, based on the parametric bootstrap

[25] Donaldson and Hornbeck (2016) state that this is their main endogeneity concern. They run robustness checks by controlling for the presence of nearby railroad tracks.

[26] I demeaned τ_{jt}^M within every year.

Table II.3: Input material distortion and supplier access

	(1)	(2)	(3)	(4)	(5)	(6)
Supplier access	-0.27***	-0.23***	-0.28**		-0.19**	-0.26**
	[-3.92]	[-3.19]	[-2.05]		[-2.43]	[-2.17]
Supplier access (lagged)				-0.31**		
				[-2.15]		
Rail km/sqkm					-0.25*	-0.21
					[-1.70]	[-1.47]
Plant level controls	No	Yes	Yes	Yes	Yes	Yes
District level controls	No	Yes	No	No	Yes	Yes
State level controls	No	Yes	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes
District FE	No	No	Yes	Yes	No	No
N	926	882	926	926	882	882
R^2	0.07	0.25	0.46	0.46	0.26	0.26

Notes: The dependent variable is the logged demeaned (by year) distortion in material inputs τ^M_{jt}. Coefficients are standardised. t-statistics in brackets are based on standard errors clustered at the district level. Plant level controls include firm age, dummies on ownership type, and whether plants are part of the census section. District level controls include population, population density and gender ratio in 2001, district population growth from 1991-2001, and whether the district was subject to left wing extremism (who sometimes target infrastructure) in the years leading up to 2009. State level controls include the male literacy rate and the share of male industrial workers in 2001. Data for district and state level controls is based on the population census and retrieved from indiastat.com.

from estimated production and demand parameters. This allows me to perform robustness checks with respect to the uncertainty in the dependent variable.

II.5 Results

II.5.1 Access to suppliers decreases material input distortions

The distortions τ^M_{jt} do not capture shipping fees, as those are accounted for in the model and observed in P^M_{jt}. In Appendix II.A.3 I show that input prices inclusive of shipping fees P^M_{jt} and supplier access SA_{dt} are negatively correlated as expected.[27] The indirect costs of trade, such as uncertainty or delays in shipping or search costs, on the other hand, will be picked up by the distortions τ^M_{jt}. I test the hypothesis that with longer routes to suppliers, i.e. a lower SA_{dt}, the distortions τ^M_{jt} are higher, most likely because the indirect trade costs are higher.

Table II.3 shows the results from estimating regression (II.4). The first two columns are without district effects $\boldsymbol{\lambda_d}$.[28] The main result is in Column (3), controlling for district fixed effects. A one standard deviation increase in supplier access reduces the material distortion by 0.28 of its standard deviation.

[27]The relationship is marginally insignificant, possibly due to heterogeneous input quality having a bigger effect on input prices than shipping fees.

[28]The OLS estimates without district fixed effects are slightly upward biased. This suggest that plants with higher distortions τ^M_{jt} tend to locate in better connected areas.

For the interpretation of the results, it is relevant whether the distortions are likely to rather capture input market power (monopsony power) or indirect trade costs. Monopsony power is proportional to the measured input distortions as shown in Appendix II.B.1. Therefore, the results in Table II.3 would only be consistent with a monopsony power story if worse access to suppliers *increased* monopsony power. We would expect, however, that if supplier access is poor, there are few suppliers around, which would suggest that there is less monopsony power. In this sense, the monopsony power story would work against the reported results. In Appendix II.B.1, I construct two proxies for monopsony power, one based on market share, one based on directly estimating the input price elasticity. The results are robust to controlling for either. In addition, I construct a measure of input distortion net of monopsony power as a dependent variable, which confirms the main results (all additional results in Table II.7). Overall, the results strongly suggest that indirect costs of trade are captured by the estimated distortions. These distortions in turn lead to misallocation of input materials with adverse aggregate consequences.

For this industry, the welfare costs of materials misallocation are large (see Chapter I), and differences in supplier access are a significant contributor. Taking the results at face value, it is worth making the following four points. First, it is important to note that the relationship between distortions and differences in supplier access are not simply features of our spatial reality. It is costly to move goods across space. However, since shipping costs are accounted for, the part in distortions that is due to differences in supplier access can be addressed without necessarily reducing shipping fees.[29] These are not the transport costs of moving goods across space, but differences in costs generated by the multitude of freight transport issues described in Section II.2.

Second, this analysis has policy implications, as it provides us with a margin that we can address to improve allocative efficiency. Reducing costs of uncertainty and delays, e.g. by strengthening the transportation infrastructure network, or reducing border checkpoints due to differential tax systems is likely to improve allocative efficiency. In fact, the GST (good and sales tax) reform in 2017 unifies the tax system and substantially reduces border checkpoints within India. Loosely speaking, if indirect trade costs are reduced, we would start to see the relationship between supplier access and input distortions disappear.

Third, while the analysis provides evidence that differences in supplier access generate misallocation, I cannot distinguish between its components, such as information, search or uncertainty costs. Fourth, while there is a significant relationship between supplier access

[29]Shipping fees can also be regarded as additional spatial frictions, quantified by Behrens et al. (2017) for the US in a setting with endogenous markups, for example.

Table II.4: Input material distortion and supplier access: robustness checks

	Excl. own dist.		Linear costs		Plant FE		Rail only	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Supplier access	-0.28**		-0.32**		-0.17*		-0.30*	
	[-2.00]		[-2.32]		[-1.76]		[-1.79]	
Supplier access (lagged)		-0.40**		-0.36**		-0.22**		-0.33*
		[-2.55]		[-2.39]		[-2.00]		[-1.93]
Plant level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Plant FE	No	No	No	No	Yes	Yes	No	No
N	926	926	926	926	924	924	926	926
R^2	0.46	0.46	0.46	0.46	0.73	0.73	0.45	0.46

Notes: The dependent variable is the logged demeaned (by year) distortion on material inputs τ^M_{jt}. The supplier access in the first two columns excludes the own district of a plant when calculating supplier access $SA_{dt} = -\sum_h T_{dh} N_{ht} \forall h \neq d$. Columns (3) and (4) assume linear costs, i.e. $T_{dh} = FP_{dh}$. Columns (5) and (6) include plant fixed effects. Columns (7) and (8) calculated the fastest path FP_{dh} based on the railway network alone instead of using roads and rails. Coefficients are standardised. t-statistics in brackets are based on standard errors clustered at the district level. Plant level controls include firm age, dummies on ownership type, and whether plants are part of the census section.

and material input distortions, there is still unexplained variation left in the material distortions as the R^2 show.

In Columns (5) and (6), I add a measure of district rail penetration, which is total rail km per district area. Both omit district fixed effects. In Column (6), I additionally instrument supplier access with a measure of supplier access that is demeaned at the district level to account for the lack of district fixed effects. The supplier access coefficient is robust to controlling for district rail penetration. The effect of district rail penetration is marginally significant as well, which suggest that the shipping constraints within districts may matter as well. I next explore the robustness of the results and provide placebo analyses to test the identification assumption.

II.5.2 Robustness and three placebo tests

Robustness checks

I perform a number of robustness checks. First, there might be some correlation between simultaneous shocks that affect the cast iron industry, but also its suppliers. In Column (4) of Table II.3, I use lagged supplier access, and the effect, if anything is slightly larger.[30] Second, I exclude the own districts of plants when calculating supplier access (Equation (II.2)) in Column (1) and (2) of Table II.4. Third, I use the fastest path directly as cost of shipping (Equation (II.3)) in Columns (3) and (4) of Table II.4. Fourth, Columns (5) and (6) include plant fixed effects. Fifth, for Columns (7) and (8) I use the railway network

[30]This suggests an upward bias in the current period supplier access. Suppose a plant receives a shock which reduces some components in its input distortion. If it increases input demand, then the well connected suppliers could benefit, introducing a reverse causality problem, which biases the current period supplier access in the reported direction.

Table II.5: Placebos: labour distortion, *irrelevant* supplier access, or market access

	τ^L_{jt}	Irrelevant inputs			Market access
	(1)	(2)	(3)	(4)	(5)
Supplier access	-0.11				
	[-1.30]				
Supplier access (textiles)		-0.15			
		[-1.16]			
Supplier access (food)			0.46		
			[1.55]		
Supplier access (rubber)				-0.33	
				[-1.26]	
Market access					-0.36
					[-1.47]
Plant level controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
N	926	926	926	926	926
R^2	0.42	0.45	0.45	0.45	0.45

Notes: The dependent variable is the logged demeaned (by year) labour distortion τ^L_{jt} in Column (1) and the material distortion τ^M_{jt} in all other columns. The supplier access in the Columns (2) -(4) is based on access to the textiles, food, and rubber industries respectively. The market access variable in Column (5) is access to output markets, i.e. to those industries that buy cast iron. Coefficients are standardised. t-statistics in brackets are based on standard errors clustered at the district level. Plant level controls include firm age, dummies on ownership type, and whether plants are part of the census section.

only, dropping all roads before calculating the fastest path.[31] Overall, the effect of supplier access differences remains robust.

To assess the effect that the dependent variable, the distortion, is estimated, I make use of the estimated covariance matrices of the underlying production and demand parameters. As described in Chapter I, I obtain a different set of τ^M_{jt} for every draw from the distribution of the estimated fundamental parameters. For every set of τ^M_{jt}, I run regression (II.4) and obtain the point estimates and t-statistics of supplier access. In total, I run 330 regressions and plot the point estimates and t-statistics in Figure II.6 in Appendix II.B.2. The average point estimate is -0.295 with a minimum to maximum of -0.34 to -0.22. All estimates are significant at the 10% level at least.

I next run three placebo tests to provide further support for the causality of the estimated relationship.

Placebo 1: No effect on labour distortions

The first placebo takes the estimated logged *labour* distortion τ^L_{jt} as dependent variable. The access to materials suppliers should not affect these distortions. Because commuting happens predominately within districts, the access to distant suppliers should also not pick

[31]In another robustness check, I use the geographic distance instead of the geodesic distance (shortest path) of the network, with similar results.

up access to labour markets, which is likely contained in τ_{jt}^{L}. As Column (1) in Table II.5 reports, the association is not statistically significant.

Placebo 2: No effect of access to irrelevant supplier

For a second placebo test, I construct a measure of supplier access to *irrelevant* supplier industries. In particular, access to textiles, food or rubber industries. Columns (2) to (4) of Table II.5 show that there is no effect. This shows that it is only differences in access to *relevant* suppliers that is contained in the input misallocation distortions τ_{jt}^{M}.

Placebo 3: No effect from access to output markets

Finally, a concern is that the input distortions τ_{jt}^{M} capture additional costs from shipping the outputs, rather than inputs. To test this, I also construct a market access variable. It is based on the size of the industries that *buy* from the cast iron plants in the sample, i.e. the downstream firms such as engine manufacturing that use cast iron as input. Column (5) in Table II.5 reports that the estimate, while negative, is not statistically significant. The data I use provides a measure of distributional (i.e. shipping costs) on the output side. I show in Table II.8 in Appendix II.B.3, that this measure of output shipping costs is significantly correlated to market access. This suggests that the measure of market access captures what we think it should capture, access to buyers, but is not significantly driving distortions on the *input* side.

On the whole, it is only the access to relevant input suppliers that affect material input distortions, and consequently misallocation losses.

II.6 Conclusion

Misallocation of input materials in Indian cast iron are severe and associated with high welfare costs, as Chapter I shows. This chapter asks what these costly material input distortions represent. I find that differences in access to suppliers through the transportation network drives the estimated input distortions. I show that it is only the access to relevant input suppliers, as access to irrelevant suppliers or access to output markets is not significantly related to the distortions on the input side.

I emphasise that the input misallocation distortions should be interpreted in terms of differences in *indirect* trade costs. Differences in shipping fees represent the spatial reality, as it is inherently costly to ship goods across space, and therefore hard to eliminate. Since input shipping fees are observed and accounted for in the model, the distortions are net of input shipping fees. On the other hand, any indirect trade costs associated with sourcing inputs, which are lower for better supplier access, are captured in the estimated distortions. This sheds more light into the black box of misallocation losses, which were found to be large in Chapter I, and more generally contributes to identify distortions in the misallocation literature.

The policy implications are that there are aggregate reallocation gains from reducing differences in indirect trade costs, without necessarily decreasing shipping costs. These include, for example, costs of delay, search and uncertainty. State border checkpoints for goods within India, for example, create shipping delays and are, for the purposes of this study, policy distortions that create input misallocation. The described relationship between supplier access and misallocation distortions is likely to have external validity in contexts of industries that require substantial input shipping, but face unreliable transport infrastructure.

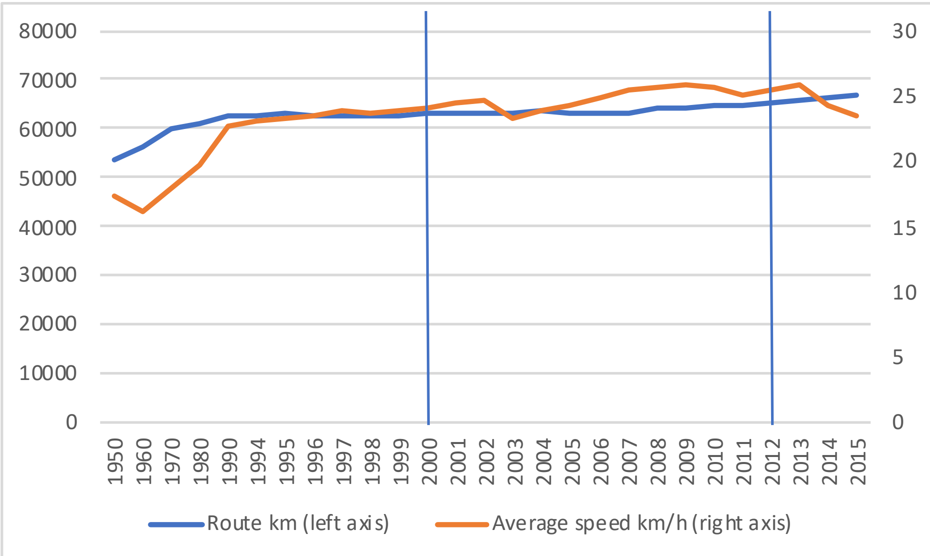
While I cannot distinguish between different types of indirect trade costs that increase with remoteness, this chapter provides an important insight into the drivers of misallocation. Especially misallocation of input materials has received little attention and we have known even less about underlying determinants. Future research aims to quantify the misallocation costs of differences in supplier access as well as other potential drivers in India and other countries.

Appendix to Chapter II

II.A Additional descriptives for the construction of supplier access

II.A.1 Railway route kilometres and speed

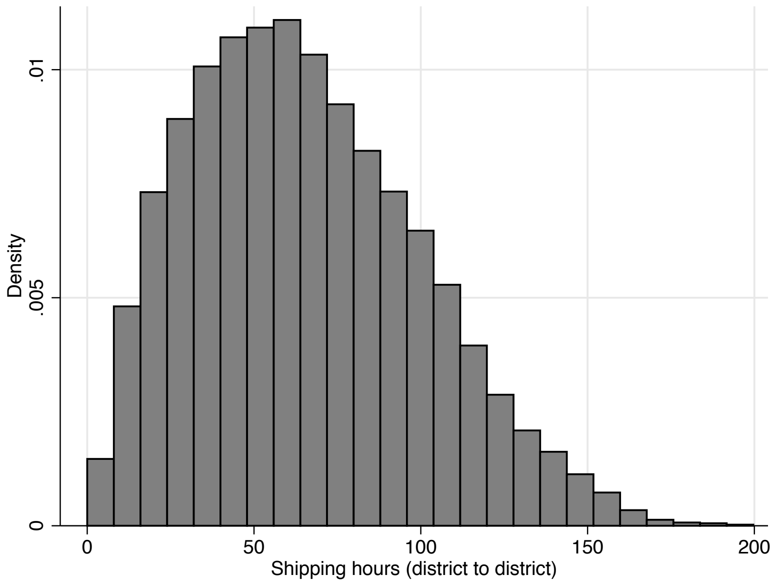
Figure II.4: Total route kilometres of Indian railways and average speed of goods trains



Notes: Vertical lines indicate sample period. Source: Calculated based on information from Ministry of Railways India, retrieved through indiastat.com

II.A.2 Estimates of the fastest path FP_{dh}

Figure II.5: Histogram of bilateral fastest paths FP_{dh}



Notes: The figure plots the histogram for bilateral estimated shipping times FP_{dh} between districts using the road and railroad transport network. Shipping times are trimmed at 200 hours. Shipping times are estimated using Dijkstra's 1959 algorithm, with speed assumptions as edge weights as shown in Table II.2.

II.A.3 Correlation between input price and supplier access

Table II.6: Input prices and supplier access

	Input price (log)			
	(1)	(2)	(3)	(4)
Supplier access	-0.22	-0.20	-0.19	-0.18
	[-1.32]	[-1.31]	[-1.41]	[-1.26]
Plant level controls	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	No	No
Plant FE	No	No	Yes	Yes
N	946	946	946	946
R^2	0.32	0.33	0.64	0.65

Notes: The dependent variable is the logged material input price $\log P^M_{jt}$. Coefficients are standardised. t-statistics in brackets are based on standard errors clustered at the district level. Plant level controls include firm age, dummies on ownership type, and whether plants are part of the census section.

II.B Robustness checks

II.B.1 Monopsony power

The first part of this Appendix is the same as Appendix I.B in Chapter I, repeated here for convenience. If plants are not price takers on the input side and have some monopsony power, the cost minimisation problem changes to:

$$\min_{L_{jt},M_{jt}} r_{jt}K_{jt} + \tau_{jt}^{L_{adj}} w_{jt}(L_{jt})L_{jt} + \tau_{jt}^{M_{adj}} P_{jt}^{M}(M_{jt})M_{jt}$$
$$\text{s.t.} \quad F(K_{jt}, L_{jt}, M_{jt})\Omega_{jt} \geq Q_{jt}$$

where $\tau^{M_{adj}}$ is the new input distortion adjusted for monopsony power, and the input prices are some functions of the input quantities. The first order condition with respect to materials is:

$$\underbrace{\tau_{jt}^{M_{adj}}(\psi_{jt}+1)}_{\equiv\tau_{jt}^{M}} = (\eta_{jt}+1)\alpha_{jt}^{M}\frac{P_{jt}Q_{jt}}{P_{jt}^{M}M_{jt}} \tag{II.5}$$

where $\psi_{jt} \equiv \frac{\partial P_{jt}^{M} M_{jt}}{\partial M_{jt} P_{jt}^{M}}$ is the inverse input price elasticity of input demand. Note that if we ignore $\tau_{jt}^{M_{adj}}$, and use that $MRPM \equiv (\eta_{jt}+1)\alpha_{jt}^{M}\frac{P_{jt}Q_{jt}}{M_{jt}}$, we can write $(\psi_{jt}+1) = \frac{MRPM}{P_{jt}^{M}}$. That is $(\psi_{jt}+1)$ is the ability to pay an input a lower price than its marginal revenue product, a common definition of market power on the input side, or monopsony power.

The measured input distortion τ_{jt}^{M} captures input market power as well as other input distortions $\tau^{M_{adj}}$. If ψ_{jt} does not vary across plants, or does not vary across time for individual plants, than it will be absorbed by year or plant fixed effects respectively (see Table II.4).

I construct an admittedly heuristic plant specific measure of ψ_{jt} by regressing logged input prices on a second order polynomial in input quantities controlling for district and time fixed effects:

$$\log(P_{jt}^{M}) = f(\log(M_{jt})) + \boldsymbol{\lambda_d} + \boldsymbol{\kappa_t} + \upsilon_{jt} \tag{II.6}$$

Based on the coefficients, I can compute ψ_{jt}. I plot the histogram of input market power $(\psi_{jt}+1)$ in Figure I.12 in Appendix I.B of Chapter I. The majority of plants has a negative elasticity ψ_{jt} and therefore an input market power $(\psi_{jt}+1)$ below one. A negative ψ_{jt} means that input prices decrease for larger quantities, which can be related to quantity discounts instead of input market power.

I use the estimated measures $(\psi_{jt}+1)$ as control variables in Column (3) and (4) in Table II.7. In Columns (5) and (6) I instead use $(\psi_{jt}+1)$ to recover the adjusted distortion $\tau^{M_{adj}}$ as dependent variable. I construct a second proxy for input market power. If larger plants can exert more market power on the input side as well, a larger market share of a plant in a given district can proxy for monopsony power. In Column (1) and (2) of Table II.7 I control for plant market shares within a district.[32] Overall, the results are robust to these additional test, and the input distortions are significantly higher for plants with worse supplier access, supporting the interpretation that this is driven by indirect trade costs.

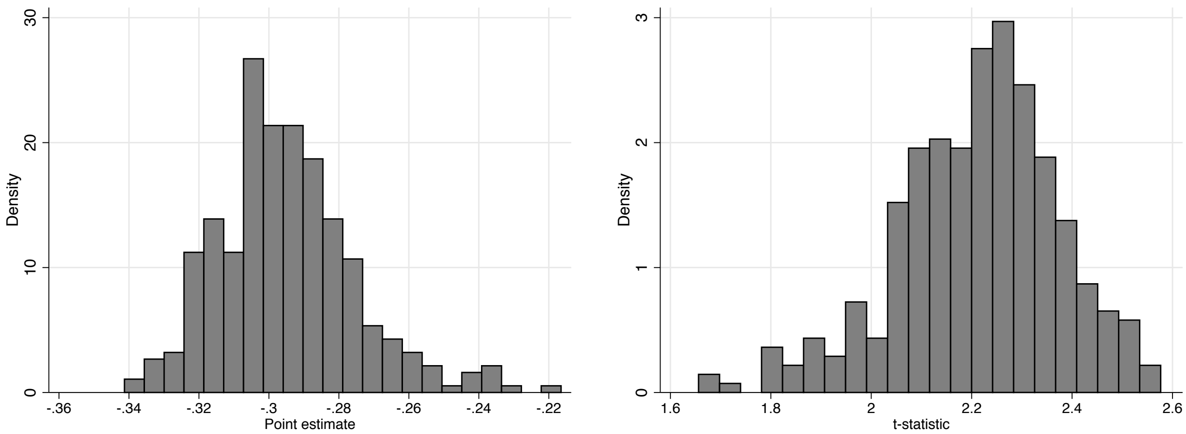
Table II.7: Monopsony power: additional proxy controls and adjusted input distortion

	Input distortion τ_{jt}^{M} (log)				Adjusted $\tau^{M_{adj}}$ (log)	
	(1)	(2)	(3)	(4)	(5)	(6)
Supplier access	-0.27***	-0.28**	-0.27***	-0.25**	-0.25***	-0.19
	[-4.76]	[-2.11]	[-4.77]	[-2.06]	[-4.45]	[-1.57]
Market share in district	0.03	-0.04				
	[0.48]	[-0.69]				
Inv. inp price elasticity + 1 (log)			0.06	0.14		
			[0.61]	[1.08]		
Plant level controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	No	Yes	No	Yes	No	Yes
N	926	925	926	925	926	925
R^2	0.20	0.46	0.20	0.46	0.31	0.54

Notes: The dependent variable is the demeaned (by year) logged material input distortion τ_{jt} in the first four columns, and the adjusted logged $\tau^{M_{adj}}$ in the last two columns. The market share is calculated at the district year level in terms of value of cast iron sold. The input price elasticity is estimated from a regression of logged input prices on a second order polynomial in logged material input use, controlling for plant and year fixed effects. I take the log of the inverse plant level input price elasticities plus one $\log(\psi_{jt}+1)$, i.e. the log of the markdown. Plant level controls include firm age, dummies on ownership type, and whether plants are part of the census section. Coefficients are standardised. t-statistics in brackets are based on standard errors clustered at the district level.

II.B.2 Accounting for estimated input distortions

Figure II.6: Distribution of point estimates and t-statistics



Notes: The left panel plots the histogram of 330 point estimates of supplier access of Equation (II.4). Each individual regression has a different τ_{jt}^{M} as dependent variable, based on 330 draws of the underlying production and demand parameters from their estimated covariance matrices. The right panel shows the histogram of t-statistics of the supplier access estimates. All estimates are significant at the 10% at least.

[32]These controls may be bad controls in the sense of Angrist and Pischke (2008), but the point is to show that the coefficient of supplier access is reasonably robust to those.

II.B.3 Market access and output shipping costs

Table II.8 shows the result from a regression of the share of output shipping costs in revenues on market access (i.e. access to buyers). Column (2) shows that the access to suppliers is in turn not relevant for *output* shipping costs.

Table II.8: Market access and share of output shipping costs in revenue

	(1)	(2)
Market access	-0.11*	
	[-1.72]	
Supplier access		0.00
		[0.02]
Year FE	Yes	Yes
N	946	946
R^2	0.04	0.02

Notes: Standardized coefficients. t-statistics in brackets based on SE clustered on districts.

Chapter III

Can lower electricity prices improve energy efficiency? Evidence from half a million Indian plant observations

III.1 Introduction

High energy prices are often regarded as a barrier to industrial development. Upgrading capital vintages to take advantage of electric power and automation are key elements to improving performance in manufacturing industries. High energy, and in particular, high electricity prices can slow this process by reducing the incentives to switch production from more traditional manufacturing processes.[1] In developed countries, energy-intensive industries are often granted exemptions and subsidies for electricity costs to withstand international competition and avoid layoffs. At the same time, countries aim to improve their energy efficiency as part of their climate goals. At least for manufacturing, improving energy efficiency is one of the principal pillars to reduce the energy and carbon intensity of GDP. Low energy prices may in turn fail to provide sufficient incentives to improve energy efficiency.

A priori, electricity prices can have ambiguous effects on electricity productivity (output divided by electricity consumed). The more obvious channel is that higher prices improve electricity productivity by inducing substitution away from electricity and optimisation of production processes which improve electricity consumption per output. Lower prices, on the other hand, can also improve electricity productivity by incentivising upgrades of production processes to those that require more electricity but also increase output. The ratio between output and consumption can fall or rise when considering choices between different production techniques. While the pricing of electricity is important from an

[1] Ryan (2018) provides evidence from a field experiment in Gujarat (India) on the complementarity of electricity and modern capital and skilled labour. Abeberese (2017) provides evidence on changes in production due to electricity prices in India.

environmental and developmental perspective, we have little causal evidence on the net effect.

This chapter examines the effect of electricity prices on electricity productivity. I use annual plant level panel data from Indian manufacturing from 1998 to 2013 which includes information on the quantity and the average price of electricity consumed. Addressing endogeneity concerns, I find that higher electricity prices *decrease* electricity productivity. While electricity consumption falls with higher prices, output decreases relatively more. This implies that lower electricity prices can not only increase output, but also *improve* energy efficiency. Policy makers often navigate trade-offs between developmental and environmental goals when it comes to energy prices. These results suggest that lowering electricity prices does not necessarily decrease energy efficiency. To my knowledge, these are the first causal estimates to show this for an entire manufacturing sector in a developing country.

I emphasise that this result is likely to be especially relevant in contexts of industrial development and where industrial electricity prices are comparatively high, which is both the case for India.[2] I find that the effect was stronger during high price periods. Furthermore, I show that this effect is unique to electricity: the effect of coal prices on coal productivity are the opposite – higher coal prices increase coal productivity and have no significant effect on different measures of firm performance.[3] Depending on the fuel mix of electricity generation, these results suggests that taxing dirtier fuels yields better economic and environmental outcomes than taxing electricity, most likely due to the special role of electricity in modern production. This finding is particularly relevant for climate policy in developing countries. Relatively lower industrial electricity prices than coal prices could deliver both, substitution from fossil fuels to electricity, and despite increasing electricity use, improving electricity productivity and output.

I begin the analysis by documenting that energy productivity in Indian manufacturing has been fairly flat since the 60s but increased sharply from around 2000. Electricity productivity increased from 2000 to 2013 by around 35%. Simultaneously, real average industrial electricity prices fell by around 45%, a robust finding across various data sources from plant level data to price indices and manually collected tariffs. Indian industries are characterised by significant cross-sectional dispersion of both electricity productivity and

[2]Compared to average industrial electricity prices in G7 countries, India's prices have been around 80% higher in 1998. They only dropped below the G7 average after 2004 and were around half the G7 prices in 2013 (but still above US prices). In PPP terms, India's industrial electricity prices have been more than double the G7 average throughout (see Figure III.20 and Table III.10).

[3]The instruments used for coal prices are different to those used for electricity prices.

electricity prices across plants, even within states within industries. With the decrease in electricity prices, there has been a convergence of electricity prices across plants. The concurrent fall in prices and secular increase in efficiency provides motivating evidence for a careful econometric analysis.

Plant level electricity prices are subject to endogeneity concerns in a regression of electricity productivity on prices. For example, most Indian states have increasing block tariffs for industry such that plants with higher consumption pay higher prices, or plants may negotiate discounts or enjoy favourable relationships with state electricity providers, which could be correlated with their economic position. I use two different instruments to address these endogeneity concerns, based on the institutional context of Indian electricity pricing. The first is based on the electricity price paid by other plants in the same state but in a different industry. I kernel weight them by the distance to the other plants in terms of the electricity quantity purchased. The second is a Bartik (1991) shift-share instrument similar to the instrument constructed in Abeberese (2017). The shares are the state level shares of coal power plants in the total installed generation capacity fixed at a pre-sample period. These shares are interacted with a representative coal price that is set by coal companies for power utilities and shifts generation costs.

The bias in the OLS estimates is sizeable. While the OLS based elasticity of electricity productivity with respect to electricity prices is 0.37, the elasticities are -0.24 and -0.78 for the two IVs, all statistically significant. This positive bias could, for example, arise from less efficient plants receiving more favourable tariffs or exemptions, perhaps through corruption. It is worth noting that from a back of the envelope calculation, the size of the causal estimates from the micro data can explain the entire secular increase in *aggregate* electricity productivity remarkably well. I provide a range of robustness checks and an analysis of heterogeneous effects by industry.

To shed more light on mechanisms, I examine further plant decisions and outcomes. The effect of prices on output outweighs the effect on electricity consumption. Since total variable costs decrease, the results suggest that plants scale down with higher electricity prices. I present evidence that higher electricity prices significantly decrease profits, plant productivity (TFP), investment, employment, machine to labour ratios and markups. This is consistent with a setting where electricity prices influence production and investment decision beyond electricity consumption. Lower prices can incentivise firms to invest in modern electricity intensive machinery, processes and products. These, in turn, improve productivity, output and performance.

While there are clear positive effects of the decrease in electricity prices on firm performance and electricity productivity, there may have also been effects on consumers. The decrease in markups suggest that there is imperfect pass-through of electricity costs to consumers. I estimate the incidence of electricity prices as share of consumer surplus in total surplus. The degree to which consumers and producers share the surplus is determined by how well producers can substitute to electricity, by their market power and demand elasticities, and how marginal costs are passed-through to prices. Ganapati et al. (2016) show how incidence can be expressed as a function of these parameters in a generalised oligopoly. I exploit the detailed information on output quantities and prices in the data to estimate the pass-through elasticities by industry, using the above instruments for marginal costs, and combine these with my estimates of plant level market power and demand elasticities to recover pass-through rates and incidence shares at the plant level. On average, two thirds of the incidence of lower electricity prices fell on consumers. The pricing of electricity for industry is therefore not only highly relevant for firms, but has substantial welfare implications for consumers.

There are a number of related papers. The closest paper is perhaps Abeberese (2017). She studies the effect of electricity prices on firms switching industries within narrow industries in India, using a similar shift-share instrument. Her analysis covers nine years beginning in 2000. She finds that higher electricity prices make firms switch to less electricity-intensive industries, which is consistent with the story of this chapter. She also finds negative effects on output, TFP, and the machine-labour ratio. The last section of this chapter can confirm these latter findings using new data over 16 years, with three times the observations and two different instruments. Similarly, Elliott et al. (2019) find that higher electricity prices induce firms to switch to less energy-intensive products in China as well.

In the literature on energy prices and industrial energy efficiency, Davis et al. (2008) is one of the first studies to use micro data on prices and electricity productivity on a national scale. They find that the correlation of electricity productivity and electricity prices in US manufacturing industries is generally positive. An IV based on fuel shares in state power generation turns the elasticities for a small number of industries negative. Nevertheless, in contrast to India, this suggests that in a developed country like the US, the effects of higher electricity prices are dominated by the first channel through cost minimisation and substitution, rather than the second channel through TFP, investment and size reduction. This is not to say that the second channel is absent. Deschenes (2011) estimates a -0.12

elasticity of employment to electricity prices using state-industry level data in the US.[4] Aldy and Pizer (2015) find a negative impact of energy prices on output using a long industry level panel in the US. Linn (2008) also finds a positive elasticity of electricity productivity to energy prices in the US (0.22).[5] Popp (2002) also uses US state level prices for a bundle of energy to show a positive effect of energy prices on innovation. However, these elasticities are with respect to an index of all energy sources, not just electricity, and rely on state level prices. State level prices ignore the substantial heterogeneity in electricity prices across plants that Davis et al. (2013) report. Kahn and Mansur (2013) show that energy-intensive industries in the US tend to cluster in low electricity price counties. Bundling energy prices mix the (potentially opposite) effects of electricity and coal prices. The Porter and Van der Linde (1995) hypothesis, which postulates firm benefits from environmental regulation, may apply to fossil fuels, but not necessarily to electricity.

In the developing context, Fisher-Vanden et al. (2004) report a positive elasticity of electricity productivity to electricity prices (0.23) for a subsample of Chinese firms in 1997-1999. This is, however, based on OLS regressions and in line with my OLS estimates.[6] Rentschler and Kornejew (2017) examine Indonesian small and micro firms in 2013. They find that firm level electricity prices reduce profitability, but increase (total) energy efficiency, based on OLS estimates. For India, Golder (2011) found that foreign firms have a higher energy productivity in 2008, and Sadath and Acharya (2015) report a negative elasticity of investment to energy prices. In the literature on decomposition analysis of energy efficiency, Mukherjee (2010, 2012) finds that energy productivity varies across Indian states, and that firms are not at their efficiency frontier. Ghani et al. (2014) report an increase in electricity productivity in the 2000s which was mainly through improvements in existing state-industry clusters.[7]

This chapter is also related to the literature studying the firm level relationships between environmental and economic performance. In the developed country context, environmental policies such as the European Emission Trading scheme or carbon taxes are often found

[4]In France, Marin and Vona (2017) find an employment elasticity with respect to electricity price of -0.26 and an electricity consumption elasticity of -0.6. See Cox et al. (2014) for broadly comparable elasticities for Germany. Marin and Vona (2017) also find a negative -0.11 TFP elasticity.

[5]His findings suggest that entrants' energy efficiency respond more to energy prices than that of incumbents. See also Pizer et al. (2002) who study technology adoption, energy prices and aggregate energy efficiency.

[6]See also their later paper (Fisher-Vanden et al., 2016). Using aggregate data, Hang and Tu (2007) find a negative elasticity of electricity productivity to electricity prices in China after 1995.

[7]For other decomposition studies of energy productivity, see e.g. Cornillie and Fankhauser (2004) for Eastern Europe and Liu and Ang (2007) for a review article on decompositions into within-energy productivity and product mix.

to improve environmental performance with little to no impact on economic performance (Martin et al., 2014, 2015; Dechezleprêtre and Sato, 2017).[8] This is not at odds with my findings. Carbon pricing increases fossil fuels prices more than electricity prices and I find a null effect of coal prices on productivity. As in Acemoglu et al. (2012), it is the relative price between clean and dirty energy that matters. The role of electricity is special, because it is a complementary input to modern machinery and production processes (Ryan, 2018). Reducing relative electricity prices can incentivise to improve production techniques and products, especially in the case of developing countries with already high electricity prices, as in India.

The chapter also contributes to the literature on energy cost pass-through and incidence. The recent theoretical literature emphasises the importance of imperfect competition in accounting for incidence (Weyl and Fabinger, 2013). De Loecker et al. (2016) estimate imperfect marginal cost pass-through from input tariff reductions in large companies in India. I estimate marginal cost pass-through allowing for imperfect competition and input substitution following Ganapati et al. (2016). They study five specific products in US manufacturing and show that incidence on consumers is lower than under perfect competition models.[9] Miller et al. (2017) study pass-through and incidence in the US cement industry based on Weyl and Fabinger (2013). They estimate that pass-through of energy costs is above unity, and the share of incidence of carbon pricing for producers only 11%. I recover the distribution of pass-through rates where some plants and industries also fully pass on costs. A small literature focuses on energy and emission cost pass-through of utilities rather than firms (Fabra and Reguant, 2014; Hausman, 2018).

While this chapter focuses on the effects of electricity prices, a related literature focuses on the *reliability* of electricity and its implications. This is important in a developing country context where shortages are frequent. Allcott et al. (2016) show that power shortages in India reduce revenues by about 5% on average, and distort the plant size distribution due to returns to scale in self-generation.[10] Due to the institutional context in India, shortages are not related to electricity prices, and I show that they are not significantly correlated. Nevertheless, I provide robustness analyses for my estimates

[8] The European Emission Trading scheme might have even spurred innovation (Calel and Dechezlepretre, 2016).

[9] In my analysis the incidence on consumers is also around 50% lower compared to an alternative perfect competition assumption.

[10] See also Alam (2013) for evidence on India using satellite data, and Reinikka and Svensson (2002) and Foster and Steinbuks (2009) using data of African countries. Ryan (2017) simulates the impact of transmission capacity improvements on the Indian electricity wholesale market.

controlling for power shortages.

The findings in this chapter have important policy conclusions. While it is more obvious that low electricity prices can spur industrial development, it is rather novel evidence that low electricity prices can also *improve* energy efficiency (i.e. electricity productivity). In this context of industrial development and high electricity prices, and in contrast to taxing fossil fuels for industry, there appears to be little trade-off in electricity pricing between economic and electricity efficiency goals. Lower prices, however, still increase the quantity of electricity used. The size of the associated negative environmental externalities depends on the source of energy in generation, but taxing fossil fuels instead of taxing electricity use in industry is likely to have net economic and environmental benefits. I contribute to the literature by exploiting a large nationally representative panel of plants with plant level information on electricity quantity and expenditure. I develop an instrument for causal interpretations that is not reliant on additional information and can thus be readily calculated in similar contexts for other studies.

The rest of the chapter sets the stage with a brief analysis of the Indian electricity market in Section III.2. While the Indian electricity sector is interesting in its own right, the insights provide context for the identification strategy. I describe the data used in Section III.2.2 and present aggregate trends and the dispersion of electricity productivity and prices in Section III.2.3. Section III.3 develops the empirical strategy. Section III.4 discusses the results, potential mechanisms and incidence before I offer a brief conclusion in Section III.5.

III.2 India's electricity sector and descriptive statistics

III.2.1 India's electricity sector

I briefly discuss the issues that are relevant for identification, interpretation or robustness checks. These include ownership, type of electricity generation, effects of deregulation, level of electricity prices and cross-subsidisation, tariff setting, coal prices for utilities, and outages and self-generation.

India's electricity generation is dominated by state and central governments. In 1998, government owned 65% and 30% of installed capacity respectively, with the remaining 5% owned privately (Ministry of Power, 1998*a*; Planning Commission, 2001). The Electricity Act of 2003 aimed to open this heavily regulated sector to more competition.[11] This led to

[11]The preamble states "An Act to consolidate the laws relating to generation [...] of electricity [...], promoting competition therein [...].".

more privately owned power plants entering the electricity generation sector. By 2013, the share of privately owned capacity rose to 31%, cutting mostly into the share of state-owned capacity (40%), while the centrally owned share remained at 29% (Planning Commission, 2014). In February 2019, the share of the private sector (46%) was almost equal to the share of the combined government owned capacity (Central Electricity Authority, 2019). From 1998 to 2013, total installed capacity rose by 143%.

Most of India's electricity is generated from thermal plants (74% in 1998, 68% in 2013), with the remainder produced by hydro (25% in 1998, 18% in 2013) and renewables (1% in 1998, 12% in 2013) (Ministry of Power, 1998*a*; Planning Commission, 2014). Of the thermal generation, the lion's share falls on coal-based generation (around 85% throughout). The share of thermal generation in a state is mainly determined by the presence of coalfields, as coal accounts for up to two-thirds of production costs in these plants (IEA, 2015; Abeberese, 2017). I collect geo-referenced data on Indian coalfields and power plants by installed capacity, ownership and year of commission.[12] Figure III.6 in Appendix III.A shows maps visualising the capacity increase and the clustering of coal-fired plants close to coalfields. In 2013, a one percent increase in the distance of a district to the nearest coalfield is associated with a 2 MW lower coal power capacity.[13]

The opening up of the power market after the Electricity Act of 2003 appears to have contributed to lower electricity prices. I examine the relationship between the median of the district level industrial electricity price and the share of installed coal fired capacity that is privately owned within a district. Table III.8 in Appendix III.B shows that the share of privately owned plants is significantly negatively associated with median electricity prices – but only after 2003.[14] A one percentage point increase in the share of privately owned plants decreases median electricity prices by 3%. I use the information on the distance to coalfields, the timing of the Electricity Act, and the share of privately owned plants for robustness checks in the empirical analysis below.

Industrial electricity prices are high in India, also due to heavy cross-subsidisation. Part of the reason why the share of privately owned plants may decrease industrial electricity prices is that they are not cross-subsidising as heavily between end-users as state and central

[12]See III.2.2 for sources.

[13]This is from a regression of installed coal capacity on logged distance to the nearest coalfield, all at the district level in 2013. This is based on 594 Indian districts. The coefficient is −191.4 with a robust t-statistic of 3.8 and R^2 of 0.066.

[14]This holds conditional on district and year fixed effects, and conditional on district and region by year fixed effects. I also control for time-varying total district level installed capacity. As the last column shows, the share of private thermal capacity is also predicted by the distance to coalfields, which I will use to construct an instrument for robustness checks.

governments.[15] Average electricity tariffs in 1998 were the equivalent of 15.7 US cents (2004 USD) for industrial users, but only 2.6 and 6.8 cents for agricultural and residential users respectively, despite cost of supply usually being lower for industry (Ministry of Power, 1998*b*).[16] While agricultural consumers made up 32% of electricity consumption in 1998, they only accounted for 3.6% of revenues from electricity sales (Planning Commission, 2002). The main reason for the heavy cross-subsidisation is political – farmers form important voting blocs that the governments try to cater to (Abeberese, 2017).

Despite efforts to reduce cross-subsidisation and depoliticize tariffs based on the Electricity Act (2003), industrial tariffs were still 7.6 US cents (2004 USD) compared to 2.2 cents for agricultural tariffs in 2013 (Ministry of Power, 2014*b*). Until 2004, India's industrial tariffs were higher than the average G7 tariff (Figure III.20 and Table III.10 in Appendix III.H) despite being a low-income country. In contrast, residential tariffs have been less than half of the G7 average.[17] While industrial tariffs have typically been above the average cost of supply, high subsidies are required for the agricultural sector. In part driven by the heavy cross-subsidisation, state electricity utilities have been loss-making almost across the board, recovering only between 73% and 89% of annual costs between 1998 and 2013 (Central Electricity Authority, 2008, 2009, 2011, 2013, 2015, 2018). The comparatively high industrial tariffs in India are important contextual information for the interpretation of the results of this chapter.

Individual states can set tariffs and cross-subsidies for different end-users and locations within their jurisdiction.[18] Industrial tariffs mostly follow increasing block tariffs, as manually collected data from government reports shows.[19] The increase of tariffs in purchase quantity in India is in contrast to block tariffs that are typically decreasing for industry. In European countries, the tariff band for the largest consumers is on average

[15] I am somewhat abstracting from generation, transmission and distribution, because utilities and generation are often integrated in India (Planning Commission, 2001; IEA, 2015).

[16] For the agricultural and residential tariffs, I calculated a simple average of state-wise average electricity tariffs, pooling consumption bands. The industrial tariffs are taken from the micro data and are comparable with reported simple averages. Values are deflated with the fuel and electricity deflator into base year 2004 and the 2004 exchange rate is applied.

[17] Based on the same data source, residential tariffs in G7 countries increased from 15.3 to 19.7 US cents (2004 USD).

[18] There is very limited regional trading of electricity. The networks across regions are in the process of getting better integrated (IEA, 2015).

[19] On average, a higher band (of five bands) is associated with a 2.5 percent increase in the tariff. This is from a regression of manually collected log deflated electricity tariffs at the state-year-band level on consumption bands, accounting for state-year fixed effects. Figure III.17 in Appendix III.G shows the average tariffs across Indian states for industrial consumers of five different sizes in 2007, using data from one of the annual government reports (Central Electricity Authority, 2008).

less than half of the tariff band for the smallest consumers (Eurostat, 2016). In any case, increasing or decreasing block tariffs are one of the challenges to identify the effect of electricity prices on firm performance that I deal with below.

Electricity prices are to be adjusted in line with cost pressures according to the 2003 Electricity Act (see also Abeberese, 2017). Coal is the dominant cost factor in coal-based electricity production. In terms of coal production, the largest public company Coal India Limited acts almost as monopoly. It supplied 81% (in 1998) and 63% (in 2013) of total domestic and imported coal (Minsitry of Coal, 2006, 2015).[20] Other public companies (mainly Singareni Collieries Company Limited) accounted for around 10% while private companies accounted for only 5% throughout this period. The production, marketing and price setting of coal is effectively controlled by the government. The prices of coal for power utilities and industry differ and are set independently.[21] Coal price adjustments for power utilities are mainly due to changes in international coal prices and the cost of production (Minsitry of Coal, 2006, 2015; Abeberese, 2017). This, in turn, affects the costs of coal-fired power plants and electricity prices, and provides a rationale for a cost shifting instrument below.

Since 2010, the coal price also contains an additional tax of 50 ₹/tonne (4% of the price) to incentivise cleaner and more energy efficient production and electricity generation. The Bureau of Energy Efficiency was created in 2002 under the Ministry of Power to coordinate policies aimed at energy efficiency. The main programmes are small credits for energy conservation, and subsidies for capital investment and energy audits (Bureau of Energy Efficiency, 2014). Ryan (2018) provides more details on these and internationally funded energy efficiency programmes in India.

Total generated electricity fell short of total required electricity by 4%-11% between 1998 and 2013 (Ministry of Power, 2018). Power shortages persist despite falling average plant load capacity factors from 79% in 2007 to 66% in 2013. In addition, India's electricity transmission losses are one of the highest in the world (IEA, 2015). Under peak times, the power shortages are higher by a few percentage points, partly driven by frequently occurring outages. Outages led to adoption of electricity generators by larger industrial plants. Importantly for the analysis below, the adoption of electricity generators is driven by smoothing over outages, not by electricity prices, since self-generation is typically

[20]Coal imports grew from 5% to 23% during this period mainly eating into the market share of Coal India Limited. The share of imported coal specifically for electricity generation was even lower (Ministry of Power, 2014*a*).

[21]This is relevant for the exclusion restriction, which I will discuss further below. See also Figure III.27 in Appendix III.J.

more expensive than buying electricity.[22] Distribution companies are not allowed to adjust electricity pricing to clear markets as a response to shortages (Allcott et al., 2016). Therefore, the correlation between annual state level electricity shortages and electricity prices is insignificant and small (see Table III.9 in Appendix III.C).[23] The main reason for shortages are problems with technical equipment or networks. Coal supply issues are only responsible for 0.2% to 3.3% of outages in thermal plants[24], and while coal supply affects electricity prices, it is thus unlikely to affect outages. These institutional features are important for the empirical analysis to identify the effect of electricity prices and not shortages. I control for shortages in robustness checks.

III.2.2 Data

Plant level data

The main data source is the Annual Survey of Industries (ASI), India's mandatory annual establishment level manufacturing survey since 1953. Its long history makes it a relatively reliable data source in the development country context. The survey divides plants into a census sector (all plants are sampled) and a sampling sector (20% within each state 4-digit-industry strata are sampled).[25] The formal firms in the ASI are representative of two-thirds of manufacturing output (Allcott et al., 2016). By combining the panel and the cross-sectional editions of the ASI, I retrieved panel identifiers as well as district codes, which are only available in the respective editions. I use an annual panel from 1998 to 2013 for the main analysis.[26]

I use the quantity and value of electricity purchased, electricity generated, and the quantity and value of coal purchased. By dividing electricity purchase value by quantity, I can calculate the average price paid for electricity at the plant level. I use further plant level data on output (sales), employees, wages, capital, investment in machineries, intermediate inputs, and other fuel expenditures (gas and oil). I construct total variable costs as the sum

[22] Bhattacharya and Patel (2008) estimate self-generation to be around 25% more expensive than buying electricity.

[23] This is in line with Allcott et al. (2016) who provide further evidence and show that a rainfall based instrument for hydro generation is also not correlated with electricity prices in India.

[24] Calculated as share of total planned and unplanned outages, annually from 1998 to 2009 using data from Allcott et al. (2016).

[25] The cutoff for the census classifier is ≥ 100 employees (until 2004 ≥ 200). The sampling frame consists of all plants ≥ 10 employees with electricity and all plants with ≥ 20 employees without electricity.

[26] The accounting year in India is from April to March. Throughout the chapter, I refer to the first year of the accounting year for ASI data and Government reports. So for example, year April 2006 to March 2007 is referred to as 2006.

of wages, input costs and other expenses, and total revenues as the sum of sales and other receipts. The difference is total profits. For the analysis of cost pass-through and incidence, I exploit the information of output sales and output quantity at the plant-product level to construct a measure of output prices and quantity.[27]

I winsorize the lowest and highest percentile of each variable within each year to reduce the sensitivity to outliers.[28] All monetary values from all sources are deflated into a common base year 2004 throughout this chapter.[29] I drop observations in non-manufacturing industries and those with a missing electricity price, electricity productivity or output. All regressions are weighted by the included sampling multiplier.

Table III.1 shows that after the cleaning steps, there are 485948 plant year observations from 160955 plants. There is considerable self-generation as the average amount of electricity self-generated is a quarter of the amount of electricity bought. This is driven by the 35% of the plants that engage in self-generation, primarily to cope with outages as discussed in the previous section. Electricity productivity is based on electricity consumed which is the sum of self-generated and purchased electricity minus electricity sold. The average electricity productivity is lower when weighting by consumed electricity, which suggests that larger electricity consumers are less electricity productive.[30] On average, electricity has the largest share in fuel expenditure (0.63).[31] Electricity expenditure constitutes on average about 6% of total average costs. The average electricity price is around seven times higher than the coal price in kWh equivalent, as coal is a rawer form of energy. Machinery is the main type of capital and investment (as opposed to e.g. buildings). The average variable cost markup (total revenues divided by total variable costs) is 20%, slightly lower than the marginal cost markup of 30%. Marginal cost markups are calculated following De Loecker and Warzynski (2012). Plant total factor productivity (TFP) are similar for different methods, following Olley and Pakes (1996), Levinsohn and Petrin (2003) or

[27] Output prices are the average of product prices, weighted by their quantities.

[28] I winsorize final variables only. That is electricity productivity (sales divided by electricity use) is winsorized before sales and electricity use are individually winsorized to avoid double winsorization.

[29] I deflate outputs and inputs using 3-digit industry deflators, investment and installed capital and machinery using a machinery deflator, wages, total revenues, total costs and total profits using a state deflator, and fuels and manually collected tariffs and prices (electricity, coal, gas, oil) using a fuel and electricity deflator.

[30] Weighting by consumption maps plant level electricity productivity into aggregate electricity productivity, comparable with Figure III.2.

[31] This is similar to the 60% that Marin and Vona (2017) report for France. Note that the share in raw energy is lower, because electricity prices are much higher per unit of energy than coal, gas or oil prices. As Figure III.16 in Appendix III.F shows, the share of electricity in the energy mix in terms of energy units has been between 16 and 20% since 1998.

Table III.1: Summary statistics from plant level data

Main variables:

	Mean
Electricity bought (GWh)	0.82
Electricity generated (GWh)	0.21
Electricity sold (GWh)	0.03
Electricity consumed (GWh)	0.99
Electricity price (₹ per kWh)	4.57
Output (in mil. ₹)	119
Electricity share in total var cost	0.06
Electricity productivity (₹ per kWh)	449
Electricity productivity (₹ per ₹)	107
Weighted by electricity consumed:	
Electricity productivity (₹ per kWh)	130
Electricity productivity (₹ per ₹)	33
Weighted by fuel consumed:	
Electricity share in fuel expenditure	0.63
Observations	485948
Firms	160955
Districts in sample	541
States in sample	32
Regions in sample	6
4-digit industries in sample	133
2-digit industries in sample	22

Additional variables:

	Mean	Obs.
Employees	72	485344
Total capital (in mil. ₹)	36	482756
Mach. capital (in mil. ₹)	21	474922
Capital investment (in mil. ₹)	8.1	483211
Mach. investment (in mil. ₹)	4.1	476043
Total revenue (in mil. ₹)	119	485867
Total variable costs (in mil. ₹)	101	485867
Total profit (in mil. ₹)	17	485867
AC-Markup (Price/AC)	1.2	485867
MC-Markup (Price/MC)	1.3	477712
TFP (Wooldridge)	7.3	477712
TFP (Levinsohn-Petrin)	9.8	477712
TFP (Olley-Pakes)	7	379040
Coal consumed (tonne)	383	485948
Coal price (₹ per tonne)	4153	49650
Coal price (₹ per kWh equivalent)	.64	49650
Coal productivity (₹ per th. tonne)	1076	49650
Coal productivity (₹ per ₹)	296	49650
Weighted by coal consumed:		
Coal productivity (₹ per th. tonne)	56	49650
Coal productivity (₹ per ₹)	23	49650

Notes: The table shows the sample means based on the pooled plant level data from 1998-2013. The means are calculated using the sampling multiplier as weights. Were indicated, the means are additionally weighted by the consumed electricity, fuel or coal to make the means more representative of aggregate productivities.

Wooldridge (2009).[32]

For robustness checks and trends in aggregate statistics, I add the 1993 and 1996 cross sectional editions of ASI micro data. I also collected aggregate ASI data at the industry by state by year level from 1967 to 1997 for long run trends.

Additional data

Coal prices for thermal power plants (as opposed to manufacturing plants) are from the Minsitry of Coal (2012, 2015). I use the published annual pit-head prices specifically for power utilities customers and inclusive of royalties and taxes, based on a representative Coal India Limited (CIL) mine and grade selected by the Minsitry of Coal (2012).[33] Shares of coal fired power plants in state installed capacity in 1998 are from the Ministry of Power

[32]See Chapter I (or Singer (2018)) for the details of an example of the TFP methodology and implementation of Wooldridge (2009). Markups are calculated following De Loecker and Warzynski (2012) after estimating production functions following Wooldridge (2009).

[33]These are the ones of Eastern Coalfields Limited of Coal India Limited, Rajmahal field, Grade E. These are also in line with those used by Abeberese (2017). After 2011, India switched the coal grading from Useful Heat Value (UHV) to Gross Calorific Value (GCV). I used the prices of the new grades G9 based on the correspondence given in Minsitry of Coal (2013). Prices are deflated with the electricity and fuel deflator from Office of the Economic Adviser (2019). Figure III.27 in Appendix III.J plots these prices in real terms.

(1998*a*, 2003).[34]

For the instrument for plant level coal prices (see Section III.3.4), I use the pit-head prices specifically for industry with the appropriate coal grades (Minsitry of Coal, 2012, 2015). Geo-located data on Indian coalfields is from Trippi and Tewalt (2011) which I combine with geo-located data of the 541 districts to calculate distances.

State-level average tariffs by consumer type and size are collected from annual reports of the Indian Central Electricity Authority (2008, 2009, 2010, 2011, 2012, 2013, 2015) and from Indiastat (2019) through Lok Sabha and Rajya Sabha questions. Data on international industrial energy prices comes from IEA (2018), and international GDP deflators, exchange rates and PPP conversion factors from World Bank (2017). Deflators for India (industry-wise, electricity and fuel, machinery) are from the Office of the Economic Adviser (2019) and the state-wise deflator is from the Reserve Bank of India (2019).

Data on state level power shortages comes from the Central Electricity Authority (2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014), and from Allcott et al. (2016) for before 2005.[35] Geo-located data on the location, capacity and ownership of coal fired power plants comes from the Center for Media and Democracy (2017), for gas plants from KAPSARC (2018), for nuclear plants from NPCIL (2015) and for hydro plants from Gupta and Shankar (2019).

III.2.3 Trends in electricity productivity and prices

To motivate the empirical analysis, I next present relevant patterns and trends in the data.

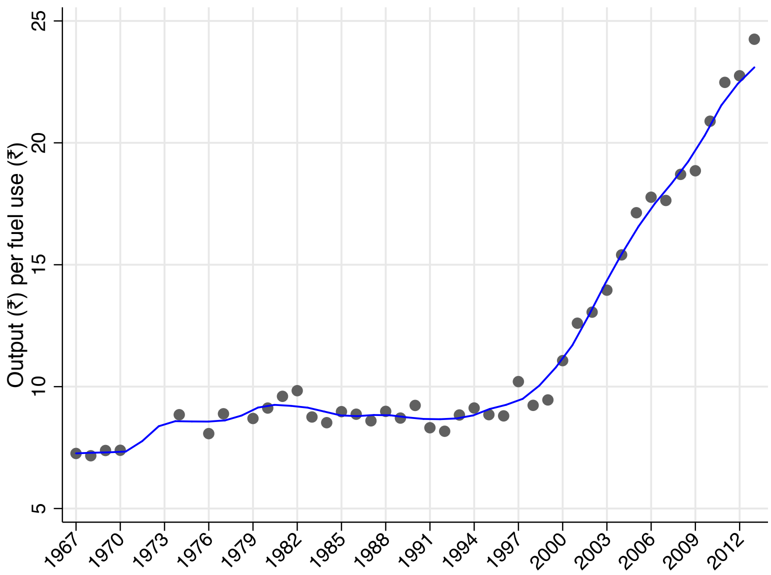
Industrial energy efficiency from 1967-2013

Combining the plant level data with sector-state level data from 1967, Figure III.1 plots the energy productivity in Indian manufacturing over 47 years. I calculate this ratio by dividing total deflated manufacturing output by total deflated fuel use (electricity, coal, oil, gas). Between 1967 and 1999, energy productivity was rather constant between 7 and 10 ₹ per ₹. From 2000, there was a remarkable increase in energy productivity, which more than doubled until 2013. This was not driven by a particular industry alone. Figure III.10 in Appendix III.E shows that a similar trend appears from 2000 for different industry groups. Furthermore, Figure III.7 in Appendix III.D shows that similar trends occurred across

[34] Thermal shares as on 31st of March 1998, one day before the beginning of the sample I use. Chhattisgarh, Jharkhand and Uttarakhand were created in 2000, and thermal shares correspond to 31st of Jan 2003, the first available data. I follow Abeberese (2017) using these shares.

[35] Data on the type of forced outages are also from Allcott et al. (2016).

Figure III.1: Indian long run energy productivity in manufacturing



Notes: The figure plots annual energy productivity ratios (value of output divided by the value of fuel and electricity used). Output is deflated at the 2-digit industry level using 2-digit industry deflators before aggregating over industries. Fuel and electricity use is deflated using a general fuel and electricity wholesale price deflator. The ratio of aggregate output to aggregate fuel and electricity consumption is displayed. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. From 1967 to 1997 the raw ASI data in pre-aggregated form is used (at industry state year aggregation). From 1998 the raw plant level ASI data is used and aggregated with sampling multipliers.

all states. This secular increase in India is entirely different than the evolution in OECD countries, as Figure III.15 in Appendix III.F shows. The increase in energy productivity is consistent with the drop in emission intensity from 1990-2010 for a subsample of large firms reported in Barrows and Ollivier (2018). The fuel with the highest share in energy costs is electricity. I next examine electricity productivity.

Industrial electricity productivity and prices 1993-2013

Figure III.2 mirrors a similar trend in aggregated electricity productivity in output per kWh from 1993 using solely micro data. From 2000 electricity productivity increased by 35%.[36] This trend did not occur because of substitution away from electricity. The share of electricity in fuel expenditure was 65% in 2000 and 63% in 2013.[37] Interestingly, electricity prices fell from 1999/2000 during this secular increase in electricity productivity, and almost halved by 2013 (right panel of Figure III.2). The purpose of this chapter is to

[36] Also the "other" fuel productivity increased considerably since 2000, as Figure III.14 in Appendix III.F shows.

[37] It was around 16 to 20% in energy unit terms, as Figure III.16 in Appendix III.F shows.

analyse whether there is a causal mechanism relating those two trends. In fact, the two figures of aggregate data visualise the causal results surprisingly well. A regression of the aggregate logged electricity productivity on aggregate logged electricity prices yields an elasticity of -0.4. This is the opposite sign of the OLS plant level estimate, but remarkably close to the IV estimates in the main analysis.

The increase in electricity productivity occurred in all sectors, except for perhaps metals and minerals (see Figure III.11 in Appendix III.E), and in most states (see Figure III.8 in Appendix III.D). Alternative data sources (IEA, 2016; UNIDO, 2016*a*) match the pattern of electricity productivity in Figure III.15 in appendix III.F. The secular electricity price decline is similar within states (Figure III.9) and within industries (Figure III.12). Two alternative sources of electricity prices confirm the price trends. Figure III.18 plots the electricity price index in real terms from the Office of the Economic Adviser (2019) (Appendix III.G), and Figure III.19 plots the average of industrial electricity tariffs collected from the reports of the Central Electricity Authority. The price trend in the 2000s is in contrast to many other countries, where electricity prices rather increased. Figure III.20 (Appendix III.H) plots industrial electricity prices for a range of OECD and non-OECD countries. While electricity prices in India almost halved during the sample period, prices in OECD countries grew by 40% (see Table III.10).[38]

In summary, the trends are similar across different data sources, and provide a rather unique setting to study their relationship.

III.2.4 Heterogeneity in electricity productivity and prices

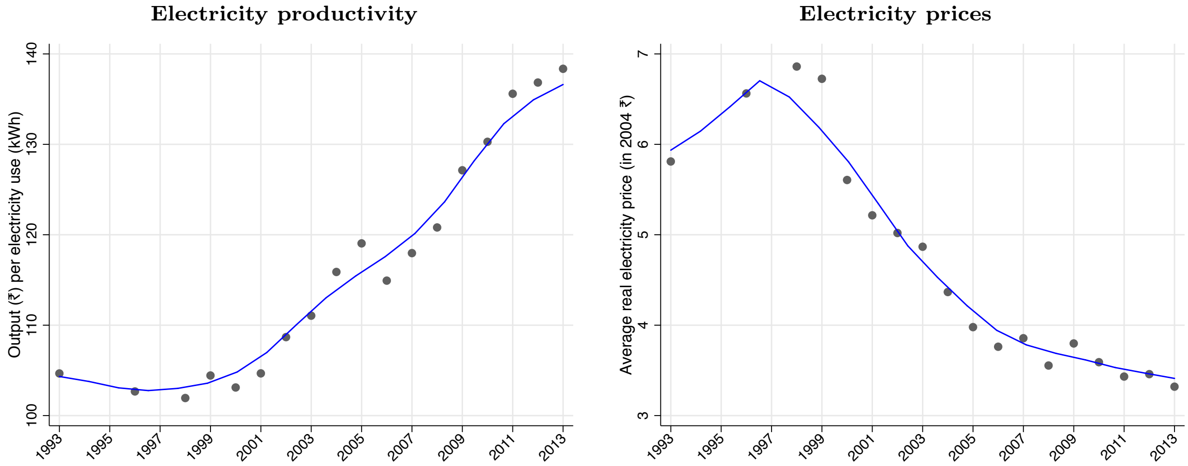
Is there much variation to explain across plants? The aggregate graphs in Figure III.2 mask substantial heterogeneity, both in terms of productivity and prices. Figure III.3 plots the histogram of electricity productivity and prices in 2003.[39] Even when partialling out state-industry (4-digit) effects, there remains substantial variation. The 90th to 10th percentile range drops from 3.5 to 2.7 for logged electricity productivity, and from 2.1 to 1.4 ₹ for electricity prices. The electricity productivity dispersion is even larger than the TFP dispersions found in the literature (Bartelsman and Doms, 2000; Syverson, 2004*b*, 2011). Plants at the 90th percentile pay around 50% higher electricity prices than those at the 10th percentile.

I decompose the variance following Davis et al. (2013) (see Figure III.4 for details).

[38]See Sato et al. (2019) for more evidence on general price trends in various countries since 1995. They show that electricity is the most important fuel when accounting for overall energy prices.

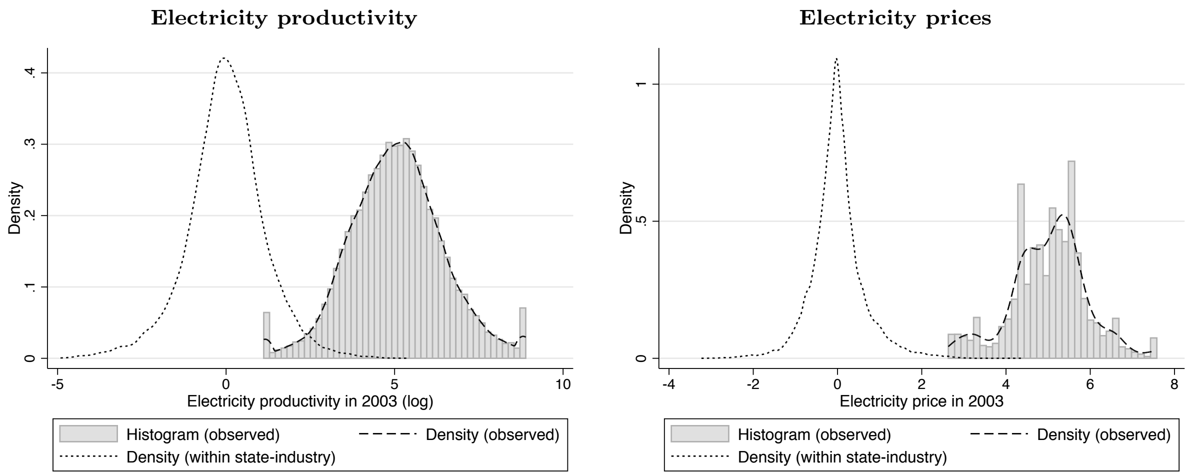
[39]Similar plots are shown in Figure III.21 and Figure III.22 in Appendix III.I for all years.

Figure III.2: Electricity productivity and electricity prices in manufacturing



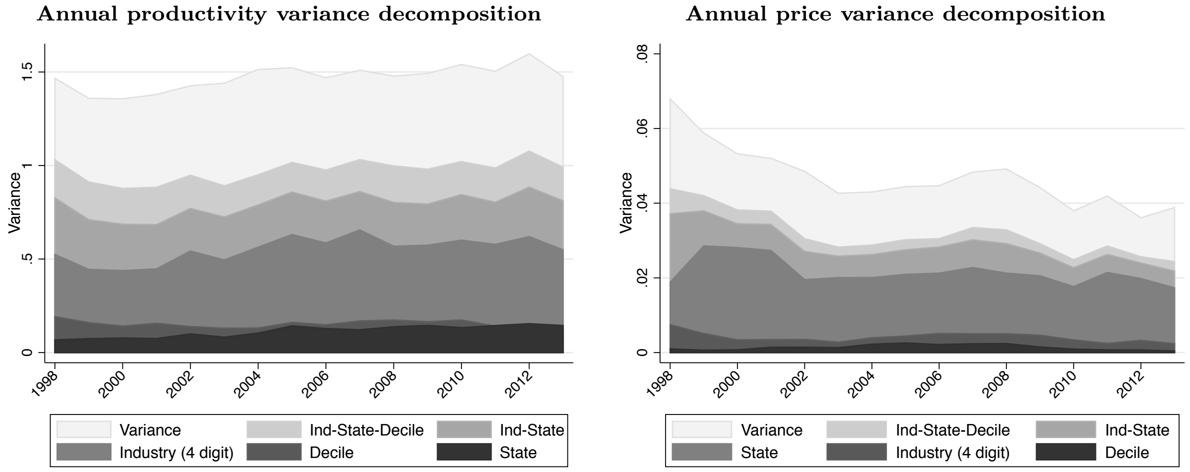
Notes: The left panel plots annual electricity productivity ratios (value of output divided by the quantity of electricity used in kWh). They are calculated by first aggregating the value of output and the quantity of electricity bought by plants, and then taking the ratio of the aggregates. The right panel plots real average electricity prices. They are calculated by first aggregating the value of electricity bought by plants and the quantity bought, and then taking the ratio of the aggregates. Plant output is deflated using 3-digit industry deflators before aggregating over industries. Electricity values are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. All data points come from the raw plant level ASI data (from 711166 observations including years before 1998) and aggregated with sampling multipliers.

Figure III.3: Heterogeneity in electricity productivity and in electricity prices



Notes: The left panel plots the histogram of plant level logged electricity productivity in 2003. Electricity productivity ratios are the value of output divided by the quantity of electricity used in kWh. The kernel density plot to the left shows the distribution of the residuals of logged electricity productivity after partialling out state by 4-digit industry by year fixed effects. The right panel plots the histogram of plant level electricity prices in 2003. The kernel density plot to the left shows the distribution of the residuals of electricity price after partialling out state by 4-digit industry by year fixed effects. Both panels are similar for all years as shown in Figure III.21 and Figure III.22 in Appendix III.I. Plant output is deflated using 3-digit industry deflators. Electricity prices are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

Figure III.4: Electricity productivity and price variance decomposition

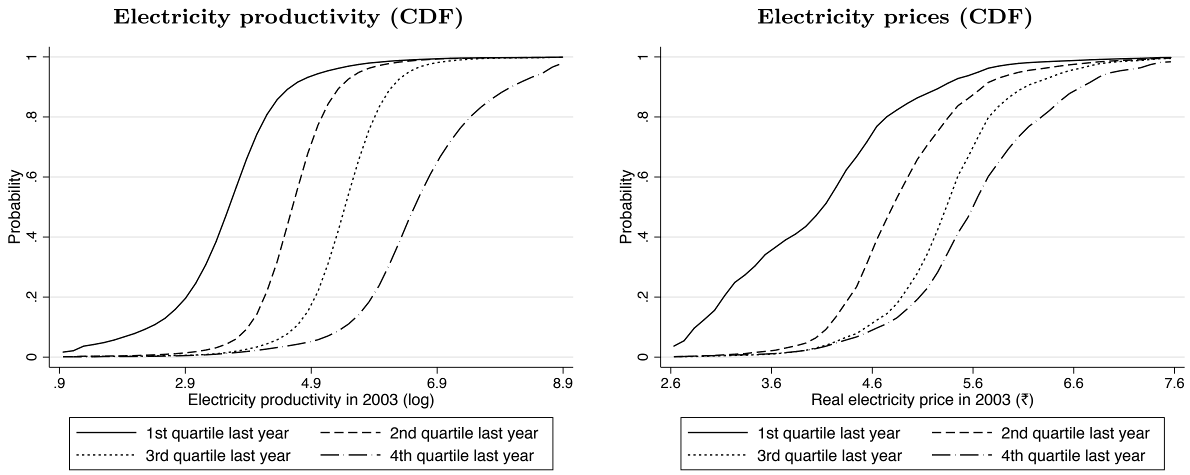


Notes: The left panel plots the annual total variance of logged electricity productivity and the variance explained by specified groups. The right panel plots the same for logged electricity prices. The annual variance is calculated as $V = \sum_e s_e \left(p_e - \overline{p}\right)^2$, where s_e are purchase weights multiplied by the sample multiplier, p_e are logged electricity productivity or prices, $\overline{p}$ the weighted average log productivity or price. I use the decomposition of Davis et al. (2013) to decompose total variance into a within "group" component V^W, and a component across "groups" V^G:

$$V = \sum_e s_e \left(p_e - \overline{p}_g\right)^2 + \sum_g s_g \left(\overline{p}_g - \overline{p}\right)^2 = V^W + V^G$$

where $s_g = \sum_{e \in g} s_e$ and $\overline{p}_g$ the weighted average of log productivity or price within group g. I calculate the decomposition separately five times for the five groups shown in the graph. The regions plot V^G and correspond to the across-group variance in the total variance V, where higher shares explain more of the variation (see also Figure III.23). Groups are deciles of electricity purchase quantity, 4-digit industries, states, and combinations. Plant output and electricity prices are deflated.

Figure III.5: CDFs of plant electricity productivity and prices in 2003 conditional on 2002 quartiles



Notes: Plotted are the CDFs in 2003, separately for each quartile of the respective values in 2002. The left panel shows the distribution of the logged electricity productivity (i.e. the value of output divided by the electricity use in kWh). The right panel shows the distribution of the electricity price. The CDFs are empirical CDFs obtained through a Gaussian kernel smoother with bandwidth 0.1. The graphs show that each higher quartile first order stochastically dominates the lower quartiles. The conditional CDF of the plants that belong to the higher *previous* year quartile lies to the right of the CDF of the plants belonging to the lower *previous* year quartile. While individual plants move up and down the ranking of electricity productivity and energy prices from one year to the next, the probability of higher productivity and prices increases in last periods productivity and prices. The conditional CDFs for other years look similar, see Figure III.25. Plant output and electricity prices are deflated.

The state-industry effects can only account for around 60% of the cross-sectional variance in electricity prices, and 50% of electricity productivity (see Figure III.23 in Appendix III.I).[40] For electricity productivity, there is more variation across industries, while for electricity prices there is more variation across states. This is intuitive, as production techniques tend to vary more across industries, while electricity price setting varies more by states as explained in Section III.2.1. The main analysis accounts for industry by year by region fixed effects to account for differences in electricity productivity across industries. Importantly, we learn from these descriptives that there is enough interesting variation left after accounting for these fixed effects.

The deciles of electricity consumption in India cannot explain much of the variance. This is in contrast to the findings for the US (Davis et al., 2013) and France (Marin and Vona, 2017) and consistent with the observation in Section III.2.1 that tariff schedules are increasing or decreasing in India. The variance in electricity prices has been decreasing from 1998 to 2013. Figure III.24 in Appendix III.I plots quantiles of the distribution over time and shows a convergence in electricity prices that accompanied the secular decline. Interestingly, when we compare the decrease in the total variance of electricity prices in Figure III.4 with the constant shares in Figure III.23, we can conclude that the convergence has not been driven by reductions across industries or states alone, but by overall convergence.

Finally, I study the persistence of electricity productivity and prices within plants. Following Farinas and Ruano (2005), I plot the conditional (on previous period values) CDF of logged electricity productivity and electricity prices in Figure III.5. I divide the sample into four quartiles based on previous period values and plot the four CDFs separately. As the CDF of the higher quartiles are to the right for every value, they first order stochastically dominate the distributions of plants ranked in lower previous period quartiles to the left. Plants from a higher previous quartile are more likely to belong to the higher quartile in the current period.[41] Both electricity productivity and electricity prices are persistent. The implication of this persistence for the analysis is that I use variation within *and* across plants, which I will discuss in the next section.

[40] Variation across districts (not plotted) can explain around 22% and 45% of electricity productivity and electricity prices respectively. Districts for the later years are not available for all observations.

[41] See Figure III.23 in Appendix III.I for the same conclusion for a different year.

III.3 Empirical strategy

There are substantial endogeneity concerns when estimating the relationship between electricity productivity and electricity prices. The baseline specification is:

$$y_{jisrt} = \beta \log(P^E)_{jisrt} + \alpha_{irt} + \epsilon_{jisrt} \tag{III.1}$$

where y_{jisrt} is the logged outcome (electricity productivity as output divided by electricity consumed in kWh) for plant j in industry i in state s in region r in year t, and P^E is the electricity price. The analysis is conditional on 4-digit industry by region by year fixed effects α_{irt}. This accounts for aggregate technology and price trends that can differ by industry.[42] India is divided into six regions, and there is poor integration of electricity markets across regions (IEA, 2015; Ryan, 2017; Ministry of Power, 2018), and therefore α_{irt} allows for differential fixed effects across regions.

III.3.1 Endogeneity concerns

Within these clusters, there are still endogeneity concerns. The exogenous component of prices $\log(P^E)_{jisrt}$ are mostly at the state-year or district-year level as discussed in Section III.2.1. These are price adjustment due to electricity generation cost pressures for example. However, $\log(P^E)_{jisrt}$ also contain endogenous variation. Suppose the endogenous elements contained in the price can be expressed as ξ_{jisrt} at the plant level and λ_{isrt} at the industry level within states. Both these elements are also contained in the composite error term $\epsilon_{jisrt} = \xi_{jisrt} + \lambda_{isrt} + \mu_{jisrt}$, where μ_{jisrt} is the true random component. Using plant fixed effects would not address any endogeneity that is time varying at the plant level. I return to additional problems associated with plant fixed effects in Section III.3.6. Industry by state by year effects would eliminate most exogenous variation as well. My strategy is to rely on instruments which are not correlated with λ_{isrt} and ξ_{jisrt} and therefore isolate the exogenous price variation. Before explaining my identification strategy I describe the main endogeneity concerns.

First, shocks to output and electricity demand (in ξ_{jisrt}) also affect electricity prices due to different tariffs for different consumption bands (see e.g. Figure III.17 in Appendix III.G). Second, plants or groups of firms within an industry may negotiate or exert pressure for lower electricity prices (in ξ_{jisrt} and λ_{isrt}). Their bargaining power in turn is likely

[42] There are 133 4-digit industries in the final sample. There are six regions, 32 states and 541 districts in the final sample.

related to their economic performance as well. This can lead to reverse causality problems at the plant level. Third, shocks to industries and regions may jointly affect economic performance, electricity productivity and electricity pricing (in λ_{isrt}). The third concern is at least partially taken care of by the industry by region by year effects. Prices may also be adjusted across the board as a response to changes in electricity productivity and electricity demand. I use lagged electricity prices to address reverse causality issues at the more aggregate level and find similar results. Fourth, even within states, plants may locate where electricity prices are low and that may be correlated to their electricity productivity and consumption (in ξ_{jisrt}). Sixth, average electricity prices at the plant level may suffer from measurement error (in ξ_{jisrt}). The two instruments discussed next aim to isolate the exogenous variation in prices from ξ_{jisrt} and λ_{isrt}.

III.3.2 An instrument based on other plants (IV^A)

The main idea of the first instrument is to extract the exogenous signal of the prices by relying on prices of other plants, which must also have been affected by exogenous electricity price changes. The exogenous part is mainly at the state-year level. Some weighted average of other plants could therefore extract the common exogenous signal. In order to avoid capturing the endogenous component λ_{isrt} in the instrument as well, I rely on information of plants in the same state, but in different industries. Specifically, I use prices of plants with similar purchase quantities in the same year, in the same state, but in different 2-digit industries i^{2d}. The underlying assumption is that the endogenous components λ_{isrt} are not correlated across 2-digit industries within a state. They are allowed to be correlated across 4-digit industries within 2-digit industries.[43] Recall that industry by region by year effects are taken out, so the element in λ_{isrt} common within regions are allowed to be correlated across 2-digit industries as well. The second assumption is that the (weighted) average of ξ_{jisrt} of plants in other industries is not correlated to the plant specific ξ_{jisrt}.

I use plants with similar purchase quantities to address the structure of tariffs which are based on purchase quantities. The instrument is a weighted average of prices of other plants, weighted by the distance in their purchase quantities, which smooths out individual

[43] There are 22 2-digit industries and 133 4-digit industries in the final sample.

shocks. I use a triangular kernel function to determine weights:

$$w_{q^*}(q_j) = \begin{cases} \frac{b_{q^*} - |\log(q_j) - \log(q^*)|}{b_{q^*}^2} & \log(q_j) \in [\log(q^*) - b_{q^*}, \log(q^*) + b_{q^*}], \\ & \forall s_j = s_{j^*}, t_j = t_{j^*}, i_j^{2d} \neq i_{j^*}^{2d}. \\ 0 & \text{otherwise} \end{cases} \tag{III.2}$$

where q^* is the electricity quantity purchased in kWh by plant j^* that we want to create the instrument for, and q_j is the electricity quantity purchased by other plants j. The cutoff b_{q^*} is the 25th percentile of the distribution of the logged ratio of the purchase quantities in absolute terms $|\log(q_j) - \log(q^*)|$, and is thus allowed to vary by plant j^* that we want to instrument for.[44] That is, the support of the kernel weights is over the 25% of plants that are closest in terms of electricity purchased, conditional on being in the same state $s_j = s_{j^*}$ and year $t_j = t_{j^*}$ and in different 2-digit industries $i_j^{2d} \neq i_{j^*}^{2d}$, and the weight decreases linearly in the distance of logged purchase quantity. The first instrument IV^A for the electricity price of plant j^* is then the average of the electricity prices of other plants P^E_{jisrt}, weighted by the triangular kernel weights:

$$IV^A_{j^*isrt} = P^E_{jisrt} \frac{w_{q^*}(q_j)}{\sum_{q_j} w_{q^*}(q_j)} \tag{III.3}$$

This instrument alleviates the concerns laid out above. It takes care of bargaining power and price distortions through corruption of a particular plant, as well as groups of plants within an industry, as only plants from all other 2-digit industries are considered. The kernel smooths over the discontinuities of different consumption and price bands. The instrument also takes care of plant location sorting within states and measurement error of prices at the plant level. What the instrument captures are price movements at the state level for similar consumption quantities, which are primarily driven by generation cost factors (see Section III.2.1) after filtering out the endogenous components described above.

The instrument is similar to the Hausman instruments in demand estimation, which instrument goods prices with prices of the same good in other cities (Hausman et al., 1994; Hausman, 1996; Nevo, 2001). They are relevant because they share the common marginal costs of producing the good (electricity). My (and the Hausman) instruments assume that there are no endogenous factors that are common across plants from *different* (2-digit) industries that affect their electricity productivity and the pricing of electricity

[44]The advantage of a bandwidth that is flexible rather than fixed is to ensure that enough observations are used for the construction of the instruments. I also tried the 10th and the 50th percentile, as well as a fixed cutoff based on the average 25th percentile, with similar results.

simultaneously. In a robustness check, I also exclude plants from the instrument which are based in the same district (IV^C). This allows for endogenous components in prices that are spatially correlated within districts, and the results are quantitatively very similar.

The advantage of this instrument is that it can be readily calculated in other settings. This facilitates comparable analyses and further explorations of the relationship between electricity prices and electricity productivities in developing vs. developed, as well as in high price vs. low price, countries. Future work will follow up on this.

III.3.3 A shift-share instrument based on electricity generation (IV^B)

The main idea for the second instrument is to use a cost shifter for electricity generation directly, following Abeberese (2017).[45] Since coal is the largest cost factor in electricity generation (see Section III.2.1), the price of coal shifts electricity generation costs, and therefore electricity prices. The instrument is based on a shift-share structure as in Bartik (1991). The shifters are nationally representative coal prices specifically for power utilities (see Section III.2.2). It is weighted by the shares of thermal coal fired installed capacity in total installed capacity at the state level:

$$IV^B_{srt} = \log(P_t^{CoalPower}) \frac{\text{coal based installed capacity}_{sr_{1998}}}{\text{total installed capacity}_{sr_{1998}}} \tag{III.4}$$

I use the pre-sample shares of installed capacity in March 1998. I provide a map of the thermal shares in Figure III.26 in Appendix III.J. As discussed in Section III.2.2, the coal price for power utilities is set independently to the coal price for industry, and is thus unlikely to directly affect manufacturing plants. Figure III.27 in Appendix III.J plots both coal prices in real terms, and shows that often one decreases while the other increases at the same time.[46]

This isolates the exogenous movements in electricity prices, driven by cost pressures from coal prices. It addresses the endogeneity concerns raised in the beginning of the section, including common endogenous movements in electricity productivity and electricity prices at the state-year level, as the coal price used in the instrument does not vary across states. While the coal prices for power utilities and industries are set independently, I also exclude industries that use coal in the sectoral analysis and find similar results.

[45] A similar shift-share instrument for energy prices relying on thermal shares in generation has been used in Abeberese (2017), Ganapati et al. (2016) and Elliott et al. (2019). Linn (2008) and Marin and Vona (2017) use national energy prices directly interacted with fixed fuel shares at the plant level.

[46] See also Abeberese (2017) for more discussion.

An advantage of instrument IV^B is that it might be less susceptible to the above described specific types of common shocks that threaten the validity of instrument IV^A, if they exist. The two disadvantages of IV^B are that it tends to be much weaker than IV^A and that it relies on external data.

III.3.4 Two similar instruments for coal prices (IV^E and IV^F)

In Section III.4.4 I compare the effect of electricity prices to the effect of coal prices. This provides additional support for the hypothesis that electricity prices can have distinct effects. Specifically, I ask whether higher electricity prices have more adverse effects than higher coal prices. Coal prices suffer from a similar endogeneity concern as electricity prices. I construct two instruments that are similar to the ones above. The first instrument, IV^E is the analogue to IV^A, using coal prices of plants in the same state, but from different 2-digit industries, without the kernel weights. The second instrument, IV^F, is a shift-share instrument like IV^B. The shares are the logged distances of district centroids to the nearest coalfields. The distance increases sourcing costs. The shifter is the nationally representative coal price (at pit heads) for industry (as opposed to power utilities), taken from the Minsitry of Coal (2012, 2015). The location of coalfields is illustrated in Figure III.6 in Appendix III.A.

III.3.5 Recovering pass-through rates and consumer incidence

While plants have to pay the electricity costs, the incidence of higher electricity prices may be shared between producers and consumers. The degree to which incidence falls on consumers depends on one hand on the degree to which electricity prices affect marginal costs ($\gamma \equiv dMC/dP^E$), which depends on the ability to substitute. On the other hand, it depends on the pass-through rate of marginal costs to output prices ($\rho_{MC} \equiv dP/dMC$), which depends on market structure and market power. I employ a partial equilibrium analysis following Ganapati et al. (2016) that allows for factor substitution, incomplete pass-through and imperfect competition. As they show, under the assumption that average variable costs are equal to marginal costs ($AVC = MC$) incidence on consumers in a generalised oligopoly, where CS and PS are consumer and producer surplus, is:

$$I \equiv \frac{dCS/dP^E}{dPS/dP^E} = \frac{\rho_{MC}}{1 - (1 - L\epsilon_D)\,\rho_{MC}} \tag{III.5}$$

where $\rho_{MC} \equiv dP/dMC$ is the pass-through rate of marginal costs to prices, $L \equiv (P - MC)/P$ is the Lerner (1934) index, and $\epsilon_D \equiv -[dQ/dP][P/Q]$ the market elasticity of demand. I next describe how I recover the three required parameters L, and ϵ_D and ρ_{MC}.

There is an established literature recovering markups μ from the production side from firm revenue and input data (Hall, 1988, 1990; Hall and Jones, 1999; De Loecker and Warzynski, 2012). The basic intuition is that if plants are cost minimising, we can use the first order condition of a variable input, which describes a relationship between markups, the output elasticity of that input, and the revenue share of that input. I follow this literature to estimate plant level markups (μ) and the plant level Lerner index L, using materials as variable input. I estimate the output elasticity along with TFP using Wooldridge (2009) building on Levinsohn and Petrin (2003).

It is well know that for standard oligopolistic environments, the first order conditions of firm profit maximisation imply a mapping between markups and demand elasticities.[47] For the market level demand elasticities ϵ_D, I take the median of the plant level demand elasticities within a 4-digit industry by year by state cluster.[48] Market demand conditions are thus allowed to vary across industries, time and space. The alternative is to estimate demand functions as e.g. in Ganapati et al. (2016). The two approaches require different assumptions. Since we need to estimate markups and production functions in any case and assume oligopolistic competition and cost minimisation already, the additional profit maximisation assumption to recover demand elasticities appears innocuous. Independently of how demand elasticities are recovered, the main challenge is to get estimates for the pass-through.

Estimating the pass-through parameter ρ_{MC} requires data on revenues and output quantity. The most direct way is to regress prices on marginal cost. Revenues and quantities are separately reported for most plants in the data, which allows me to calculate average sales prices at the plant-product level. I calculate the plant level average price across products, weighted by the quantity of each product. From the estimated plant level price marginal cost markups μ, I can back out plant level marginal costs with these prices. I recover prices and marginal costs for 87% of the 485948 observations, covering 121 of the 133 4-digit industries. Since I also construct total variable cost, I can recover AVC by dividing total variable costs by quantity. This allows me to examine the validity of the underlying assumption ($AVC = MC$) for Equation (III.5). A regression of logged AVC on logged MC yields a coefficient of 0.98 and an R^2 of 0.95, which suggests that the

[47] For example, see the first chapter of this thesis.

[48] Plant level markups (and demand elasticities) can diverge from the market demand elasticities due to distortions for example. The first chapter of this thesis provides a good example of such distortions (alternatively see Singer (2018)). Taking the median or mean of production or demand elasticities is common in the literature, see e.g. Asker et al. (2014). The median is more robust to outliers.

assumption is not unreasonable.

The pass-through parameter ρ_{MC} is likely to differ by industry and firms, depending for example on the market structure, concentration or market power. I estimate a pass-through *elasticity* for each 4-digit industry separately, regressing prices ($\log(P)$) on marginal costs ($\log(MC)$). I instrument for the endogenous marginal costs using the two instruments for the electricity price IV^A and IV^B described above.[49] The pass-through elasticity is converted into the pass-through rate ρ_{MC} by multiplying it with the plant level markup μ. To summarise, the empirical components are:

$$\widehat{L}_{jisrt} = 1 - \frac{1}{\widehat{\mu_{jisrt}}} \tag{III.6}$$

$$\widehat{\epsilon}_{D,isrt} = \text{MEDIAN}_{isrt}\left(\frac{1}{1 - 1/\widehat{\mu_{jisrt}}}\right) \tag{III.7}$$

$$\widehat{\rho}_{MC,jisrt} = \widehat{\mu}_{jisrt}\widehat{\frac{d\log(P_{jisrt})}{d\log(MC_{jisrt})}} \tag{III.8}$$

Finally, the incidence of consumer surplus as share of total incidence is:

$$I^{share} = I/(1+I) \tag{III.9}$$

III.3.6 Specification choice and estimation

I conclude this section by making a few remarks about model specifications and estimation. First, I do not include state by year effects for the baseline specification. This is because IV^B only varies at the state by year level and most of the exogenous variation is also at the state by year level.

Second, I do not include plant fixed effects for the baseline specification. This is primarily because the IVs address time varying plant unobservables, while plant fixed effects cannot address those. On the contrary, plant fixed effects could introduce bias because of violations of the strict exogeneity assumption that comes with it. Past shocks to output and electricity productivity are likely correlated with current electricity prices, as block tariffs increase or decrease with consumption, violating the strict exogeneity condition for fixed effects. Moreover, much of the interesting variation is between plants. I showed in Section III.2.4 that electricity productivity and prices are persistent within

[49] Endogeneity concerns arise for example because marginal costs are estimated leading to measurement error. I use the instruments separately. For each industry, I take the weighted average of the two IV coefficients, where the weights are the t-statistics.

plants. A regression of logged electricity productivity on plant fixed effects can explain 80% of the variation (R^2). Additionally, including plant fixed effects can be thought of as exploiting shorter-run variation, as in Ganapati et al. (2016) for example. The mechanisms in this chapter, e.g. scaling and upgrading production processes, are likely to be more relevant in the medium to longer run. In robustness checks, I included state trends or plant fixed effects, which leaves us with broadly similar conclusions.

Third, I exploit the panel structure for calculation of standard errors in all specifications. I two-way cluster standard errors at the plant level, and at the state by year level, since one of the instruments varies at that level. I provide robustness checks clustering at the district, and the region by year level with similar results. Since I am running the same model with multiple outcomes, I apply the Holm (1979) Bonferroni correction for multiple hypothesis testing in Table III.21 in Appendix III.M. Finally, I use the two instruments separately to enable comparisons, but provide an over-identified IV-regression with two instruments as robustness check.

III.4 Results

I first present the main results, along with robustness checks. Then I explore mechanisms and calculate incidence towards the end of this section.

III.4.1 Electricity prices and electricity productivity, use and output

First stages

The first stage coefficients, standard errors and Kleibergen Paap F-statistic are reported in each table for each regression separately. For the main specifications, Table III.2 shows that both instruments are strong and shift the endogenous electricity price in the expected direction.

Lower electricity prices improve electricity productivity

The correlation between electricity prices and electricity productivity is positive. An OLS regression of logged electricity productivity on logged electricity prices suggests an elasticity of 0.37 (Column (1) in Table III.2). The endogeneity bias in these estimates is large, however. The causal IV estimates in Column (2) and (3) are of opposite sign and statistically highly significant. A one percent increase in electricity prices is associated with a 0.24 or 0.78 percent decrease in electricity productivity for the IV^A based on other plants and the shift-share IV^B respectively. The positive bias in the OLS estimates suggests

Table III.2: Electricity prices and electricity productivity

	Electricity productivity (log)		
	(1)	(2)	(3)
$\log(P^E)$	0.366***	-0.239***	-0.776***
	(0.044)	(0.070)	(0.105)
OLS/IV	OLS	IV^A	IV^B
Observations	485948	485948	485948
Ind-region-year FE	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***
First stage SE	-	0.005	0.003
F-stat (Kleib.-Paap)	-	43147.813	296.255
SE clustered by	Plant	Plant	Plant
No. of first clusters	160955	160955	160955
SE clustered by	State-year	State-year	State-year
No. of second clusters	501	501	501

Notes: The dependent variable is logged electricity productivity (value of output divided by the quantity of electricity used in kWh). Each column represents a separate regression at the plant level. The first column reports the results from an OLS regression on logged electricity prices. The second column uses the IV^A based on the electricity prices of similar plants. The third column uses the shift-share IV^B. The first stage statistics are reported. All regressions contain industry by year by region fixed effects. Regressions are weighted by the recorded sampling multiplier. Standard errors in parentheses are two-way clustered at the plant and the state by year level. Plant output is deflated using 3-digit industry deflators and electricity prices are deflated using a general fuel and electricity wholesale price deflator.

that less efficient plants also manage to obtain lower electricity prices through deliberate exemptions, negotiations, corruption or location choices, for example. The effect is more strongly negative for IV^B, which could be due to heterogeneous local average treatment effects, but it is reassuring that both instruments significantly correct the OLS bias in the same direction.

As documented in Section III.2.3, there was a secular increase in aggregate electricity productivity (35%) with a concurrent reduction in electricity prices of 45% during the sample period . How well can the causal estimates from micro data explain this aggregate phenomenon? In a back of the envelope calculation taking the average of the IV^A and IV^B estimates as -0.51, the documented reduction of electricity prices predicts a $(1-0.45)^{-0.51} = 36\%$ increase in electricity productivity. Considering that the simple OLS correlation is of opposite sign, these estimates can explain the aggregate secular trends remarkably well.

Electricity prices affect electricity consumption and output

Why have lower electricity prices improved electricity productivity in India? Higher electricity prices still do reduce electricity consumption. Table III.3a presents the regressions split up into the components of electricity productivity, with logged electricity consumption (in kWh) or logged output as dependent variables. In both the OLS and IV regressions, electricity prices reduce electricity consumption, with the causal effect being slightly larger.

Table III.3: Electricity prices, output, electricity use, and lagged electricity prices

(a) Electricity prices, output and electricity use

	Output (log))			Electricity consumption (log)		
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(P^E)$	-0.0265	-0.743***	-1.597***	-0.385***	-0.479***	-0.797***
	(0.073)	(0.143)	(0.153)	(0.064)	(0.155)	(0.148)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	485948	485948	485948	485948	485948	485948
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***	-	0.97***	0.06***
First stage SE	-	0.005	0.003	-	0.005	0.003
F-stat (Kleib.-Paap)	-	43147.813	296.255	-	43147.813	296.255
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes

(b) Lagged electricity prices and electricity productivity

	Electricity productivity (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\log(P^E)$	0.296***	-0.272***	-0.735***			
	(0.049)	(0.062)	(0.087)			
Lagged $\log(P^E)$				0.0177	-0.274***	-0.727***
				(0.042)	(0.060)	(0.086)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A(lag)	IV^B(lag)
Observations	225833	225833	225833	225833	225833	225833
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.98***	0.06***	-	0.98***	0.07***
First stage SE	-	0.005	0.003	-	0.005	0.003
F-stat (Kleib.-Paap)	-	46140.249	421.264	-	39687.361	405.830
SE clustered by	Plant	Plant	Plant	Plant	Plant	Plant
No. of first clusters	67834	67834	67834	67834	67834	67834
SE clustered by	State-year	State-year	State-year	State-year	State-year	State-year
No. of second clusters	469	469	469	469	469	469

Notes: The dependent variable in panel (a) is logged output or logged electricity consumption (in kWh) as indicated. The dependent variable in panel (b) is logged electricity productivity (value of output divided by the quantity of electricity used in kWh). Each column represents a separate regression at the plant level. The first three columns in panel (b) restrict the sample to the same observations as in the last three columns, where lagged logged electricity prices (and lagged instruments) are used. Regressions are weighted by the recorded sampling multiplier. Standard errors in parentheses are two-way clustered at the plant and the state by year level. The rest of the table layout follows the same structure as the main Table III.2.

A one percent increase in electricity prices reduces physical electricity consumption by 0.49 to 0.81 percent.

The OLS effect of electricity prices on output is close to zero. In contrast, the IV estimates of the output elasticity are large and negative (between -0.74 and -1.59). The endogeneity bias operates through both electricity consumption and output, but mainly through the latter. Tariff schedules that increases in size biases OLS output estimates upwards. The effects on electricity consumption and output are close to the ones in Abeberese (2017). While she does not examine electricity productivity at the plant level, she finds that firms with high electricity prices produce products that are typically less electricity intensive on average, suggesting product switching as a channel. I will explore further mechanisms in Section III.4.4.

Table III.4: Electricity prices and electricity productivity in high price periods

	Electricity productivity (log)		
	(1)	(2)	(3)
$\log(P^E)$	0.471***	0.00847	-0.732***
	(0.061)	(0.094)	(0.168)
$\log(P^E) \cdot \mathbf{1}(year < 2006)$	-0.217**	-0.531***	-0.0926
	(0.084)	(0.128)	(0.193)
OLS/IV	OLS	IV^A	IV^B
Observations	485948	485948	485948
Ind by region by year FE	Yes	Yes	Yes
First stage coef. 1/1	-	0.96***	0.06***
First stage SE 1/1	-	0.006	0.005
First stage coef. 1/2	-	0.03***	0.01
First stage SE 1/2	-	0.009	0.007
First stage coef. 2/1	-	-0.00***	-0.00***
First stage SE 2/1	-	0.000	0.000
First stage coef. 2/2	-	0.99***	0.06***
First stage SE 2/2	-	0.007	0.005
F-stat (Kleibergen-Paap)	-	11055.255	68.011
Two-way cluster plant state-year	Yes	Yes	Yes

Notes: The dependent variable is logged electricity productivity (value of output divided by the quantity of electricity used in kWh). Each column represents a separate regression at the plant level. The independent variables are the logged electricity price, and an interaction with a dummy that is one for all years before 2006. Instruments are interacted in the same way. The first stage statistics refer to variable 1 and corresponding instrument 1 etc. Note that mainly the corresponding instruments shift the variables (i.e. 1/1 and 2/2). Regressions are weighted by the recorded sampling multiplier. Standard errors in parentheses are two-way clustered at the plant and the state by year level. The rest of the table layout follows the same structure as the main Table III.2.

III.4.2 Stronger effect during high price periods

Next, I examine whether the effect was stronger for high price periods. It is likely that decreasing electricity prices have particularly strong effects on output when electricity prices are at high levels already. This particularly discourages plants from using electricity associated with modern productive production processes. The nature of the comparatively high Indian electricity prices (see Section III.2.1 and Appendix III.H) and the subsequent halving of prices during the sample period (Figure III.2) lend itself to test this hypothesis. I interact the electricity price with an indicator for the first eight years of the sample periods in Table III.4. The average real price in the first eight years was 5.5 ₹ per kWh compared to 3.8 ₹ per kWh in the second eight years. The interaction term is negative for both IV and the OLS specifications. For the IV^A, the entire effect is driven by the period where electricity prices were high. For IV^B, the interaction effect is negative as well, but insignificant.[50] This suggests that the negative implications of high electricity prices on output and electricity productivity are particularly relevant in contexts with high electricity prices.

[50] The conclusions are similar when looking at three periods as in Table III.17 in Appendix III.K.

III.4.3 Robustness and further analysis

I conduct a range of robustness checks, with most of the results in Appendix III.K. First, I use lagged prices and lagged instruments to allow for some time to adjust to prices. This also addresses potential remaining reverse causality concerns. Using lags cuts the sample in around a half as spells of firm observations are required. Table III.3b first shows the contemporaneous effects for the smaller sample and then the lagged effects. The IV estimates reassuringly hardly change. The positive bias in the OLS estimates is substantially reduced when using lags.

Second, I use two alternative instruments. The first, IV^C, is similar to IV^A except that I exclude plants in the same districts for the construction of the instrument. The second one, IV^D, is similar to IV^B and is also a shift-share instrument. The shift uses the timing of the 2003 Electricity Act and the shares are the calculated distance of district centroids to coalfields. The rationale for the second instruments builds on the finding in Section III.2.1 and Table III.8 in Appendix III.B that the share of private power capacity can explain lower electricity prices, but only after 2003. Since local changes in private power share are likely to be endogenous, I use the distance to coalfields. Table III.8 also shows that the distance of districts to coalfields predicts shares in the private power capacity. Therefore, I use the distance to coalfields interacted with the post 2003 dummy as an instrument, controlling for the distance to coalfields. Table III.11 in Appendix III.K shows that the estimate using IV^C are very close to IV^A. The estimate for IV^D is -0.48, in magnitude similar to the other three instruments, but is insignificant, and also rather weak (F=7.2).

Third, I restrict the sample to electricity intensive sectors, loosely defined as the 2-digit sectors with an above average electricity intensity. The effects are marginally smaller as Table III.12 in Appendix III.K shows. Fourth, I run the analysis by six broad industry groups in Table III.16a and Table III.16b. Only for metals and minerals, the estimates are non-negative, but insignificant, and there is still significant upward bias in the OLS estimates. The null effect for this sector might be explained with the basic metals industry predominately relying on coal across many production techniques,[51] such that there is less scope to move to electricity based production. Figures III.10 and III.11 in Appendix III.E support this. While energy productivity rose in this sector, electricity productivity remained fairly stable.

Fifth, I run an over-identified model using both IV^A and IV^B simultaneously in Table III.13. The effects are again similar, mainly driven by the stronger IV^A. The Sargan-

[51] See also Chapter I of this thesis.

Hansen J-test rejects that both instruments have the same effect. This is not too surprising given the difference in the estimates, which can, however, also be due to heterogeneous local average treatment effects.

Sixth, I control for the distance from districts to coalfields, for state-year level power shortages, and for both in Table III.18. The estimates remain negative and are similar in magnitude. I already showed in Table III.9 in Appendix III.C that shortages can not explain electricity prices. Both, the distance to coalfields and shortages are significant when explaining electricity productivity, however. Seventh, I control for state fixed effects, state trends, and then for plant fixed effects in Table III.14. One of the estimates turns insignificant, but as discussed in Section III.3.6, plant fixed effects can introduce bias. Eighth, I two-way cluster at the district and the region year level, allowing more generously for arbitrary correlation in errors, with slightly larger standard errors but still significant results (Table III.15). Finally, I adjust all p-values upwards to account for multiple hypothesis testing in Table III.21 in Appendix III.M. Almost all estimates remain statistically significant at conventional levels.

Overall, these checks reinforce the conclusion that the OLS estimates are significantly upward biased and higher electricity prices reduce electricity productivity.

III.4.4 Mechanisms and incidence

How do high electricity prices affect plants? In this section I explore the impacts of electricity prices on a range of outcomes to shed more light on mechanisms, as well as calculate incidence and contrast the effects with the impact of coal prices.

Plants scale up with lower electricity prices and substitute from coal

We have seen that higher electricity prices reduce output. Table III.5a shows the effect on total revenues, total variable costs and profits (in levels).[52] Electricity prices reduce total revenues, but they also reduce total variable costs. A one percent increase in electricity prices reduces revenues by 1.3-1.4 million ₹, but also reduce total costs by 1.1 million ₹, and profits go down by 0.2 million ₹. This strongly suggest that plants scale down with rising electricity prices. It is difficult to think of an alternative mechanism that brings total costs down when electricity prices rise. Substitution to other fuels should generally increase costs. Table III.5c shows that there is some substitution. The IV estimate of the effect of prices on the share of electricity expenditure in total fuel expenditure is near zero

[52]See Section III.2.2 for their description.

and insignificant. With these constant shares, electricity consumed must decrease with increasing prices (as shown in Table III.3a). Using plants that report physical electricity and coal units, the ratio between electricity to coal energy inputs decreases with rising electricity prices in Columns (8-9), as plants substitute to coal. Finally, employment also decreases with higher electricity prices as Table III.19 in Appendix III.K shows. These results are consistent with plant size increasing with lower electricity prices.

Electricity prices affect investment, productivity and markups

If electricity is complementary to modern production techniques, then lower electricity prices (compared to other fuels) can incentivise switching to these production techniques and scaling up. Table III.5c reports the impact of higher electricity prices on investment in machinery and plant total factor productivity (TFP). Both investment and TFP decline. The effects on investment are sizeable.[53] The effects on TFP are small, but highly significant and robust to different methodologies to estimate TFP.[54] This is in line with Abeberese (2017), who found reductions in employment, investment and TFP to support her main analysis of product switching. My results are consistent with the finding of firms switching to products that require less electricity but also reduce performance. Vice versa, lower electricity prices incentivise to a switch to production techniques that rely on electricity, but improve performance. Ryan (2018) uses experimental variation through consulting services to show a positive effect of higher electricity productivity on the modernisation of plants' input mix in India.[55] Since higher efficiency reduces the de-facto price of electricity per unit produced, these results are both consistent with the effects of lower electricity prices on production technique. Table III.19 in Appendix III.K provides support for this and shows that the machine to labour ratio falls with higher electricity prices, even though employment falls. The effect is large with an elasticity between -0.6 and -1.5.[56] There is also some evidence that lower electricity prices increase product scope measured as the number of products (Table III.19).

[53] I use the inverse hyperbolic sine instead of the log of machinery investments to deal with zeros. The effects can be interpreted as elasticity.

[54] The baseline effects are on TFP measured via Wooldridge (2009) using deflated revenue data, so should be interpreted as revenue TFP. Since markups shrink, we would expect the impact on physical TFP to be larger. Table III.20 in Appendix III.K provide the effects on TFP measured via Olley and Pakes (1996), Levinsohn and Petrin (2003) or Ackerberg et al. (2015).

[55] He runs a field experiment of energy audits in the state of Gujarat. He also finds that plants use more energy as a response to energy efficiency improvements, due to a rebound effect, consistent with a de-facto reduction in prices.

[56] The machine intensity of output also falls.

Table III.5: Electricity prices and firm performance: scale, substitution, productivity and markups

(a) Electricity prices, profits, revenues and costs (levels)

	Profits (mil. ₹)			Total revenues (mil. ₹)			Avg. variable costs (AVC) (mil. ₹)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	-4.952***	-20.63***	-22.43***	-30.18***	-132.6***	-139.9***	-24.12***	-109.1***	-114.3***
	(1.518)	(3.258)	(4.043)	(8.858)	(19.749)	(21.231)	(7.398)	(16.539)	(17.469)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	485867	485867	485867	485867	485867	485867	485867	485867	485867
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***	-	0.97***	0.06***	-	0.97***	0.06***
First stage SE	-	0.005	0.003	-	0.005	0.003	-	0.005	0.003
F-stat (Kleib.-Paap)	-	43124.701	296.290	-	43124.701	296.290	-	43124.701	296.290
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(b) Electricity prices and substitution: shares in fuel expenditure and ratio of electricity to coal consumption

	Electricity share in fuel expenditure			Other fuels' share in output			Ratio electricity to coal quantity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	0.0251***	0.0144	-0.0233	0.00442***	0.0135***	0.0234***	-10.20***	-17.54***	-21.84*
	(0.006)	(0.013)	(0.020)	(0.001)	(0.002)	(0.003)	(3.099)	(5.790)	(12.354)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	485948	485948	485948	485948	485948	485948	48015	48015	48015
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***	-	0.97***	0.06***	-	0.96***	0.05***
First stage SE	-	0.005	0.003	-	0.005	0.003	-	0.016	0.004
F-stat (Kleib.-Paap)	-	43147.813	296.255	-	43147.813	296.255	-	3705.137	157.253
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(c) Electricity prices, TFP, investment and markups

	TFP (log) (Wooldridge, 2009)			Investment in machinery (IHS)			Price marginal cost markups $\log(\mu)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	-0.00699***	-0.0156***	-0.0330***	0.162	-0.846**	-2.877***	-0.0184***	-0.0404***	-0.106***
	(0.002)	(0.003)	(0.006)	(0.204)	(0.390)	(0.442)	(0.006)	(0.011)	(0.019)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	477697	477697	477697	476042	476042	476042	485548	485548	485548
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***	-	0.97***	0.06***	-	0.97***	0.06***
First stage SE	-	0.005	0.003	-	0.004	0.003	-	0.005	0.003
F-stat (Kleib.-Paap)	-	44391.045	297.573	-	46975.370	309.613	-	43180.457	296.198
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column represents a separate regression at the plant level. The dependent variables are indicated and described in Section III.2.2. In panel (a), the regressions are reported in levels because profits can be negative. In panel (b), other fuels refer to gas, coal and oil. The ratio of electricity to coal is in quantity terms in MWh per tonne. In panel (c), the inverse hyperbolic sine (IHS) of investment is taken to deal with zeros in investment. Regressions are weighted by the recorded sampling multiplier. Standard errors in parentheses are two-way clustered at the plant and the state by year level. The rest of the table layout follows the same structure as the main Table III.2.

There is no prior evidence on how electricity prices affect price over marginal cost markups $\mu \equiv P/MC$ in the Indian context. Table III.5c shows that markups decrease 0.04 percent with a one percent increase in electricity prices. The reduction in profitability comes with a loss in market power. The adjustment of markups also suggests that there is imperfect pass-through of costs to consumers. This raises the important question of the incidence of electricity price changes which I examine next.

The incidence of electricity price changes

The degree to which firms can pass on increases or reductions in electricity prices to consumers determines the incidence of the electricity price changes. As described in Section III.3.5 I estimate pass-through elasticities by industry. The cumulative distribution function of these pass-through elasticities, as well as two example regressions are presented in Figure III.28 in Appendix III.L. The vast majority of pass-through elasticities is between 0.8 and 1.1. A pass-through elasticity of greater than one means that costs are disproportionately passed through to consumers.[57] This can be the case if producers fail to collude in an oligopoly. An increase in costs can help to solve the coordination problem of raising prices, which can explain pass-through rates greater than one.

The pass-through elasticities are combined with the plant level markups ($\widehat{\mu}$) into the pass-through rates $\widehat{\rho}_{MC}$. The three components to calculate incidence I^{share}, the Lerner index $\widehat{L}$, the market demand elasticity $\widehat{\eta}_D$ and the marginal cost to price pass-through rate $\widehat{\rho}_{MC}$ are reported in Table III.6. The estimates shown are the median, the 25th and 75th percentile of the distribution across plants, sectors and years.[58] The estimates for the Lerner index are in line with the descriptive statistics of markups reported in the beginning.

Table III.6 reports the median of I^{share} over all sectors and the whole sample period. The incidence on consumer surplus is 63%. The decline of electricity prices not only improved profits and electricity productivity, but also disproportionately affected consumer surplus. This implies that electricity pricing for industry is important for industrial development and consumer welfare alike. The reduction in the severe cross-subsidisation from industry to agriculture (see Section III.2.1) may thus have also benefited non-industrial consumers.

There is some heterogeneity across industries and years. The 25th and 75th percentile

[57] While the pass-through elasticity is smaller than one for the five industries studied in Ganapati et al. (2016), the pass-through *rate* ρ_{MC} is also greater than one for three of the five industries and in some of the studies cited therein.

[58] $\widehat{L}$ and $\widehat{\rho}_{MC}$ vary at the plant-year level, and $\widehat{\eta}_D$ varies at the industry-state-year level.

Table III.6: Electricity prices and the share of incidence on consumers

Incidence	Oligopolistic competition	Monopoly	Perfect competition
Median	0.63	0.54	1.17
25th to 75th percentile	[0.53 - 0.79]	[0.50 - 0.59]	[0.99 - 1.45]
Components	$\widehat{L}$	$\widehat{\eta}_D$	$\widehat{\rho}_{MC}$
Median	0.18	3.21	1.17
25th to 75th percentile	[0.03 - 0.34]	[2.48 - 4.34]	[0.99 - 1.45]

Notes: The table shows the share of incidence on consumers from electricity price changes, according to $I^{share} = I/(1+I)$. The quantiles are across all plants and all periods, using the sampling multipliers as frequency weights. The reported components ($\widehat{L}_{jisrt}$, $\widehat{e}_{D,isrt}$ and $\widehat{\rho}_{MC,jirst}$) for the calculation are described in the text. The monopoly case corresponds to $\widehat{L}_{jisrt} = 1/\widehat{e}_{D,isrt}$, and the perfect competition case to $\widehat{L}_{jisrt} = 0$.

in Table III.6 are 53% and 79% respectively. Even at the 5th percentile, the share of consumer incidence was a quarter of the total. Figure III.29 in Appendix III.L plots the incidence share over time for six aggregate industries. There has been a few percentage points decline of incidence over time. I also calculate the incidence under the extreme conduct assumptions of monopolies and perfect competition, where $L = 1/\epsilon_D$ and $L = 0$ respectively. As in Ganapati et al. (2016), the monopoly estimate is below the oligopolistic estimate, and the perfect competition higher than the oligopoly counterpart.[59]

How large was the gain in producer surplus (profits) and consumer surplus in terms of Rupees or USD? In a back of the envelope calculation, I start with the semi-elasticity of profits to electricity prices of -21.53. This is the average of the causal estimates (-20.63 and -22.43) in Table III.5a. A 45% reduction of electricity prices over the sample period corresponds to an increase of $\log((1-0.45)^{-21.53}) = 12.87$ mil. ₹ for the average plant. In 1998, there were 113065 plants in the manufacturing sector sampling frame.[60] The gains in profits for the entire manufacturing sector from the electricity price reduction were thus 1.46 trillion ₹ (in constant 2004 terms). The halving of industrial electricity prices from its comparatively high level had substantial effects on the Indian economy. According to this simple calculation, it has contributed the equivalent of 32 billion USD to producer surplus, equivalent to 2% of Indian real GDP in 2013. The gains in consumer surplus has accordingly been 55 billion USD.

[59] For the perfect competition case, the incidence share is equivalent to the pass-through rate as $L = 0$ (see Equation (III.5)).

[60] Since not all plants are sampled every year, I recover the number of plants by summing over the sampling multiplier within a year. The regression estimates are weighted by the sampling multiplier and account for this. The number of plants in the analysis spanning 16 years is larger due to entry and exit.

Table III.7: The contrary effects of coal prices on coal productivity and firm performance

(a) Coal prices and coal productivity, output, coal use and electricity use

	Coal productivity (log)			Output (log)			Coal consumption (log)			Electricity consumption (log)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\log(P^C)$	0.846***	1.487***	1.612***	0.0899***	-0.300	-0.135	-0.756***	-1.843***	-1.796***	-0.0413	-0.426	0.734*
	(0.025)	(0.179)	(0.213)	(0.031)	(0.248)	(0.344)	(0.036)	(0.272)	(0.384)	(0.036)	(0.269)	(0.428)
OLS/IV	OLS	IV^E	IV^F	OLS	IV^E	IV^F	OLS	IV^E	IV^F	OLS	IV^E	IV^F
Observations	45009	45009	45009	45009	45009	45009	45009	45009	45009	45009	45009	45009
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.57***	0.01***	-	0.57***	0.01***	-	0.57***	0.01***	-	0.57***	0.01***
First stage SE	-	0.046	0.001	-	0.046	0.001	-	0.046	0.001	-	0.046	0.001
F-stat (Kleib.-Paap)	-	155.090	86.217	-	155.090	86.217	-	155.090	86.217	-	155.090	86.217
SE clustered by	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
No. of first clusters	16277	16277	16277	16277	16277	16277	16277	16277	16277	16277	16277	16277
SE clustered by	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year
No. of second clusters	426	426	426	426	426	426	426	426	426	426	426	426

(b) Coal prices and profits, revenues, costs and TFP

	Profits (mil. ₹)			Total revenues (mil. ₹)			Avg. variable costs (AVC) (mil. ₹)			TFP (log) (Wooldridge, 2009)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\log(P^C)$	-5.917***	-5.745	-7.108	-19.99**	-18.74	-0.843	-14.36**	-27.76	4.644	-0.000544	-0.0198	-0.0306
	(1.624)	(15.031)	(25.898)	(7.990)	(85.440)	(128.629)	(6.583)	(70.784)	(103.547)	(0.002)	(0.013)	(0.020)
OLS/IV	OLS	IV^E	IV^F	OLS	IV^E	IV^F	OLS	IV^E	IV^F	OLS	IV^E	IV^F
Observations	45006	45006	45006	45006	45006	45006	45006	45006	45006	44582	44582	44582
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.57***	0.01***	-	0.57***	0.01***	-	0.57***	0.01***	-	0.57***	0.01***
First stage SE	-	0.046	0.001	-	0.046	0.001	-	0.046	0.001	-	0.046	0.001
F-stat (Kleib.-Paap)	-	155.060	86.214	-	155.060	86.214	-	155.060	86.214	-	153.047	88.672
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column represents a separate regression at the plant level. Reported are results from OLS regression on logged coal prices, and IV regressions. The IV^E is based on the coal prices of similar plants. In the shift-share IV^F, the share is the logged distance of a district to the nearest coal mine and the shift is the logged raw coal price for industry at a representative mine. The dependent variables are indicated and described in Section III.2.2. In panel (a), coal productivity is the value of output divided by the quantity of coal used in tonnes. In panel (b), the regressions are reported in levels except for TFP because profits can be negative. The first stage statistics are reported. All regressions contain industry by year by region fixed effects. Regressions are weighted by the recorded sampling multiplier. Standard errors in parentheses are two-way clustered at the plant and the state by year level. Plant output is deflated using 3-digit industry deflators and coal prices are deflated using a general fuel and electricity wholesale price deflator.

The contrary effects of coal prices and implications for climate policy

So far the analysis has been about electricity prices. The most plausible mechanism is that electricity prices affect output in particular because higher electricity prices deter from upgrading to more modern electricity using production processes. If this is the case, then the effect of higher coal prices should be different, as coal (or oil and gas) fuel is not generally associated with more modern productive production processes. To further test this hypothesis, I run regressions where the main independent variable is plant level coal prices for the roughly 45000 observations of plant-years that use coal. As these suffer from similar endogeneity problems as electricity prices, I construct two instruments as described in Section III.3.4.

In contrast to electricity, higher coal prices significantly *improve* coal productivity. In the first three columns in Table III.7a, both the OLS and the IV estimates are significantly positive, with the IV estimates being roughly double in magnitude.[61] While coal prices significantly reduce coal consumption, they only have a small and insignificant effect on output in the IV specifications, also shown in Table III.7a. The impact on electricity use is either insignificant or positive. There is a small insignificant effect on profits and revenues and an ambiguous effect on costs (Table III.7b). There is no similar scaling down effect with higher coal prices as there is with higher electricity prices. Contrary to electricity prices, higher coal prices also have no effect on TFP (Table III.7b).

This is somewhat good news for climate policy in developing countries. In contrast to electricity prices, the results suggest that taxing dirtier fuels has little effect on firm performance. Of course, this also depends on the fuel mix in electricity generation. With an increasing share of low carbon electricity generation from hydro, nuclear and renewables, carbon pricing may have limited effect on firms when taking these results at face value. On the other hand taxing electricity use may have perverse consequences in some circumstances, as it may lower industrial electricity productivity through its adverse impact on output. This is likely to be especially relevant in the context of developing industries in the process of adopting modern electricity intensive production techniques, and in contexts with already high electricity prices as in India in the late 90s and early 2000s.

[61] The first stage is reported and the F-stats sufficiently high.

III.5 Conclusion

In this chapter, I estimate the causal effect of industrial electricity prices on electricity productivity using two alternative instruments and a large panel of Indian manufacturing firms. The effects are negative, especially during the high price period in the late 1990s and early 2000s. While higher electricity prices reduce electricity consumption, they also decrease electricity productivity. This is driven by the significant negative effects on output. The mechanisms that I explore support the hypothesis that electricity is a complementary production input to modern high performance production techniques. Investment and productivity are deterred by higher electricity prices, which may hold back industrial development.

I document a secular increase in aggregate manufacturing electricity productivity. My causal estimates of electricity prices can quantitatively explain this secular rise in electricity productivity remarkably well. Decreasing electricity prices may thus have significantly helped to improve efficiency through technology upgrading and performance improvements. The main message of this chapter is that lower industrial electricity prices can actually *improve* the electricity intensity of output. I argue and provide some evidence that negative impacts of high electricity prices on firm performance are especially relevant in the context of industrial development and when electricity prices are already high.

Markups decrease as a result of higher electricity prices. I estimate marginal cost to price pass-through rates under imperfect competition and calculate the welfare incidence on producers and consumers. The share of incidence of consumers is around two thirds on average. The reduction in the severe cross-subsidisation from industry to agriculture and residential electricity users may thus have also benefited non-industrial consumers.

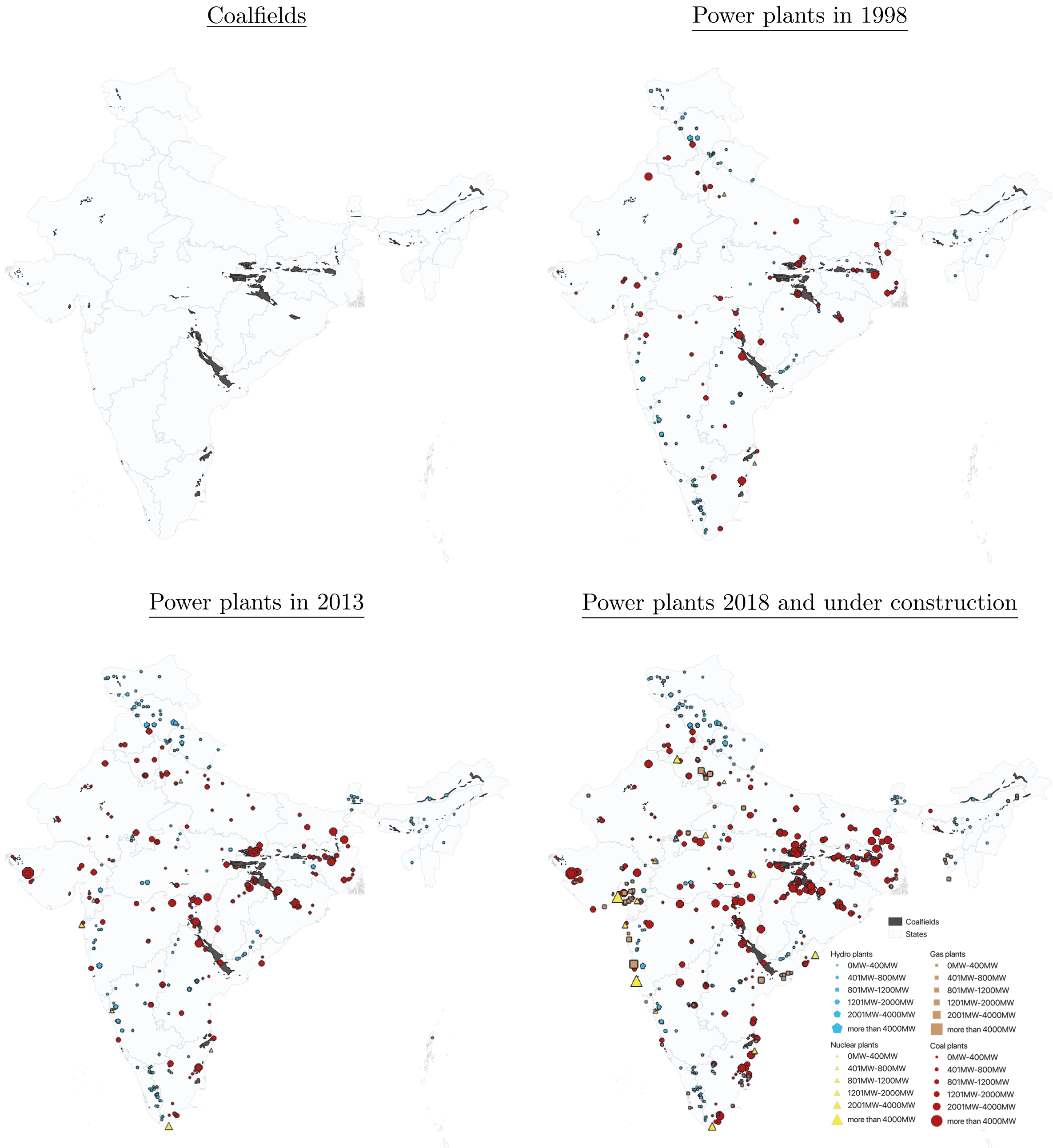
I end the chapter by comparing the impacts of electricity prices to the impacts of coal prices. Higher coal prices improve coal productivity and hardly affect output, profits and productivity. This is consistent with the hypothesis that electricity is distinctive as a complement to modern production techniques. This has important implications for climate policy. Taxing electricity for industry harms firms and consumers, and may increase the electricity intensity of output. Taxing carbon and coal, on the other hand, improves energy efficiency and has limited impact on firm performance. Naturally, the fuel mix in electricity generation, as well as the pass-through rates of power utilities, determines how much electricity prices are affected by taxing carbon. Nevertheless, the relative price of coal to electricity would increase in any case. In the described contexts, relatively lower electricity prices for industry could deliver both: substitution from fossil fuels to electricity, and despite increasing electricity use, improving electricity productivity of output. Both are essential components of reducing industrial carbon emissions.

One of the instruments that I develop can readily be calculated in other settings, which I hope can foster more research on this important link. Future research aims to estimate the relationship between electricity pricing and production in a structural model and to explore the nuanced relationships in the context of other developing and developed countries.

Appendix to Chapter III

III.A Maps of coal reservoirs and power plants

Figure III.6: Maps of coalfields and powerplants by year



Notes: The maps plot the coalfields (time invariant) and the stock of power plants in the corresponding years. The size of the markers corresponds to installed capacity. Coal plants are built near coalfields. Hydro plants near rivers especially in the mountainous region. Nuclear plants are typically built near the sea or rivers. Gas plants are built near ports and the major gas pipelines (e.g. in the north east). Data sources are described in Section III.2.2.

III.B Electricity prices and privately owned share in installed capacity

Table III.8: Electricity prices and privately owned share in district installed capacity

	Electricity price				Priv. share
	(1)	(2)	(3)	(4)	(5)
Share private capacity	0.09	0.11	0.06	0.00	
	(0.96)	(1.17)	(0.63)	(0.03)	
Share private capacity x After 2003	-0.24***	-0.24***	-0.19**	-0.19**	
	(-2.93)	(-2.93)	(-2.34)	(-1.98)	
Distance to coalfield ('00 km) x After 2003				-0.09***	-0.03***
				(-3.15)	(-2.85)
N	7994	7994	7994	7994	7994
Total capacity	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Region-year FE	No	No	Yes	Yes	Yes

Notes: The table shows estimates from OLS regressions at the district year level with the median electricity price within a district as dependent variable in the first 4 columns. The Indian Electricity Act was introduced in 2003. The share of privately owned capacity in district level installed capacity includes private/state and private/central ownership categories. The total capacity covariate controls for total installed capacity at the district year level. The distance to coalfields at the district level is in hundreds of km. Column 5 has the share of privately owned capacity as dependent variable. Regressions are weighted by the sampling multipliers and by the number of plants within a district year cluster. Standard errors in parentheses are clustered at the district level. The coefficients on the interaction in column (1) and (2) correspond to a semi-elasticity of 0.03. Stars indicate p-values: * < 0.1, **< 0.05, ***< 0.01.

III.C No significant correlation between shortages and electricity prices

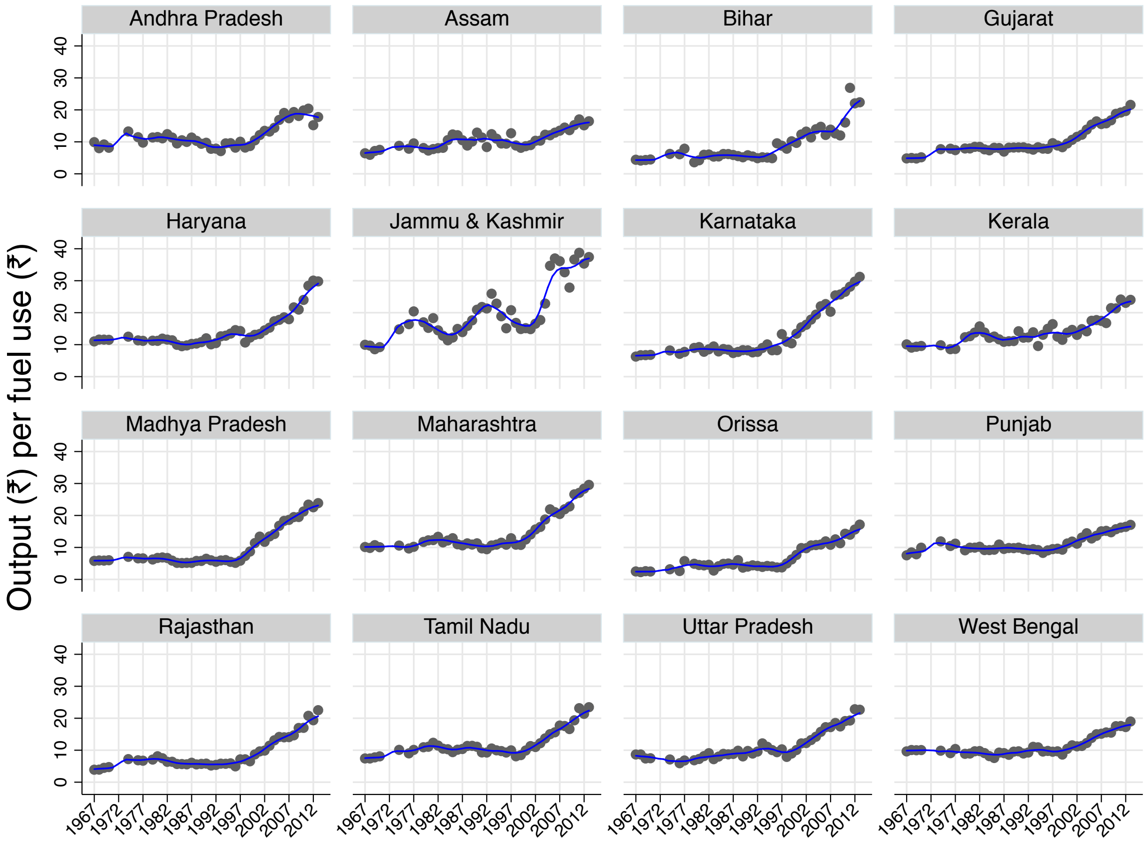
Table III.9: Electricity prices and shortages

	Plant level			State level		
	(1)	(2)	(3)	(4)	(5)	(6)
Shortages	0.34	-0.02	0.12	1.08	-0.01	0.11
	(1.58)	(-0.20)	(0.94)	(1.02)	(-0.01)	(0.64)
N	475809	475809	475809	458	458	458
Year FE	No	Yes	Yes	No	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes
Region-year FE	No	No	Yes	No	No	Yes

Notes: The table shows estimates from OLS regressions of the logged electricity price on shortages. The first three columns are using logged electricity prices at the plant level. The second three columns are regressions at the state year level with logged median electricity prices. Regressions are weighted by the sampling multipliers. The second three regressions are also weighted by the number of plants within a state year cluster. Shortages are at the state year level. Standard errors in parentheses are clustered at the state year level. Stars indicate p-values: * < 0.1, **< 0.05, ***< 0.01.

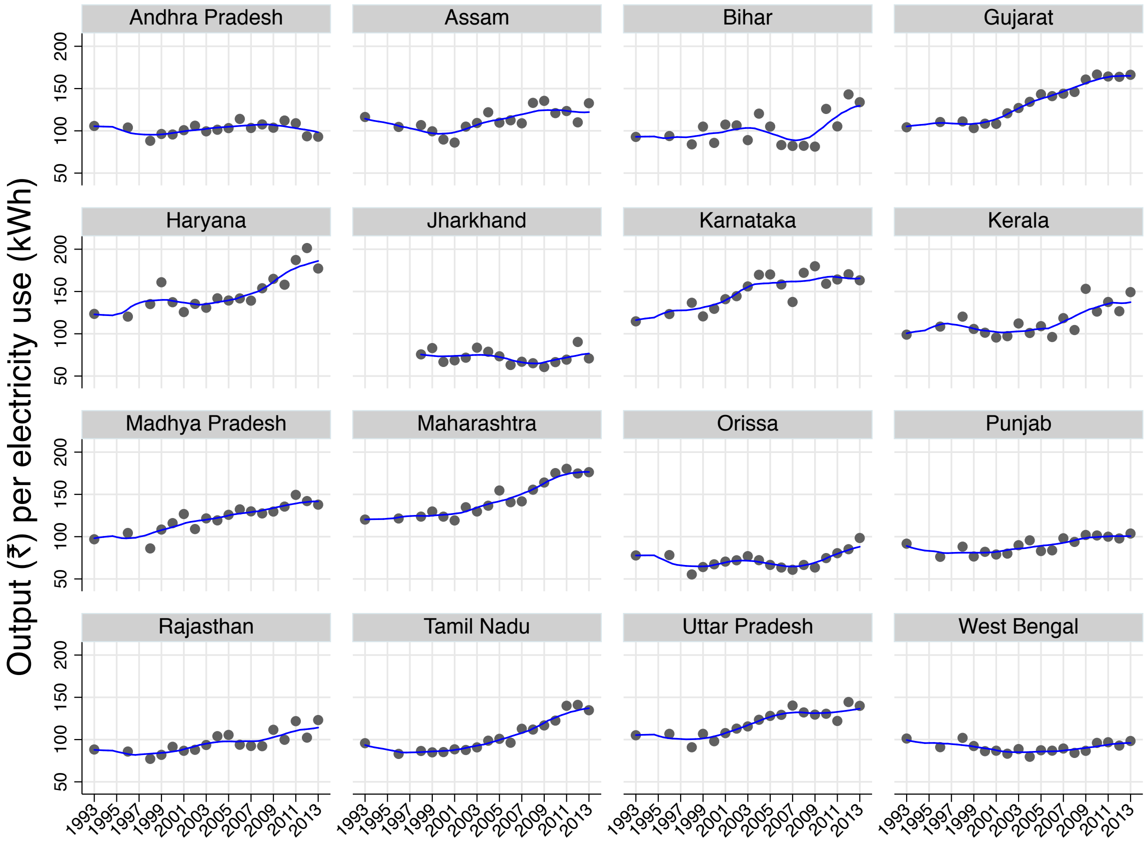
III.D State level trends

Figure III.7: Energy productivity (per ₹) by state



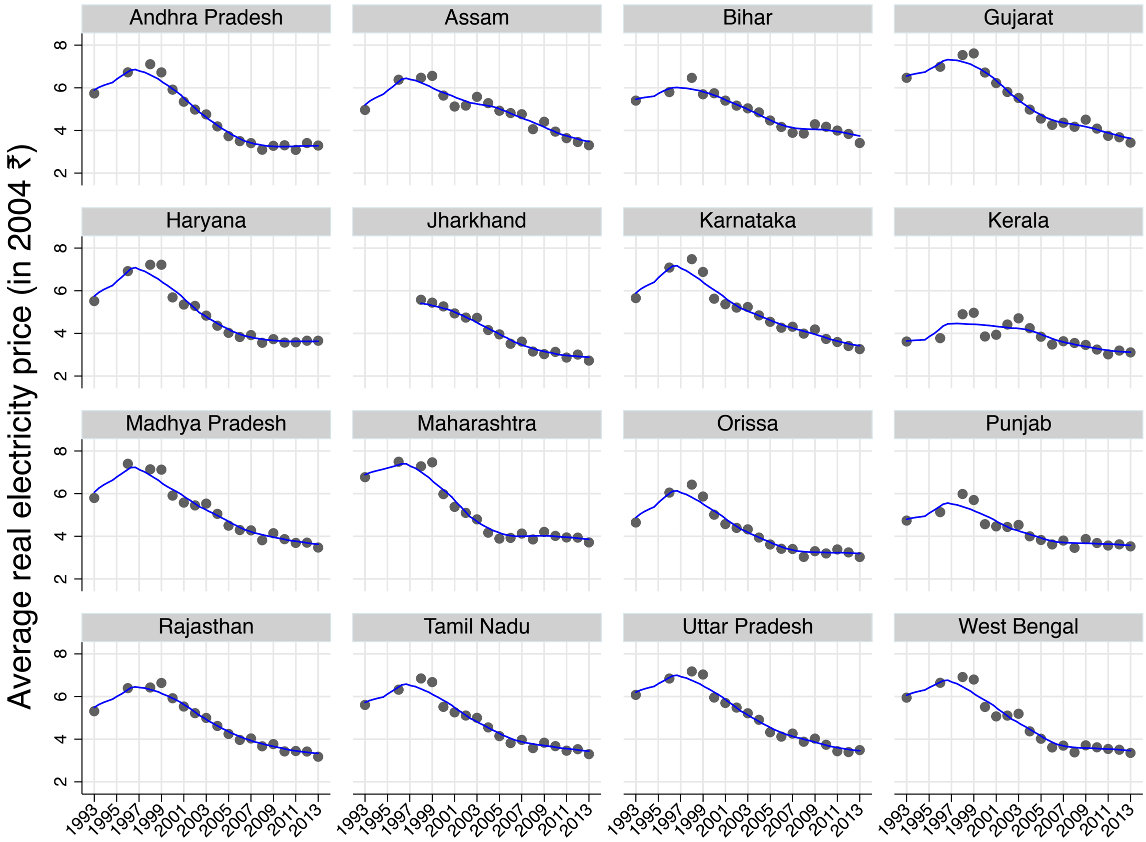
Notes: The figure plots the annual energy productivity ratios (value of output divided by the value of fuel and electricity used). Sixteen of the largest states are displayed in this figure. Output is deflated at the 2-digit industry level using 2-digit industry deflators before aggregating over industries. Fuel and electricity use is deflated using a general fuel and electricity wholesale price deflator. The ratio of aggregate output to aggregate fuel and electricity consumption is displayed. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. From 1967 to 1997 the raw ASI data in pre-aggregated form is used (at industry state year aggregation). From 1998 the raw plant level ASI data is used and aggregated with sampling multipliers.

Figure III.8: Electricity productivity (per kWh) by state



Notes: The figure plots the annual electricity productivity ratios by states (value of output divided by the quantity of electricity used in kWh). Sixteen of the largest states are displayed in this figure. Plant output is deflated using 3-digit industry deflators before aggregating over industries. The ratio of aggregate output to aggregate electricity use is displayed. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. All data points come from the raw plant level ASI data and aggregated with sampling multipliers.

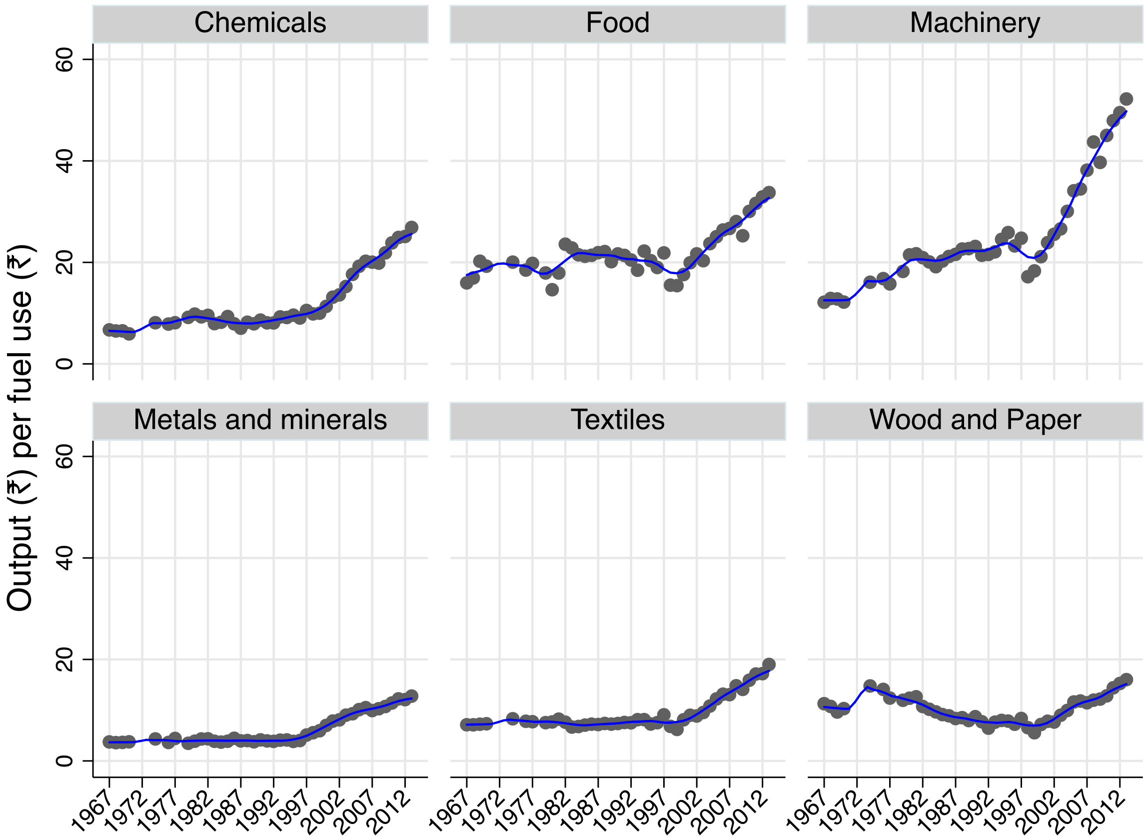
Figure III.9: Electricity prices by state



Notes: The figure plots the real average electricity prices by states. Sixteen of the largest states are displayed in this figure. They are calculated by first aggregating the value of electricity bought by plants and the quantity bought, and then taking the ratio of the aggregates. Electricity values are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. All data points come from the raw plant level ASI data and aggregated with sampling multipliers.

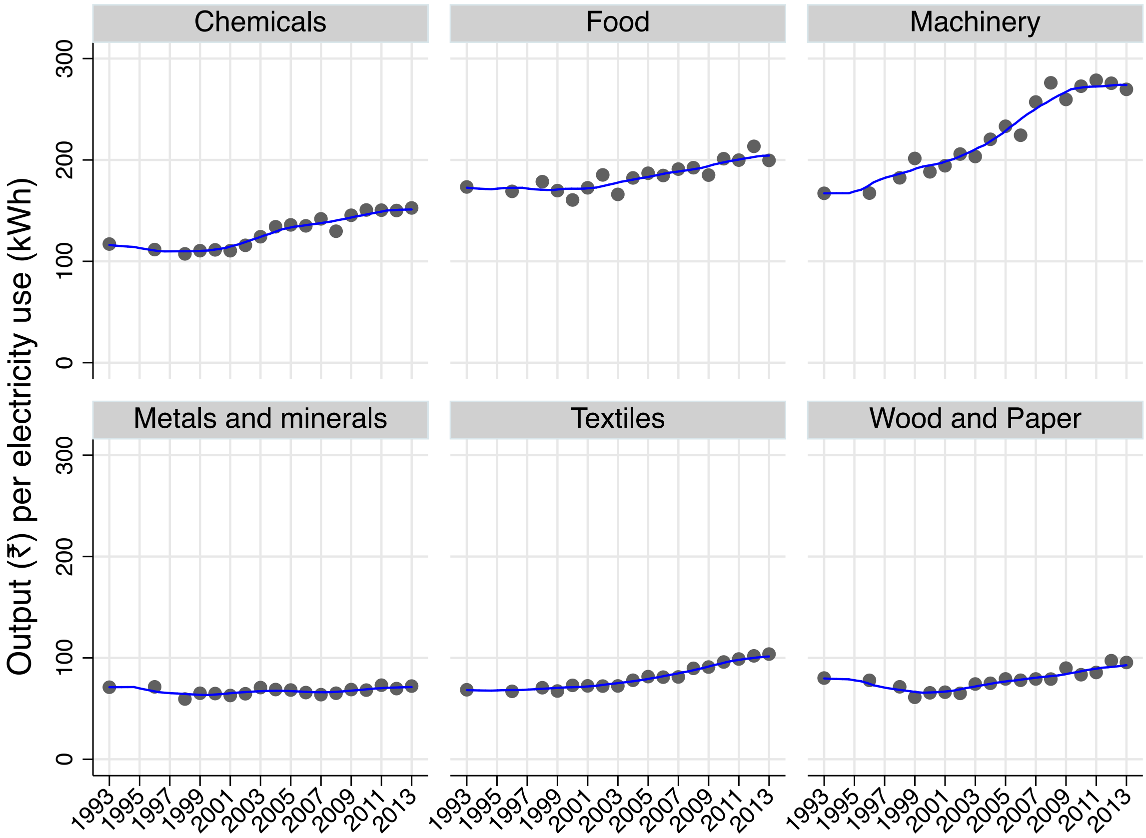
III.E Industry level trends

Figure III.10: Energy productivity (per ₹) by industry



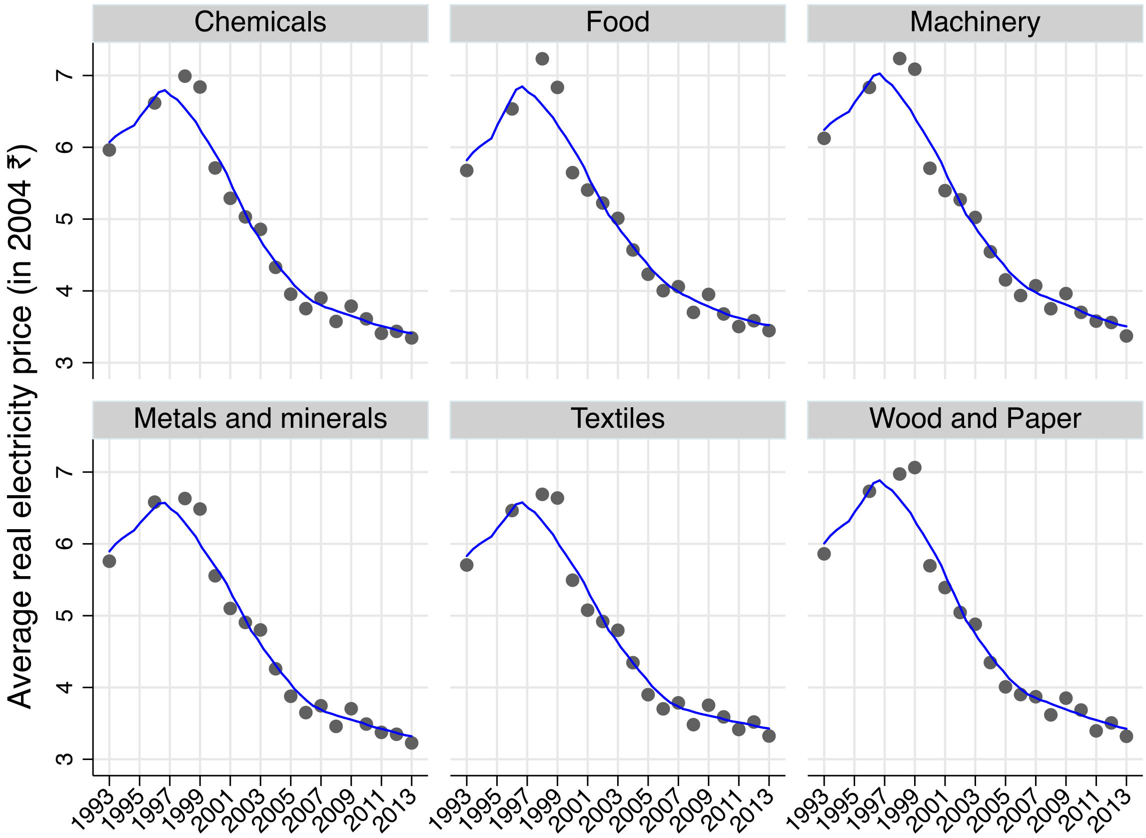
Notes: The figure plots the annual energy productivity ratios by industry (value of output divided by the value of fuel and electricity used). The industries are broad: chemicals includes rubber and plastics, machinery includes metal products, and textiles includes leather. Output is deflated at the 2-digit industry level using 2-digit industry deflators before aggregating over industries. Fuel and electricity use is deflated using a general fuel and electricity wholesale price deflator. The ratio of aggregate output to aggregate fuel and electricity consumption is displayed. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. From 1967 to 1997 the raw ASI data in pre-aggregated form is used (at industry state year aggregation). From 1998 the raw plant level ASI data is used and aggregated with sampling multipliers.

Figure III.11: Electricity productivity (per kWh) by industry



Notes: The figure plots the annual electricity productivity ratios by industry (value of output divided by the quantity of electricity used in kWh). The industries are broad: chemicals includes rubber and plastics, machinery includes metal products, and textiles includes leather. Plant output is deflated using 3-digit industry deflators before aggregating over industries. The ratio of aggregate output to aggregate electricity use is displayed. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. All data points come from the raw plant level ASI data and aggregated with sampling multipliers.

Figure III.12: Electricity prices by industry



Notes: The figure plots the real average electricity prices by industry. The industries are broad: chemicals includes rubber and plastics, machinery includes metal products, and textiles includes leather. They are calculated by first aggregating the value of electricity bought by plants and the quantity bought, and then taking the ratio of the aggregates. Electricity values are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. All data points come from the raw plant level ASI data and aggregated with sampling multipliers.

III.F Additional figures for energy and electricity productivity trends

Figure III.13: Electricity productivity (per ₹)

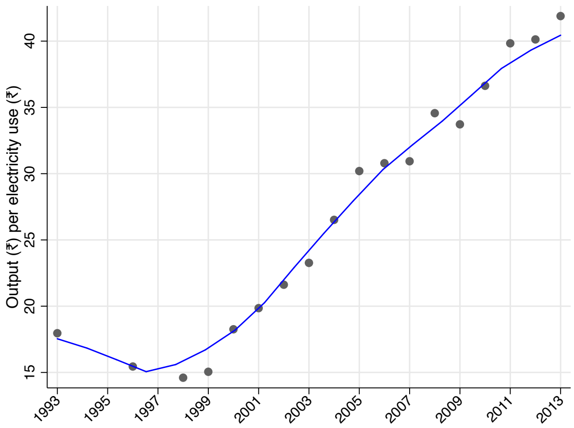
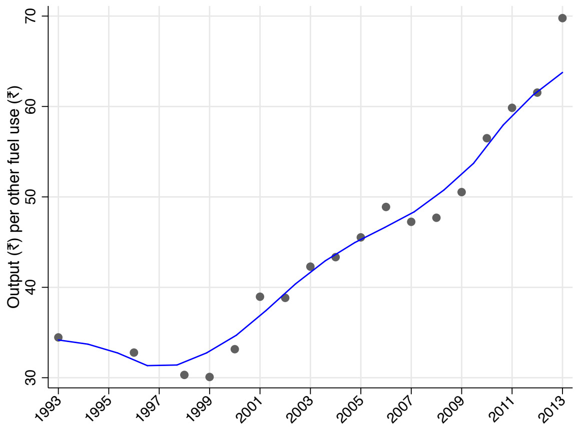


Figure III.14: Other fuel productivity (per ₹)



Notes: The figure plots the annual electricity productivity ratios (value of output divided by the value of electricity used) and the other fuel productivity ratios (value of output divided by the value of fuel other than electricity used). Plant output is deflated using 3-digit industry deflators before aggregating over industries. Electricity and fuel values are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. All data points come from the raw plant level ASI data and aggregated with sampling multipliers.

Figure III.15: Electricity productivity (per kWh)

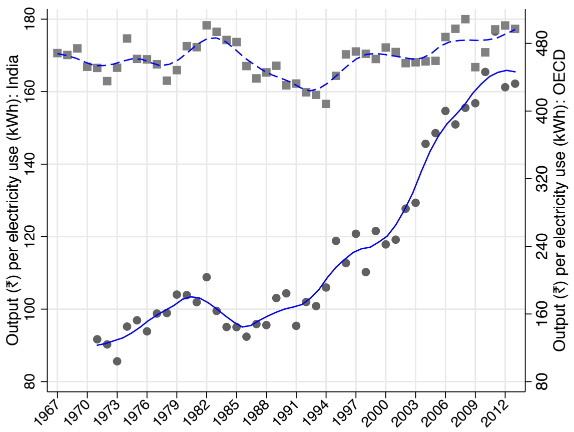
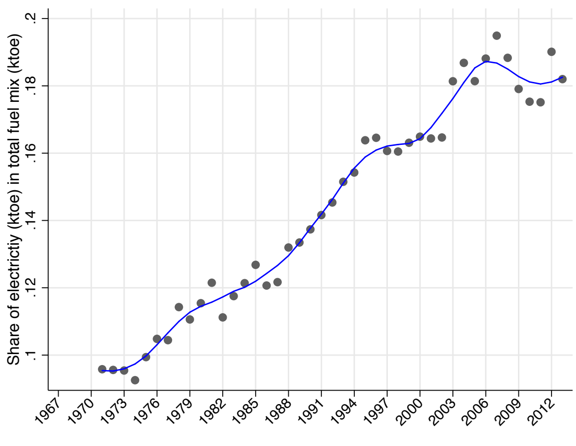


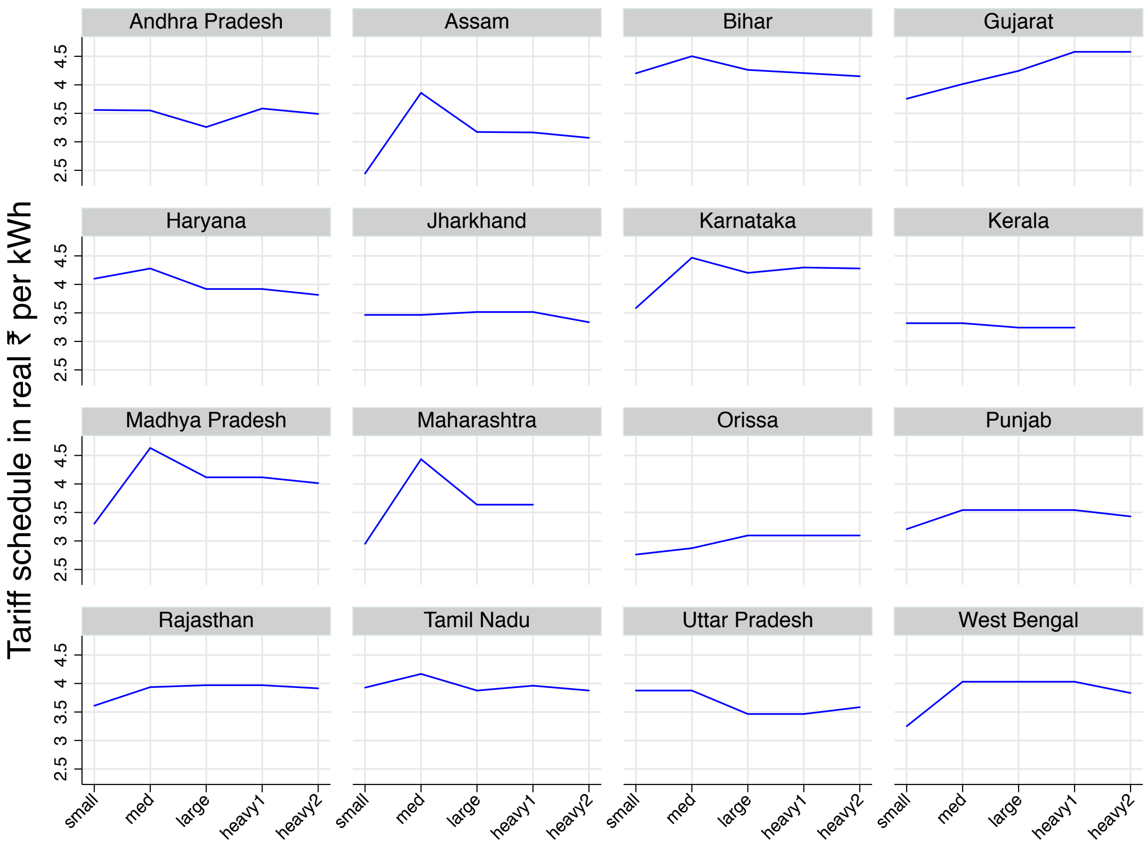
Figure III.16: Share of electricity in fuel mix



Notes: The left figure plots the annual electricity productivity ratios (value of output divided by the quantity of electricity used (in kWh)). Both quantities are for manufacturing only. Output is from UNIDO (2016*a*), deflated with GDP deflators from World Bank (2017), and electricity consumption from the IEA (2016). The base year for deflation is 2004 throughout this chapter. Plotted are the values and kernel smoother for India with the solid line, corresponding to the left axis. The values and kernel smoother for OECD countries are the dashed lines, corresponding to the right axis. The right figure plots the share of electricity consumption in total fuel consumption in India (both in ktoe) using data from IEA (2016).

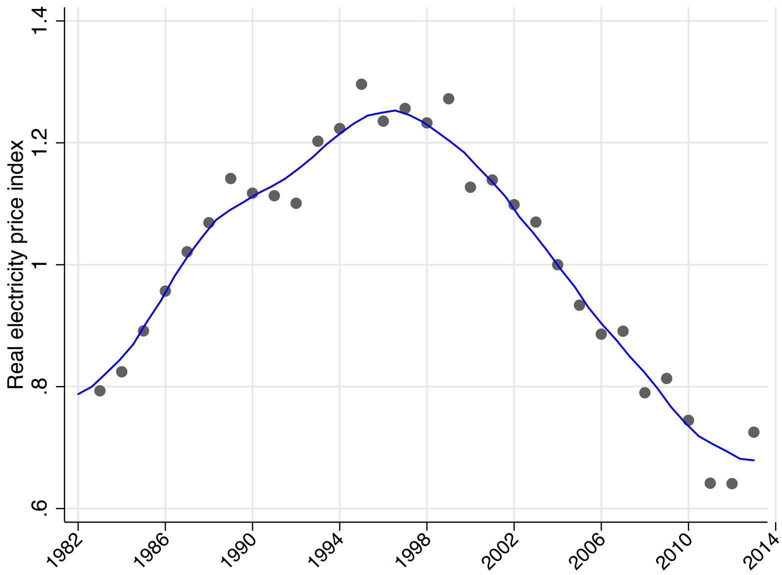
III.G Additional figures for electricity tariffs and price trends

Figure III.17: Reported industrial average tariff schedules in large states in 2007



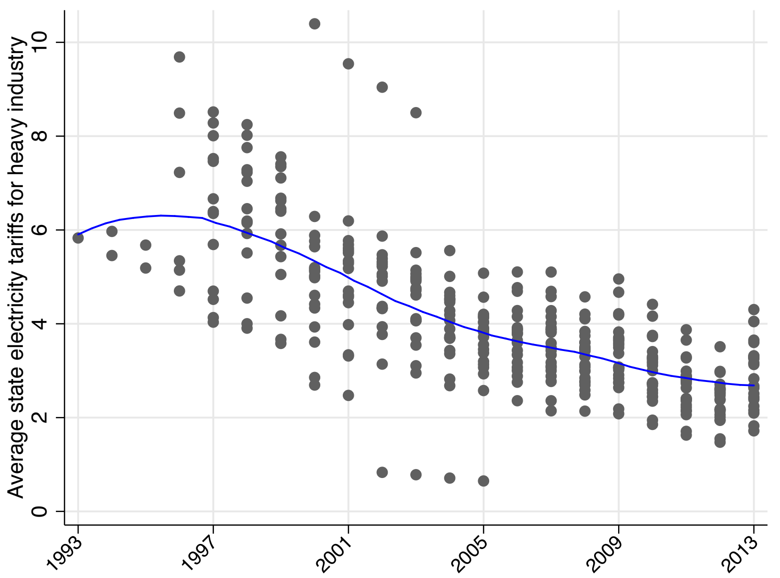
Notes: Plotted are the estimated average tariffs by state by size of industrial consumer. There are five categories increasing in electricity consumption from *small* to *heavy2*. The reported average tariffs are taken from the Indian Central Electricity Authority (2008). The tariffs are deflated with the general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

Figure III.18: Real electricity price index



Notes: Plotted is the real electricity price index for industry. It is based on the wholesale price index for electricity for industrial purposes. The wholesale price index for electricity is deflated with the general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

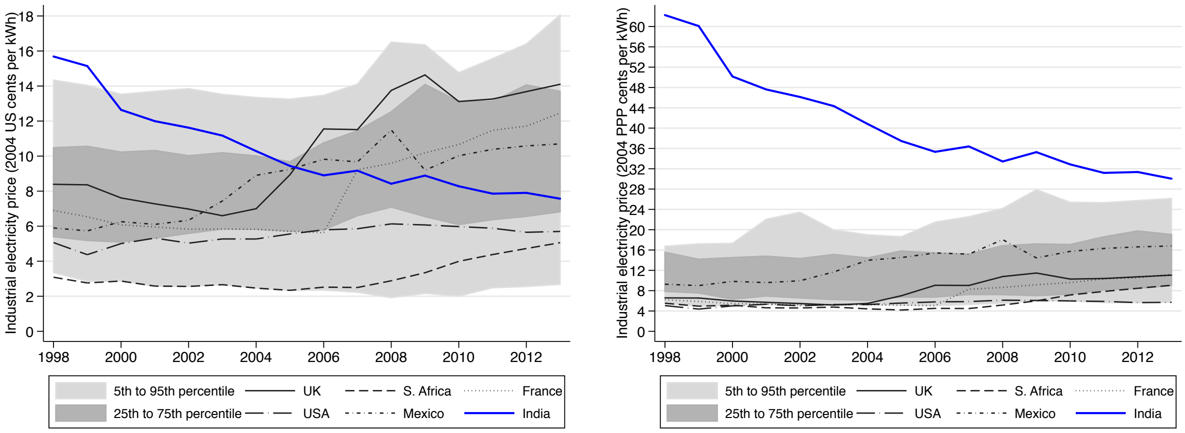
Figure III.19: Average real state tariffs for heavy industry



Notes: Plotted is the real electricity tariff for heavy industry. The tariffs are manually collected from publications of the Indian Central Electricity Authority (2008, 2009, 2010, 2011, 2012, 2013, 2015) and from Indiastat (2019) through Lok Sabha and Rajya Sabha questions. Individual data points correspond to state level average tariffs for heavy industry. Tariffs are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

III.H International electricity price comparison

Figure III.20: Industrial electricity prices in an international context (USD and PPP)



Notes: The figures plot real industrial electricity prices for six individual countries. The left figure is based on market exchange rates, the right figure is based on PPP conversion factors. The shaded areas correspond to the interquartile range and the 5th to 95th percentile of a given year. This is based on a consistent set of 26 countries for which data for all years was available (see below). Raw price data comes from IEA (2018), except for India, where the prices are based on the micro data in the main text. For India, IEA (2018) data is only available from 2006, which is similar to the plotted data. Prices are deflated with national GDP deflators and turned into USD or PPP-USD with exchange rates and PPP conversion factors from World Bank (2017). For India, prices are deflated using a general fuel and electricity wholesale price deflator as in the main text. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. The 26 countries used for the percentiles are: Algeria, Canada, Czech Republic, Denmark, France, Germany, Hungary, India, Ireland, Israel, Italy, Japan, Kazakhstan, Mauritius, Mexico, New Zealand, Paraguay, Poland, Portugal, Slovak Republic, South Africa, Spain, Switzerland, Turkey, United Kingdom, United States.

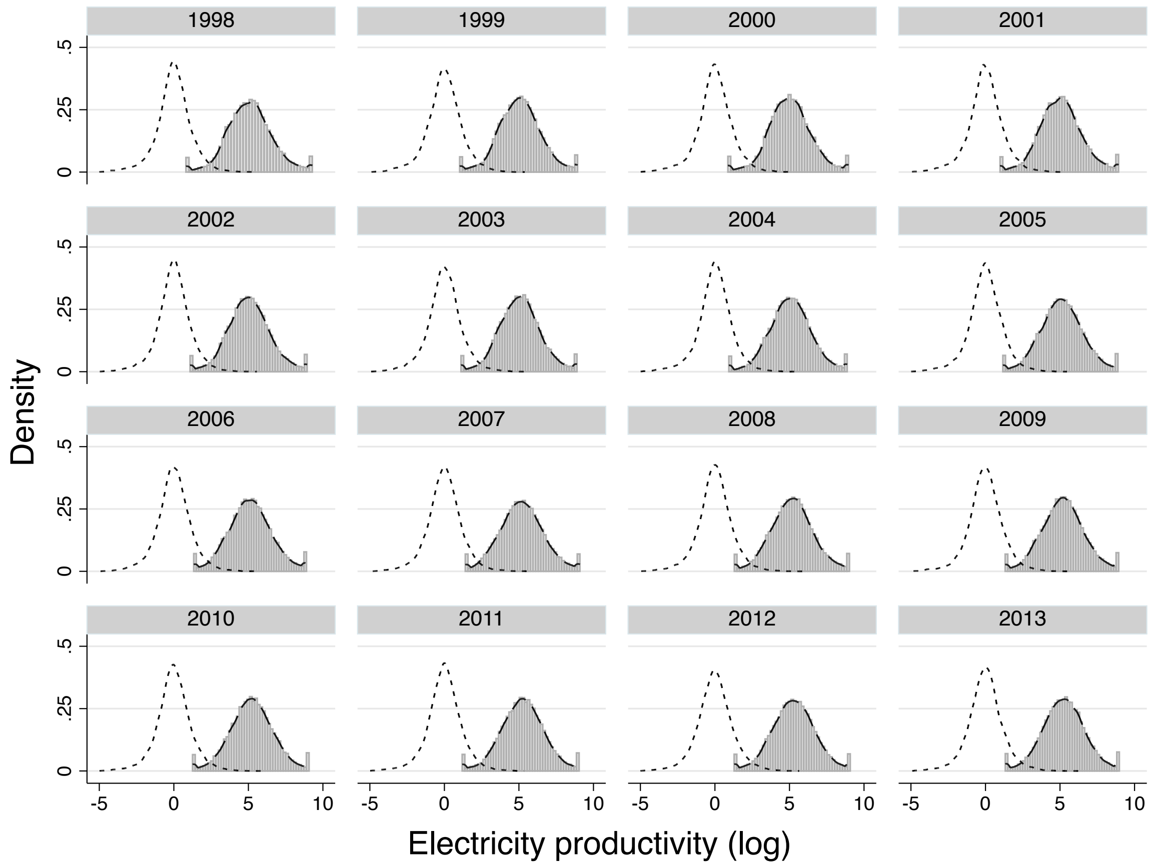
Table III.10: Industrial electricity prices in US-cents: India and G7 average (USD and PPP)

	Market exchange rates					PPP				
	India	G7	OECD	% of G7	% of OECD	India	G7	OECD	% of G7	% of OECD
1998	15.69	8.91	8.96	176	175	62.25	8.24	10.40	756	598
1999	15.14	8.42	8.57	180	177	60.09	7.76	10.03	774	599
2000	12.64	8.36	8.43	151	150	50.16	7.75	9.94	648	504
2001	12.00	8.97	8.81	134	136	47.61	8.36	10.40	570	458
2002	11.62	8.68	8.89	134	131	46.13	8.08	10.49	571	440
2003	11.17	9.01	9.11	124	123	44.34	8.41	10.78	527	411
2004	10.28	9.00	9.07	114	113	40.82	8.38	10.77	487	379
2005	9.44	9.55	9.43	99	100	37.46	8.88	11.16	422	336
2006	8.90	10.58	10.03	84	89	35.33	9.79	11.77	361	300
2007	9.17	11.25	10.30	82	89	36.39	10.41	12.11	350	301
2008	8.42	10.88	11.02	77	76	33.43	9.98	13.05	335	256
2009	8.89	11.59	11.46	77	78	35.27	10.61	13.70	332	257
2010	8.28	11.42	11.11	72	74	32.86	10.50	13.24	313	248
2011	7.86	12.20	11.50	64	68	31.18	11.24	13.60	278	229
2012	7.90	12.79	12.18	62	65	31.36	11.77	14.38	266	218
2013	7.57	13.53	12.43	56	61	30.04	12.45	14.56	241	206

Notes: The table shows the real industrial electricity prices for India, the simple average of the G7 nations, and the simple average of OECD countries, for which data in all years were available. The left part is based on market exchange rates, the right part is based on PPP conversion factors. Raw price data comes from IEA (2018), except for India, where the prices are based on the micro data in the main text. For India, IEA (2018) data is only available from 2006, which is similar to the reported data. Prices are deflated with national GDP deflators and turned into USD or PPP-USD with exchange rates and PPP conversion factors from World Bank (2017). For India prices are deflated using a general fuel and electricity wholesale price deflator as in the main text. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India. The included OECD countries are: Canada, Czech Republic, Denmark, France, Germany, Hungary, Ireland, Israel, Italy, Japan, Mexico, New Zealand, Poland, Portugal, Slovak Republic, Spain, Switzerland, Turkey, United Kingdom, United States.

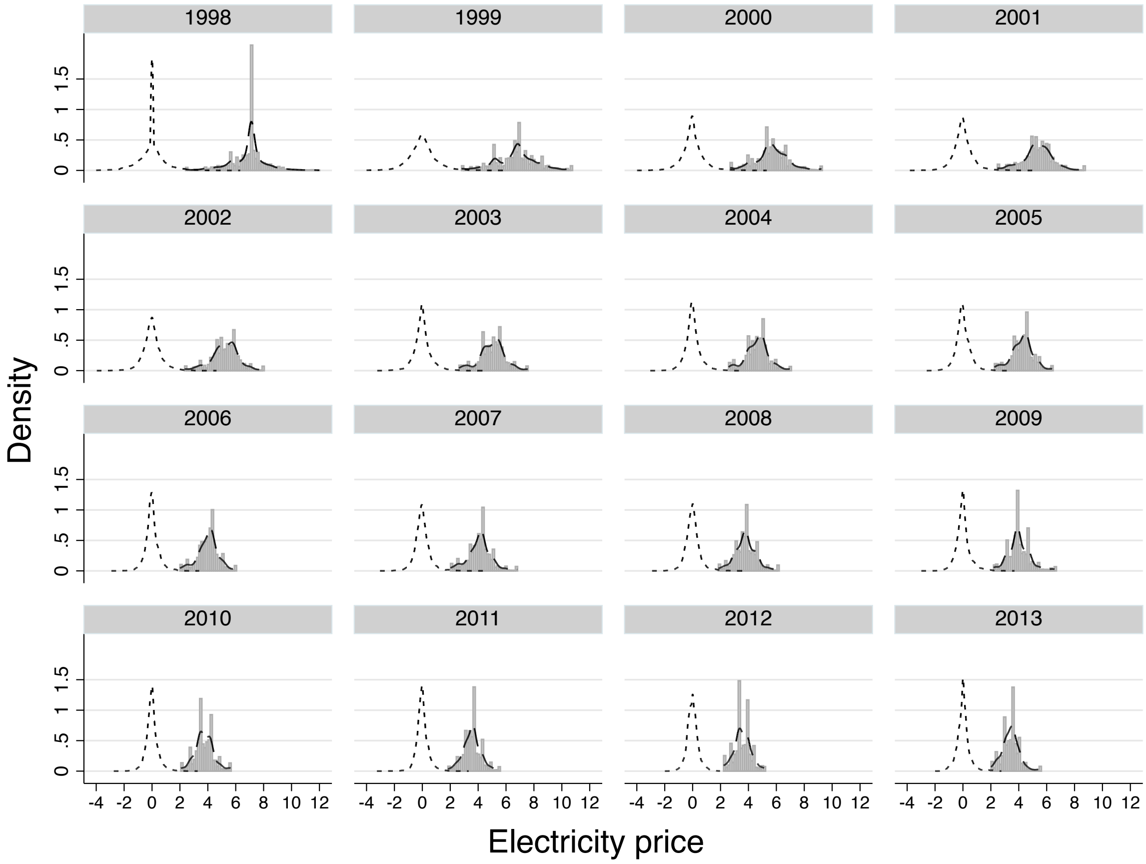
III.I Dispersion in electricity productivity and prices throughout the years

Figure III.21: Heterogeneity in electricity productivity



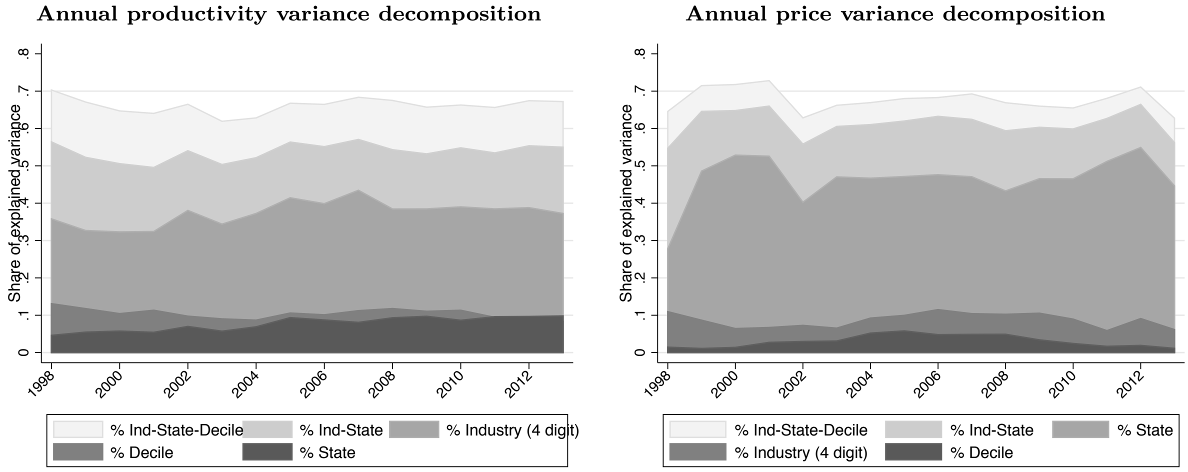
Notes: The figure plots the histograms of plant level logged electricity productivity by year. The left kernel density plot shows the distribution of the residuals of logged electricity productivity after partialling out state by 4-digit industry by year fixed effects. Electricity productivity ratios are the value of output divided by the quantity of electricity used in kWh. Plant output is deflated using 3-digit industry deflators. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

Figure III.22: Heterogeneity in electricity prices



Notes: The figure plots the histograms of plant level electricity prices by year. The left kernel density plot shows the distribution of the residuals of electricity prices after partialling out state by 4-digit industry by year fixed effects. Electricity prices are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

Figure III.23: Electricity productivity and price variance decomposition: percentage shares

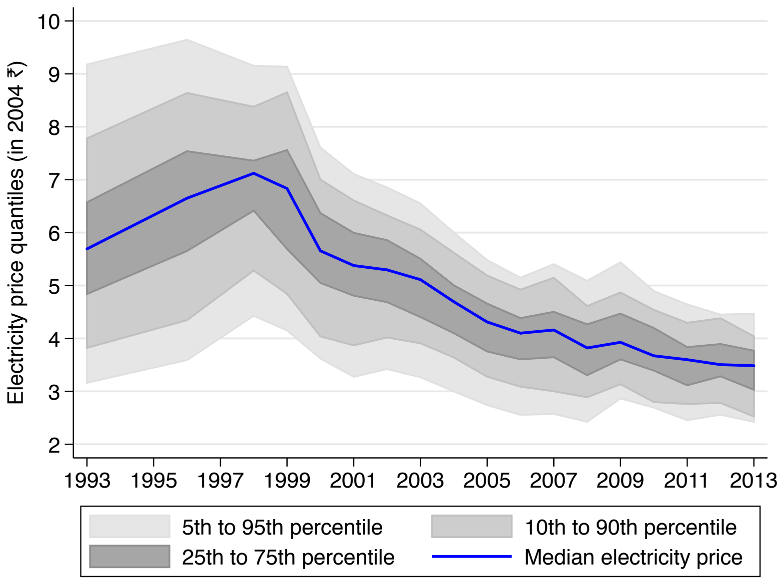


Notes: The left panel plots the share of the annual total variance of logged electricity productivity explained by specified groups. The right panel plots the same for logged electricity prices. The annual variance is calculated as $V = \sum_e s_e \left(p_e - \overline{p}\right)^2$, where s_e are purchase weights multiplied by the sample multiplier, p_e are logged electricity productivity or prices, $\overline{p}$ the weighted average log productivity or price. I use the decomposition of Davis et al. (2013) to decompose total variance into a within "group" component V^W, and a component across "groups" V^G:

$$V = \sum_e s_e \left(p_e - \overline{p}_g\right)^2 + \sum_g s_g \left(\overline{p}_g - \overline{p}\right)^2 = V^W + V^G$$

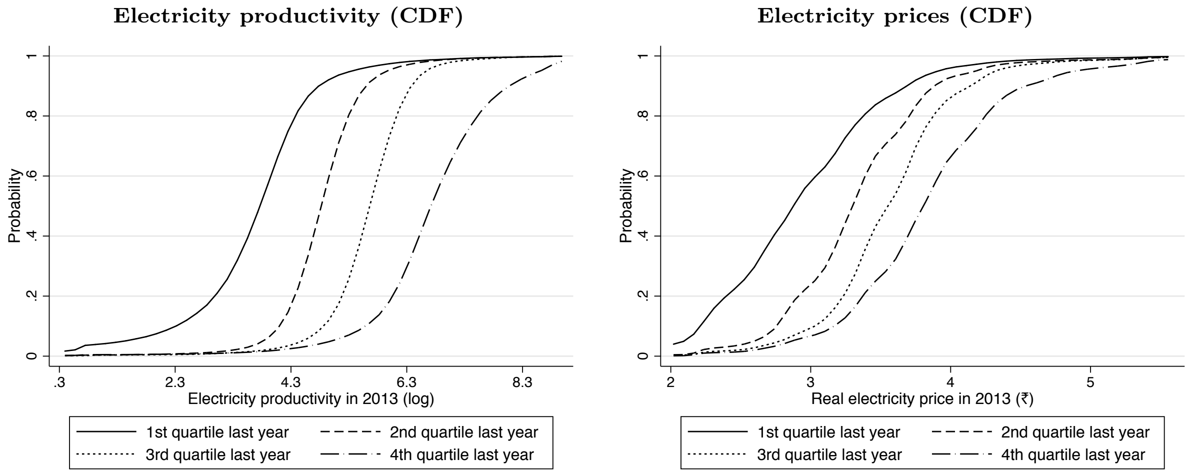
where $s_g = \sum_{e \in g} s_e$ and $\overline{p}_g$ the weighted average of log productivity or price within group g. I calculate the decomposition separately five times for the five groups shown in the graph. The regions plot the share of V^G in V (V^G/V), where higher shares explain more of the variation. Groups are deciles of electricity purchase quantity, 4-digit industries, states, and combinations. Plant output and electricity prices are deflated.

Figure III.24: Convergence in electricity prices



Notes: Plotted are the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentile of the annual plant level electricity prices. Electricity prices are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

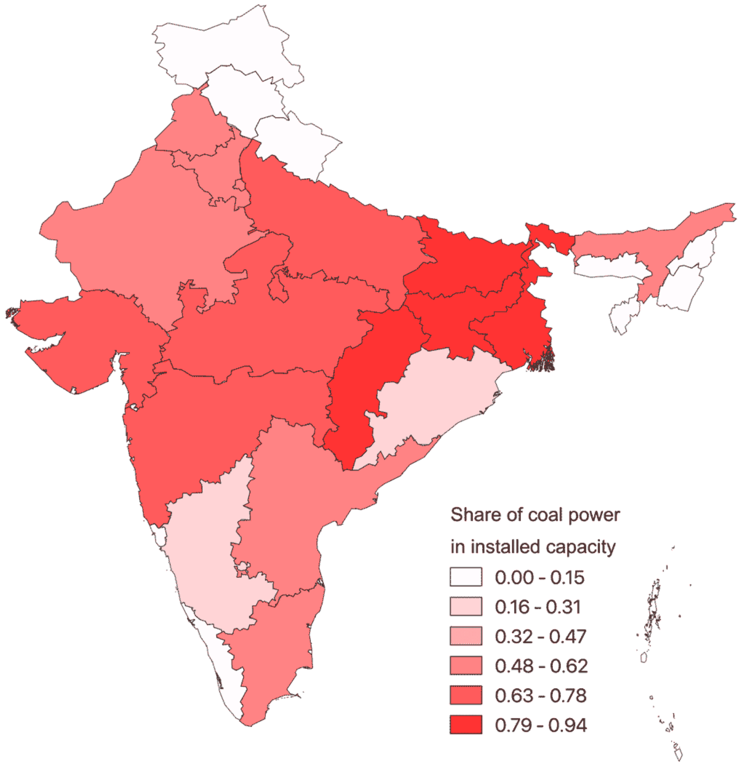
Figure III.25: CDFs of plant electricity productivity and prices in 2013 conditional on 2012 quartiles



Notes: Plotted are the CDFs in 2013, separately for each quartile of the respective values in 2012. The left panel shows the distribution of the logged electricity productivity (i.e. the value of output divided by the electricity use in kWh). The right panel shows the distribution of the electricity price. The CDFs are empirical CDFs obtained through a Gaussian kernel smoother with bandwidth 0.1. The graphs show that each higher quartile first order stochastically dominates the lower quartiles. The conditional CDF of the plants that belong to the higher *previous* year quartile lies to the right of the CDF of the plants belonging to the lower *previous* year quartile. While individual plants move up and down the ranking of electricity productivity and energy prices from one year to the next, the probability of higher productivity and prices increases in last periods productivity and prices. Plant output is deflated using 3-digit industry deflators. Electricity prices are deflated using a general fuel and electricity wholesale price deflator. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

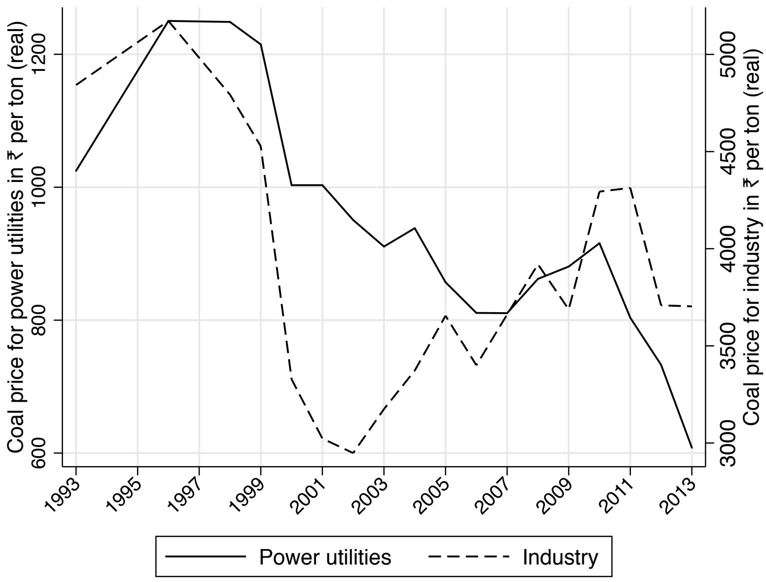
III.J Coal share in installed capacity and coal price for power utilities and industry

Figure III.26: Share of coal power in total installed capacity



Notes: The shading indicates the share of coal fired thermal power generation capacity in total installed capacity at the state level in March 1998. Data comes from Ministry of Power (1998*a*, 2003).

Figure III.27: Coal price for power utilities and industry



Notes: The solid line plots the coal prices for thermal power plants and are from Minsitry of Coal (2012, 2015) as described in Section III.2.2. Prices for coal used in manufacturing industries are plotted with the dashed line. These are averages of the coal prices at the plant level in the ASI micro data (see Section III.2.2). All coal prices are in real terms and deflated using a general fuel and electricity wholesale price deflator. In nominal terms, coal prices have been mostly increasing. The base year for deflation is 2004 throughout this chapter. Wholesale price deflators are from the Office of the Economic Adviser from the Government of India.

III.K Robustness checks and additional regressions

Table III.11: Electricity prices and electricity productivity with two alternative instruments IV^C and IV^D

	OLS	IV^C	IV^D
	(1)	(2)	(3)
$\log(P^E)$	0.366***	-0.267***	-0.475
	(0.044)	(0.071)	(0.679)
Distance to coalfield			-0.0126
(in ’00 km)			(0.009)
OLS/IV	OLS	IV^C	IV^D
Observations	485948	444952	444952
Ind-region-year FE	Yes	Yes	Yes
First stage coef.	-	0.97***	0.02***
First stage SE	-	0.005	0.008
F-stat (Kleib.-Paap)	-	37708.429	7.194
Two-way cluster plant state-year	Yes	Yes	Yes

Notes: See Table III.2 for notes. The main difference in this table is the use of two alternative instruments, IV^C and IV^D.

Table III.12: Electricity prices and electricity productivity in electricity intensive sectors

	OLS	IV^A	IV^B
	(1)	(2)	(3)
$\log(P^E)$	0.323***	-0.208***	-0.582***
	(0.047)	(0.074)	(0.102)
OLS/IV	OLS	IV^A	IV^B
Observations	260900	260900	260900
Ind-region-year FE	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***
First stage SE	-	0.005	0.004
F-stat (Kleib.-Paap)	-	32789.655	324.114
Two-way cluster plant state-year	Yes	Yes	Yes

Notes: See Table III.2 for notes. The main difference is that the sample is restricted to electricity intensive sectors only.

Table III.13: Electricity prices and electricity productivity: using both IVs

	OLS	IV^A & IV^B	IV^C & IV^B
	(1)	(2)	(3)
$\log(P^E)$	0.366***	-0.256***	-0.288***
	(0.044)	(0.068)	(0.069)
IV 1	-	IV^A	IV^C
IV 2	-	IV^B	IV^B
Observations	485948	485948	444952
Ind by region by year FE	Yes	Yes	Yes
State FE	No	No	No
Plant FE	No	No	No
State trends	No	No	No
State by year FE	No	No	No
First stage coef. 1/1	-	0.94***	0.94***
First stage SE 1/1	-	0.007	0.008
First stage coef. 1/2	-	0.00***	0.00***
First stage SE 1/2	-	0.001	0.001
F-stat (Kleibergen-Paap)	-	23320.712	20389.385
Anderson-Rubin F	-	0.000	0.000
J-statistic	-	26.12	28.81
Two-way cluster plant state-year	Yes	Yes	Yes

Notes: See Table III.2 for notes. The main difference is that both instruments are used simultaneously. The Sargan-Hansen J statistic is reported. The difference in the instrument is consistent with heterogeneous LATEs.

Table III.14: Electricity prices and electricity productivity: additional fixed effects and trends

	OLS		IV^A		IV^B
	(1)	(2)	(3)	(4)	(5)
$\log(P^E)$	0.708***	0.684***	-0.545*	0.0229	-1.809*
	(0.030)	(0.018)	(0.291)	(0.056)	(0.982)
OLS/IV	OLS	OLS	IV^A	IV^A	IV^B
Observations	485948	425794	485948	425794	425794
Ind-year FE	Yes	Yes	Yes	Yes	Yes
Ind-region-year FE	Yes	No	Yes	No	No
State FE	Yes	Yes	Yes	Yes	Yes
Plant FE	No	Yes	No	Yes	Yes
State trends	Yes	No	Yes	No	No
First stage coef.	-	-	0.89***	0.92***	0.16**
First stage SE	-	-	0.015	0.009	0.069
F-stat (Kleib.-Paap)	-	-	3499.772	9379.183	5.204
SE clustered by	Plant	Plant	Plant	Plant	Plant
No. of first clusters	160955	100418	160955	100418	100418
SE clustered by	State-year	State-year	State-year	State-year	State-year
No. of second clusters	501	501	501	501	501

Notes: See Table III.2 for notes. The main difference is the inclusion of different fixed effects as indicated.

Table III.16: Electricity prices and electricity productivity by industry groups

(a) Electricity prices and electricity productivity (Chemicals, food, machinery))

	Chemicals			Food			Machinery		
	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	0.178***	-0.389***	-0.765***	0.572***	0.0436	-1.546***	0.217***	-0.629***	-1.250***
	(0.064)	(0.085)	(0.106)	(0.073)	(0.162)	(0.404)	(0.066)	(0.093)	(0.133)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	73838	73838	73838	96601	96601	96601	89944	89944	89944
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.98***	0.08***	-	0.91***	0.04***	-	1.01***	0.07***
First stage SE	-	0.007	0.003	-	0.014	0.003	-	0.007	0.004
F-stat (Kleib.-Paap)	-	17799.309	533.240	-	4339.858	115.564	-	23887.783	337.618
SE clustered by	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
No. of first clusters	26826	26826	26826	33492	33492	33492	29046	29046	29046
SE clustered by	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year
No. of second clusters	472	472	472	500	500	500	440	440	440

(b) Electricity prices and electricity productivity (Metals and minerals, textiles, wood and paper)

	Metals and minerals			Textiles			Wood and Paper		
	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	0.476***	0.0885	0.210	0.410***	-0.177	-0.949***	0.342***	-0.227**	-0.733***
	(0.053)	(0.102)	(0.191)	(0.078)	(0.156)	(0.257)	(0.067)	(0.096)	(0.138)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	104738	104738	104738	71166	71166	71166	36352	36352	36352
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.96***	0.05***	-	0.99***	0.07***	-	0.99***	0.06***
First stage SE	-	0.009	0.003	-	0.012	0.005	-	0.009	0.004
F-stat (Kleib.-Paap)	-	11445.114	181.778	-	6266.604	196.845	-	11169.861	273.140
SE clustered by	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
No. of first clusters	40261	40261	40261	23117	23117	23117	13346	13346	13346
SE clustered by	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year	State-year
No. of second clusters	486	486	486	443	443	443	499	499	499

Notes: See Table III.2 for notes. The main difference is that regressions are run individually by industry groups.

Table III.15: Electricity prices and electricity productivity: clustering at district and region year

	OLS	IV^A	IV^B
	(1)	(2)	(3)
log(P^E)	0.340***	-0.264*	-0.818***
	(0.117)	(0.154)	(0.218)
OLS/IV	OLS	IV^A	IV^B
Observations	444952	444952	444952
Ind-region-year FE	Yes	Yes	Yes
First stage coef.	-	0.98***	0.06***
First stage SE	-	0.018	0.010
F-stat (Kleib.-Paap)	-	3057.138	38.818
SE clustered by	District	District	District
No. of first clusters	541	541	541
SE clustered by	Region-year	Region-year	Region-year
No. of second clusters	96	96	96

Notes: See Table III.2 for notes. The main difference is that the standard errors are clustered at a higher level, at the district level and the region-year level.

Table III.17: Electricity prices and electricity productivity interacted with three periods

	OLS	IV^A	IV^B
	(1)	(2)	(3)
log(P^E)	0.506***	0.0573	-0.719***
	(0.067)	(0.111)	(0.200)
log(P^E) $\cdot$ $\mathbf{1}(year < 2003)$	-0.275***	-0.729***	-0.124
	(0.098)	(0.163)	(0.234)
log(P^E) $\cdot$ $\mathbf{1}(year \geq 2003$ or $year \leq 2007)$	-0.177*	-0.272*	-0.0682
	(0.104)	(0.147)	(0.247)
OLS/IV	OLS	IV^A	IV^B
Observations	485948	485948	485948
Ind by region by year FE	Yes	Yes	Yes
First stage coef. 1/1	-	0.95***	0.05***
First stage SE 1/1	-	0.007	0.005
First stage coef. 1/2	-	0.04***	0.01
First stage SE 1/2	-	0.012	0.008
First stage coef. 1/3	-	0.03***	0.01
First stage SE 1/3	-	0.010	0.008
First stage coef. 2/1	-	0.00	0.00
First stage SE 2/1	-	0.000	0.000
First stage coef. 2/2	-	0.99***	0.06***
First stage SE 2/2	-	0.009	0.006
First stage coef. 2/3	-	0.00	-0.00
First stage SE 2/3	-	0.000	0.000
First stage coef. 3/1	-	-0.00***	0.00
First stage SE 3/1	-	.	0.000
First stage coef. 3/2	-	0.00***	-0.00
First stage SE 3/2	-	.	0.000
First stage coef. 3/3	-	0.98***	0.06***
First stage SE 3/3	-	0.007	0.007
F-stat (Kleibergen-Paap)	-	3875.232	35.761
Two-way cluster plant state-year	Yes	Yes	Yes

Notes: See Table III.4 for notes. The main difference is that prices are interacted with three different periods (one baseline omitted).

Table III.18: Electricity prices and electricity productivity: controlling for distance to coalfields and shortages

	OLS			IV^A			IV^B		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	0.343***	0.472***	0.459***	-0.255***	-0.130	-0.121	-0.827***	-0.938***	-0.980***
	(0.045)	(0.043)	(0.044)	(0.070)	(0.085)	(0.088)	(0.101)	(0.149)	(0.148)
Distance to coalfield	-0.0181***		-0.0192***	-0.0141**		-0.0178**	-0.0102		-0.0157*
(in '00 km)	(0.007)		(0.007)	(0.007)		(0.007)	(0.008)		(0.008)
Shortage		0.397*	0.282		0.644***	0.515***		0.976***	0.860***
		(0.226)	(0.239)		(0.187)	(0.192)		(0.198)	(0.201)
OLS/IV	OLS	OLS	OLS	IV^A	IV^A	IV^A	IV^B	IV^B	IV^B
Observations	444952	474029	433262	444952	474029	433262	444952	474029	433262
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	-	-	0.98***	0.97***	0.98***	0.06***	0.05***	0.05***
First stage SE	-	-	-	0.005	0.006	0.006	0.003	0.004	0.004
F-stat (Kleib.-Paap)	-	-	-	41022.067	25440.719	26150.603	307.715	173.552	176.792
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See Table III.2 for notes. The main difference is that control variables are added as indicated.

Table III.19: Electricity prices, employment, machine labour ratio and product scope

	Employees (log)			Ratio machinery to employees (log)			Number of products (log)		
	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	0.0119	-0.339***	-0.518***	-0.160**	-0.627***	-1.517***	0.0456***	-0.00288	-0.0960***
	(0.041)	(0.076)	(0.079)	(0.065)	(0.114)	(0.151)	(0.012)	(0.023)	(0.036)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	485342	485342	485342	467686	467686	467686	485067	485067	485067
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.97***	0.06***	-	0.97***	0.06***	-	0.97***	0.06***
First stage SE	-	0.005	0.003	-	0.005	0.003	-	0.005	0.003
F-stat (Kleib.-Paap)	-	43194.635	296.507	-	46754.073	308.855	-	43038.018	296.577
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See Table III.2 for notes. The main difference is that the dependent variables are different as indicated.

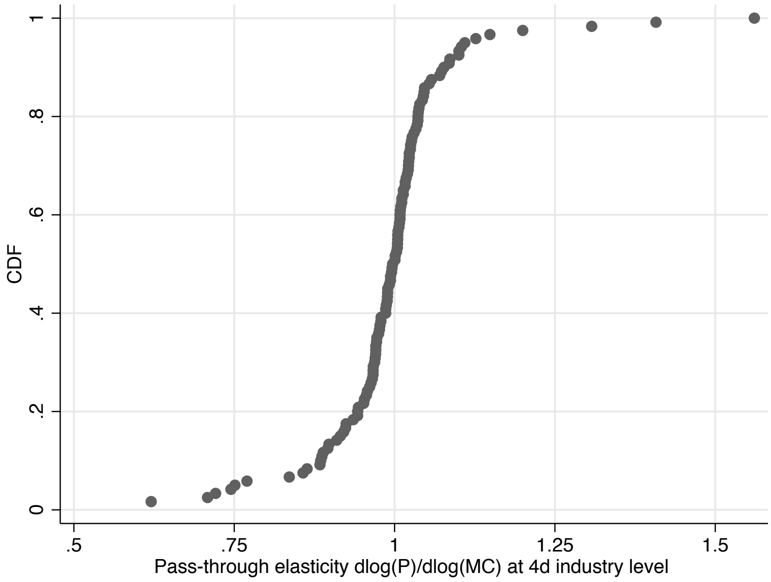
Table III.20: Electricity prices and productivity (TFP): alternative methodologies

	log(TFP) (Olley and Pakes, 1996)			log(TFP) (Levinsohn and Petrin, 2003)			log(TFP) (Ackerberg et al., 2015)		
	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log(P^E)$	-0.00735***	-0.0273***	-0.0387***	-0.000566	-0.0168***	-0.0321***	-0.00414**	-0.00761***	-0.0233***
	(0.002)	(0.004)	(0.005)	(0.002)	(0.004)	(0.007)	(0.002)	(0.003)	(0.006)
OLS/IV	OLS	IV^A	IV^B	OLS	IV^A	IV^B	OLS	IV^A	IV^B
Observations	378824	378824	378824	477697	477697	477697	477697	477697	477697
Ind-region-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage coef.	-	0.98***	0.06***	-	0.97***	0.06***	-	0.97***	0.06***
First stage SE	-	0.004	0.003	-	0.005	0.003	-	0.005	0.003
F-stat (Kleib.-Paap)	-	51023.623	390.549	-	44391.045	297.573	-	44391.045	297.573
Two-way cluster plant state-year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See Table III.2 for notes. The main difference is that different methods to recover TFP are used, and TFP used as dependent variable.

III.L Pass-through elasticities and incidence on consumers over time for aggregated industries

Figure III.28: The distribution of pass-through elasticities

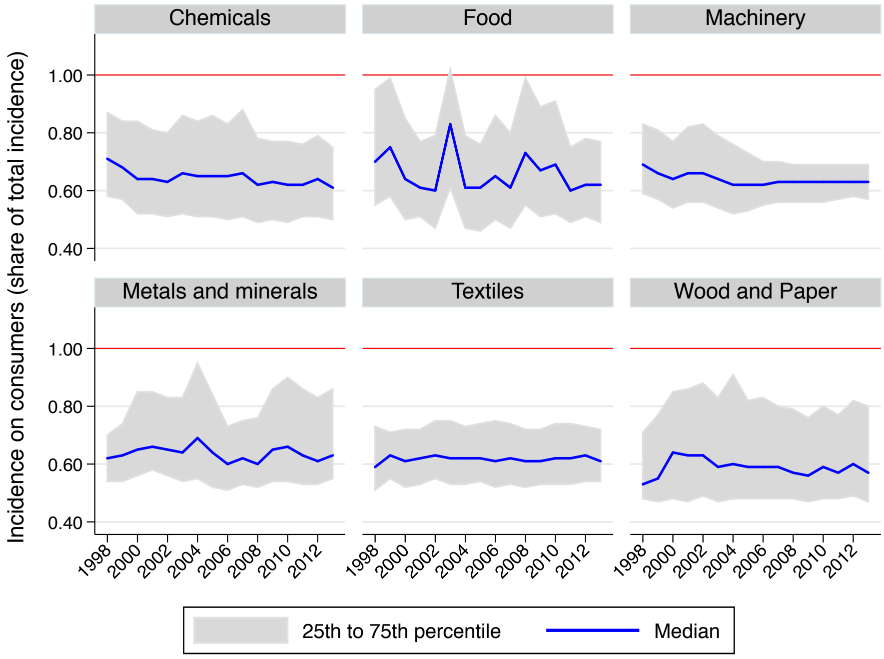


Notes: The figure plots the cumulative distribution function of the pass-through elasticities ($d\log(P)/d\log(MC)$). The pass-through elasticities vary at the 4-digit industry level: there are 121 different pass-through elasticities. The pass-through elasticities are the coefficient on a regression of log prices on log marginal costs at the plant level for each 4-digit industry separately. Prices are calculated as average prices for the different products sold at the firm level, weighted by the quantity sold of each product. Marginal costs are recovered from the estimated markups and the average prices. The marginal costs in the regressions are instrumented with IV^A and IV^B, and regressions are weighted by the sampling weights. Therefore, there are two coefficients per pass-through elasticity per industry. The reported pass-through elasticities are weighted averages, for each pair of coefficients, where the weights are the t-statistics from the IV regression. Here are two example regressions for two different 4-digit industries of log prices on log marginal costs with different IVs:

	Manufacture of:	
	Grain mill products	Structural non-refractory clay and ceramic products
$\log(MC)$	0.997***	0.730***
	(0.0130)	(0.0555)
OLS/IV	IV^A	IV^B
Observations	21812	6208
Region-year FE	Yes	Yes
F-stat (Kleib.-Paap)	35.65	28.98
SE clustered by	Plant	Plant
No. of first clusters	11707	3577
SE clustered by	State-year	State-year
No. of second clusters	435	220

Notes above table.

Figure III.29: Share of incidence on consumers from electricity price changes



Notes: The figure plots the median share of incidence on consumers I^{share} from electricity price changes for each year within each industry. The 25th and 75th percentiles are plotted as well. The industries are broad: chemicals includes rubber and plastics, machinery includes metal products, and textiles includes leather.

III.M Holm-Bonferroni q-values for multiple hypothesis testing

Table III.21 applies the Holm (1979) Bonferroni correction to the p-values to adjust for multiple hypothesis testing.

Table III.21: Holm (1979) Bonferroni correction for multiple hypotheses testing

	OLS			IV^A			IV^B		
	Coef.	p-value	q-value (adj. pval)	Coef.	p-value	q-value (adj. pval)	Coef.	p-value	q-value (adj. pval)
Independent variable: log(electricity price)									
Electricity productivity (log)	0.366	8.5e-16***	1.8e-14***	-0.239	6.9e-04***	0.0041***	-0.776	5.3e-13***	6.4e-12***
Output (log)	-0.027	0.718	1	-0.743	2.7e-07***	2.7e-06***	-1.597	3.7e-23***	5.6e-22***
Electricity consumption (log)	-0.385	3.1e-09***	6.2e-08***	-0.479	0.0021***	0.0103**	-0.797	1.2e-07***	6.1e-07***
Profits	-4.952	0.0012***	0.0153**	-20.634	5.3e-10***	6.4e-09***	-22.429	4.7e-08***	2.8e-07***
Total revenues	-30.182	7.1e-04***	0.0113**	-132.586	5.2e-11***	7.2e-10***	-139.858	1.1e-10***	1.3e-09***
Avg. variable costs (AVC)	-24.118	0.0012***	0.0153**	-109.134	1.1e-10***	1.4e-09***	-114.291	1.5e-10***	1.4e-09***
Electricity share in fuel expenditure	0.025	7.0e-05***	0.0013***	0.014	0.265	0.53	-0.023	0.241	0.241
Other fuels' share in output	0.004	8.6e-04***	0.013**	0.014	1.2e-11***	1.7e-10***	0.023	6.3e-16***	8.2e-15***
Ratio electricity to coal quantity	-10.203	0.0011***	0.015**	-17.542	0.0026***	0.0103**	-21.836	0.0778*	0.156
TFP (log)	-0.007	0.0031***	0.0339**	-0.016	5.0e-06***	4.5e-05***	-0.033	2.9e-07***	1.1e-06***
Investment in machinery (IHS)	0.162	0.428	1	-0.846	0.0305**	0.0916*	-2.877	1.8e-10***	1.4e-09***
Price marginal cost markup log(μ)	-0.018	0.0035***	0.035**	-0.040	3.9e-04***	0.0028***	-0.106	3.4e-08***	2.4e-07***
Employees (log)	0.012	0.771	1	-0.339	1.1e-05***	8.7e-05***	-0.518	1.3e-10***	1.3e-09***
Ratio machinery to employees (log)	-0.160	0.0138**	0.102	-0.627	5.3e-08***	5.9e-07***	-1.517	8.3e-22***	1.2e-20***
Number of products (log)	0.046	1.8e-04***	0.0032***	-0.003	0.9	0.9	-0.096	0.0078***	0.0234**
Independent variable: log(coal price)									
Coal productivity (log)	0.846	0***	0***	1.487	1.5e-15***	1.2e-14***	1.612	2.1e-13***	1.7e-12***
Output (log)	0.090	0.0036***	0.035**	-0.300	0.226	0.903	-0.135	0.694	1
Coal consumption (log)	-0.756	0***	0***	-1.843	4.2e-11***	3.0e-10***	-1.796	3.9e-06***	2.7e-05***
Electricity consumption (log)	-0.041	0.246	1	-0.426	0.114	0.685	0.734	0.0873*	0.524
Profits	-5.917	3.0e-04***	0.0051***	-5.745	0.703	1	-7.108	0.784	1
Total revenues	-19.988	0.0127**	0.102	-18.739	0.827	1	-0.843	0.995	1
Avg. variable costs (AVC)	-14.357	0.0297**	0.178	-27.758	0.695	1	4.644	0.964	1
TFP (log)	-0.001	0.764	1	-0.020	0.124	0.685	-0.031	0.128	0.642

Notes: The table contains the coefficients and p-values from the original regressions in the main text. The q-values are the adjusted p-values for multiple hypothesis testing using the procedure outlined in Holm (1979). The correction procedures are separetly applied by model (OLS, IV^A, IV^B) and by independent variable log(electricity price) and log(coal price).

Bibliography

Abeberese, A. B. (2017), 'Electricity cost and firm performance: Evidence from india', *Review of Economics and Statistics* **99**(5), 839–852.

Acemoglu, D., Aghion, P., Bursztyn, L. and Hemous, D. (2012), 'The environment and directed technical change', *American economic review* **102**(1), 131–66.

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N. and Kerr, W. (2018), 'Innovation, reallocation, and growth', *American Economic Review* **108**(11), 3450–91.

Ackerberg, D. A., Caves, K. and Frazer, G. (2015), 'Identification properties of recent production function estimators', *Econometrica* **83**(6), 2411–2451.

Ackerberg, D., Benkard, C. L., Berry, S. and Pakes, A. (2007), 'Econometric tools for analyzing market outcomes', *Handbook of econometrics* **6**, 4171–4276.

Adamopoulos, T. (2011), 'Transportation costs, agricultural productivity, and cross-country income differences', *International Economic Review* **52**(2), 489–521.

Adamopoulos, T., Brandt, L., Leight, J. and Restuccia, D. (2017), 'Misallocation, selection and productivity: A quantitative analysis with panel data from china', *NBER working paper* .

Adamopoulos, T. and Restuccia, D. (2014*a*), 'Land reform and productivity: A quantitative analysis with micro data', *mimeo, University of Toronto* .

Adamopoulos, T. and Restuccia, D. (2014*b*), 'The size distribution of farms and international productivity differences', *The American Economic Review* **104**(6), 1667–1697.

Adao, R., Costinot, A. and Donaldson, D. (2017), 'Nonparametric counterfactual predictions in neoclassical models of international trade', *American Economic Review* **107**(3), 633–89.

Akcigit, U., Alp, H. and Peters, M. (2016), 'Lack of selection and imperfect managerial contracts: Firm dynamics in developing countries', *MIT working paper* .

Akcigit, U., Baslandze, S. and Lotti, F. (2018), 'Connecting to power: political connections, innovation, and firm dynamics', *mimeo* .

Alam, M. (2013), 'Coping with blackouts: Power outages and firm choices', *Department of Economics, Yale University* .

Alder, S. (2017), 'Chinese roads in india: The effect of transport infrastructure on economic development', *Work. Pap., Univ. North Carolina, Chapel Hill* .

Aldy, J. E. and Pizer, W. A. (2015), 'The competitiveness impacts of climate change mitigation policies', *Journal of the Association of Environmental and Resource Economists* **2**(4), 565–595.

Alfaro, L. and Chari, A. (2014), 'Deregulation, misallocation, and size: Evidence from india', *The Journal of Law and Economics* **57**(4), 897–936.

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

Allen, T. (2014), 'Information frictions in trade', *Econometrica* **82**(6), 2041–2083.

Allen, T. and Arkolakis, C. (2014), 'Trade and the topography of the spatial economy', *The Quarterly Journal of Economics* **129**(3), 1085–1140.

Allen, T. and Arkolakis, C. (2019), 'The welfare effects of transportation infrastructure improvements', *NBER working paper* .

Allen, T. and Atkin, D. (2016), 'Volatility and the gains from trade', *NBER working paper* .

Amaral, P. S. and Quintin, E. (2010), 'Limited enforcement, financial intermediation, and economic development: A quantitative assessment*', *International Economic Review* **51**(3), 785–811.

Amiti, M. and Konings, J. (2007), 'Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia', *American Economic Review* **97**(5), 1611–1638.

Anderson, J. E. and Van Wincoop, E. (2004), 'Trade costs', *Journal of Economic literature* **42**(3), 691–751.

Angrist, J. D. and Pischke, J.-S. (2008), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Arellano, M. and Bond, S. (1991), 'Some tests of specification for panel data: Monte carlo evidence and an application to employment equations', *The review of economic studies* **58**(2), 277–297.

Arkolakis, C., Costinot, A., Donaldson, D. and Rodríguez-Clare, A. (2018), 'The elusive pro-competitive effects of trade', *The Review of Economic Studies* .

Armstrong, T. B. (2016), 'Large market asymptotics for differentiated product demand estimators with economic models of supply', *Econometrica* **84**(5), 1961–1980.

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

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

Atalay, E. (2014), 'Materials prices and productivity', *Journal of the European Economic Association* **12**(3), 575–611.

Atkin, D. and Donaldson, D. (2015), 'Who's getting globalized? the size and implications of intra-national trade costs', *NBER working paper* .

Au, C.-C. and Henderson, J. V. (2006), 'Are chinese cities too small?', *The Review of Economic Studies* **73**(3), 549–576.

Baily, M. N., Hulten, C., Campbell, D., Bresnahan, T. and Caves, R. E. (1992), 'Productivity dynamics in manufacturing plants', *Brookings papers on economic activity. Microeconomics* **1992**, 187–267.

Banerjee, A. V. and Duflo, E. (2005), 'Growth theory through the lens of development economics', *Handbook of economic growth* **1**, 473–552.

Banerjee, A. V., Duflo, E. and Qian, N. (2012), 'On the road: Access to transportation infrastructure and economic growth in china', *NBER Working Paper* (w17897).

Baptist, S. and Hepburn, C. (2013), 'Intermediate inputs and economic productivity', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**(1986).

Baqaee, D. R. and Farhi, E. (2017), 'Productivity and misallocation in general equilibrium.', *NBER working paper* .

Barrows, G. and Ollivier, H. (2018), 'Cleaner firms or cleaner products? how product mix shapes emission intensity from manufacturing', *Journal of Environmental Economics and Management* **88**, 134–158.

Barseghyan, L. and DiCecio, R. (2011), 'Entry costs, industry structure, and cross-country income and tfp differences', *Journal of Economic Theory* **146**(5), 1828–1851.

Bartelsman, E. J. and Doms, M. (2000), 'Understanding productivity: Lessons from longitudinal microdata', *Journal of Economic literature* **38**(3), 569–594.

Bartik, T. J. (1991), 'Who benefits from state and local economic development policies?'.

Baum-Snow, N., Henderson, J. V., Turner, M. A., Zhang, Q. and Brandt, L. (2018), 'Does investment in national highways help or hurt hinterland city growth?', *NBER working paper* .

Bayer, C., Mecikovsky, A. M. and Meier, M. (2018), 'Misallocation, markups, and technology', *mimeo* .

Behrens, K., Mion, G., Murata, Y. and Suedekum, J. (2017), 'Spatial frictions', *Journal of Urban Economics* **97**, 40–70.

Behrens, K., Mion, G., Murata, Y. and Suedekum, J. (2018), 'Quantifying the gap between equilibrium and optimum under monopolistic competition', *Mimeo* .

Bento, P. and Restuccia, D. (2017), 'Misallocation, establishment size, and productivity', *American Economic Journal: Macroeconomics* .

Berndt, E. R. and Christensen, L. R. (1973), 'The translog function and the substitution of equipment, structures, and labor in us manufacturing 1929-68', *Journal of econometrics* **1**(1), 81–113.

Berry, S., Levinsohn, J. and Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica* **63**(4), 841–890.

Berry, S., Levinsohn, J. and Pakes, A. (1999), 'Voluntary export restraints on automobiles: Evaluating a trade policy', *American Economic Review* **89**(3), 400–430.

Berry, S. T. (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* **25**(2), 242–262.

Berry, S. T. and Haile, P. A. (2014), 'Identification in differentiated products markets using market level data', *Econometrica* **82**(5), 1749–1797.

Bhagwati, J. and Panagariya, A. (2014), *Reforms and Economic Transformation in India*, Oxford University Press.

Bhattacharya, S. and Patel, U. R. (2008), The power sector in india: An inquiry into the efficacy of the reform process, *in* S. Bery, B. Bosworth and A. Panagariya, eds, 'India Policy Forum', Vol. 4, National Council of Applied Economic Research, pp. 211–283.

Bils, M., Klenow, P. J. and Ruane, C. (2017), Misallocation or mismeasurement?, *in* '2017 Meeting Papers', number 715, Society for Economic Dynamics.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D. and Roberts, J. (2013), 'Does management matter? evidence from india *', *The Quarterly Journal of Economics* **128**(1), 1–51.

Blundell, R. and Bond, S. (1998), 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of econometrics* **87**(1), 115–143.

Blundell, R. and Bond, S. (2000), 'Gmm estimation with persistent panel data: an application to production functions', *Econometric reviews* **19**(3), 321–340.

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

Bollard, A., Klenow, P. and Sharma, G. (2013), 'India's mysterious manufacturing miracle', *Review of Economic Dynamics* **16**(1), 59–85.

BP (2016), *Statistical review of world energy*, BP.

Brown, M. B. and Forsythe, A. B. (1974), 'Robust tests for the equality of variances', *Journal of the American Statistical Association* **69**(346), 364–367.

Bryan, G., Chowdhury, S. and Mobarak, A. M. (2014), 'Underinvestment in a profitable technology: The case of seasonal migration in bangladesh', *Econometrica* **82**(5), 1671–1748.

Bryan, G. and Morten, M. (Forthcoming), 'The aggregate productivity effects of internal migration: Evidence from indonesia', *Journal of Political Economy* .

Buera, F. J., Kaboski, J. P. and Shin, Y. (2011), 'Finance and development: A tale of two sectors', *American Economic Review* **101**(5), 1964–2002.

Buera, F. J., Moll, B. and Shin, Y. (2013), 'Well-intended policies', *Review of Economic Dynamics* **16**(1), 216–230.

Bureau of Energy Efficiency (2014), *Annual Report 2013-2014*, Ministry of Power, Government of India.

Byrd, R. H., Gilbert, J. C. and Nocedal, J. (2000), 'A trust region method based on interior point techniques for nonlinear programming', *Mathematical Programming* **89**(1), 149–185.

Calel, R. and Dechezlepretre, A. (2016), 'Environmental policy and directed technological change: evidence from the european carbon market', *Review of economics and statistics* **98**(1), 173–191.

Caselli, F. and Gennaioli, N. (2013), 'Dynastic management', *Economic Inquiry* **51**(1), 971–996.

Catherine, S., Chaney, T., Huang, Z., Sraer, D. A. and Thesmar, D. (2018), 'Quantifying reduced-form evidence on collateral constraints', *Mimeo* .

Center for Media and Democracy (2017), 'Sourcewatch: Global coal plant tracker'. [Online; accessed 20-April-2019].
URL: *https://www.sourcewatch.org/*

Central Electricity Authority (2006), *Annual report 2005-06*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2007), *Annual report 2006-07*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2008), *Annual report 2007-08*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2009), *Annual report 2008-09*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2010), *Annual report 2009-10*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2011), *Annual report 2010-11*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2012), *Annual report 2011-12*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2013), *Annual report 2012-13*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2014), *Annual report 2013-14*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2015), *Annual report 2014-15*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2018), *Annual report 2017-18*, Central Electricity Authority, Ministriy of Power, Government of India.

Central Electricity Authority (2019), *Monthly report on installed capacity (February 2019)*, Central Electricity Authority, Ministriy of Power, Government of India.

Chamberlain, G. (1987), 'Asymptotic efficiency in estimation with conditional moment restrictions', *Journal of Econometrics* **34**(3), 305–334.

Chari, A., Liu, E. M., Wang, S.-Y. and Wang, Y. (2017), 'Property rights, land misallocation and agricultural efficiency in china', *NBER working paper* (24099).

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

Chen, C., Restuccia, D. and Santaeulàlia-Llopis, R. (2017), 'The effects of land markets on resource allocation and agricultural productivity', *NBER working paper* (24034).

Cheng, W. and Morrow, J. (2018), 'Firm productivity differences from factor markets', *The Journal of Industrial Economics* **66**(1), 126–171.

Christensen, L. R., Jorgenson, D. W. and Lau, L. J. (1971), 'Conjugate duality and the transcendental logarithmic production function', *Econometrica* **39**(4), 255–256.

Christensen, L. R., Jorgenson, D. W. and Lau, L. J. (1973), 'Transcendental logarithmic production frontiers', *The Review of Economics and Statistics* **55**(1), 28–45.

Cirera, X., Jaef, R. F., Gonne, N. and Maemir, H. (2018), 'The elusive missing middle: The size distribution of firms in sub-saharan africa', *mimeo* .

Collard-Wexler, A. and De Loecker, J. (2014), 'Reallocation and technology: evidence from the us steel industry', *The American Economic Review* **105**(1), 131–171.

Cornillie, J. and Fankhauser, S. (2004), 'The energy intensity of transition countries', *Energy Economics* **26**(3), 283–295.

Costinot, A. and Donaldson, D. (2016), 'How large are the gains from economic integration? theory and evidence from us agriculture, 1880-1997', *NBER working paper* .

Cox, M., Peichl, A., Pestel, N. and Siegloch, S. (2014), 'Labor demand effects of rising electricity prices: Evidence for germany', *Energy Policy* **75**, 266–277.

Da-Rocha, J.-M., Tavares, M. M. and Restuccia, D. (2017), 'Policy distortions and aggregate productivity with endogenous establishment-level productivity', *NBER working paper* .

Database of Global Administrative Areas (GADM) (2016), 'Database of global administrative areas'.
URL: *https://gadm.org/*

David, J. M., Hopenhayn, H. A. and Venkateswaran, V. (2016), 'Information, misallocation, and aggregate productivity', *The Quarterly Journal of Economics* **131**(2), 943–1005.

David, J. M. and Venkateswaran, V. (2017), 'The sources of capital misallocation', *NBER working paper* (23129).

Davis, P. and Schiraldi, P. (2014), 'The flexible coefficient multinomial logit (fc-mnl) model of demand for differentiated products', *The RAND Journal of Economics* **45**(1), 32–63.

Davis, S. J., Grim, C. and Haltiwanger, J. (2008), 'Productivity dispersion and input prices: The case of electricity', *US Census Bureau Center for Economic Studies Paper* (CES-WP-08-33).

Davis, S. J., Grim, C., Haltiwanger, J. and Streitwieser, M. (2013), 'Electricity unit value prices and purchase quantities: Us manufacturing plants, 1963–2000', *Review of Economics and Statistics* **95**(4), 1150–1165.

De Loecker, J. (2011), 'Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity', *Econometrica* **79**(5), 1407–1451.

De Loecker, J. (2014), 'Firm performance in a global economy', *Annual Review of Economics* **6**, 201–227.

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

De Loecker, J. and Scott, P. T. (2016), 'Estimating market power evidence from the us brewing industry', *NBER working paper* .

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

Deaton, A. and Muellbauer, J. (1980), 'An almost ideal demand system', *The American economic review* **70**(3), 312–326.

Dechezleprêtre, A. and Sato, M. (2017), 'The impacts of environmental regulations on competitiveness', *Review of Environmental Economics and Policy* **11**(2), 183–206.

Deschenes, O. (2011), Climate policy and labor markets, *in* 'The design and implementation of US climate policy', University of Chicago Press, pp. 37–49.

Dhingra, S. and Morrow, J. (Forthcoming), 'Monopolistic competition and optimum product diversity under firm heterogeneity', *Journal of Political Economy* .

Dias, D. A., Marques, C. R. and Richmond, C. (2016), 'Misallocation and productivity in the lead up to the eurozone crisis', *Journal of Macroeconomics* **49**, 46–70.

Dijkstra, E. W. (1959), 'A note on two problems in connexion with graphs', *Numerische mathematik* **1**(1), 269–271.

Dixit, A. K. and Stiglitz, J. E. (1977), 'Monopolistic competition and optimum product diversity', *American Economic Review* **67**(3), 297–308.

Donaldson, D. (2018), 'Railroads of the raj: Estimating the impact of transportation infrastructure', *American Economic Review* **108**(4-5), 899–934.

Donaldson, D. and Hornbeck, R. (2016), 'Railroads and american economic growth: A "market access" approach', *The Quarterly Journal of Economics* **131**(2), 799–858.

Doraszelski, U. and Jaumandreu, J. (2013), 'R&d and productivity: Estimating endogenous productivity.', *Review of Economic Studies* **80**(4).

Dougherty, S. M. (2009), 'Labour regulation and employment dynamics at the state level in india', *Review of Market Integration* **1**(3), 295–337.

Dubé, J.-P. H., Fox, J. T. and Su, C.-L. (2012), 'Improving the numerical performance of blp static and dynamic discrete choice random coefficients demand estimation', *Econometrica* **80**(5), 2231–2267.

Duranton, G., Ghani, E., Grover, A. and Kerr, W. (2015), 'The misallocation of land and other factors of production in india', *World Bank Policy Research Working Paper* (7221).

Eberhardt, M., Helmers, C. et al. (2010), 'Untested assumptions and data slicing: A critical review of firm-level production function estimators', *Department of Economics, University of Oxford, Working paper series* .

EC (2016), *in* European Commission, Joint Research Centre (JRC)/PBL Netherlands Environmental Assessment Agency, ed., 'Emission Database for Global Atmospheric Research (EDGAR)', release version 4.3.2. http://edgar.jrc.ec.europe.eu.

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

Edmond, C., Midrigan, V. and Xu, D. Y. (2018), 'How costly are markups?', *NBER working paper* .

EIA (2015), *International energy statistics*, EIA.

Elliott, R., Sun, P. and Zhu, T. (2019), 'Electricity prices and industry switching: Evidence from chinese manufacturing firms', *Energy Economics* **78**, 567–588.

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

Ericson, R. and Pakes, A. (1995), 'Markov-perfect industry dynamics: A framework for empirical work', *The Review of Economic Studies* **62**(1), 53–82.

Eslava, M. and Haltiwanger, J. (2019), 'The life-cycle growth of plants: The role of productivity, demand and distortions', *Mimeo* .

European Comission (2013), *Sectoral Resource Maps*, European Comission, DG Environment.

Eurostat (2016), *Electricity prices for non-household consumers - bi-annual data*, Eurostat.

EY (2013), *Movement of goods in India*, Ernst and Young.

EY (2014), *Indian steel. Strategy to ambition*, Ernst and Young.

Faber, B. (2014), 'Trade integration, market size, and industrialization: evidence from china's national trunk highway system', *Review of Economic Studies* **81**(3), 1046–1070.

Fabra, N. and Reguant, M. (2014), 'Pass-through of emissions costs in electricity markets', *American Economic Review* **104**(9), 2872–99.

Faccio, M. (2006), 'Politically connected firms', *The American economic review* **96**(1), 369–386.

Fackler, P. L. and Goodwin, B. K. (2001), 'Spatial price analysis', *Handbook of agricultural economics* **1**, 971–1024.

Fajgelbaum, P. D., Morales, E., Suárez Serrato, J. C. and Zidar, O. (2018), 'State taxes and spatial misallocation', *The Review of Economic Studies* **86**(1), 333–376.

Fajgelbaum, P. D. and Schaal, E. (2017), 'Optimal transport networks in spatial equilibrium', *NBER working paper* .

Farinas, J. C. and Ruano, S. (2005), 'Firm productivity, heterogeneity, sunk costs and market selection', *International Journal of Industrial Organization* **23**(7), 505–534.

Firth, J. (2017), 'I've been waiting on the railroad: The effects of congestion on firm production', *MIT mimeo* .

Fisher-Vanden, K., Hu, Y., Jefferson, G., Rock, M. and Toman, M. (2016), 'Factors influencing energy intensity in four chinese industries', *The Energy Journal* **Volume 37**(China Special Issue).

Fisher-Vanden, K., Jefferson, G. H., Liu, H. and Tao, Q. (2004), 'What is driving china's decline in energy intensity?', *Resource and Energy economics* **26**(1), 77–97.

Fogel, R. W. (1964), *Railroads and American economic growth*, Johns Hopkins Press Baltimore.

Forlani, E., Martin, R., Mion, G. and Muûls, M. (2016), 'Unraveling firms: demand, productivity and markups heterogeneity'.

Foster, L., Haltiwanger, J. C. and Krizan, C. J. (2001), Aggregate productivity growth. lessons from microeconomic evidence, *in* 'New developments in productivity analysis', University of Chicago Press, pp. 303–372.

Foster, L., Haltiwanger, J. and Syverson, C. (2008), 'Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?', *American Economic Review* **98**(1), 394–425.

Foster, V. and Steinbuks, J. (2009), *Paying the price for unreliable power supplies: in-house generation of electricity by firms in Africa*, The World Bank.

Fredman, M. L. and Tarjan, R. E. (1987), 'Fibonacci heaps and their uses in improved network optimization algorithms', *Journal of the ACM (JACM)* **34**(3), 596–615.

Galle, S. (2016), 'Competition, financial constraints and misallocation: Plant-level evidence from indian manufacturing', *Job market paper. University of California, Berkeley* .

Ganapati, S., Shapiro, J. S. and Walker, R. (2016), Energy prices, pass-through, and incidence in us manufacturing, Technical report, National Bureau of Economic Research.

Gandhi, A. and Houde, J.-F. (2016), 'Measuring substitution patterns in differentiated products industries', *mimeo* .

Gandhi, A., Navarro, S. and Rivers, D. (2016), 'On the identification of production functions: How heterogeneous is productivity?', *Working paper.* .

Gandhi, A., Navarro, S., Rivers, D. et al. (2017), 'How heterogeneous is productivity? a comparison of gross output and value added', *Mimeo* .

Garicano, L., LeLarge, C. and Van Reenen, J. (2016), 'Firm size distortions and the productivity distribution: Evidence from france', *American Economic Review* **106**(11), 3439–79.

Ghani, E., Goswami, A. G. and Kerr, W. R. (2014), 'Spatial dynamics of electricity usage in india', *World Bank Policy Research Working Pape* .

Goldberg, P. (1995), 'Product differentiation and oligopoly in international markets: The case of the us automobile industry', *Econometrica* **63**(4), 891–951.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N. and Topalova, P. (2010), 'Imported intermediate inputs and domestic product growth: Evidence from india', *The Quarterly Journal of Economics* **125**(4), 1727–1767.

Golder, B. (2011), 'Energy intensity of indian manufacturing firms: effect of energy prices, technology and firm characteristics', *Science, Technology and Society* **16**(3), 351–372.

Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L. and Villegas-Sanchez, C. (2017), 'Capital allocation and productivity in south europe', *The Quarterly Journal of Economics* **132**(4), 1915–1967.

Gorodnichenko, Y., Revoltella, D., Svejnar, J. and Weiss, C. T. (2018), 'Resource misallocation in european firms: The role of constraints, firm characteristics and managerial decisions', *NBER working paper* (24444).

Griliches, Z. and Mairesse, J. (1999), Production functions: The search for identification, *in* S. Strøm, ed., 'Econometrics and Economic Theory in the 20th Century', Cambridge University Press, pp. 169–203.

Griliches, Z. and Regev, H. (1995), 'Firm productivity in israeli industry 1979–1988', *Journal of econometrics* **65**(1), 175–203.

Griliches, Z. and Ringstad, V. (1971), *Economies of Scale and the Form of the Production Function*, North-Holland Publishing Company.

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

Gupta, R. and Shankar, H. (2019), *Global Energy Observatory*, Database: http:// globalenergyobservatory.org/.

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

Hall, R. E. (1988), 'The relation between price and marginal costs in US industry', *Journal of Political Economy* **96**(5).

Hall, R. E. (1990), Invariance properties of solow's productivity residual, *in* P. Diamond, ed., 'Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday', Cambridge, MA: MIT Press, pp. 71–112.

Hall, R. E. and Jones, C. I. (1999), 'Why do some countries produce so much more output per worker than others?', *The Quarterly Journal of Economics* **114**(1), 83–116.

Haltiwanger, J., Kulick, R. and Syverson, C. (2018), 'Misallocation measures: The distortion that ate the residual', *NBER working paper* (24199).

Hang, L. and Tu, M. (2007), 'The impacts of energy prices on energy intensity: Evidence from china', *Energy policy* **35**(5), 2978–2988.

Harberger, A. C. (1954), 'Monopoly and resource allocation', *The American Economic Review* **44**(2), 77–87.

Harris, C. D. (1954), 'The, market as a factor in the localization of industry in the united states', *Annals of the association of American geographers* **44**(4), 315–348.

Harrison, A. E., Martin, L. A. and Nataraj, S. (2013), 'Learning versus stealing: how important are market-share reallocations to india's productivity growth?', *The World Bank Economic Review* **27**(2), 202–228.

Hasan, R. and Jandoc, K. R. (2014), Labor regulations and firm size distribution in indian manufacturing, *in* J. Bhagwati and A. Panagariya, eds, 'Reforms and Economic Transformation in India', Oxford University Press, pp. 15–48.

Hausman, C. (2018), 'Shock value: Bill smoothing and energy price pass-through', *NBER working paper* .

Hausman, J. A. (1996), Valuation of new goods under perfect and imperfect competition, *in* T. Bresnahan and R. Gordon, eds, 'The economics of new goods', University of Chicago Press, pp. 207–248.

Hausman, J., Leonard, G., Zona, J. D. et al. (1994), 'Competitive analysis with differentiated products', *Annals of Economics and Statistics* (34), 143–157.

Ho, S.-J. and Ruzic, D. (2017), 'Returns to scale, productivity measurement, and trends in us manufacturing misallocation', *Mimeo* .

Holm, S. (1979), 'A simple sequentially rejective multiple test procedure', *Scandinavian Journal of Statistics* **6**(2), 65–70.

Hopenhayn, H. A. (1992), 'Entry, exit, and firm dynamics in long run equilibrium', *Econometrica* **60**(5), 1127–50.

Hopenhayn, H. A. (2014*a*), 'Firms, misallocation, and aggregate productivity: A review', *Annu. Rev. Econ.* **6**(1), 735–770.

Hopenhayn, H. A. (2014*b*), 'On the measure of distortions', *NBER working paper* (20404).

Hopenhayn, H. A. and Rogerson, R. (1993), 'Job turnover and policy evaluation: A general equilibrium analysis', *Journal of Political Economy* **101**(5), 915–38.

Hornbeck, R. and Rotemberg, M. (2019), 'Engines of productivity growth: Railroads, reallocation, and the rise of american manufacturing', *Presentation (at LSE)* .

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

Hsieh, C.-T., Hurst, E., Jones, C. I. and Klenow, P. J. (Forthcoming), 'The allocation of talent and us economic growth', *Econometrica* .

Hsieh, C.-T. and Klenow, P. J. (2009), 'Misallocation and manufacturing tfp in china and india', *The Quarterly Journal of Economics* **124**(4), 1403–1448.

Hsieh, C.-T. and Klenow, P. J. (2014), 'The life cycle of plants in india and mexico', *The Quarterly Journal of Economics* **129**(3), 1035–1084.

Hsieh, C.-T. and Moretti, E. (2019), 'Housing constraints and spatial misallocation', *American Economic Journal: Macroeconomics* **11**(2), 1–39.

Huang, Y. and Xiong, W. (2018), 'Geographic distribution of firm productivity and production: A "market access" approach', *Mimeo Harvard University* .

IEA (2015), *India Energy Outlook*, International Energy Agency.

IEA (2016), *World Energy Balances*, Mimas, University of Manchester. DOI: http://dx.doi.org/10.5257/iea/web/2013.

IEA (2018), *World Energy Prices (2018 Edition)*, International Energy Agency.

Indiastat (2019), *State/Utility-wise Estimated Average Rates of Electricity in India, various years*, Indiastat Database: indiastat.com.

IPCC (2007), *in* B. Metz, O. Davidson, P. Bosch, R. Dave and L. Meyer, eds, 'Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change', Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jevons, W. (1865), The coal question: can britain survive?, *in* A. Flux, ed., 'The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-mines.', Augustus M. Kelley, New York.

Jones, C. (2013), Misallocation, economic growth, and input-output economics misallocation, economic growth, and input-output economics, *in* D. Acemoglu, M. Arellano and E. Dekel, eds, 'Advances in Economics and Econometrics: Tenth World Congress', Vol. 2, Cambridge University Press, p. 419.

Jones, C. I. (2011), 'Intermediate goods and weak links in the theory of economic development', *American Economic Journal: Macroeconomics* **3**(2), 1–28.

Kahn, M. E. and Mansur, E. T. (2013), 'Do local energy prices and regulation affect the geographic concentration of employment?', *Journal of Public Economics* **101**, 105–114.

KAPSARC (2018), *India Gas Power Plants*, King Abdullah Petroleum Studies And Research Center.

Katayama, H., Lu, S. and Tybout, J. R. (2009), 'Firm-level productivity studies: illusions and a solution', *International Journal of Industrial Organization* **27**(3), 403–413.

Kimball, M. (1995), 'The quantitative analytics of the basic neomonetarist model', *Journal of Money, Credit and Banking* **27**(4), 1241–77.

Klenow, P. J. and Willis, J. L. (2016), 'Real rigidities and nominal price changes', *Economica* **83**(331), 443–472.

Kmenta, J. (1967), 'On estimation of the ces production function', *International Economic Review* **8**(2), 180–189.

Kothari, S. (2014), 'The size distribution of manufacturing plants and development', *IMF Working Paper* **14**(236).

Kugler, M. and Verhoogen, E. (2012), 'Prices, plant size, and product quality', *Review of Economic Studies* **79**(1), 307–339.

Lagos, R. (2006), 'A model of tfp', *The Review of Economic Studies* **73**(4), 983–1007.

Lenzu, S. and Manaresi, F. (2018), 'Do marginal products differ from user costs? micro-level evidence from italian firms', *Mimeo* .

Lerner, A. P. (1934), 'The concept of monopoly and the measurement of monopoly power', *Review of Economic Studies* **1**(157), 75.

Levene, H. (1960), Robust tests for equality of variances, *in* I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow and H. B. Mann, eds, 'Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling', Stanford University Press, pp. 278–292.

Levinsohn, J. and Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The Review of Economic Studies* **70**(2), 317–341.

Liang, Y. (2017), 'Imperfect competition and misallocations', *mimeo* .

Linn, J. (2008), 'Energy prices and the adoption of energy-saving technology', *The Economic Journal* **118**(533), 1986–2012.

Liu, N. and Ang, B. (2007), 'Factors shaping aggregate energy intensity trend for industry: Energy intensity versus product mix', *Energy Economics* **29**(4), 609–635.

Marin, A. G. and Voigtländer, N. (Forthcoming), 'Exporting and plant-level efficiency gains: It's in the measure', *Journal of Political Economy* .

Marin, G. and Vona, F. (2017), 'The impact of energy prices on employment and environmental performance: Evidence from french manufacturing establishments', *FEEM Working Paper* .

Marschak, J. and Andrews, W. H. (1944), 'Random simultaneous equations and the theory of production', *Econometrica, Journal of the Econometric Society* **12**(3/4), 143–205.

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

Martin, R., De Preux, L. B. and Wagner, U. J. (2014), 'The impact of a carbon tax on manufacturing: Evidence from microdata', *Journal of Public Economics* **117**, 1–14.

Martin, R., Muûls, M. and Wagner, U. J. (2015), 'The impact of the european union emissions trading scheme on regulated firms: What is the evidence after ten years?', *Review of Environmental Economics and Policy* **10**(1), 129–148.

Mayer, T., Melitz, M. J. and Ottaviano, G. I. (2014), 'Market size, competition, and the product mix of exporters', *The American Economic Review* **104**(2), 495–536.

Melitz, M. J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica* **71**(6), 1695–1725.

Meunier, G., Ponssard, J.-P. and Thomas, C. (2016), 'Capacity investment under demand uncertainty: The role of imports in the us cement industry', *Journal of Economics & Management Strategy* **25**(2), 455–486.

Midrigan, V. and Xu, D. Y. (2014), 'Finance and misallocation: Evidence from plant-level data', *American Economic Review* **104**(2), 422–458.

Miller, N. H., Osborne, M. and Sheu, G. (2017), 'Pass-through in a concentrated industry: empirical evidence and regulatory implications', *The RAND Journal of Economics* **48**(1), 69–93.

Ministry of Power (1998*a*), *Annual Report 1997-1998*, Ministry of Power, Government of India.

Ministry of Power (1998*b*), *State/Utility-wise Estimated Average Rates of Electricity in India (As on 01.04.1998)*, Ministry of Power, Government of India, Accessed through indiastat.com.

Ministry of Power (2003), *Annual Report 2002-2003*, Ministry of Power, Government of India.

Ministry of Power (2014*a*), *Annual Report 2013-2014*, Ministry of Power, Government of India.

Ministry of Power (2014*b*), *State/Utility-wise Estimated Average Rates of Electricity in India (As on 01.04.2014)*, Ministry of Power, Government of India, Accessed through indiastat.com.

Ministry of Power (2018), *Annual Report 2017-2018*, Ministry of Power, Government of India.

Ministry of Road Transport and Highways (2016), *Basic Road Statistics of India*.

Minsitry of Coal (2006), *Provisional Coal Statistics 2005-2006*, Ministry of Coal, Government of India.

Minsitry of Coal (2012), *Coal Directory of India 2011-2012*, Ministry of Coal, Government of India.

Minsitry of Coal (2013), *Coal Directory of India 2012-2013*, Ministry of Coal, Government of India.

Minsitry of Coal (2015), *Coal Directory of India 2014-2015*, Ministry of Coal, Government of India.

Moll, B. (2014), 'Productivity losses from financial frictions: Can self-financing undo capital misallocation?', *The American Economic Review* **104**(10), 3186–3221.

Morlacco, M. (2019), 'Market power in input markets: Theory and evidence from french manufacturing', *mimeo* .

Mukherjee, K. (2010), 'Measuring energy efficiency in the context of an emerging economy: The case of indian manufacturing', *European Journal of Operational Research* **201**(3), 933–941.

Mukherjee, S. (2012), 'Statistical analysis of the road network of india', *Pramana* **79**(3), 483–491.

Munshi, K. and Rosenzweig, M. (2016), 'Networks and misallocation: Insurance, migration, and the rural-urban wage gap', *The American Economic Review* **106**(1), 46–98.

NCAER (2015), *The Indian Steel Industry: Key Reforms for a Brighter Future*, The National Council of Applied Economic Research.

Nevo, A. (2000*a*), 'Mergers with differentiated products: The case of the ready-to-eat cereal industry', *RAND Journal of Economics* **31**(3), 395–421.

Nevo, A. (2000*b*), 'A practitioner's guide to estimation of random-coefficients logit models of demand', *Journal of Economics & Management Strategy* **9**(4), 513–548.

Nevo, A. (2001), 'Measuring market power in the ready-to-eat cereal industry', *Econometrica* **69**(2), 307–342.

Nishida, M., Petrin, A. and Polanec, S. (2014), 'Exploring reallocation's apparent weak contribution to growth', *Journal of Productivity Analysis* **42**(2), 187–210.

Nishida, M., Petrin, A., Rotemberg, M. and White, T. K. (2015), 'Are we undercounting reallocation's contribution to growth?', *Harvard working paper* .

Nocedal, J. and Wright, S. J. (2006), *Sequential quadratic programming*, Springer.

NPCIL (2015), *Indian Nuclear Power Plants*, Nuclear Power Corporation of India.

NTDPC (2014), *India Transport Report: Moving India to 2032*, National Transport Development Policy Committee. Routledge.

OECD (2015), *Material Resources, Productivity and the Environment*, OECD Green Growth Studies, OECD Publishing, Paris.

Office of the Economic Adviser (2019), *Index Files for WPI Series, various years*, Ministry of Commerce and Industry, Government of India.

Olley, G. S. and Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.

OSMF (2016), 'Openstreetmap'.
URL: *https://www.openstreetmap.org*

Pakes, A. (1996), Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interaction, *in* 'Advances in Econometrics, Sixth World Congress', Vol. 2, C. Sims, pp. 171–259.

Peters, M. (2013), 'Heterogeneous mark-ups, growth and endogenous misallocation', *Working paper. The London School of Economics and Political Science* .

Petrin, A. (2002), 'Quantifying the benefits of new products: The case of the minivan', *Journal of Political Economy* **110**(4), 705–729.

Petrin, A. and Levinsohn, J. (2012), 'Measuring aggregate productivity growth using plant-level data', *The RAND Journal of Economics* **43**(4), 705–725.

Pizer, W. A., Harrington, W., Kopp, R. J., Morgenstern, R. D. and Shih, J.-S. (2002), 'Technology adoption and aggregate energy efficiency', *Resources for the Future Discussion Paper* (02–52).

Planning Commission (2001), *Annual Report on The Working of State Electricity Boards and Electricity Departments*, Planning Commission (Power and Energy Division), Government of India.

Planning Commission (2002), *Annual Report on The Working of State Electricity Boards and Electricity Departments*, Planning Commission (Power and Energy Division), Government of India.

Planning Commission (2014), *Annual Report on The Working of State Power Utilities and Electricity Departments*, Planning Commission (Power and Energy Division), Government of India.

Popp, D. (2002), 'Induced innovation and energy prices', *American economic review* **92**(1), 160–180.

Porter, M. E. and Van der Linde, C. (1995), 'Toward a new conception of the environment-competitiveness relationship', *Journal of economic perspectives* **9**(4), 97–118.

Pozzi, A. and Schivardi, F. (2016), 'Demand or productivity: what determines firm growth?', *The RAND Journal of Economics* **47**(3), 608–630.

Raghuram, G. (2004), 'Integrating coastal shipping with the national transport network in india', *The International Association of Maritime Economists: Annual Conference 2004* pp. 1389–99.

Rangaraj, N. and Raghuram, G. (2007), 'Viability of inland water transport in india', *Asian Development Bank Policy Brief* .

Redding, S. J. (2010), 'The empirics of new economic geography', *Journal of Regional Science* **50**(1), 297–311.

Redding, S. J. and Rossi-Hansberg, E. (2017), 'Quantitative spatial economics', *Annual Review of Economics* **9**, 21–58.

Redding, S. and Venables, A. J. (2004), 'Economic geography and international inequality', *Journal of international Economics* **62**(1), 53–82.

Reinikka, R. and Svensson, J. (2002), 'Coping with poor public capital', *Journal of development economics* **69**(1), 51–69.

Rentschler, J. and Kornejew, M. (2017), 'Energy price variation and competitiveness: Firm level evidence from indonesia', *Energy Economics* **67**, 242–254.

Reserve Bank of India (2019), *Database on Indian Economy: Net State Domestic Product at Factor Cost - State-Wise*, Reserve Bank of India.

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

Restuccia, D. and Rogerson, R. (2013), 'Misallocation and productivity', *Review of Economic Dynamics* **16**(1), 1–10.

Restuccia, D. and Rogerson, R. (2017), 'The causes and costs of misallocation', *Journal of Economic Perspectives* **31**(3), 151–74.

Restuccia, D. and Santaeulalia-Llopis, R. (2017), 'Land misallocation and productivity', *NBER working paper* .

Reynaert, M. and Verboven, F. (2014), 'Improving the performance of random coefficients demand models: the role of optimal instruments', *Journal of Econometrics* **179**(1), 83–98.

Roberts, M., Deichmann, U., Fingleton, B. and Shi, T. (2012), 'Evaluating china's road to prosperity: A new economic geography approach', *Regional Science and Urban Economics* **42**(4), 580–594.

Roeger, W. (1995), 'Can imperfect competition explain the difference between primal and dual productivity measures? estimates for us manufacturing', *Journal of political Economy* **103**(2), 316–330.

Rotemberg, M. (2014), Equilibrium effects of firm subsidies, Technical report, Mimeo, Harvard University.

Rotemberg, M. and White, T. K. (2017), 'Measuring cross-country differences in misallocation', *mimeo* .

Ryan, N. (2017), 'The competitive effects of transmission infrastructure in the indian electricity market', *NBER working paper* .

Ryan, N. (2018), 'Energy productivity and energy demand: Experimental evidence from indian manufacturing plants', *NBER working paper* .

Sadath, A. C. and Acharya, R. H. (2015), 'Effects of energy price rise on investment: Firm level evidence from indian manufacturing sector', *Energy Economics* **49**, 516–522.

Sato, M., Singer, G., Dussaux, D. and Lovo, S. (2019), 'International and sectoral variation in industrial energy prices 1995–2015', *Energy Economics* **78**, 235–258.

Schmalensee, R. (1976), 'Another look at the social valuation of input price changes', *The American Economic Review* **66**(1), 239–243.

Shenoy, A. (2015), 'Market failures and misallocation', *Mimeo, UC Santa Cruz* .

Shenoy, A. (2016), 'Estimating the production function when firms are constrained', *Mimeo, UC Santa Cruz* .

Singer, G. (2018), 'Endogenous markups, input misallocation and geographical supplier access', *LSE mimeo* .

Small, K. and Rosen, H. (1981), 'Applied welfare economics with discrete choice models', *Econometrica* **49**(1), 105–30.

Startz, M. (2018), 'The value of face-to-face: Search and contracting problems in nigerian trade', *Mimeo* .

Stone, R. (1954), 'Linear expenditure systems and demand analysis: an application to the pattern of british demand', *The Economic Journal* **64**(255), 511–527.

Syverson, C. (2004*a*), 'Market structure and productivity: A concrete example', *Journal of Political Economy* **112**(6), 1181–1222.

Syverson, C. (2004*b*), 'Product substitutability and productivity dispersion', *Review of Economics and Statistics* **86**(2), 534–550.

Syverson, C. (2011), 'What determines productivity?', *Journal of Economic Literature* **49**(2), 326–365.

Tortarolo, D. and Zarate, R. D. (2018), 'Measuring imperfect competition in product and labor markets. an empirical analysis using firm-level production data', *CAF – Working paper N° 2018/03* .

Trippi, M. H. and Tewalt, S. J. (2011), Geographic information system (gis) representation of coal-bearing areas in india and bangladesh, Technical report, US Geological Survey.

UNFCCC (2016), *Greenhouse Gas Inventory Data*, UNFCCC.

UNIDO (2016*a*), *INDSTAT2 Industrial Statistics Database*, United Nations Industrial Development Organization.

UNIDO (2016*b*), 'Industrial statistics database', *Indstat4: Industrial Statistics Database* .

Van Leemput, E. (2016), 'A passage to india: Quantifying internal and external barriers to trade', *Mimeo* .

Vries, G. J. (2014), 'Productivity in a distorted market: the case of brazil's retail sector', *Review of Income and Wealth* **60**(3), 499–524.

Walker, W. R. (2013), 'The transitional costs of sectoral reallocation: Evidence from the clean air act and the workforce', *The Quarterly journal of economics* **128**(4), 1787–1835.

Wang, W. (2017), 'Essays on growth and input misallocation in china', *University of Western Ontario Electronic Thesis and Dissertation Repository* (4879).

Weyl, E. G. and Fabinger, M. (2013), 'Pass-through as an economic tool: Principles of incidence under imperfect competition', *Journal of Political Economy* **121**(3), 528–583.

Wooldridge, J. M. (2009), 'On estimating firm-level production functions using proxy variables to control for unobservables', *Economics Letters* **104**(3), 112–114.

World Bank (2005*a*), *Enterprise Survey India.*

World Bank (2005*b*), *India: Road Transport Service Efficiency Study.*, World Bank, Energy and Infrastructure Operations Division.

World Bank (2017), *World Development Indicators*, Washington, DC: World Bank.

World Coal Association (2014), *Coal and Steel Statistics.*

WSA (2010), *Steel Statistical Yearbook*, World Steel Association.

WSA (2016*a*), *Steel and coal factsheet*, World Steel Association.

WSA (2016*b*), *Steel Statistical Yearbook*, World Steel Association.

WSA (2017), *World Steel Association FAQ*, Vol. retrieved from http://www.worldsteel.org/faq, World Steel Association.