Essays in Public and Development Economics

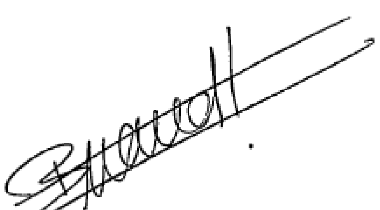
Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Océane Briand

vorgelegt im Herbstsemester 2016

Abteilungssprecher: Prof. Dr. Carsten Trenkler
Referent: Prof. Dr. Eckhard Janeba
Korreferent: Prof. Dr. Sebastian Siegloch
Tag der mündlichen Prüfung: 16. Januar 2017

To Amandine, Erwan, Julien, Thalassio and Hardy.



Acknowledgements

First of all, I would like to thank my advisor Eckhard Janeba for his excellent guidance and encouraging support. I am grateful for the confidence he showed in me and for the freedom he gave me throughout my doctoral studies. I am especially grateful to Markus Frölich for giving me the opportunity to work with him as a consultant for the International Labour Organization (ILO) in Morocco and for his constant availability and guidance. I am also thankful to Johannes Voget and Sebastian Siegloch for enthusiastic advice and constructive criticism.

Special thanks to Mohammad Hoseini and Claudio Baccianti for their work in co-writing the third and fourth chapter respectively of this dissertation. I am thankful to all the members of the Department of Economics and the Center for Doctoral Studies in Economics for their comments and their constant support in achieving my academic goals. In particular, I would like to thank Maria Isabel Santana for supporting my efforts in finding a field experience in India. I was also fortunate to have the best officemates that one could think of in Giuseppe Corbisiero and Christoph Esslinger.

I further wish to thank Peter Hilliges and Kurt Strasser for giving me the privilege to work at the German Bank of Development in India and for many insightful discussions on the world of international development. Furthermore, I would like to show my appreciation to Rukhmini Parthasarathy and Martina Stamm: their personality and great professionalism during our evaluation mission still inspire me today.

I am greatly indebted to my friends Valérie Pfister, Julien Pénasse, Pamela Dissak-Delon, Laura Benveniste and Fabienne Rasel, as well as to my brother Thalassio for their encouragements, cheerfulness, and patience all along. All equations included in this dissertation have survived due to the magnanimity of Valérie Pfister. At last, a warmhearted thanks to Claudio for his unconditional support and for always showing me the bright side of life.

Mannheim, 2016

Océane Briand

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

This thesis addresses two issues at the crossroads of Public and Development Economics: the impact of policies on the informal sector on the one hand and on the level of negative externalities caused to the environment by economic activities on the other. All empirical analyses use Indian data and exploit the Indian economic and political context to test the theoretical predictions developed in the different chapters of this dissertation. India has become the focus of much attention in the last decade. First, it is the largest democracy in the world but it is also home to the largest absolute number of poor in 2012.[1] Second, the environmental consequences of economic development in this country have already triggered migrations, worsened poverty and exacerbated ethnic conflicts. Third, India's ability to face the challenges in its path towards sustainable development will impact the global economy: the size of its consumer market and the growing role taken by Indian firms in the international value chain simultaneously represent an opportunity and a threat to existing players. This thesis contributes to a better understanding of how Indian fiscal policies impacted the inclusion of the most disadvantaged into the formal economy and influenced the environmental footprint of economic agents. However, the policy implications of the results presented here are not limited to India or to developing countries, but are also pertinent for advanced economies facing growing social inequalities, fiscal constraints, and greater awareness of environmental problems.

All the chapters of this dissertation are written as independent papers and each of them contains its own introduction and appendices that provide the proofs of the main theoretical results as well as descriptive statistics, robustness checks, and supplementary information on the data used to obtain the empirical results. References from all three can be found in a bibliography at the end of this dissertation. I provide a more detailed summary of each chapter below.

[1] World Bank; International Monetary Fund. 2016. Global Monitoring Report 2015/2016 : Development Goals in an Era of Demographic Change. Washington, DC: World Bank.

The presence of a large informal sector has been blamed for lower-than-expected tax revenues following value-added tax (VAT) reforms, but the divergence between the tax bases of economies with similar levels of informality in South America or South-East Asia after the adoption of a VAT remained unexplained. Chapter 2[2] shows that the structure of the value chain is another key determinant of the success of the VAT in increasing formal participation and tax revenues.

The first contribution made in this chapter is theoretical: we demonstrate that the level of distortion introduced by a cascading sales tax in a value chain increases with the length and the complexity of the supply chain leading to the inputs used by a given industry. The second contribution is empirical. We exploit a tax reform in India between 2003 and 2008 to estimate empirically the impact of replacing a distorting first-point sales tax by a VAT on production, investments, and informality. Using representative surveys of the Indian formal and informal sectors, we exploit exogenous variations in the timing of VAT adoption across Indian states to identify the impact of the tax reform. Moreover, we capture variations in industries' position in the value chain in a simple but comprehensive way. We disentangle the increase in tax enforcement caused by the introduction of the VAT credit and refund system from the production efficiency gains caused by removing the distorting sales tax, and show a heterogeneous effect of these two aspects of the reform across industries. First, the increase in tax compliance is larger in upstream firms with a larger share of formal businesses among their customers. Second, production efficiency gains are larger in downstream industries with high added value inputs.

We take into account in our analysis of the VAT reform that tax compliance is not equivalent to tax registration due to the VAT mandatory registration threshold (MRT). Non-tax registered firms below the MRT are also compliant if registered under another authority than tax authorities. These results lay the ground for the research question raised in Chapter 3: it points out that formality is not a dual phenomenon dividing fully compliant firms from their informal counterparts, but rather a continuous process through which firms gradually become compliant with all regulations.

In Chapter 3[3], I investigate further the complexity of the phenomenon of informality by highlighting the causes and consequences of unreported labour within formally registered firms. I develop a theoretical model where monitoring of compliance is imperfect and heterogeneity in productivity across and within firms generates informality both at the intensive and at the extensive margin of production. The intensive margin of informality arises when formal enterprises fail to register their

[2]Chapter 2 is based on Briand and Hoseini (2016).
[3]Chapter 3 is based on Briand (2014).

workers or to abide by their legal obligations towards them in order to increase their profitability. An important contribution of Chapter 3 is to show that the two margins of formality do not necessarily evolve hand-in-hand in reaction to reforms of tax rates or labour regulations. Indeed, the intensive margin of informality decreases the costs of formal entrepreneurs. Hence, it increases the incentives to register formally at the extensive margin. Welfare analyses of policy or tax reforms susceptible to impact informality must then take both margins into account.

In the model presented in Chapter 3, agents self-select into entrepreneurship and employment based on their entrepreneurial skills according to a pattern that is consistent with empirical evidence on the intensive margin of informality in India. Indeed, I use three rounds of a representative survey of the Indian workforce in an innovative way to measure non-compliance with social security and pension schemes by formal employers. I explore some key patterns of non-compliance by formally registered employers with the most important Indian welfare schemes and underline their consistency with the assumptions and results of this dissertation chapter. Both types of occupation can be either formal or informal and this status implies different production functions or wages. On the one hand, I show that informal entrepreneurship is the occupation bringing the highest income to agents in the middle of the productivity distribution. Hence, they "choose" to exit from the formal sector. On the other hand, agents at the lower end of the productivity distribution are "forced" into informal employment within formal firms because it is the only available occupation to them due to their limited comparative advantage on the labour market. The main result of this chapter is that the impact of a given tax reform on welfare - captured by the real wage - depends on the industry's diversification and on fiscal capacity, which is defined as the government's ability to increase the share of taxes it collects from formal firms' revenues.

Chapter 4[4] investigates the interaction between the two main investments targeted by energy policies: energy efficient technologies and self-generation of electricity. In both emerging and advanced economies, their combination is expected to support economic growth by increasing access to electricity while minimizing its environmental impact by lowering pollution intensity of production. In emerging countries like India, where the public power grid is subject to frequent outages, the legislator intends to reduce demand with greater energy efficiency on the one hand, and increase supply with incentives to self-generate on the other hand. In advanced economies, the ongoing fast expansion of distributed electricity generation through small-scale electricity generators using renewable

[4] Chapter 4 is based on Baccianti and Briand (2016).

energy is reviving the attention[5]on self-generation as an important component of energy demand. In Chapter 4, we develop a theoretical model to illustrate that incentives to self-generate electricity in the short-run through fiscal advantages or power shortages can reduce investments in energy efficiency in the long-run. First, we highlight a direct crowding out of energy conservation measures by self-generation, when limited productivity or credit constraints force firms to choose between either one or the other of these investments. Both energy conservation efforts and self-generation of electricity involve high adoption fixed costs, while their benefits are proportional to energy expenditures and productivity. However, only self-generation insures producers against power outages and firms do not internalize the negative externalities caused by higher pollution intensity. Hence, self-generation is preferred to energy efficiency by firms in the middle of the productivity distribution. Second, we show that the most productive firms self-select into self-generation when the costs of grid purchased electricity are higher than self-generated electricity or when power outages are very frequent in the economy. Under monopolistic competition, this comparative advantage allows them to charge lower prices, which increases the intensity of competition in the economy and increases the minimum level of productivity required to adopt energy efficient technologies. Hence, we underline the existence of a determinant of investments in energy efficient technology that is neither linked to behavioral aspects nor to market failures.

We test the predictions of this theoretical model by exploiting an exogenous change in the costs of self-generation caused by the replacement of a Retail Sales tax by a Value-Added tax in India between 2003 and 2008. Thus, we address several endogeneity problems linked to the relationship between the adoption of self-generation and energy efficient technologies. On the one hand, higher energy efficiency reduces required self-generation capacity and thus reduces its adoption costs. On the other hand, generators are not (or less) dependent from the power grid, which lowers their incentives to invest in energy conservation measures as long as the price of self-generated electricity remains low. Finally, our theoretical model shows that higher productivity drives investments in both self-generation, self-generation capacity and energy efficiency. Using a repeated cross-section representative of the Indian manufacturing sector, we remediate to the lack of data on investments in energy saving technologies and estimate the stochastic frontier of energy efficiency in India for this period. Hence, we build an indicator to rank Indian plants according to their energy efficiency that we use as a proxy for investments in energy-efficiency. In line with the predictions of the model, our

[5]Severin Borenstein, Professor of Business Administration and Public Policy in the Economic Analysis and Policy Group of the Haas School of Business at the University of California, Berkeley summarizes the contemporaneous political debate on residential solar in the following blog post

empirical results suggest that self-generation crowds out investment in energy-saving technologies.

Chapter 2

Production Efficiency and Self-Enforcement under the VAT: Empirical evidence from India[1]

2.1 Introduction

Implemented in 160 countries, the value-added tax (VAT) is the source of more than 20% of global tax revenues. VAT has been at the center of many tax reforms due to its reputation as a particularly effective tax instrument for raising fiscal revenues, while being consistent with production efficiency as defined by (Diamond and Mirrlees, 1971). Indeed under the VAT credit and refund system, tax-registered producers are refunded taxes paid on inputs upon display of a receipt. Relative and absolute input prices are left unaltered. On the contrary, the practical implementation of retail sales taxes relies on the difficult distinction between sales for production or for consumption purposes. Thus, consumption taxes eventually apply to many production steps (Friedlaender, 1967). Additionally, business-to-business transactions are "self-enforcing" (Pomeranz, 2015; Slemrod, 1996) as input suppliers would like to under-report the value of the transaction and remit less taxes to the government, while input purchasers would like to over-report it and to be refunded a higher amount of taxes.

Nevertheless, results on the impact of VAT reforms on fiscal revenues and formality are mixed. Analyzing a large panel of countries, Keen and Lockwood (2010) report a positive impact of VAT adoption on the total-taxes-to-GDP ratio but large variations in the sign and the size of this effect across countries. The informal sector has been blamed by Emran and Stiglitz (2005); Piggott and Whalley (2001) for the counter-performances of the VAT. In countries where informality is "too" important, either entire production chains operate outside of the VAT system (De Paula and

[1] Chapter 2 is based on Briand and Hoseini (2016).

Scheinkman (2011)), or trade between tax registered and informal firms leads to interruptions in the crediting VAT chain. Hence, self-enforcement remains limited to a narrow tax base and tax registration creates a competitive disadvantage with respect to firms operating in fully informal value chains. However, the literature remains silent on the reasons for which VAT reforms successfully or unsuccessfully increased tax registration in countries that exhibited comparable levels of informality prior to the tax reform. It follows that the size of the formal sector cannot be the sole factor at play. This paper offers a complementary answer to cross-country differences in the performance of VAT reforms in terms of tax compliance, production, and investments. We disentangle empirically the two main mechanisms that occur when a VAT is introduced: the increase in tax enforcement and the gains of production efficiency. We show that the intensity of these two mechanisms varies across industries as it is respectively determined by the share of output used by other businesses and by the share of added-value embedded in inputs. Therefore, economies where industries differ, or are not connected in the same way within the production chain, react differently to the adoption of a VAT. We develop a theoretical model to prove the validity of our empirical approach when capturing the variations across industries in production efficiency gains. Moreover, we demonstrate empirically that the intensity of these two mechanisms also changes above and below the VAT mandatory registration threshold within an industry, which implies a redistribution of firms within industries across the three existing compliance options: tax registration, non-registration, and registration under a public body other than tax authorities (*simple registration* thereafter). Independently of fiscal considerations, simple registration strengthens the rule of law and the ability of policy makers to support small and medium enterprises. Furthermore, cross-country differences in the success of VAT reforms in raising tax revenues can also be explained by the interaction between industry characteristics with the level of the mandatory registration threshold (Brockmeyer et al., 2015; Kanbur and Keen, 2014).

We analyse the replacement of a first-point sales tax (inappropriately labelled "retail sales tax") by a VAT in Indian states between 2003 and 2008. In practice, first-point sales taxes are charged at several steps of the chain value, due to the difficulty of distinguishing between sales for consumption and for production purposes. Bagchi (1994) reports that at least 30% of Indian states sales tax revenues are collected on business inputs. VAT self-enforcement properties affect detection probabilities in the informal sector and evasion opportunities for tax registered firms. Additionally, the removal of a distorting sales tax positively impacts production and investments[2] through production efficiency

[2]For example, Smart and Bird (2009) find a 12% increase in investments after the replacement of a retail sales tax by a VAT in Canada.

gains.

Using a cross-section of two representative formal and informal enterprise surveys in India over the period between 1989 and 2010, we measure the consequences of the Indian VAT reform on four different outcomes: production, investments, and the two different types of VAT compliance: tax registration and simple registration. We estimate a differences-in-differences model exploiting the fact that Indian states implemented the VAT at different points in time due to exogenous variations in the timing of elections for the states Legislative Assemblies. Based on the methodology developed by Rasmussen (1957), we construct two indices to exploit two different sources of heterogeneity across Indian industries. First, VAT self-enforcement is tighter when businesses account for a larger share of an industry's sales. To capture variations in the increase in tax enforcement, we use the *forward-linkages* (FL) index that represents the dependence of producers in the economy on each industry's output. Second, production efficiency gains caused by removing a distorting sales tax are larger in downstream industries that rely on longer chains of input suppliers. To capture variations in production efficiency gains across industries, we use the *backward linkages* (BL) index that increases with the number of production steps leading to each industry's inputs. We develop a theoretical model of a hierarchical economy to show that the BL index appropriately measures the decrease in distortions caused by the replacement of a cascading sales tax with a VAT. We are then able to show that the introduction of a self-enforcing VAT and the removal of a distorting sales tax both have a heterogeneous effect across industries. Finally, we estimate a triple differences model to take into account that tax registration is mandatory only above a certain turnover threshold and that the intensity of the two properties of the sales tax reform previously mentioned will consequently also differ above and below this threshold.

We find a positive impact of the VAT reform on aggregate investments and production. Crucially, we show large discrepancies between industries in the effects of the VAT implementation resulting from various sensitivity of industries to the two mechanisms mentioned above: tighter enforcement and higher production efficiency. For example, the VAT reform increased tax registration by +4.9% in the transport industry whose inputs embed a large share of added value and whose output is largely used by businesses. Indeed, production efficiency gains compensated for the loss in profitability caused by stricter tax enforcement at the intensive margin. Crucially, we show that synergies between these two industry characteristics constitute an increase in enforcement at the extensive margin of tax registration. This arises from the fact that business customers of the transport industry experience production efficiency gains that are at least as large as their suppliers. Hence, they register for taxes and pressure their input providers into doing the same. On the contrary, tax registration decreased by -8.2% in the non-ferrous and non-metallic minerals industry that experienced limited gains in production efficiency but large increase in tax enforcement at the intensive margin of tax evasion. Consequently, tax registration costs increased without any corresponding tightening of enforcement at the extensive margin or any mitigating boost in profitability following enhanced production efficiency.

The first source of heterogeneity across industries originates from the removal of the distorting retail sales tax in Indian states. With a model of inter-sectoral linkages similar to Hoseini (2015), we show theoretically that the cascading effect of sales tax increases with the backward linkages of an industry. Intuitively, first-point sales taxes are said to *cascade* down the chain value, because they are in practice charged at every - or at least several - step of the production process. Hence, they represent a larger share of prices at each step of the value chain. As a result, input price distortions introduced by a cascading tax increase with the length and the complexity of the value chain. Our theoretical contribution refines Friedlaender (1967)'s results by showing that distortions vary across manufacturers in the presence of a producer tax[3] when the production process involves more than two industries. In this context, VAT increases production efficiency when implemented in place of a distorting tax as it re-establishes pre-tax input prices. Although the theoretical and empirical literature on the relationship between VAT and informality is quite extensive (Boadway and Sato, 2009; Davies and Paz, 2011; Keen, 2008; Keen and Lockwood, 2010), it focuses on the relative properties of tariffs and VAT.[4] However, the VAT replaced a manufacturer tax or a cascading retail sales tax in several countries such as Canada in 1991, Mexico in 1980 or Bangladesh in 1981. We test our theoretical predictions empirically and find that the larger the number of steps leading to an industry's inputs, the larger the distortions faced by this industry. Indeed, the VAT reform induced relatively larger investment and production levels in industries where production efficiency increased more significantly due to the larger share of tax cascading embedded in their inputs under the former tax system.

The second source of heterogeneity across industries comes from the fact that the value of business-to-tax-registered-business transactions are known from tax authorities under the VAT, while it remains difficult to control transactions with final consumers. Without auditing firms, invoices provided by

[3] A producer or a manufacturer tax is a cascading sales tax when the production process involves more than two steps.

[4] In their empirical analysis, (Keen and Lockwood, 2010) consider the impact of a VAT reform dummy on the tax revenues-to-GDP ratio and thus do not differentiate between the specifics (rates, thresholds, previous tax) of the considered VAT reforms.

input suppliers and input purchasers can be cross-checked to detect fraud. We empirically test the theoretical predictions of Hoseini (2015) and find that the larger the share of an industry's output purchased by tax registered business customers, the tighter the VAT self-enforcement properties for this industry. For instance, tobacco manufacturers, which sell mainly to final consumers, are little affected by VAT self-enforcement properties. On the contrary, upstream industries like the non-ferrous metal industry producing alloys used in all electronic components, mainly sell to other firms, and are thus very sensitive to the tightening of enforcement introduced by the VAT. Our empirical results confirm that the implementation of the VAT created compliance costs that adversely impacted production and investments in industries with a large share of business customers. We show that this tightening of enforcement at the intensive margin results in lower registration rates in the industry, unless it is accompanied by production efficiency gains or by an increase in enforcement at the *extensive* margin of tax evasion.

Section 2.2 enters into the details of the tax system in India before and after the reform and underlines the production inefficiencies created by the sales tax that was replaced by a VAT. It also explains why the VAT reform provides us with an exogenous source of variations in the tax systems of Indian states. The theoretical contribution of this paper is presented in section 2.3. In subsection 2.3.1, we focus on explaining the most important concepts used in this paper: the measure of industries forward and backward linkages using the indices developed by Rasmussen (1957). In subsection 2.3.2, we provide an intuition for Hoseini (2015)'s results that the VAT self-enforcement properties are increasing with industries' forward linkages, as measured by Rasmussen (1957)'s forward linkages index. In subsection 2.3.3, we present the intuition behind the nexus between distortion and backward linkages under a distorting sales tax. In subsection 2.3.4, we show rigorously that the extent of distortions in production and investments under a cascading tax is increasing in the backward linkages of an industry. Data and methodology used to measure informality and linkages are described in section 2.4. Empirical strategy and results are explained in section 2.5.

2.2 The Indian VAT Reform

2.2.1 Prior to the Reform: States Sales Tax

Before its replacement by a VAT, the retail sales tax (ST) was the main source of tax revenue for Indian state governments. It explains the reluctance of the states to reform it in spite of its flaws being regularly underlined (Bagchi, 1994; Purohit, 1975, 1982, 1993). In particular, we refer to Purohit (1993) for a good description of the retail sales tax system in India and the distortions it created in the economy prior to the adoption of the VAT. The multiplicity of levies and the complexity of its structure are illustrated in figure A.9 of the appendix for the state of Madhya Pradesh. In spite of being widely referred to as a *retail* sales tax, ST in Indian states was a first-point sales tax. From an administrative point of view, it is less costly to collect a tax from large-scale manufacturers or wholesalers rather than from a myriad of retailers. From a political perspective, politicians may also prefer to place the tax burden on economic agents other than final consumers. Over the years, most Indian states have then gradually switched the point of levy of their sales tax to the first point of production. Some exemptions were possible under two general rules to differentiate between final and intermediate consumption. According to the *physical ingredient* rule, an input is a material or component that is physically incorporated in the goods destined to be sold. According to the *direct use* rule, items directly used in the production of goods are exempted. However, these rules are neither easy to apply nor guaranteeing the non-taxation of business inputs.[5] Consequently, the Indian states ST was a cascading tax, and Bagchi (1994) evaluated the share of sales tax revenues paid on inputs to be 30% on average across states[6].

2.2.2 Adoption of the VAT by Indian States

Starting with Haryana in 2003, the VAT was gradually adopted in Indian states (see a map of Indian states in figure A.1b), with 21 states in 2005, 5 states in 2006 and finally, Tamil Nadu and Pondicherry in 2007, and Uttar Pradesh in 2008. Map A.1a shows the different years of VAT adoption across states. Variations in the adoption of the VAT seem to have been driven by differences across states' elections schedules. Map A.2c pictures these variations caused by the accumulation of past interruptions in the five-year terms of the legislative assembly and are thus exogenous to our outcomes of interest[7]. Figures A.2a and A.2b illustrate the fact that electoral stakes were higher in 2003 than in 2005, except for Tamil Nadu, Gujarat, and Uttar Pradesh for which elections were

[5] For instance, Van Brederode (2009) gives the following example: lumber for manufacturing of furnitures will be exempted under the *physical ingredient* rule, while machines use to assemble the furnitures will be exempted only through the *direct use* rule. However, machines used to store the items in a warehouse will be subject to taxation under both rules. Due to the widespread adoption of the *physical ingredient* rule, fuel, tools, machinery and equipments were not considered as inputs in the majority of states.

[6] It is to be noted that this feature is due to the nature of the RST, rather than to some defaults of the Indian tax system, even if the complexity of the latter of course exacerbates the problem. The estimate RST revenue come from the taxation of business inputs are estimated 40% and 43% in the United States and Canada respectively (Ring, 1999; Smart and Bird, 2009)

[7] There was no changes in the planning of these elections during the period of interest in this paper, between 2003 and 2010, indicating that the reform did not impact the political game in that respect.

scheduled for 2006 and 2007 respectively.

Table 2.1: **Elections timing and VAT implementation.** States ruled by the Congress party, also in power at the central level, implemented the VAT reform after the elections for states legislative assemblies. States ruled by the opposition implemented it before elections to fuel anger against the Congress, their political opponent.

Reform implemented ...		
	BEFORE Elections	AFTER Elections
Opposition	4 (100%)	0 (0%)
Congress Party	3 (16.7%)	15 (83.3%)
Coalition	2 (50%)	2 (50%)

Traders associations engaged in strong lobbying against the reform. Newspapers[8] reported that this antipathy could only come from the decrease in evasion opportunities not only for sales taxes but also for turnover taxes that represents the VAT credit and refund system in comparison to the previous system. The VAT reform was also largely unpopular among consumers, because it was perceived as a regressive tax. As illustrated by table 2.1, parties in the opposition (to the central government) have an incentive to implement the reform before the elections for states legislative assemblies and blame their political opponents for it. Parties of the same color as the central government have an incentive to wait and implement the reform after the elections to avoid being associated to politicians in the central government. The Bharatiya Janata Party (BJP) took advantage of this widespread resistance of voters to the VAT by blaming its main political opponent - namely, the Congress party ruling the central government - for the reform's flaws (real or perceived). In reaction to this, the central government announced in 2003 that the VAT would be adopted only when all states will be prepared for it, which threw a general confusion on whether the central government would financially support the VAT implementation. Finally, the reform implementation was postponed until after the elections in 2005 and 2006 in most states. Tamil Nadu and Uttar Pradesh first delayed the reform because of the uncertainty created by electoral conflicts between the Congress and their political opponents in other states. In turn, they faced electoral tensions after 2006 which pushed the VAT implementation further back in time. The fact that political opponents to the Congress may have benefited from the unpopularity of the VAT reform does not imply that states having implemented the reform at different points in time are fundamentally different from one another. As shown by the maps in figure A.3, political volatility is high in India while informality is a persistent phenomenon. Thus, no political party would stay in power long enough to affect states institutions in such a determinant way that the VAT would have a radically different impact on our outcomes of interest.

The central government emitted detailed guidelines for states VAT Acts in order to foster harmonized

[8]See the article "The VAT Game" in *Economic and Political Weekly*, Vol. 38, No. 18 (May 3-9, 2003), pp. 1735-1736

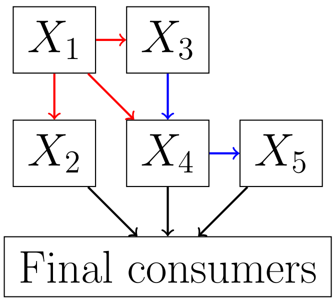
tax schedules across the country. The general tax rates were: 0% (or exempted) for agricultural product and basic necessities, 1% for precious metals, 4% (5% after 2009) for common consumption goods, and 12.5% for the rest. For special items like fuel and petrol, states were allowed to tax at 20% or more. Nevertheless, state governments retained some discretionary powers in setting the level of the mandatory registration threshold, penalties, and VAT exemptions. Mandatory registration turnover thresholds adopted in different states are shown in table A.2. Finally, the central government provided a technical and financial support to states for the modernization of their tax administration and the organisation of awareness campaigns. The central government's guarantee to provide for a computerised tax administration gives confidence in the fact that tax authorities indeed performed cross-checks of invoices provided by businesses. For example, in the state of Haryana, there is evidence of an audit team operating on the basis of risk management studies (Mukhopadhyay (2005)). Furthermore, all medium and big traders are audited at least once in five years. Both self-enforcement and input tax refunds are then expected to be uniformly present in Indian states after the reform.
The central government approved a compensation package for any revenue loss on account of VAT introduction at the rate of 100% of revenue loss during 2005-06, 75% during 2006-07, and 50% during 2007-08. The comparison of the maps presenting the variations in the VAT adoption timing in figure A.1a and the importance of sales tax revenues for states in figure A.4b shows that there is no correlation between the level of collected taxes and the timing of the reform. Thus, it does not seem that states with lower level of revenues implemented the reform in the view of obtaining a better bargain with the central government concerning compensations for potential losses.

2.3 Theoretical framework

In this section, we start by defining Rasmussen (1957)'s forward and backward linkages. In the next step, we discuss how the linkages indices are associated with the self-enforcing feature of the VAT and the cascading effect of sales tax. We give an intuition of why they are adapted to capture exogenous variations in the two main aspects of the Indian VAT reform across industries: first, the variations in the intensity of self-enforcement (forward looking index), and second, the variations in the extent of production efficiency gains (backward looking index). We then rigorously show that the extent of distortions in production and investments under a cascading tax is increasing with an industry's backward linkages.

2.3.1 Definition: Forward and Backward Linkages Indices

Intuitively, the forward linkages (FL thereafter) index of an industry measures the level of sales to other industries and the backward linkages (BL thereafter) index of an industry gauges the level of purchases from other industries. Figure 2.1 shows a simple example of direct and indirect forward and backward linkages. In order to quantify all direct and indirect linkages of an industry, following Hoseini (2015), we use the forward and backward linkages indices introduced by Rasmussen (1957).



$$\mathbf{A} = \begin{bmatrix} 0 & \alpha_{12} & \alpha_{13} & \alpha_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_{34} & 0 \\ 0 & 0 & 0 & 0 & \alpha_{45} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Figure 2.1: **Direct (red) and indirect (blue) forward linkages on industry 1 and the input-output table corresponding to this value chain.** Industry X_1 is directly linked with X_2, X_3, and X_4, and it is indirectly linked with industry X_4 (through X_3) and X_5 (through X_4). Industry 5 is backwardly linked to industry 4, industry 3, and industry 1. Industry 1 does not exhibit any backward linkages. α_{kn} denotes units of the output of industry k which is used as the intermediate input for industry n.

Definition of the FL index The FL index of a given industry n expresses the increase in output of industry n needed to cope for a 1 unit increase in the final demand for the output of all other industries including itself. In an economy with N different industries, the forward linkages index of industry n is given by:

$$fl_n = 1 + \underbrace{\sum_{k=1}^{N} \alpha_{nk} \frac{y_k}{y_n}}_{\text{direct forward linkages}} + \overbrace{\sum_{k=1}^{N} \sum_{j=1}^{N} \alpha_{nk} \alpha_{kj} \frac{y_j}{y_n}}^{\text{indirect forward linkages}} + \dots \tag{2.1}$$

where y_n denotes the value of total production of industry n, and α_{nk} required number of units of the output of industry n to produce one unit of good k.

The first term of the sum, 1, corresponds to the increase in output n necessary to satisfy the increase in its own final demand[9], while the other terms of the sum quantify the share of industry n's output that goes to other sectors $m = \{1, \dots, N; m \neq n\}$ to be used as an intermediate input so that these

[9]In our empirical analysis, we will use $\mathrm{FL}_n - 1$.

industries can satisfy the increase in their respective final demand[10]. The more forwardly linked an industry is, the more of its output is used as an input by other industries either directly (by this industry's direct customers) or indirectly (by the customers of this industry's customers).

Definition of the BL index In a similar way, the backward linkages index of an industry n gauges the amount of inputs that is required directly or indirectly from other industries for producing a single unit of good n. In an economy with N different industries, it is defined as:

$$bl_n = 1 + \underbrace{\sum_{k=1}^{N} \alpha_{kn}}_{\text{direct backward linkages}} + \overbrace{\sum_{k=1}^{N}\sum_{j=1}^{N} \alpha_{jk}\alpha_{kn}}^{\text{indirect backward linkages}} + \dots \tag{2.2}$$

where the same notation as above applies[11].

The important difference between fl_n and bl_n is that in the former n is the first subscript in all terms reflecting the output flow to other industries, while in the latter n is the last subscript reflecting the inputs requirement from other industries. The advantage of Rasmussen (1957)'s indices is that they can be computed from the Input-Output (I-O) tables of the Indian economy. In vector form, the two linkages indices can be written as:

$$\text{FL} = (\mathbf{I} + \mathbf{Y}^{-1}\mathbf{AY} + \mathbf{Y}^{-1}\mathbf{A^2Y} + \dots)J = \mathbf{Y}^{-1}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{Y}J \tag{2.3}$$

$$\text{BL} = (\mathbf{I} + \mathbf{A}' + \mathbf{A}'^{\mathbf{2}} + \dots)J = (\mathbf{I} - \mathbf{A}')^{-1}J \tag{2.4}$$

where FL and BL are vectors of fl_n and bl_n, $\mathbf{Y}$ is the diagonal matrix of total production value with diagonal elements y_n, $\mathbf{A}$ is the I-O matrix composed of $\mathbf{A}_{(i,j)} = \alpha_{ij}$ as elements, $\mathbf{I}$ is the identity matrix and J is a vector of ones (summation vector)[12]. In the simple example of Figure 2.1, the FL and BL indices of each industry are given in Table 2.2.

[10] Input requirements of each industry n is weighted by its by its relative size to industry ($\Delta y_k = y_k/y_n$) for the sake of comparability with the BL index. As we will discuss in section 2.3, because α_{kn} is the flow of inputs from k to n, non-increasing return to scale in production of n results in $\sum_k \alpha_{kn} < 1$. Therefore, all single summation of BL index lie between 0 and 1. To have similar pattern for FL, because $\sum_k \alpha_{nk} y_k \leq y_n$, we weigh each element in the summations by y_k/y_n.

[11] In our empirical analysis, we will use $\text{BL}_n - 1$.

[12] For more detailed discussion about the definition and construction of the linkages indices see Hoseini (2015).

Table 2.2: **FL and BL indices for the simple example of Figure 2.1**

Industry	Forward Linkages (FL)	Backward linkages (BL)
1	$FL_1 = 1 + \alpha_{12}(y_2/y_1) + \alpha_{13}(y_3/y_1) + (\alpha_{14} + \alpha_{13}\alpha_{34})(y_4/y_1) + (\alpha_{14}\alpha_{45} + \alpha_{13}\alpha_{34}\alpha_{45})(y_5/y_1)$	$BL_1 = 1$
2	$FL_2 = 1$	$BL_2 = 1 + \alpha_{12}$
3	$FL_3 = 1 + \alpha_{34}(y_4/y_3) + \alpha_{43}\alpha_{45}(y_5/y_3)$	$BL_3 = 1 + \alpha_{13}$
4	$FL_4 = 1 + \alpha_{45}(y_5/y_4)$	$BL_4 = 1 + \alpha_{14} + \alpha_{34} + \alpha_{13}\alpha_{34}$
5	$FL_5 = 1$	$BL_5 = 1 + \alpha_{45} + \alpha_{14}\alpha_{45} + \alpha_{34}\alpha_{45} + \alpha_{13}\alpha_{34}\alpha_{45}$

2.3.2 VAT self-enforcement and Forward Linkages: Intuition

VAT self-enforcement in a particular industry is conditional on third-party reporting, and thus conditional on the existence of tax registered business customers for this industry. The intensity of VAT self-enforcement is expected to increase with the number of these business customers and the share of total sales they represent. By considering an interlinked economy with formal and informal sectors, Hoseini (2015) highlights that the intensity of the self-enforcing properties of the VAT in a given industry is an increasing function of its FL index.

Stronger *direct* FL implies that a larger share of the industries' customers are businesses. Tax evasion is more difficult on business-to-formal-business transactions under the VAT, because formal buyers report the transaction to the tax administration in order to be refunded of the VAT paid on their input. Therefore, informal firms are reluctant to deliver invoice that could potentially lead to tax audits. A larger share of tax registered business customers increases the incentive to formally register under the VAT, because they require a formal invoice for every purchase. Therefore, it is straightforward that the enforcement properties of the VAT are tighter in industries with stronger direct forward linkages.

Moreover, stronger *indirect* FL implies that the business customers of an industry are themselves strong suppliers of other industries. The self-enforcing properties of the VAT are therefore relatively tougher on them, which increases their incentive to register formally. This in turn increases the incentive of their suppliers to register formally by the same argument as above. For example, in figure 2.1, the VAT increases enforcement in industry 4 through its direct forward linkages, assuming that at least some firms in industry 5 are formal. Therefore, a larger share of firms in industry 4 will register formally, which will increase enforcement in industry 1 as well.

Hence, tax compliance and tax registration are expected to increase relatively more in forwardly-linked industries following the introduction of a VAT with a credit and refund system. Finally, the government obtains more information on formal firms' true level of output and employment.

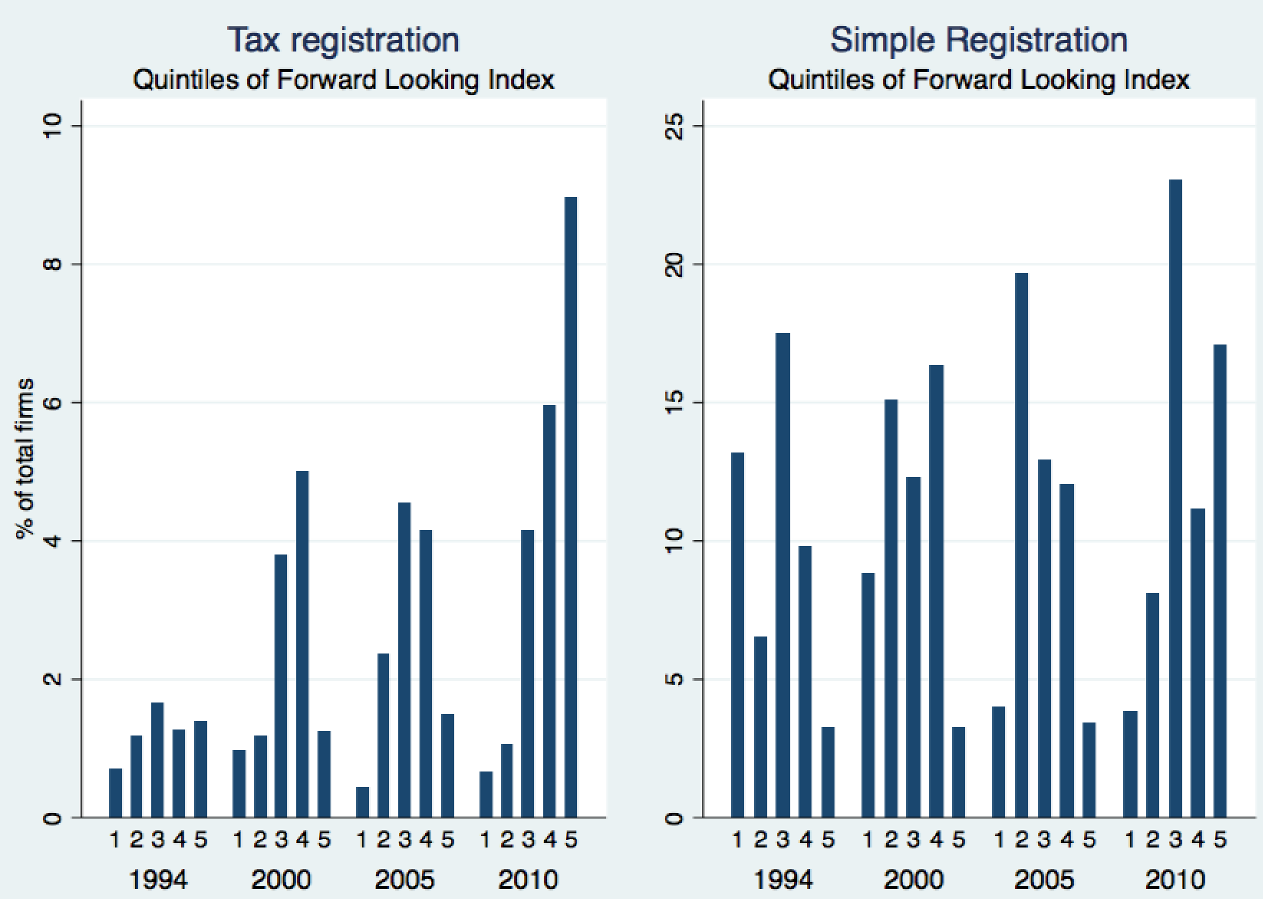


Figure 2.2: **Evolution of simple registration, and tax registration in India over the quintiles of the Forward Looking (FL) Index.** We see that the proportion of tax registered firms increased relatively more between 2005 and 2010 in the upper quintile of the FL index. The pattern is not as clear for the proportion of simple registration firms, suggesting a different mechanism at play, as explained in section 2.5.

Therefore, the VAT reform could also induce a larger compliance with other regulations such as labour laws.

2.3.3 Cascading tax and Backward Linkages: Intuition

Tax cascading interacts with variations in linkages and tax rates across industries to distort relative prices of business inputs. Consider again the simple economy of figure 2.1: in the absence of any tax, this economy can be represented by its input-output matrix. For simplicity, assume the economy is a Leontief type in the sense that all factor coefficients α_{ij} are constant and each industry n uses constant α_n units of labour at the numéraire wage in addition to the intermediate inputs to produce one unit of its final good. Then, by denoting $\tau = 1 + t$ as tax multiplier, the output prices are given in table 2.3. By comparing these prices with table 2.2, it becomes clear that distortions in prices $p_i^{\mathrm{ST}} - p_i^{\mathrm{VAT}}$ are increasing in the BL index of industry i due to the cascading tax. On the contrary, the VAT does not distort relative prices, and therefore does not affect relative input prices. In

Table 2.3: **Tax exclusive price of output under VAT and cascading ST in the simple economy of Figure 2.1. ($\tau = 1 + t$ is the tax multiplier.)**

Price	VAT	Cascading ST
p_1	α_1	α_1
p_2	$\alpha_2 + \alpha_1\alpha_{12}$	$\alpha_2 + \tau\alpha_1\alpha_{12}$
p_3	$\alpha_3 + \alpha_1\alpha_{13}$	$\alpha_3 + \tau\alpha_1\alpha_{13}$
p_4	$\alpha_4 + \alpha_1\alpha_{14} + \alpha_3\alpha_{34} + \alpha_1\alpha_{13}\alpha_{34}$	$\alpha_4 + \tau\alpha_1\alpha_{14} + \tau\alpha_3\alpha_{34} + \tau^2\alpha_1\alpha_{13}\alpha_{34}$
p_5	$\alpha_5 + \alpha_4\alpha_{45} + \alpha_1\alpha_{14}\alpha_{45} + \alpha_3\alpha_{34}\alpha_{45} + \alpha_1\alpha_{13}\alpha_{34}\alpha_{45}$	$\alpha_5 + \tau\alpha_4\alpha_{45} + \tau^2\alpha_1\alpha_{14}\alpha_{45} + \tau^2\alpha_3\alpha_{34}\alpha_{45} + \tau^3\alpha_1\alpha_{13}\alpha_{34}\alpha_{4}$

order to take this into account, in the next section, we model an input-output economy with flexible coefficients based on Hoseini (2015), and compare the distortions between a VAT and a cascading sales tax (ST).

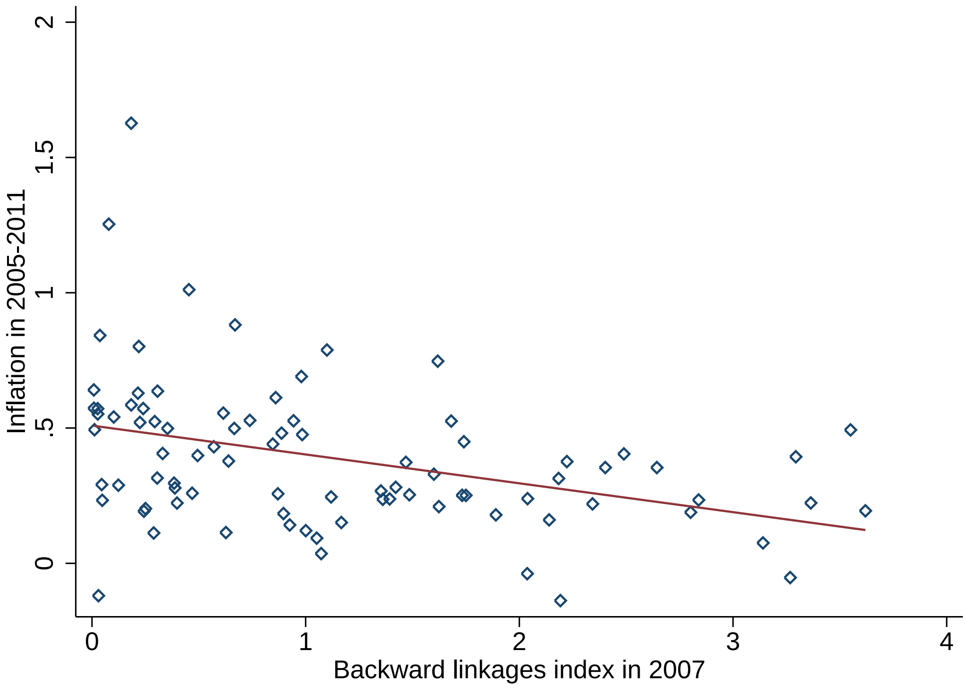


Figure 2.3: **Inflation versus backward linkages of different commodity groups.** The figure shows that the inflation rate before and after state-VAT adoption has had a lower pace in commodity groups with more backward linkages, suggesting the cascading nature of the former sales tax. The slope of the fitted line is equal to -0.17 and is significant at 1% level. The wholesale price index published by the Reserve Bank of India is used to compute the inflation rate, which is equal to the logarithm of the price index in 2011 subtracted by the logarithm of price index in 2005 (the earliest available year). The backward linkages index is computed using I-O tables. The list of commodity groups and the data is available in the Appendix table A.22.

2.3.4 Cascading tax and backward linkages: Generalization

This section aims to show formally that the distortions created by a cascading sales tax for an industry n increase with the BL index of this industry. Our model is based on the framework used by Hoseini (2015). He considers an interlinked economy under the VAT, comprising formal and informal sectors, and shows the variation in VAT self-enforcement across industries. In this paper, we consider the formal part[13] of the economy in the same framework and model production efficiency gains arising when replacing cascading sales tax by a VAT. Compared to Hoseini (2015), who addresses the impact of the self-enforcement properties of the VAT on informality, our model focuses on the differences in distortions between the VAT and cascading sales tax without accounting for the presence of an informal sector. Relative prices between different goods are unchanged under the VAT. In a cascading sales tax, however, relative prices are distorted in the presence of inter-industry linkages and the extent of this distortion is increasing in the BL index of the industry. Recall that the BL index of an industry expresses the degree of dependence of this industry on others: the number and the value of inputs it requires from the rest of the economy to produce 1 unit of its own.

Consider an economy with N competitive sectors. Each sector produces a homogeneous and unique good using intermediate products from other sectors and value-adding inputs (capital, labour, etc.). The production technology of sector n is a Cobb-Douglas as follows:

$$x_n = v_n^{\alpha_n} \prod_{m=1}^{N} x_{mn}^{\alpha_{mn}}, \quad \sum_{m=1}^{N} \alpha_{mn} + \alpha_n = 1 \tag{2.5}$$

where x_n is the production of sector n, v_n are the value-adding inputs,[14] and x_{mn} is the amount of product m as an intermediate input of n. By taking the input v as the numéraire, we can write the profit of a representative firm as:

$$\pi_n = p_n x_n - \sum_{m=1}^{N} p_m x_{mn} - v_n \tag{2.6}$$

where p_n is the price of product n. Regarding final consumption, we assume a Cobb-Douglas utility and exogenous wealth w for a representative final consumer, in the sense that, the utility maximization problem becomes:

$$\max u(x_1, \ldots, x_N) = \prod_{i=1}^{N} x_i^{\beta_i}, \qquad \text{s.t.} \sum_{i=1}^{N} p_i x_i = w \tag{2.7}$$

In this framework, by defining $y_n = p_n x_n$ as the value of total production in sector n, in Appendix B,

[13] In this section, 'formal' is the label of tax registered firm. Solving a model with both formal and informal firms is not mathematically possible, thus we only focus on formal firms which are the primary target of the indirect tax.

[14] By value-adding input, we mean all factors of production that the firm does not pay sales tax at the time of purchase (labour, tax-exempted capital, etc.). In a more general framework, one can differentiate between the value-adding factors by assuming $v = l^{\beta_1} k^{\beta_2} \ldots$, but this does not change our general results.

we show that the market equilibrium in industry n results in:[15]

$$y_n = w\beta_n + \sum_{m=1}^{N} \alpha_{nm} y_m \tag{2.8}$$

Equation (2.8) has a simple interpretation: In each sector, the total value of production is equal to the sum of final consumers demand ($w\beta_n$ is the net spending on product n in households' consumption bundle) and the intermediate demand ($\alpha_{nm} y_m$ is the value of sales from sector n to sector m). In this setting, in a sector with high degree of forward linkages, the intermediate demand term comprises the bulk of total production. In comparison, a sector with high level of backward linkages, appears in the right-hand side of the equation (2.8) for many other sectors. This simple framework enables us to introduce VAT and ST while we take into account the linkages between sectors.

2.3.5 Comparing VAT and cascading ST

Under the VAT and its credit and refund system, tax registered firms are refunded of the taxes they paid on their inputs, and remit to tax authorities the VAT collected on their sales. Thus, VAT is charged only on the added-value component of the good at each production step. The profits of a representative firm are given by:

$$\pi_n^{\text{VAT}} = (1+t)p_n x_n - \sum_{m=1}^{N}(1+t)p_m x_{mn} - v_n - \underbrace{tp_n x_n}_{\text{charged tax}} + \underbrace{t\sum_{m=1}^{N} p_m x_{mn}}_{\text{input refund}} = p_n x_n - \sum_{m=1}^{N} p_m x_{mn} - v_n \tag{2.9}$$

In this system, the net tax charged by the firm is equal to the net payment to the government and the optimization decision of the firm is the same as in equation (2.6). As this tax shifting goes along the production chain, only final consumers pay the tax at the end. Therefore, all above results for the case of no tax are also valid for the VAT system if final consumers' wealth is measured in net terms, i.e. w is replaced by $w/(1+t)$. In other words, the VAT is indeed a tax on final consumption and its only effect in the model is the decline in final consumer demand. This decline directly decreases the production of the taxed product and indirectly the intermediate demand of other products, in the sense that the total reduction is proportional to $1/(1+t)$ in all sectors.[16]

To model the cascading ST system, we assume that the tax is levied on all stages of production and

[15] For more details about the assumptions and derivations of the model see Hoseini (2015).

[16] This result is based on the assumption that all sectors are taxed at the same rate and final consumers' preference (β coefficients) are not affected by tax. see Appendix B for details.

no input refund is available[17]. Then the profit becomes:

$$\pi_n^{\mathrm{ST}} = p_n x_n - \sum_{m=1}^{N} (1+t) p_m x_{mn} - v_n \tag{2.10}$$

In contrast to the VAT, the profit is distorted under a cascading ST. This distortion is intensified at each stage of the production process and is maximized at the end of the production chain. The following proposition expresses the impact of backward linkages on price distortion in the ST relative to the VAT.

Proposition 1. *The price distortion of the ST in each industry is proportional to its backward linkages*

$$\ln p_n^{ST} - \ln p_n^{VAT} = \ln(1+t)\Big(\underbrace{1 - \alpha_n + \sum_{m=1}^{N}(1-\alpha_m)\alpha_{mn} + \sum_{m=1}^{N}\sum_{i=1}^{N}(1-\alpha_m)\alpha_{mj}\alpha_{jn} + \dots}_{\textit{weigthed backward linkage}} \Big)$$

Proof. Appendix B. □

To understand Proposition 1, assume that an upstream industry u is at the beginning of a production chain and has no backward linkages (e.g. a raw material supplier). For this industry $\alpha_u = 1$ and $\forall m, \alpha_{mu} = 0$, and we get $p_u^{ST}/p_u^{\mathrm{VAT}} = 1$. This means that u does not experience any extra distortion under ST. Next, consider a sector k whose only intermediate input is the output of industry u. For industry k, we have $\alpha_{uk} + \alpha_k = 1$ and Proposition 1 gives the price ratio as $p_k^{ST}/p_k^{\mathrm{VAT}} = (1+t)^{\alpha_{uk}}$. In other words, the price distortion in ST is proportional to the intermediate factor share in production (α_{uk}). Similarly, assume industry j uses the product of k as its only intermediate input ($\alpha_{kj} + \alpha_j = 1$), then the price distortion in ST becomes $p_j^{ST}/p_j^{\mathrm{VAT}} = (1+t)^{(1+\alpha_{uk})\alpha_{kj}}$. Here there are two sources of distortion, one is from the tax of purchasing input k, the other from the tax paid by k for purchasing u, which is partly burdens j. In a general framework, with several interlinked industry, Proposition 1 gives a reduced form formula for the level of distortion under ST burdened by each industry. When such distortion is removed from the economy – cascading ST is replaced by the VAT – we expect the use of intermediate and taxable capital goods to increase, and this increase to be larger in backwardly linked sectors. This increase comes from two channels: on the one hand, reduced input price increases firms production (income effect); on the other hand, compared to ST, taxable inputs

[17] This is an extreme case, but more tractable and has simpler mathematics. One can split the sectors into with and without cascading tax and follow the procedure of Appendix B by use a 2 × 2 decomposition of the factor coefficient matrix A. Still, the qualitative results of Proposition 1 holds with small modifications in backward linkage formula.

become cheaper relative to non-taxed inputs under the VAT, and the firm demands more taxable input (substitution effect).

In the next step, we derive the change in the output vector when the tax system evolves from a cascading sales tax to a VAT:

Proposition 2. *The output distortion of cascading ST in industry n is given by:*

$$y_n^{ST} - y_n^{VAT} = \frac{-tw}{1+t} ll_n$$

where ll_n is the length of linkages to final consumers defined as

$$ll_n = \underbrace{\sum_{k=1}^{N} \alpha_{nk}\beta_k}_{\text{direct linkages to final consumers}} + 2 \overbrace{\sum_{k=1}^{N}\sum_{i=1}^{N} \alpha_{nk}\alpha_{ki}\beta_i}^{\text{linkages through one intermediate good}} + 3 \underbrace{\sum_{k=1}^{N}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_{nk}\alpha_{ki}\alpha_{ij}\beta_j}_{\text{linkages through two intermediate goods}} + \dots \tag{2.11}$$

Proof. Appendix B. □

In Proposition 2, the right-hand side is a measurement of the length of linkages (LL henceforth) to final consumers. This index is similar to FL in terms of reflecting the flow of product to other sectors because n is the first subscript in each term. The difference between LL and FL is that LL weighs indirect linkages more than direct linkages and it highlights the length of linkages, in the sense that, the weight of a chain of production ending to final consumers in the LL index is proportional to its length. Proposition 2 suggests that under a cascading ST, the distortion in the value of output measured by y_n depends on the length of linkages of the sector n to final consumers reflected in LL. If industry n has no forward linkage and only households as costumers ($\forall i, \alpha_{ni} = 0$ and $x_n = x_{0n}$), one can show that in both VAT and ST, we have $y_n = w\beta_n$. For this industry, the distortion in price p_n is fully compensated by the reduction in final demand x_{0n} in the sense that y_n remains unchanged. But, when a sector also has business customers, the value of total output reduces in ST. The intuition behind this comes from the shifts in the use of inputs from taxable intermediate goods to non taxable factors such as labour in downstream firms, which as a result, reduces the demand for the intermediate product of the forwardly linked sectors with a long length of linkages. The rate of reduction in the value of output in ST depends on LL. From (B.1), we have $v_n = \alpha_n y_n$, which means that distortions in the amount of value-added created in each sector is proportional to the output distortion.

To summarize, we have three theoretical predictions for the effect of replacing a cascading ST with a VAT which lay the groundwork for our empirical strategy:

1. *less informality in forwardly linked industries*: This result is due to the self-enforcing property of the VAT which is modeled in Hoseini (2015).

2. *more reduction of price distortion in backwardly linked industries*: After replacing a cascading ST by the VAT, the use of intermediate and taxable capital goods increases, specially in backwardly linked sectors.

3. *more reduction of output distortion in industries with long length of linkages to final consumers*: when the cascading ST is replaced by the VAT, the value of output and value-added increases more in sectors with longer length of forward linkages to final consumers (high LL index).

2.4 Data

To identify the relationship between VAT adoption and informality in India, we will utilize sub-national data from different sources. The main data sources are 5 rounds of two nation-wide enterprise surveys in India, over the period 1989-2010[18], published by the Ministry of Statistics and Programme Implementation:[19] **Annual Survey of Industries** (ASI) and **Unorganized Manufacturing Enterprises** (UME). The ASI consists of manufacturing establishments registered under the Factories Act[20]. It is mandatory for all manufacturing firms employing at least 20 workers, or more than 10 workers while using industrial power, to be registered under this act. Each year, all establishments with more than 100 employees, and at least 12 percent of the rest, are surveyed with a representative sample at the state and 4-digit industry code level. The UME surveys are conducted by the National Sample Survey Organization (NSSO) and include all manufacturing units that are not covered by the ASI. In contrast, the UME survey is not conducted annually and its sample is representative at the state and village/town level. Its sample design is based on the latest Economic Census: a complete count of all economic units in India synchronized with the house listing operations of the Population Census.[21] Between 1989 and 2010, the UME survey was conducted in five

[18] Data are not expressed in calendar year but in fiscal year. Therefore, year "2005" should be read as "April 2005 through March 2006".

[19] These datasets are widely used in the literature. See for example Hsieh and Klenow (2009, 2014).

[20] The Factories Act of 1948 is a social legislation which has been enacted for occupational safety, health and welfare of workers at work places. It states that all manufacturing firms require a license to be delivered by the district Chief Inspector of Factories and certifying that the building in which its activities take place are conform to some health and safety standards. Furthermore, it also states the conditions under which the firm can employ labour.

[21] For the detailed information about the sample design of each survey, visit the official website of the Ministry of Statistics and Programme Implementation.

rounds: rounds 45 (1989-90), 51 (1994-95), 56 (2000-01), 62 (2005-06), and 67 (2009-10). Table 2.4 shows the sample size of the surveys. Both ASI and UME surveys contain a variety of information such as the type of activity carried on, ownership and registration status, as well as employment, fixed assets, production, revenues, costs and value-added. Using the ASI and UME surveys, we can then compute aggregate variables at the state-year-industry level: **production** is the value of total gross output of the industry in the given state and year, and **value-added** is the difference between the value of gross output and gross input. **Investment** is measured as the sum of new addition to fixed assets, expenses on repair and maintenance of fixed assets, and changes in current assets (inventory, cash in hand and bank) during the last year. Similarly, we compute total production, value-added and investment of firms registered under any act, registered under tax, and non-registered firms in each industry in a given state and year. In addition to the ASI and UME surveys, we use state-

Table 2.4: **List of surveys and Sample Size**

year	1989-90	1994-95	2000-01	2005-06	2010-11
No. observations ASI	49,323	57,908	37,055	49,637	46,843
No. observations NSS	123,321	192,029	222,529	80,637	99,243

and industry-level data from other sources. Data on the political cycle of Indian states (elections time and elected parties), and on their respective tax systems, come from the official websites of state governments. Data on the fiscal revenues of state governments come from the Indian Ministry of Finance website. Finally, the data used for the computation of the forward and backward looking indices are described in subsection 2.4.2. Appendix A.3 enters into the details of the data cleaning process, and presents additional information on the ASI and UME surveys.

Table 2.5: **Formal Registration and Sample size (ASI and NSS)**

year	Total	Above CENVAT threshold	Between VAT and CENVAT thresholds	Below VAT threshold
1989-90	168,015	27,688	16,506	123,821
1994-95	245,285	33,593	17,477	194,215
2000-01	260,687	17,153	35,099	208,435
2005-06	136,382	16,168	31,797	88,417
2010-11	150,420	22,217	29,666	98,537

2.4.1 Measuring informality: General registration and Tax registration

We construct the measures of informality based on firm registration status under various acts of Indian law. First, we count all firms in the ASI as formal since they are registered both under the Factories act and with tax authorities. Second, in the UME survey, firms are inquired whether they are registered under any act at all. In case of registration, they are asked to report the authority of registration.[22] Our first definition of formality is **general registration** which concerns all types of registration under any act or formal authority. Our second definition is **tax registration** that takes into account only registration under tax authorities. We define as **simple registration** the status of firms that are generally registered, but not tax registered. Although the identity of establishments in the ASI and UME surveys are strictly confidential, firms may pretend to be registered for taxes for fear of information leakages. Thus, in the UME survey, we determine the tax registration status of the firm by checking whether it has paid any sales tax or VAT (listed as distributive expenses) to the government during the last year. The advantage of using the general registration as a criteria for formality is that most of the listed formal authorities are designed to help SME growth by providing technological and financial assistance. Therefore, firms have no incentive to be dishonest about their general registration status on the contrary of their tax registration status. Sample size is shown in table 2.4. Our dependent variables are computed using the enterprise surveys and in order to compute them at industry, state, and year level, we collapse the firm-level observations using the sample weights provided in the survey. Overall, we have 32 industries, 33 states in 5 rounds of enterprise surveys. Some smaller states do not host all industries and the total number of observations is 1942.

Table 2.6: **Mean of measures of informality at the extensive margin in India**

Name	Discription	1989	1994	2000	2005	2010
Reg	Share of firms formally registered under any act	8.3	11.8	10.7	10.4	15.1
Treg	Share of firms formally registered with tax authorities		1.2	1.7	1.8	3.3
Ereg	Employment share of firms formally registered under any act	24.3	32.8	32.7	33.4	46.5
Etreg	Employment share of firms registered with tax authorities		20.3	19.1	21.3	32.7

Table 2.6 shows us that there was an increase in formal registration in the period of interest in this paper. Nevertheless, a large majority of Indian firms is still informal: around 10-15% for general registration and as low as 1-3% for tax registration between 1989 and 2010. Moreover, formal

[22] There are at least ten listed governmental or semi-governmental agencies under which a firm can be registered. Some examples are District Industries Centre for Small Scale Industry (SSI), Khadi and Village Industries Commission, Council for Technological Up-gradation, Small Industries Development bank, Development Commissioners of handicraft/handloom/jute, Coir/Silk board, Panchayat and other local bodies.

firms are at the origin of most of the employed workforce: around half (46.5%) in 2010 for general registration and almost a third (32.7%) for tax registration. Table A.6 summarizes these measures of informality at the industry level. Not surprisingly, we find higher levels of informality in industries where sub-contracting is widespread (the textile or tobacco industries for example). On the contrary, industries where there is a strong control by the states (rail or mining industries for example) exhibit higher levels of formality.

2.4.2 Linkages indices

For estimating the forward and backward linkages in the Indian economy, we use the I-O tables of Indian economy published by Central Statistical Organization (CSO). The I-O tables are available online for 1993, 1998, 2003 and 2007 and include the "use and make" matrices[23] and the associated (commodity $\times$ commodity) and Leontief matrices which provide us with the elements for computing the linkages indices in equation (2.3). Time-varying linkages indices enable us to capture structural changes, such as technological innovations, taking place in each industry over time. To fit the timing of the I-O tables with the surveys, we use the years with minimum time distance. Specifically, linkages indices of 1989, 1993, 1998, 2003, and 2007 are respectively matched with survey years 1989, 1994, 2000, 2005, and 2010.

The 1993 and 1998 tables consist of 115 sectors, but the 2003 and 2007 tables have 130 listed sectors. To build a consistent classification over time, we aggregate the sectors in 60 divisions based on the classification Dholakia et al. (2009).[24] The methodology of aggregation is described in Appendix A.3. Using the aggregated matrices $\mathbf{A}$ and $\mathbf{Y}$, we computed the FL and BL indices using equations (2.3) and (2.4). The matrix $\mathbf{A}$ of 1989 adjusted with prices of 1993 is available in Dholakia et al. (2009) and by assuming a similar production matrix $\mathbf{Y}$ for 1989-90 and 1993-94, we estimate the linkages indices for 1989. The sectoral divisions and their respective linkages indices over time are presented in Table A.23.

[23] Although, in general, use (commodity $\times$ industry) and make (industry $\times$ commodity) matrices are not necessarily square and one industry can produce more than one commodity, the classification of I-O tables published by CSO are based on a one-to-one relationship between commodities and industries and all "use and make" matrices are square. For more information see the CSO's manual of Input-Output Transaction Tables: http://mospi.nic.in/Mospi_New/upload/report&publication/ftest10/appendix202.pdf

[24] Dholakia et al. (2009) provide the price-adjusted (commodity $\times$ commodity) matrix ($\mathbf{A}$ in equation 2.3) of the I-O tables of India between 1968 and 2003 for 60 sectors. However, because they do not report total production ($\mathbf{Y}$ in equation 2.3), we cannot estimated the linkages indices using their data.

2.5 Empirical Results

2.5.1 Empirical Model

To estimate the heterogeneous impact across industries of the replacement of a distorting sales tax with the VAT on informality, production and investments in Indian states, we estimate the following differences-in-differences model:

$$g_{ist} = \alpha + \beta_0 \text{VAT}_{st} + \beta_1 \text{VAT}_{st} \times \bar{FL}_{i(t-1)} + \beta_2 \text{VAT}_{st)} \times \bar{BL}_{i(t-1)} + \\ \beta_3 \text{VAT}_{st)} \times \bar{BL}_{i(t-1)} \times \bar{FL}_{i(t-1)} \\ + \boldsymbol{\beta^T} X_{ist} + a_i + b_s + c_t + \varepsilon_{ist} \quad (2.12)$$

where g_{ist} is one of our dependent variables (production, investments, or the proportion of firms that are tax registered, simply registered and non registered[25]) in industry i, state s and time t as described in section 2.4. VAT_{st} is a reform dummy taking the value 1 when the sales tax has been replaced by the VAT in state s and time t. $\bar{FL}_{i(t-1)}$ and $\bar{BL}_{i(t-1)}$[26] are respectively the forward and backward linkages indices of industry i at time $t-1$ centered around the minimum value respectively taken by $\text{FL}_{i(t-1)}$ and $\text{BL}_{i(t-1)}$. Centering the indices allows for an easier interpretation of the coefficients. First, we capture variations in tax enforcement across industries by interacting the VAT reform dummy with the centered FL index that measures the importance of business-to-business transactions for each industry i at time $t-1$. Second, we capture differences in production efficiency gains across industries with the interaction term between the VAT reform dummy and the centered BL index that represents the length and complexity of the production chain leading to industry i's inputs at time t. Finally, we interact the VAT reform dummy with $\bar{BL}_{it} \times \bar{FL}_{it}$, to investigate whether the combined impact of self-enforcement and production efficiency gains differs from the sum of their respective impact if there were considered in isolation. Figure A.8 shows that the distributions of BL_{it} and FL_{it} across Indian industries are skewed[27]. For this reason, and to reduce the effects of outliers on our results, we also estimate this model using the logarithm of the forward and backward linkages indices. Results are reported in table A.10 and table A.11.

[25] An OLS estimation does not account for the fact that some of our dependent variables are proportions, and are therefore bounded between 0 and 1. Furthermore, the proportions of tax-registered, simply registered and non-registered firms must sum to 1. As a robustness check, we estimate by maximum likelihood a Dirichlet distribution, the multivariate generalization of the beta distribution, and find qualitatively similar results.

[26] The Spearman correlation coefficient between the FL and BL indices is 0.15, and the Pearson correlation coefficient is 0.11. None of these coefficient is significant at the 5% significance level.

[27] In 2000, the skewness (measured by the difference between the mean and the median) is -1.023 for the distribution of BL_{it}, and 2.44 for the distribution of FL_{it}. Tests of skewness reject that both variables are normally distributed.

Following the existent literature, the vector of control variables X_{ist} includes different variables that impact the extensive measure of informality. $log(\text{SDP})_{st}$ is the logarithm of state domestic product (SDP), $log(\text{Credit/SDP}))_{st}$ is the logarithm of the ratio of available credit to SDP, RZ_i is the Rajan-Zingales index measuring the financial dependence of each industry i. We also include the ratio of tax expenditures to SDP and different measures of labour enforcement like the number of inspections per 1000 firms in state s at time t, or the number of inspectors per 10.000 firms in state s at time t. $\bar{FL}_{i(t-1)}$, $\bar{BL}_{i(t-1)}$, and $\bar{BL}_{i(t-1)} \times \bar{FL}_{i(t-1)}$ are also included in the vector of control variables. Industry and state fixed effects, respectively a_i and b_s, are included to control for all time invariant industry and state characteristics that can potentially affect the dependent variables. Additionally, year fixed effects c_t are included to control for year-specific effects common across all states and industries. Standard errors are clustered at the state level.

$\boldsymbol{\beta^T}$ is the transpose of the vector of coefficients associated with these control variables. β_0 represents the main effect of the VAT reform on firms in an industry where forward and backward linkages are fixed at the minimum respective values taken by the FL and BL indices in the economy at time $t-1$. Hence, β_0 is the impact of the VAT reform through the change from a distorting sales tax to the VAT on an industry that experienced the lowest possible production efficiency gains and changes in tax enforcement. It represents the main impact of the VAT in the economy: it captures the differences in the ST and the VAT rate schedules and the subsequent changes in demand. β_1 represents the additional impact of the VAT self-enforcement properties on the dependent variable abstracting from any production efficiency gains simultaneously created by the reform. β_2 represents the additional impact of the production efficiency gains caused by the removal of the sales tax on the dependent variable in the absence of tax enforcement. Finally, β_3 will capture the incremental effect resulting from the combination of production efficiency gains and tighter enforcement. If β_3 is positive, there are synergies between an increase in self-enforcement and production efficiency gains: the combined impact of these two effects is larger than the sum of both effects taken in isolation. If β_3 is negative, tighter self-enforcement and production efficiency gains are mitigating each other's respective impact: the sum of both effects taken in isolation is larger than their combined impact.

2.5.2 Indian VAT Reform and Informality

In this subsection, we investigate the impact of the VAT reform on tax compliance in Indian states. We disentangle different effects of the VAT reform: i) the increase in production efficiency gains that could trigger larger tax compliance through improved profitability or greater scale of production,

ii) the increase in tax enforcement at the intensive margin through the VAT credit and refund system, and iii) the increase in tax enforcement at the extensive margin through the VAT credit and refund system, and finally iv) the changes in final demand caused by the differences in the former and the new tax rate schedules. In the presence of the VAT mandatory registration threshold (MRT), tax compliance is a broader concept than tax registration. Indeed, simply registered firms are tax compliant if they maintain their turnover below the MRT. Thus, we measure the changes in the distribution of firms across the three possible compliance profiles: tax registration, simple registration and non-registration. Table 2.5.3 summarizes the results explained in this section.

Impact of the VAT reform on Tax Registration Column (3) of table A.8 shows a relatively larger positive impact of the VAT reform on tax registration in industries with larger backward linkages. Therefore, relatively larger production efficiency gains created in those industries induced firms to register for taxes. Indeed, tax registered firms are refunded of VAT paid on their inputs (*direct* decrease in distortions). On the contrary of non-registered firms, they can also trade with other tax registered firms without restrictions. These suppliers experience production efficiency gains themselves, and may pass then on to their customers (*indirect* decrease in distortions). This would be consistent with the idea that many firms are *forced* into informality, because they cannot afford registration and tax compliance costs. The VAT would induce higher tax registration rates by increasing firms profitability. Column (3) of table A.8 also shows a non-significant coefficient for the interaction term VAT $\times$ $\bar{FL}$. This is not surprising, because self-enforcing VAT properties could impact incentives to register for taxes in two opposite ways. On the one hand, the VAT credit and refund system allows tax authorities to know the true amount of business-to-business transactions, which implies lower evasion opportunities at the intensive margin. It potentially lowers firm profitability, and increases the costs of tax registration. On the other hand, firms with a large share of business customers would be relatively more induced to register for taxes. First, production efficiency gains caused an increase in production and investments in the rest of the economy, and by definition forward looking industries will see an increase in the demand for their products. Second, backward looking business customers that are tax registered prefer other tax registered suppliers in order to obtain VAT credits on from the purchase of their inputs. A large share of tax registered customers within a large share of customers thus represents an increase in tax enforcement at the extensive margin. The extensive margin of tax evasion would thus be better enforced under a VAT, because firms need to register for taxes in order to retain their tax registered business customers.

The interaction term VAT $\times \bar{FL} \times \bar{BL}$ is added to the model in order to disentangle between these two different mechanisms. Indeed, as explained above, the VAT credit and refund system increases incentives to register for taxes when industries exhibit a large share of tax registered customers, and not only of business customers. Hence, we need to differentiate between industries whose business customers have a larger incentive to register for taxes and the others. Since production efficiency gains are larger in industries with higher backward linkages, and that customers of those industry must be at least as backward looking as them, it follows that industries with both high forward and backward linkages are those with a larger share of tax registered business customers. Column (4) of table A.8 accounts for potential synergies between higher tax enforcement and production efficiency gains. First, the VAT reform led to relatively larger decrease in tax registration in industries with higher forward linkages, that is to say industries with larger share of business customers, but not necessarily a larger share of tax registered customers. Higher enforcement at the intensive margin leads to a significant decrease in tax registration rates. This result indicates low enforcement of the extensive margin of tax evasion (tax registration) outside of the VAT credit and refund system. Thus, firms can easily avoid tax registration, if they cannot or do not want to afford tax compliance costs. Second, we see that production efficiency gains contributed to increase tax registration only when combined with tighter tax enforcement. Without changes in enforcement, the decrease in distortions caused by the removal of distorting sales tax did not significantly increase tax registration. However, the VAT credit and refund system increases relatively more tax enforcement at the extensive margin in industries whose business customers experienced high production efficiency gains themselves. Indeed, those customers are then more likely to register for taxes and to favor tax registered suppliers. Our results imply that the increase of tax enforcement at the intensive margin of business-to-business transactions introduced by a VAT in an economy is not sufficient in itself, but must be combined with an increase in enforcement at the extensive margin. Positive incentives to register for taxes are effective only if they are combined with negative incentives such as the self-enforcement properties of the VAT. It is consistent with the idea that informal firms can also be *opportunistic*, and avoid formal compliance costs to increase their profitability. Thus, the self-enforcing properties of the VAT do not necessarily imply that a government can raise larger tax revenues with lower level of expenditures on tax enforcement. This was previously pointed out as a common caveat of VAT reforms in developing countries.

Once we control for synergies between the self-enforcing and the efficiency aspects of the VAT reform, the main effect of the VAT reform on tax registration - abstracting from self-enforcing and production efficiency aspects - switches from being negative and significant to being small, positive and non significant. As explained in subsection 2.5.1, the main effect of the VAT reform captures the changes in the tax schedule, as highlighted in figure A.9, caused by the replacement of a first-point sales tax by a VAT. First, the effect on final demand depends on the difference between the former and new tax rate schedules. Second, final consumers' awareness of the share of taxes included in commodities prices could have increased following the change from a first-point sales tax to a value-added tax. In column (3) of table A.8, where we implicitly assume perfect enforcement at the extensive margin of tax registration, the negative significant coefficient is consistent with the *forced* non-compliance hypothesis. Higher tax salience or smaller tax pass-through diverts consumers from tax registered firms, which lowers their profitability and prevents them from registering. However, once we remove this assumption in column (4), the changes in the rate schedule does not play a role anymore in tax registration rates. In consistency with the *opportunistic* non-compliance hypothesis, firms deciding whether to register for taxes or not are sufficiently profitable to absorb negative shock on their margin or in the demand for their product.

To conclude, figure 2.4 illustrates the heterogeneous impact of the VAT reform on tax registration across industries as a function of their forward and backward linkages. On the one hand, we see that, in industries in which both production efficiency and tax enforcement increased, tax registration grew up to +4.9% (non-ferrous metal industry). This number is relatively large, because the standard deviation of the share of tax registered firms (**Treg**) subtracted by its averages in each state, year, and industry is 2.02. On the other hand, the VAT reform also caused a decrease in tax registration as large as −8.1% in industries where only the self-enforcement properties of the VAT have played a role, and tax enforcement became tighter without any production efficiency gains to compensate for it (other minerals industry) or without additional incentive to register at the extensive margin to attract tax register customers. Figure A.7 shows the share of added-value respectively created by each Indian industry before the VAT reform in 2000. The bulk of created added-value in Indian states comes from industries in which the impact of the VAT reform was moderate: either slightly negative (in the textiles and the food industry that respectively represent 8.1% and 10.4% of created added-value) or slightly positive (in the jute and the pesticides industry that respectively represent 8% and 10.7% of created added-value). A large share of the added-value in the economy thus remained as difficult to tax as before the implementation of the VAT. Furthermore, figure A.6 illustrates that the employment shares of industries in which the VAT reform caused a decrease in tax registration were significant in 2000: 7.2% of total employment in the non-metallic minerals industry, and 8.8%

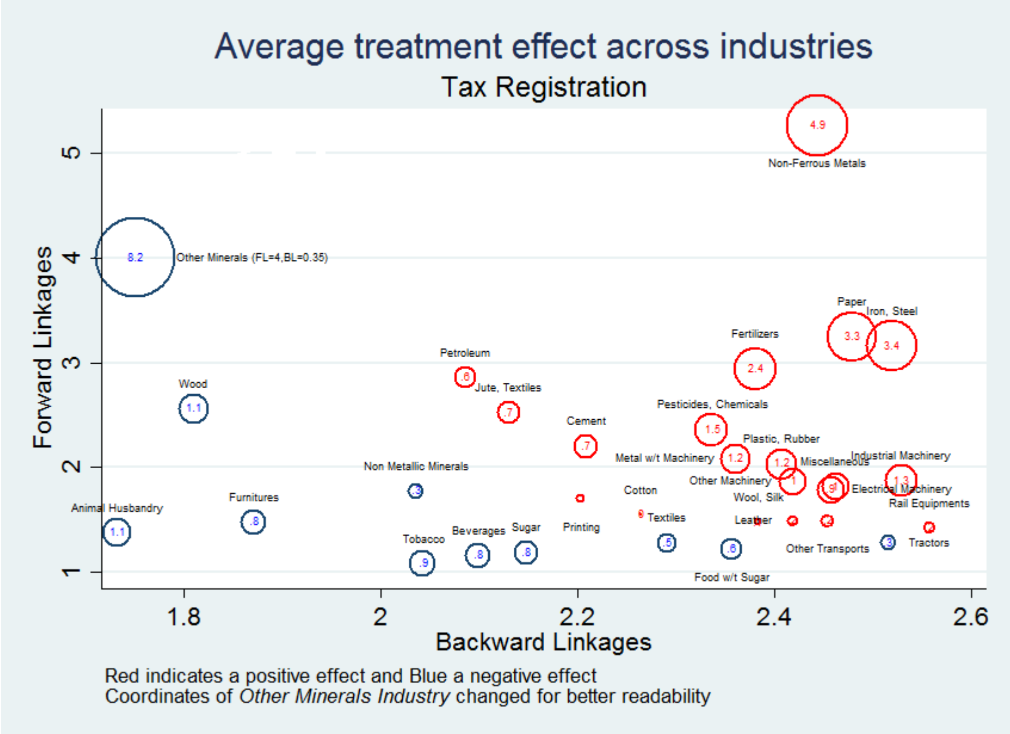


Figure 2.4: **Heterogeneous impact of the VAT reform on tax registration in Indian states across industries.** Each bubble represents an industry i and has coordinates $(\text{FL}_{i2000}, \text{BL}_{i2000})$. The size of each bubble is proportional to the average treatment effect for industry i across Indian states.

in the tobacco industry. At a later stage of this paper, we will investigate the consequences of these variations in tax enforcement on compliance with labour regulations and workers registration rates in the social security system.

Impact of the VAT reform on Simple Registration Column (1) and (2) of table A.8 show that the effects of the VAT reform on the share of simply registered firms differ from its effects on tax registration. First, we find a relatively larger increase in simple registration in industries with higher backward linkages *even after controlling* for potential synergies between the self-enforcing and the production efficiency aspects of the VAT reform. It can sound surprising, because simply registered firms are not directly part of the VAT credit and refund system: they do not experiment an increase in enforcement at the intensive margin of tax evasion. However, being compliant if their turnover remains under the VAT mandatory registration threshold, they do not face constraints in trading with tax registered firms. Production efficiency gains are thus *indirectly* accessible to simply registered firms, because they benefit of lower input prices if they trade with tax registered firms that are refunded of taxes paid on their inputs. They can then benefit *indirectly* from the decrease in distortions caused by the removal of a distorting sales tax in the economy. Furthermore, table A.10 and table A.11 indicate a shift from non-registration to simple registration in industries that benefited from relatively larger production efficiency gains: non-registration decreased more in industries with higher backward linkages, while simple registration increased simultaneously relatively more in those same industries. It is consistent with the idea that informal firms can also be *forced* out of the formal economy, if formal costs are too important. Formal participation rates improve once barriers to formality are lowered by the implementation of the VAT reform. Table A.6 indicates that the interquartile of BL is 0.32. From column (3) of table A.10, the production efficiency gains induced by the removal of a distorting sales tax would cause a 17.644*0.32=5.64% larger increase in simple registration in the metal industry at the 75th percentile of BL than in the beverages industry at the 25th percentile.

Second, in table A.8 table A.10, and table A.11, we find a switch from tax and non-registration to simple registration in industries with relatively larger share of business customers. On the one hand, higher demand from producers in the economy triggers larger opportunities for their suppliers and leads to higher formal participation (lower *forced* formality[28] in the form of lower non-registration). On the other hand, tax enforcement at the intensive margin has increased, deterring firms to register for taxes (higher *opportunistic* non-compliance in the form of lower tax registration). From the same tables, we see that the increase in tax registration in industries exhibiting both higher forward and backward linkages originates from a decrease in simple registration, and not from a decrease in non-registration. As explained above, the VAT makes tax registered suppliers more attractive for other tax registered business customers in industries where production efficiency gains are the largest. Thus, simply registered firms with both high forward and backward linkages become less competitive with respect to their tax registered counterparts.

Finally, the main effect of the VAT reform - abstracting from self-enforcing and production efficiency aspects - caused a shift from simple to non registration.

2.5.3 Indian VAT reform, Production and Investments

Our main theoretical result, presented in Proposition 9, predicts that replacing a cascading sales tax with a VAT reduces inputs price distortions. Furthermore, the size of this production efficiency gain is increasing in the share of added-value contained in an industry input, that is to say increasing with

[28] Note that the estimation of the model in equation 2.12 does not allow us to conclude on whether this increase in formality corresponds to an increase in compliance as the compliance status of simply registered firms depends on their turnover. We look at this issue in subsection 2.5.4.

the backward linkages of an industry. Table 2.5.3 summarizes the results explained in this section. Concordantly, the coefficient associated to the interaction term VATxBL is positive in columns (1) to (4) in tables A.12 and table A.15. Industries that experienced larger production gains invested and produced relatively more following the VAT reform.

The negative coefficient associated to the interaction term VATxFL in columns (1) and (2) of table A.14 and table A.16 indicates that the VAT reform had a relatively larger negative impact on total investment and production in industries with larger forward linkages. On the one hand, firms with a large share of business customers would be relatively more induced to register for taxes, when tax registered business customers prefer suppliers that are also tax registered in order to obtain VAT credits. The extensive margin of tax evasion would thus be better enforced under a VAT, because firms need to register for taxes in order to retain their tax registered business customers. On the other hand, the VAT credit and refund system allows tax authorities to know the true amount of business-to-business transactions, which implies lower evasion opportunities at the intensive margin, and potentially lower firm profitability.

The interaction term VAT $\times$ FL $\times$ BL is added to the model in order to disentangle between these two different mechanisms. Columns (3) and (4) of table A.14 and table A.16 account for potential synergies between higher enforcement and higher production efficiency. First, as expected, industries experiencing larger production efficiency gains invest and produce relatively more after the VAT reform. Investment goods become cheaper, and profitability is also increased by lower input prices. Second, investment and production also increased relatively more in industries with a larger share of business customers. Indeed, production efficiency gains increase demand from the rest of the economy. Forward-looking forms are thus indirectly benefiting from them. Third, if those industries simultaneously rely on a large share of business customers, tax registration costs increased relatively more in industries where production efficiency gains and the share of business customers are both relatively higher, because it means that a large share of these customers would be tax registered and would pressure their suppliers to register for taxes too. It explains why the relatively larger production increased caused by the VAT reform in those industries was not as important in industries that also exhibited higher forward linkages.

These results are consistent with those presented in subsection 2.5.2, where we show that tax registration - and the costs associated to it - increased relatively more in industries where production efficiency gains and the share of business customers are high, because it means that a large share of these customers would be tax registered and would pressure their suppliers to register for taxes too. It explains why the relatively larger production increase caused by the VAT reform in industries with higher backward linkages was not as important in industries that also exhibited higher forward linkages.

	Registration			
	Tax Act	Other Act	None	Investment
VAT	n.s.	- -	+++	- -
VAT × FL	-	++	-	+++
VAT × BL	n.s.	++	- -	++
VAT × FL × BL	+++	- -	n.s.	-

+ + + or − − − , ++ or −−, and +or − indicate significance at 1%-level, 5%-level and 10%-level, respectively, as well as the sign of the coefficient. n.s. indicates a non-significant coefficient. Standard errors are clustered at the state level.

Table 2.7: Summary of the results of the Diff-in-diff estimations for the effects of the VAT reform on tax registration, simple registration, non-registration and investments

2.5.4 Heterogeneous Impact of the VAT Reform across Industries and Above/Below the VAT Mandatory Registration Threshold (MRT)

In subsection 2.5.2, we analysed the effects of the VAT reform on the distribution of firms across the different registration status (tax, simple, and non-registration). However, the model in equation 2.12 does not allow us to conclude on the consequences of the effects of the reform on compliance because it does not differentiate between non-tax registration below (compliant firms) and above (non compliant firms) the MRT. The extent of non-compliance in the economy is important because it can allow for a better enforcement strategy of the government and a more accurate evaluation of the "real" tax base in the economy and of the distortions created by taxing only a part of it. Furthermore, non-compliant firms will be reluctant to engage into activities or to respond to incentives designed by the government to support economic growth, in particular in small and medium enterprises. In this subsection, our contribution is to highlight how the extent and the drivers of non-compliance vary in firms above and below the threshold. Forced non-compliance calls for improvements in the business environment and reduction in the tax burden carried by formal firms. On the contrary, opportunistic informality must be curbed by improving enforcement and increasing penalties.

In order to estimate the heterogeneous effect of the VAT reform above and below the mandatory registration threshold *and* across industries, we estimate a triple differences model as described in equation 2.13.

$$\begin{aligned} g_{ist}^{above/below} = \alpha + \beta_0 \text{VAT}_{st} + \beta_1 \text{VAT}_{ist}^{above} + \beta_2 \text{VAT}_{st} \times \text{FL}_{i(t-1)} + \beta_3 \text{VAT}_{st} \times \text{BL}_{i(t-1)} + \\ \beta_4 \text{VAT}_{ist}^{above} \times \text{FL}_{i(t-1)} + \beta_5 \text{VAT}_{ist}^{above} \times \text{BL}_{i(t-1)} + \\ + \beta_6 \text{VAT}_{st} \times \text{FL}_{i(t-1)} + \times \text{BL}_{i(t-1)} + \beta_7 \text{VAT}_{ist}^{above} \times \text{FL}_{i(t-1)} + \times \text{BL}_{i(t-1)} \\ + \boldsymbol{\beta^T} X_{ist} + a_i + b_s + c_t + a_i \times \text{AboveMRT}_{ist} + b_s \times \text{AboveMRT}_{ist} + c_t \times \text{AboveMRT}_{ist} + \varepsilon_{ist} \end{aligned} \tag{2.13}$$

where g_{ist} is one of our dependent variables in industry i, state s and time t but computed for the sample of firms above and below the MRT[29] respectively. AboveMRT_{ist} is a dummy variable equal to 1 when we are considering the sample of firms above the MRT in industry i, state s and time t. $\text{VAT}_{ist}^{above} = \text{VAT}_{st} \times \text{AboveMRT}_{ist}$ is then a newly defined treatment dummy that takes the value 1 when the cascading sales tax has been replaced by the VAT in state s and time t and we are considering the sample of firms above the MRT. This new treatment dummy accounts for the fact that the VAT reform will impact differently firms for which tax registration is mandatory (above the threshold) and firms for which simple registration only is mandatory (below the threshold). In addition to state, industry and year fixed effects, the model provides full non-parametric control for state-specific time effects that are common across firms above the MRT: $b_s \times \text{AboveMRT}_{ist}$, time-varying effects above the MRT: $c_t \times \text{AboveMRT}_{ist}$, and industry-specific effects above the MRT: $a_i \times \text{AboveMRT}_{ist}$. $\text{BL}_{it} \times \text{AboveMRT}$ and $\text{FL}_{it} \times \text{AboveMRT}$ are also included in the set of control variables. Standard errors are clustered at the state level. The interpretation of coefficients in this regression is made explicit in table 2.8.

Results are reported in table A.17, table A.20, and table A.18. We discuss first the effects of the VAT reform above the MRT when tax registration is mandatory and all other registration status imply non-compliance of firms.

Columns (3) and (7) in table A.17, along with column (3) in table A.20, show that the increase in enforcement at the intensive margin followed by the VAT reform caused a relatively larger shift from tax to simple and non-registration in industries with relatively larger share of business customers. In other words, tighter enforcement decrease tax compliance above the VAT registration threshold in the absence of production efficiency gains. Nevertheless, in industries with relatively larger share of

[29] We compute firms turnover from the balance sheet information provided in the ASI and NSS surveys. Total turnover is defined as the total value of manufacturing production (sum of value of products and by-products) plus other receipts for non-manufacturing activities.

Coefficient	Interpretation
β_0	Main effect of the VAT reform on the sample of firms below the MRT due to changes from the ST to the VAT rate schedule
β_1	Deviation from the main effect of the VAT on the sample of firms above the MRT due to changes from the ST to the VAT rate
Additional deviations from the main effect of the VAT	
β_2	**VAT self-enforcement properties at the intensive margin** Effect of removing evasion on transactions between tax registered firms across variations in FL
β_3	**Production efficiency gains** Effect of removing cascading in the value chain across variations in BL
β_6	**VAT self-enforcement properties at the extensive margin** across variations in BL × FL
Deviations specific to firms above the MRT	
β_4	**VAT self-enforcement properties at the intensive margin** Effect of removing evasion on transactions between tax registered firms across variations in FL
β_5	**Production efficiency gains** Effect of removing cascading in the value chain across variations in BL
β_7	**VAT self-enforcement properties at the extensive margin** across variations in BL × FL

Table 2.8: **Interpretation of the coefficients of the triple differences model in equation 2.13.** The **main effect of the VAT reform** corresponds to the effect of the VAT on the sample of firms below the registration threshold in an industry where both FL and BL would be equal to 0. In other words, an industry with no business customers and that does not purchase any input. Such an industry could be thought of as the subsistence sector that is affected only through the shift in demand from final consumers following changes in the rate schedule.

tax registered business customers, this negative shift in tax compliance was largely compensated for by a reverse shift in tax registration/compliance. This is consistent with the mechanism explained in subsection 2.5.2: the VAT exhibits self-enforcement properties at the extensive margin only in industries with a large business customers and significant production efficiency gains.

Columns (7) and (8) of table A.17 show a relatively larger decrease in tax registration in industries with higher backward linkages. This indicates that tax enforcement at the extensive margin is limited in India, and that some non-compliant firms are opportunistic: they evade taxes as much as possible. Nonetheless, production efficiency gains and self-enforcement are complementary in inducing relatively larger tax compliance in industries that combine a relatively larger share of business customers with relatively more sophisticated inputs.

We discuss next the the effects of the VAT reform below the MRT when tax registration is not mandatory, and thus simple and tax registered firms are compliant. Columns (1) and (2) of table A.17 show that the VAT self-enforcement properties do not impact simple registration below the VAT mandatory registration threshold. Only production efficiency gains are at play and they cause a significantly larger increase in simple registration below the registration threshold. Hence, lower distortions in the economy lead to larger formal participation, which is consistent with the fact that some informal firms are excluded from the formal economy by excessive costs (*forced informality*). It is interesting to see that production efficiency gains induce relatively larger tax compliance below the registration threshold through simple or tax registration, but relatively less tax compliance above the threshold if they are not combined with higher enforcement (*opportunistic informality*).

2.5.5 Robustness tests

State x Year fixed effects To check the robustness of our results, we estimate the model described above and include state x year fixed effects, that is to say the term $b_s \times c_t$, in the empirical models presented above. Note that the term VAT_{st} is no longer included in the equation regression once we control for state x year fixed effects. The model presented in equation 2.12 becomes:

$$g_{ist} = \alpha + \beta_1 \text{VAT}_{st} \times \text{FL}_{it} + \beta_2 \text{VAT}_{st} \times \text{BL}_{it} + +\beta_3 \text{VAT}_{st} \times \text{BL}_{it} \times \text{FL}_{it} \\ + \boldsymbol{\beta}^{\boldsymbol{T}} X_{ist} + b_s \times c_t + b_s + a_i + c_t + \varepsilon_{ist} \quad (2.14)$$

This model allows treatment and control states to follow different trends over time.

Fixed Effect versus Lagged Dependent Variables Fixed effects models and differences-in-differences estimations rely on the assumption of time invariant or state/industry invariant omitted variables. We would like to control for past levels of informality within a certain state and industry to account for the fact that the timing of the VAT adoption was driven by past values of our outcomes of interest. We would like to estimate the following model:

$$g_{ist} = \alpha + \beta_{-1} g_{is(t-1)} + \beta_0 \text{VAT}_{st} + \beta_1 \text{VAT}_{st} \times \text{FL}_{it} + \beta_2 \text{VAT}_{st} \times \text{BL}_{it} \\ + \boldsymbol{\beta}^{\boldsymbol{T}} X_{ist} + c_t + \varepsilon_{ist} \quad (2.15)$$

where $g_{is(t-1)}$ is the lagged dependent variable and is included in the list of regressors. Unfortunately, the OLS estimates of such an equation are not consistently estimated as the differenced residuals

$\Delta\varepsilon_{ist}$ are necessarily correlated with $g_{is(t-1)}$. However, fixed effects estimates from equation 2.12 and lagged dependent variables estimates from equation 2.15 have a useful bracketing property. If the model in equation 2.15 is correct, then we will under-estimate a positive treatment, while we will over-estimate a positive treatment if the model in equation 2.12 is correct. Therefore, both estimates "bound" the real causal effect.

Category variables to capture variations in forward and backward linkages We estimate the differential impact of the VAT reform across industries using dummy variables to reference whether industries have relatively lower or higher forward and backward linkages. We estimate the following model:

$$g_{ist} = \alpha + \beta_0 \text{VAT}_{st} + \beta_1 \text{VAT}_{st} \times \text{FL}_{it}^{High} + \beta_2 \text{VAT}_{st} \times \text{BL}_{it}^{High} + \beta_3 \text{VAT}_{st} \times \text{BL}_{it}^{High} \times \text{FL}_{it}^{High} + \boldsymbol{\beta^T} X_{ist} + a_i + b_s + c_t + \varepsilon_{ist} \quad (2.16)$$

where g_{ist} is one of our dependent variables i.e total production and total investment in industry i, state s, and time t as described in section 2.4. FL_{it}^{High} (respectively BL_{it}^{High}) is a dummy variable equal to 1 if FL_{it} (resp. BL_{it}) is superior to the median of FL (resp. BL) at time t. Figure A.8 shows the distribution of industries across these respective categories. All other notations are similar to those discussed in equation 2.12.

$$g_{ist} = \alpha + \beta_1 \text{VAT}_{st} \times \text{FA}_{is} + \beta_2 \text{VAT}_{st} \times \text{FA}_{is} \times \text{FL}_{it} + \beta_3 \text{VAT}_{st} \times \text{FA}_{is} \times \text{BL}_{it} + \boldsymbol{\beta^T} X_{ist} + a_i + b_s + c_t + \varepsilon_{ist} \quad (2.17)$$

where FA_{is}, $\text{FA}_{is} \times \text{BL}_{it}$ and $\text{FA}_{is} \times \text{FL}_{it}$ are included in the vector of control variables.

Other Reforms in the Indian Tax System

The Central VAT The Gandhi's government pushed for the adoption in 1986 of a primitive form of VAT with provision of input tax credit (but not covering capital goods) named MODVAT and raised by the central government. It includes manufacturing products and thus affects the outcomes of interest in this paper. In 1993, it was converted to an ad-valorem rate and extended to capital goods and petroleum products. The credit and refund system on manufacturing goods was simplified and extended to a larger number of commodities in 2002[30]. It was then renamed Central VAT (CENVAT). These successive changes are expected to reduce distortions in the economy in particular in backwardly-linked industries. Fortunately for our analysis, this reform was simultaneously implemented in all states and industries. Our empirical strategy is then identifying only the changes due to the states VAT reform and the validity of our results is not affected.

VAT and Inter-States Trade The treatment of inter-states trade is an important issue when implementing a VAT at the state level in a federal system. Ideally, importing producers should be granted a VAT credit equivalent to taxes paid in the exporting states. Inter-states trade would not be distorted in the sense that producers face the same input prices in all states (notwithstanding transportation costs that are significant in India, even within states). We explain here how the modalities of the VAT reform in Indian states regarding the treatment of inter-states sales differ from this ideal set-up. We argue that it does not affect the validity of our results. States sales tax were initially destination based, which means that exports were zero-rated. However, a Central Sales Tax (CST) was subsequently levied on cross-border transactions, which led to an origin based tax system. Prior to the VAT reform, the CST was charged at a 4% rate by the Central government that redistributed the corresponding revenues to exporting states. The VAT reform provides for the refund of input taxes for intra-states transactions only. This creates distortions in inter-states trade as some tax cascading remains for cross-border sales. To minimize this loss in production efficiency, the rate of the CST was decreased for all states in 2007 from 4% to 3% or the rate of ST/VAT applicable in the exporting state and reduced again to 2% in 2008.

Our first concern linked to this aspect of the reform is that it could have determined the timing of the VAT reform, which could have driven our results and/or be symptomatic of fundamental differences between states in the different waves of the VAT adoption. Maps A.5a and A.5b show the importance of exports for states measured by the share of collected CST in total tax revenues in 2000 and 2005. We can see that there is no correlation between the importance of inter-states transactions and the timing of the reform implementation. Thus, we conclude that there are no fundamental differences between our treatment and control groups linked to exporting activities. The importance of import is more difficult to quantify given available data, but Rao (2003) compares the share of consumption

[30] For example, it reduces the complexity of the procedure allowing for refunds of taxes paid on capital goods.

in total expenditures with the share of collected sales tax across states, and concludes that the poorest states are the largest importers and bear the largest burden of the cascading ST and CST. Additionally, Rao (2003) compares absolute levels of collected CST and underlines that only 4 “rich” states (Maharashtra, Gujarat, Tamil Nadu and West Bengal) account for 50% of the CST paid by all other states. In fact, these 4 states all belong to different adoption waves of the VAT reform.

Our second concern is that the VAT reform could have affected inter-states flows and that those changes would have driven our results on the positive impact of the reform on aggregate levels of production and investments. For example, differences in the timing of the VAT adoption could have caused differences in states tax systems that would have benefited either the early or the late adopters of the VAT. Specifically, firms and final consumers in states having implemented the VAT later could have taken advantage of differences in tax regimes (ST versus VAT) by importing larger quantities or different types of goods. Additionally, firms and final consumers in states having implemented the VAT earlier could also have substituted imported goods for locally produced goods. If early adopters were importing states prior to the reform, we would then expect a decrease (or even a reversal) in inter-states flows as producers and consumers react to production efficiency gains by re-focusing on their domestic markets[31]. On the contrary, if early adopters were exporting states, we would expect an increase in inter-states flows. To conclude, in both cases, we expect an increase in domestic and non-domestic demand for industries located in states where VAT is implemented. Our results of a positive impact of the reform on production and investments would reflect a switch from consumption in states with a ST to consumption in states with a VAT, in addition to the purely domestic effect of production efficiency gains. However, in the presence of the CST that decreased only in 2007 (that is to say 4 years after the first state adopted the VAT), we would argue that this is not very far off from the effects of replacing a cascading sales tax by a VAT in a small open economy.

The key question is whether the impact of the VAT reform on inter-states transactions would drive our results on the differential effect of the reform across industries. Indeed, it could be that industries using more sophisticated inputs are more prone to export their goods, or that trade flows of these goods are more volatile. Therefore our result of a relatively larger positive impact of the reform on backward looking industries would be driven by the fact that the reform increased relatively more non-domestic demand for these industries, while production efficiency gains were in fact constant across industries. Finally, the analysis above confirms our result of a relatively larger negative impact

[31] The reduction of the CST in 2007 and 2008 would have further strengthened the increase in imports from early adopters, but would have mitigated (potentially reversed, but it seems unlikely with such a small decrease in the tax rate) the switch from imports to local production. It is therefore not possible to conclude on the combined effects of the reduction in the CST and VAT adoption on inter-states flows.

of the VAT reform on investments and production in industries for which business customers represent a large share of their output. Furthermore, it would be difficult to explain the heterogeneous effect of the VAT reform across forward-looking industries by variations in inter-states trade. There will be no reason why such variations in trade flows would simultaneously cause i) a decrease in investments and production, ii) an increase in tax registration above the MRT, and iii) a switch from simple to non-registration above the MRT in backward looking industries.

2.6 Conclusion

This paper exploits a natural experiment formed by the replacement of a distorting retail sales tax with a VAT in India. Using a differences-in-differences approach and a large representative survey of informal and formal Indian firms, we have shown that the impact of VAT on production, investment and tax compliance is closely related to inter-industry linkages of industries in the economy. We account for a common feature of the VAT, its mandatory registration threshold, by considering two different types of tax compliance: tax registration above the threshold and other types of registration (*simple* thereafter) below the threshold.

The VAT decreases profitability in tax registered firms that are strongly dependent on business customers following a VAT reform. Thus, the implementation of a VAT credit and refund system broadens the tax base in a particular industry only if it is accompanied by an increase at the extensive margin of tax evasion. In developing countries, enforcement of tax registration above the VAT mandatory registration threshold is weak due to limited availability of resources. Accordingly, we find a large shift from tax registration to simple registration above this threshold. However, we also show empirically that the replacement of a distorting sales tax by a VAT represents significant production efficiency gains in industries whose inputs embed a large share of added-value. Under the VAT credit and refund system, the possibility not to pay VAT on added-value intensive inputs increases enforcement at the extensive margin of tax evasion in the form of a positive incentive to register. Additionally, suppliers of such industries will in turn be induced to register for taxes if they want to retain them as customers.

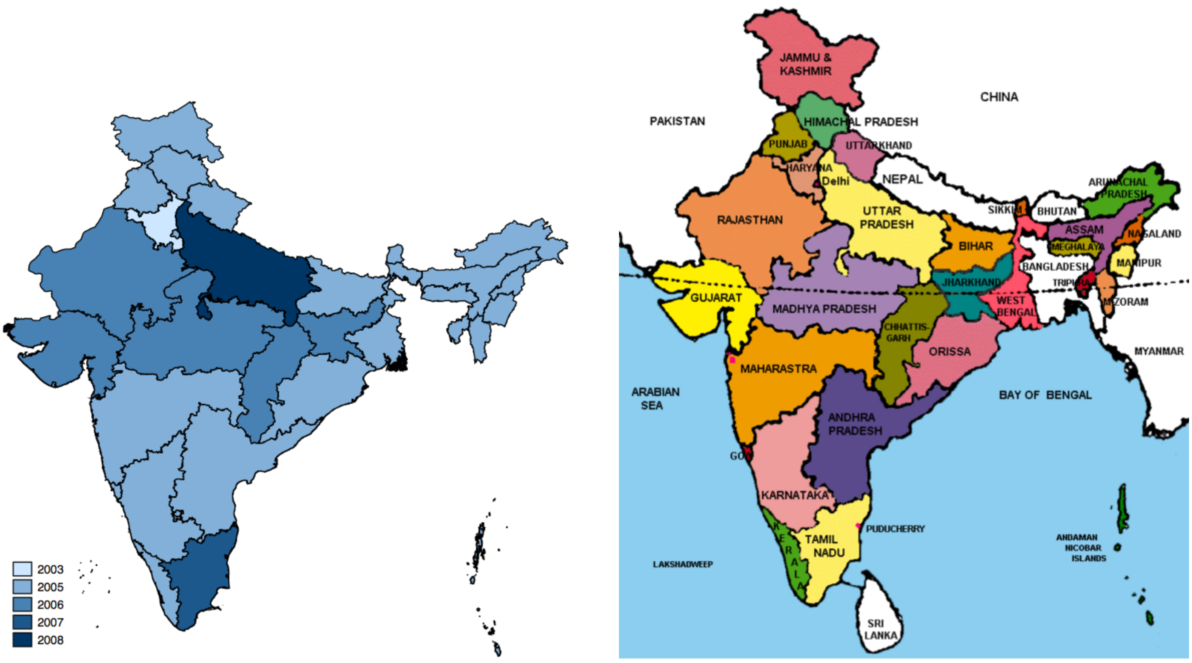
Our results have important policy consequences. First, we show that interactions between industries and between the formal and informal sectors matter as much as the size of the latter in determining the success of a VAT reform. A correct understanding of synergies or undermining effects between different enforcement strategies is crucial to design more efficient tax administrations in the future. We shed light on how to allocate enforcement expenditures across industries because tax authorities can rely on VAT self-enforcement properties only in industries where both the share of business customers and production efficiency gains are important. Second, we show that the VAT reform has increased formal participation even though tax compliance has decreased. Indeed, in the presence of a VAT mandatory registration threshold and limited enforcement at the extensive margin, simple registration offers production efficiency gains and a limited risk of detection in case of fraud. Fiscal capacity may only be indirectly increased by those firms but the capacity of the government to enforce labour regulations or to support them could also be improved, for example through subsidies or micro-credits. Finally, this paper allows for better predictions of the impact of the VAT reform on formality and fiscal revenues across Indian states. Thus, it has important policy consequences on the forthcoming fusion of Indian states VAT with the Service tax into a national Good and Services Tax (GST) that will be collected by the central government and re-distributed across states.

Appendices

Appendix A

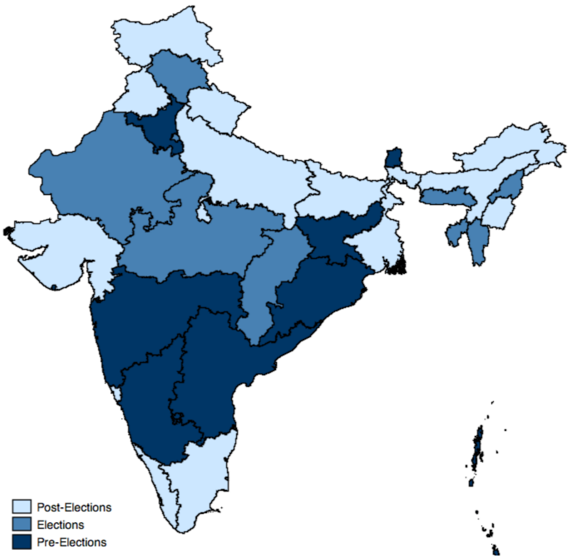
Data and Empirical Results

A.1 Descriptive Statistics

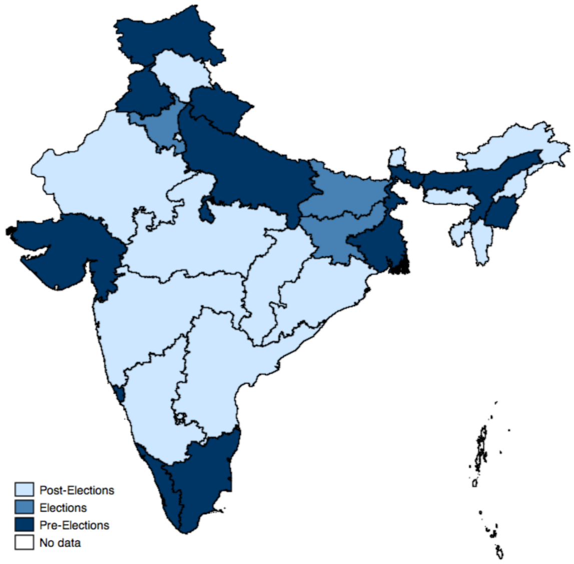


(a) Timing of the VAT Adoption (b) Indian States and Union Territories in 2000

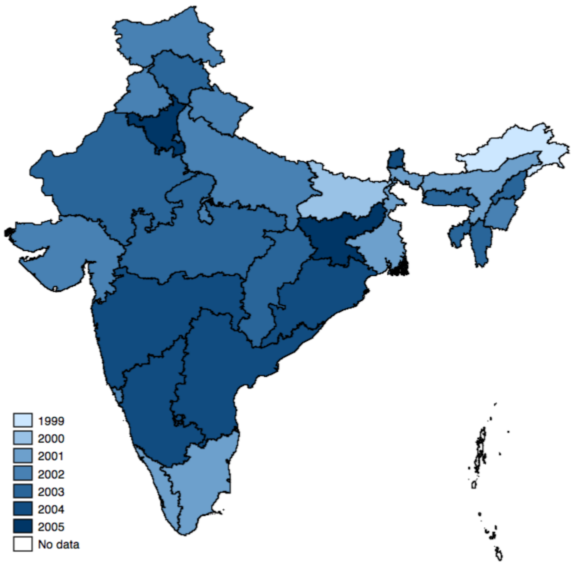
Figure A.1: **VAT Implementation across Indian States.** In 2000, there were 28 states and 7 Union Territories. In 2014, the states of Telangana was carved out of the state of Andhra Pradesh, which makes 29 states. Unlike states which have their own elected governments, union territories are ruled directly by the central government. Finally, Delhi and Pondichérry were given partial statehood, and have their own elected legislative assemblies.



(a) Electoral Stakes in 2003

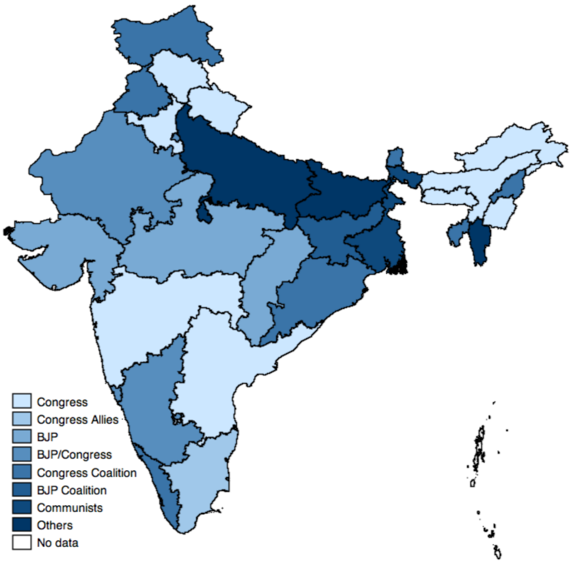


(b) Electoral Stakes in 2005 (Elections were planned in 2006 for Tamil Nadu and in 2007 for Gujarat and Uttar Pradesh.)

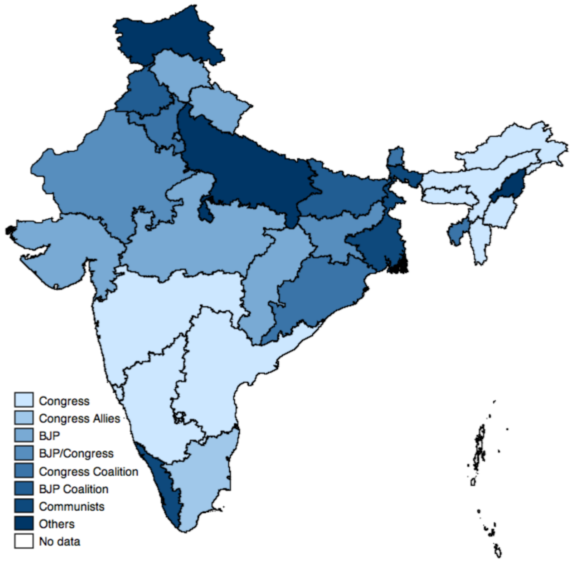


(c) Timing of Elections for States Legislative Assemblies

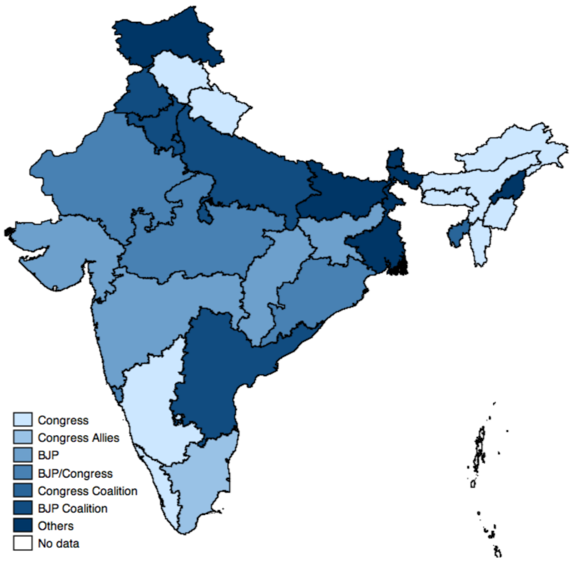
Figure A.2: **Timing of States Legislative Elections**



(a) Pre-Reform Elections (Elected between 1999 and 2005 as shown in map A.2c)

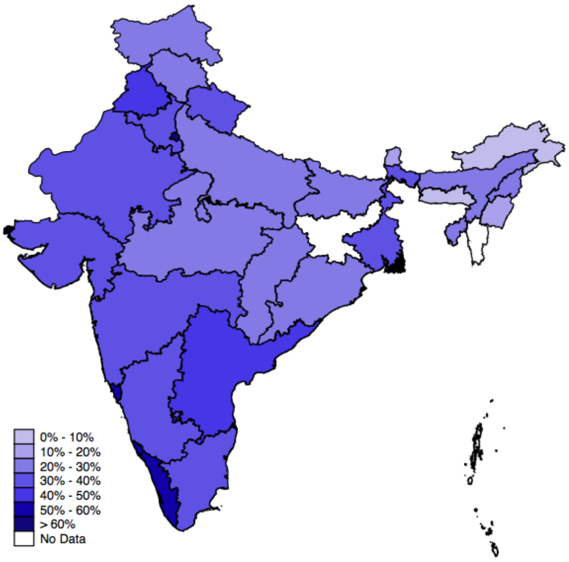


(b) During Reform Elections (5 years later)

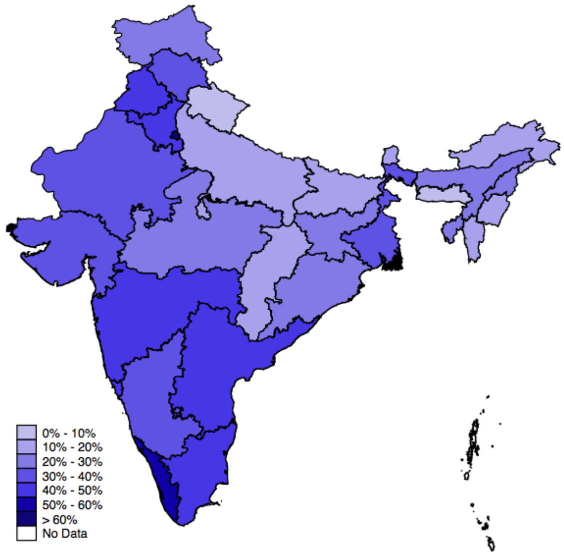


(c) Post-Reform Elections (10 years later)

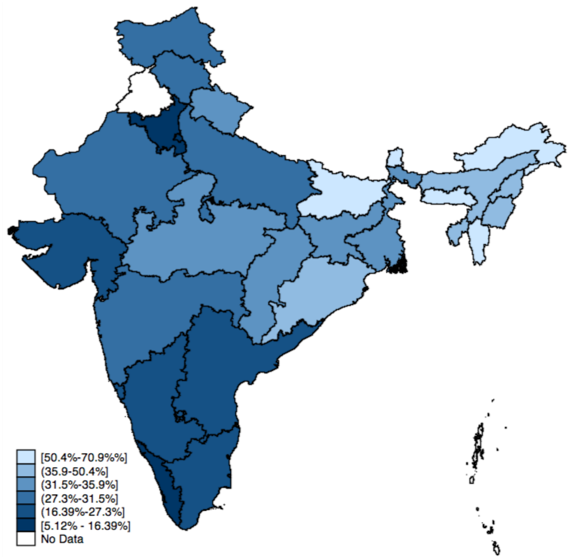
Figure A.3: **Political Parties in States Legislative Assemblies**



(a) 2000

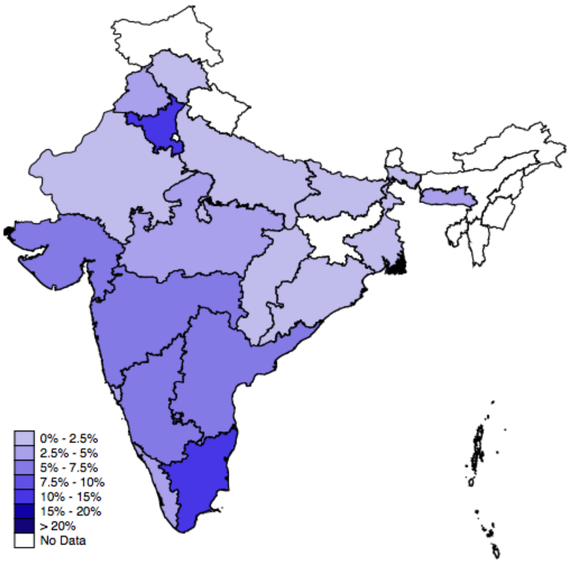


(b) 2005

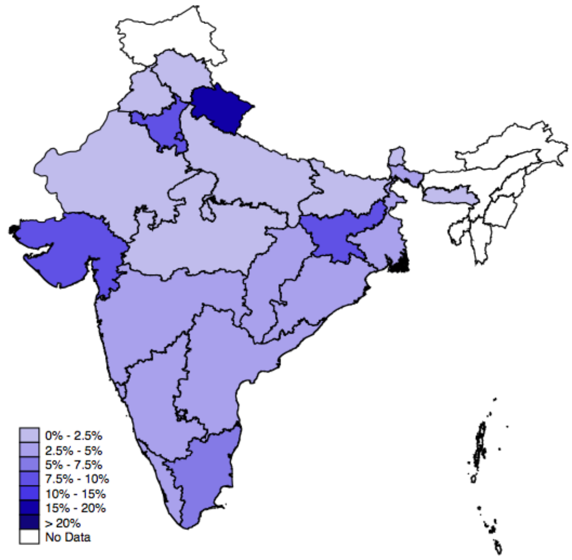


(c) 2010

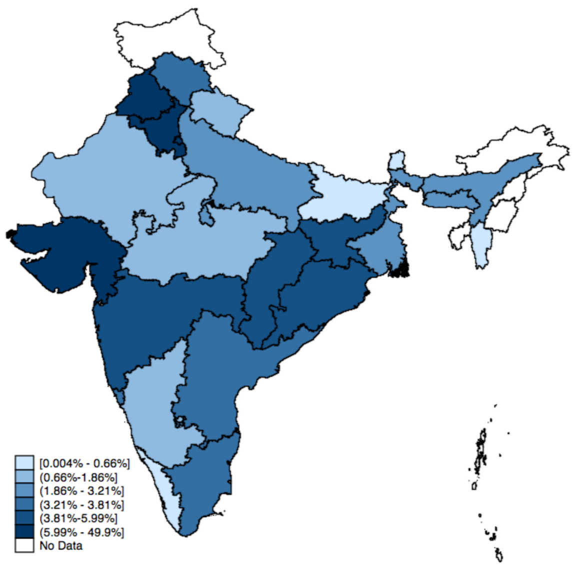
Figure A.4: **Share of Sales Tax in States Total Tax Revenues.** The state of Jharkhand for which data are missing in 2000 was carved out of the state of Bihar in 2000.



(a) 2000



(b) 2005



(c) 2010

Figure A.5: **Share of Central Sales Tax in States Total Tax Revenues.** The state of Jharkhand for which data are missing in 2000 was carved out of the state of Bihar in 2000.

Table A.1: **VAT adoption year among states and the ruling party.** The central government ruling party was BJP in 2003 and Indian National Congress afterwards. The incumbent party in Tamil Nadu (Dravida Munnetra Kazhagam) was a part of National Democratic Alliance (in coalition with BJP) in the 2001 election, but it joined United Progressive Alliance (coalition with Indian National Congress) in the 2006 election.

State	**year**	**Party**	**Gov/ Opp**
Haryana	2003	Indian National Lok Dal	Gov
Andhra Pradesh	2005	Indian National Congress	Gov
Arunachal Pradesh	2005	Indian National Congress	Gov
Bihar	2005	Rashtriya Janata Dal	Gov
Goa	2005	BJP	Opp
Himachal Pradesh	2005	Indian National Congress	Gov
Jammu and Kashmir	2005	People's Democratic Party	Gov
Karnataka	2005	Indian National Congress	Gov
Kerala	2005	Indian National Congress	Gov
Maharashtra	2005	Indian National Congress	Gov
Mizoram	2005	Mizo National Front	Opp
Nagaland	2005	Nagaland People's Front/BJP	Opp
Orissa	2005	Biju Janata Dal	Opp
Punjab	2005	Indian National Congress	Gov
Sikkim	2005	Sikkim Democratic Front	Gov
West Bengal	2005	Communist Party of India (Marxist)	Gov
Assam	2005	Indian National Congress	Gov
Manipur	2005	Indian National Congress	Gov
Tripura	2005	Communist Party of India (Marxist)	Gov
Uttarakhand	2005	Indian National Congress	Gov
Chhattisgarh	2006	BJP	Opp
Gujarat	2006	BJP	Opp
Jharkhand	2006	BJP	Opp
Madhya Pradesh	2006	BJP	Opp
Meghalaya	2006	Indian National Congress	Gov
Rajasthan	2006	BJP	Opp
Tamil Nadu*	2007	Dravida Munnetra Kazhagam	Opp/Gov
Uttar Pradesh	2008	Samajwadi Party	Opp

Table A.2: **Thresholds of mandatory registration under states VAT.** The threshold is for total turnover per year and is obtained from Bagchi (1994). The thresholds for excise duties is 100. All number are in Rs. lakh (1 lakh=100,000)

State		State		State		State	
Andhra Pardesh	5	Gujrat	5	Madhya Pradesh	10	Rajasthan	2
Arunachal Pradesh	5	Haryana	5	Maharastra	5	Sikkim	3
Assam	10	Himachal Pradesh	2	Meghalaya	5	Tamil Nadu	5
Bihar	5	Jammu & Kashmir	7.5	Mizoram	10	Tripura	10
Chhattisgarh	20	Jharkhand	5	Nagaland	10	Uttar Pradesh	5
Delhi	20	Karnataka	5	Orissa	1	Uttaranchal	5
Goa	1	Kerala	5	Punjab	1	West Bengal	5

Table A.3: **Formal status of firms in India (Share of total number of firms)** **Note:** Before the VAT reform, the indicated share of fraud to VAT is only the share of firms in-between the thresholds that is not registered for taxes. The VAT reform was implemented at different time by different states. Only disaggregated statistics at the state level or aggregated statistics for 2010 indicate the share of VAT fraud in India.

	Above CENVAT threshold	Between VAT and CENVAT thresholds	Registered under any act	Registered for taxes	Voluntary registration CENVAT	Voluntary registration VAT	Frauding VAT (Formal and informal)
1989	.21	1.72	8.30	.	.	.	.
1994	.39	3.02	11.79	1.19	.90	.34	2.55
2000	.36	3.90	10.69	1.70	1.40	.76	3.32
2005	.21	4.69	10.42	1.79	1.59	.64	3.74
2010	.36	6.16	15.16	3.31	2.95	1.09	4.29

Table A.4: **Formal status of firms in-between the VAT and the CENVAT thresholds (Share of firms between the VAT and CENVAT thresholds).** **Note:** Before the VAT reform, the indicated share of fraud to VAT is only the share of firms in-between the thresholds that is not registered for taxes. The VAT reform was implemented at different time by different states. Only disaggregated statistics at the state level or aggregated statistics for 2010 indicate the share of VAT fraud in India.

	Formal fraud to VAT	Registered under any act
1989	.	42.77
1994	31.71	50.34
2000	39.06	55.78
2005	36.83	57.16
2010	33.00	63.33

Table A.5: **Employment share of different categories of firms in India.** **Note:** Before the VAT reform, the indicated share of fraud to VAT is only the share of firms in-between the thresholds that is not registered for taxes. The VAT reform was implemented at different time by different states. Only disaggregated statistics at the state level or aggregated statistics for 2010 indicate the share of VAT fraud in India.

	Registered under any act	Registered for taxes	Above CENVAT threshold	In between VAT and CENVAT thresholds	Frauding VAT (Formal and informal)
1989	24.52	.	14.24	3.19	.
1994	33.02	20.58	16.28	7.35	5.80
2000	32.90	19.34	14.25	11.88	8.15
2005	33.57	21.56	14.65	15.35	9.39
2010	46.7	32.9	21.7	17.9	8.47

Table A.6: **Summary statistics at the industry level.** The numbers are averaged over years and states using sample weight. The bold numbers denote 25, 50, and 75 percentiles.

Industry	No. firms	Reg	Treg	Ereg	Etreg	Rinv	Uinv	VaReg	VaUreg	FL	BL	LL
Animal Husbandry	65309	30.8	32.69	88.64	75.38	82.79	0.13	63.41	1.36	1.406	1.655	1.0265
Sugar	515983	10.13	2.48	45.51	31.27	462.21	3.22	255.11	63.7	1.222	2.212	1.0022
Food Products Excluding Sugar	11578642	21.27	1.97	36.3	15.83	673.08	33.71	1698.37	472.21	1.287	2.359	1.0147
Beverages	1041632	11.87	2.33	29.73	17.46	195.8	2.36	329.06	30.46	1.201	**2.24**	1.0013
Tobacco products	10320280	3.13	0.48	16.25	11.62	80.39	1.87	351.91	176.25	1.082	1.896	1.0003
Cotton textiles	1477731	14.07	2.59	40.64	24.33	516.04	5.86	552.02	79.6	1.476	2.287	1.0098
Wool, Silk and synthetic fibre textiles	773639	9.15	0.93	20.31	8.71	46.38	2.8	106.88	52.52	1.605	2.504	1.0076
Jute, hemp, mesta textiles	6672641	8.11	1.62	33.39	17.72	751.16	7.74	1166.55	224.56	2.959	2.425	**1.0044**
Textile Products & wearing apparel	14476485	6.91	0.6	21.46	11.24	393.5	26.91	766.64	599.31	1.289	**2.407**	1.0077
Wood and wood products	11825137	3.37	0.72	7.84	2.35	55.85	9.04	157.85	350.97	2.683	1.75	1.0108
Furniture and fixtures-wooden	4698167	20.61	2.26	29.35	6.42	204.38	11.19	397.71	319.42	**1.467**	1.818	1.0022
Paper, paper prods. & newsprint	520355	12.91	5.94	54.81	39.88	332.05	1.14	473.87	22.53	3.213	2.496	1.0254
Printing and publishing	630409	50.72	8.9	69.86	29	172.01	3.42	363.09	37.07	1.898	2.336	1.0069
Leather and leather products	883008	10.34	2.94	37.41	25.31	86.84	1.92	132.36	58.36	1.447	2.251	1.003
Plastic and Rubber Products	601683	34.17	13.4	72.19	50.19	465.61	3.04	796.51	37.23	2.111	2.514	**1.0293**
Petroleum products	23070	45.05	22.15	88.17	70.12	630.77	0.04	974.01	1.21	**2.642**	2.174	1.0769
Organic heavy chemicals	5365	46.41	34.3	84.83	74.66	35.67	0.01	55.34	0.17	4.134	2.488	1.0349
Fertilizers	16316	30.28	18.11	93.19	71.82	506.99	0.01	615.14	0.9	2.905	2.586	1.0239
Paints, varnishes and lacquers	34346	78.33	64.15	93.3	73.09	109.8	0.04	187.95	1.24	3.012	2.435	**1.011**
Pesticides, Drugs and other chemicals	1112082	10.73	5.74	64.68	52.09	1190.07	0.62	2361.84	22.78	2.436	2.412	1.0774
Cement	181548	31.83	11.82	72.74	52.15	557.62	0.79	852.05	13.87	2.284	2.281	1.0087
Non-mettalic mineral products	3682799	7.98	4.1	38.07	22.98	370.08	4.7	695.91	177.76	**1.966**	2.258	1.0122
Iron and steel indsutries and foundries	63691	62.12	38.02	96.87	88.57	2063.25	0.13	1734.96	3.4	2.82	2.575	1.121
Non-ferrous basic metals	107628	33.26	11.63	77.13	56.75	650.22	0.15	526	6.44	5.209	2.574	1.0673
Metal products except transport equip.	2761363	23.28	4.6	48.17	22.82	563.02	8.02	924.94	185.78	2.359	**2.56**	1.0371
Tractors and agri. implements	662992	7.06	1.95	26.96	15.76	77.09	0.95	158.03	23.58	1.663	2.716	1.0026
Industrial machinery (F & T)	341631	40.5	10.79	79.07	54.56	539.86	1.17	678.74	20.32	1.996	2.805	1.0034
Other machinery	199007	63.71	18.19	86.98	54.98	297.94	1.42	482.8	13.56	2.012	2.573	1.0375
Electronic machinery and appliances	313922	40.28	14.12	85.23	62.44	951.06	1.31	1357.96	21.48	1.916	2.671	1.0452
Rail equipments	3196	90.69	50.45	99.86	60.92	9.88	0	73.91	0.01	2.475	2.6	1.0067
Other transport equipment	775185	30.42	8.45	74.58	54.33	1268.8	2.17	2131.06	56.01	1.58	2.347	1.0251
Miscellaneous manufacturing industries	5628783	8.97	0.91	20.87	7.66	242.54	9.17	433.01	196.85	1.843	2.108	1.0252

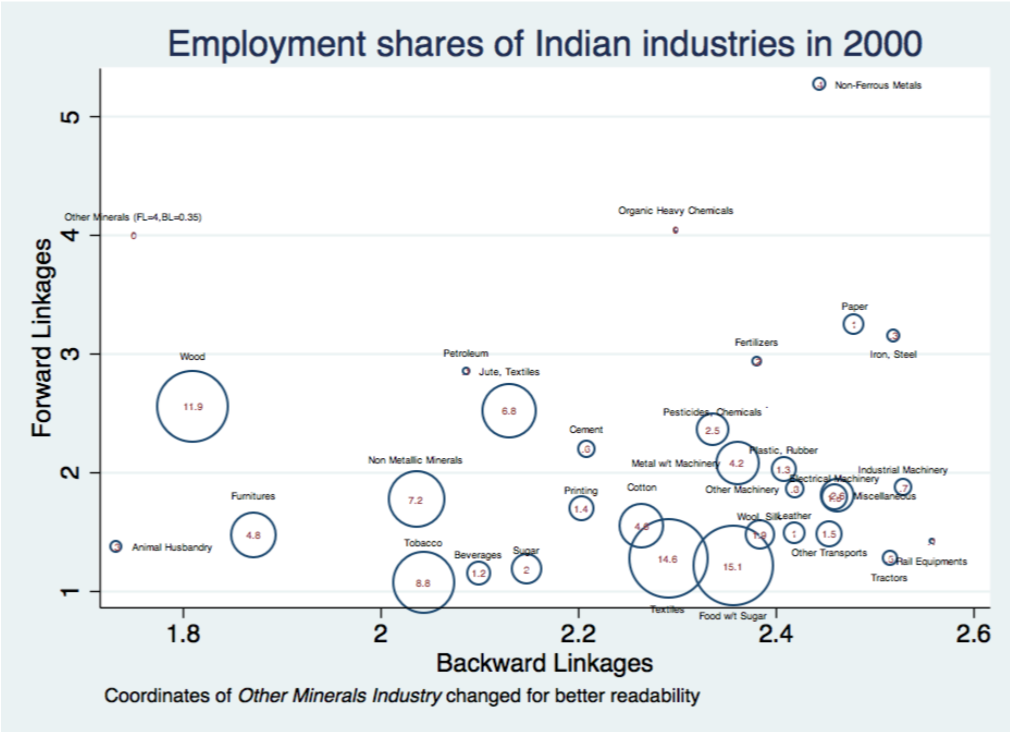


Figure A.6: **Share of each Indian industry in total employment (formal and informal) in 2000 prior to the VAT reform.**

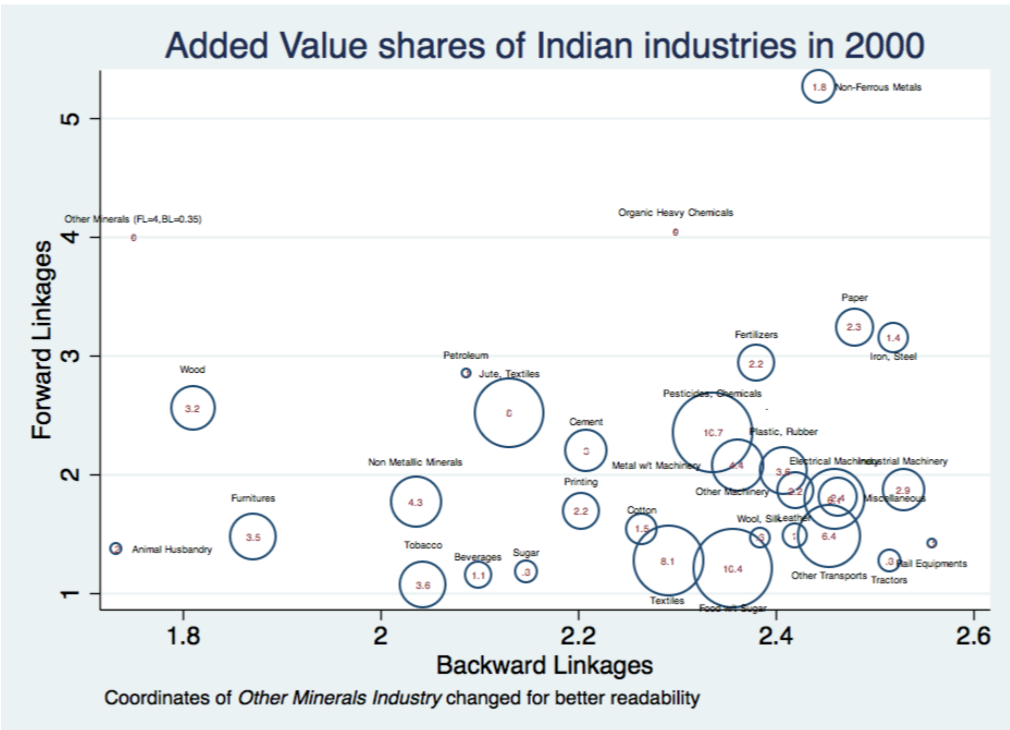


Figure A.7: **Share of each Indian industry in total created value (formal and informal) in 2000 prior to the VAT reform.**

Table A.7: **Value Added share of different categories of firms in India.** **Note:** Before the VAT reform, the indicated share of fraud to VAT is only the share of firms in-between the thresholds that is not registered for taxes. The VAT reform was implemented at different time by different states. Only disaggregated statistics at the state level or aggregated statistics for 2010 indicate the share of VAT fraud in India.

Total Value added Share	Registered under any act	Registered for taxes	Above CENVAT threshold	In between VAT and CENVAT thresholds	Frauding VAT (Formal and informal)
1989	79.36	.	71.19	4.87	.
1994	82.37	74.34	79.12	5.89	6.07
2000	80.93	71.25	67.75	10.72	7.7
2005	87.24	80.60	74.39	12.54	6.64
2010	92.49	88.57	83.17	8.19	3.10

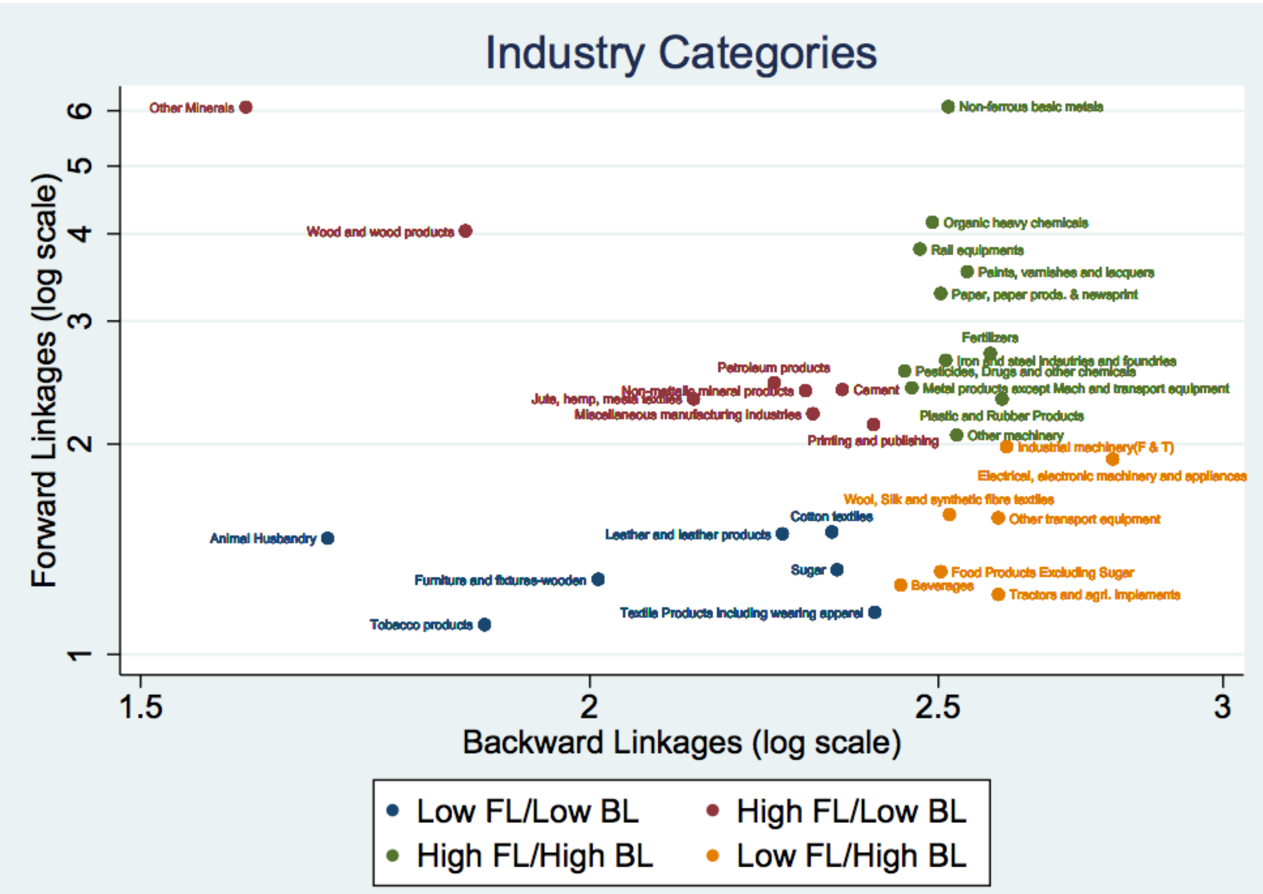


Figure A.8: **Forward and Backward Linkages of Indian Industries in 2005.** Each category corresponds to the intersection of the two groups formed by industries above and below the median for the FL and BL index respectively.

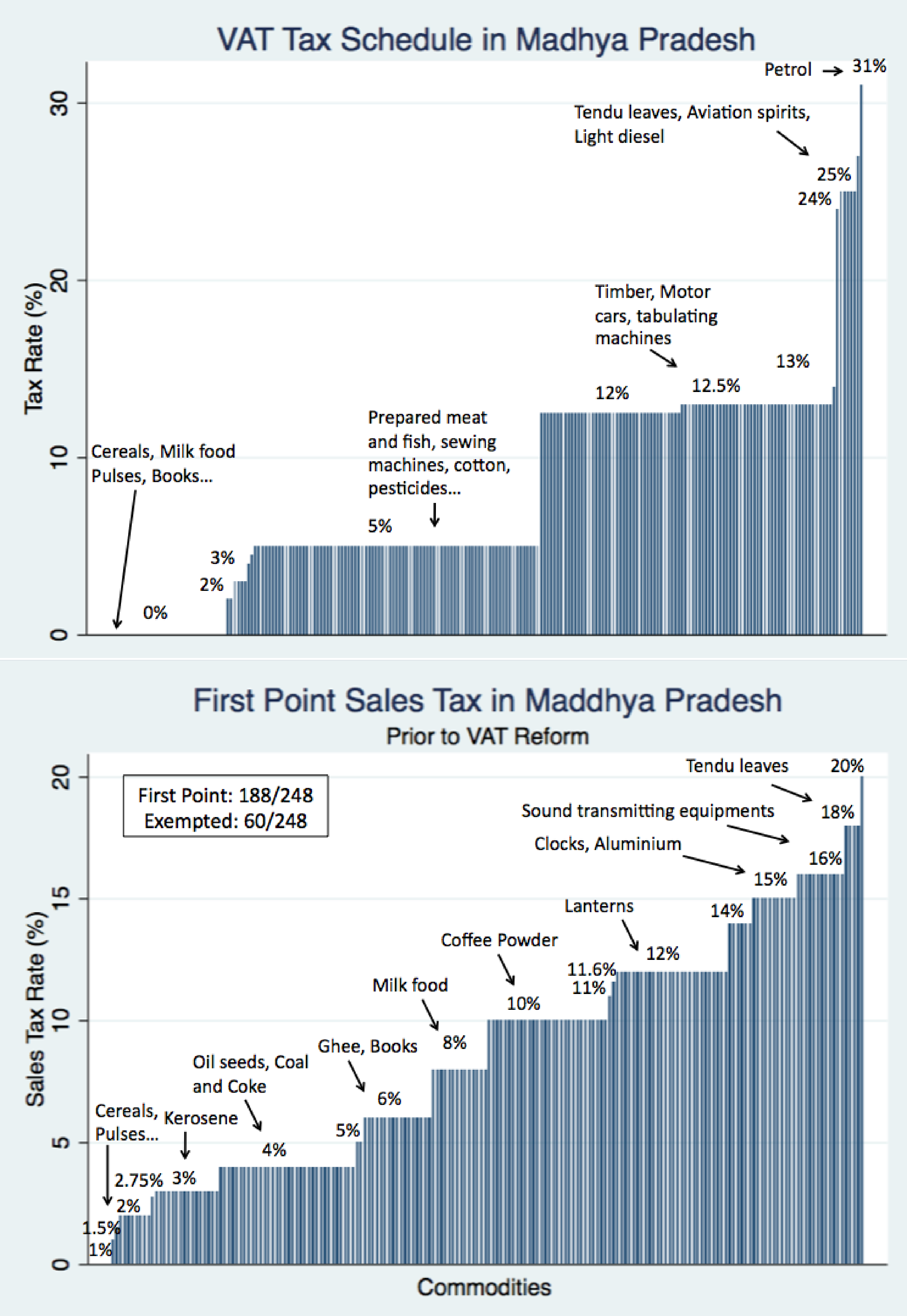


Figure A.9: **Retail Sales Tax and Value Added Tax schedules in Madhya Pradesh.**

A.2 Empirical Results

Table A.8: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on the proportion of tax registered firms (**Treg**), on the proportion on-registered firms (**Nreg**), and on the proportion of simply registered firms (**Sreg**). Standard errors are in parenthesis and are clustered at the state level. **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.12.

	Sreg	Sreg	Treg	Treg	Nreg	Nreg
	(1)	(2)	(3)	(4)	(5)	(6)
VAT dummy	-10.277*	-21.663*	-4.295*	0.664	14.572**	21.000*
	(5.370)	(11.574)	(2.430)	(2.864)	(6.111)	(11.594)
VAT x lag log FL	-2.069	18.370	1.025	-8.313**	1.044	-10.057
	(2.718)	(15.981)	(0.782)	(3.212)	(2.803)	(16.358)
VAT x lag log BL	11.955*	26.598*	5.470*	-1.422	-17.425**	-25.176
	(5.941)	(15.480)	(2.882)	(3.406)	(6.481)	(15.274)
VAT x lag log FL		-26.924		13.464***		13.459
x lag log BL		(20.550)		(4.715)		(21.848)
Constant	1.442	9.605	30.415**	32.639***	68.144**	57.756*
	(27.090)	(28.604)	(11.084)	(10.569)	(31.278)	(33.335)
Observations	1942	1942	1942	1942	1942	1942
R-squared	0.521	0.523	0.453	0.458	0.573	0.574
Includes State, Year, Industry and Political Party FE						

* p< 0.10 ** p< 0.05 *** p< 0.01

Table A.9: Results based on panel differences-in-differences estimations with **State x Year** fixed effects (**FE**) of the VAT reform's impact on the proportion of tax registered firms (**Treg**), on the proportion on-registered firms (**Nreg**), and on the proportion of simply registered firms (**Sreg**). Standard errors are in parenthesis and are clustered at the state level. **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.14.

	Sreg	Sreg	Treg	Treg	Nreg	Nreg
	(1)	(2)	(3)	(4)	(5)	(6)
VAT x lag log FL	-0.334	23.692	0.985	-8.237**	-0.651	-15.455
	(2.530)	(15.439)	(0.842)	(3.196)	(2.448)	(15.687)
VAT x lag log BL	14.064**	31.508**	5.288*	-1.601	-19.352**	-29.907*
	(6.679)	(15.167)	(2.810)	(3.070)	(7.099)	(14.996)
VAT x lag log FL x lag log BL		-32.150		13.243***		18.908
		(19.467)		(4.624)		(20.248)
Constant	-235.201***	-368.326***	-51.304	6.638	386.505***	461.688***
	(59.432)	(123.622)	(30.272)	(31.048)	(56.390)	(110.425)
Observations	1942	1942	1942	1942	1942	1942
R-squared	0.593	0.596	0.472	0.476	0.633	0.634
Includes State, Year, Industry and Political Party FE						
Includes State x Year FE						

* p< 0.10 ** p< 0.05 *** p< 0.01

Table A.10: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on the proportion of simply registered firms (**Sreg**). Standard errors are in parenthesis and are clustered at the state level. $log\bar{F}L$: Log of forward linkages index centered around the minimum value of $log(FL)$. $log\bar{B}L$: Log of backward linkages index centered around the minimum value of $log(BL)$. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.12.

	Proportion of Simply Registered Firms (**Sreg**)			
	(1)	(2)	(3)	(4)
VAT	-2.505		-6.540**	
	(2.556)		(2.492)	
VAT x lag $log\bar{F}L$	-3.172	-0.334	5.183*	6.545*
	(2.817)	(2.530)	(2.773)	(3.371)
VAT x lag $log\bar{B}L$	6.855	14.064**	19.290**	26.514**
	(4.614)	(6.679)	(7.677)	(10.638)
VAT x lag $log\bar{F}L$ x lag $log\bar{B}L$			-25.936**	-26.991**
			(12.257)	(12.926)
Constant	-5.522	-98.890***	-3.109	-89.527***
	(33.474)	(6.848)	(32.439)	(7.847)
Observations	1942	1942	1942	1942
R-squared	0.521	0.593	0.523	0.596
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry FE				

* p< 0.10 ** p< 0.05 *** p< 0.01

Table A.11: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on the proportion of non registered firms (**Nreg**). Standard errors are in parenthesis and are clustered at the state level. $log\bar{F}L$: Log of forward linkages index centered around the minimum value of $log(FL)$. $log\bar{B}L$: Log of backward linkages index centered around the minimum value of $log(BL)$. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.12.

	Proportion of Non-Registered Firms (**Nreg**)			
	(1)	(2)	(3)	(4)
VAT	3.665		7.368***	
	(2.755)		(2.605)	
VAT x lag $log\bar{F}L$	2.125	-0.651	-5.974*	-7.564*
	(3.045)	(2.448)	(2.979)	(3.703)
VAT x lag $log\bar{B}L$	-11.929**	-19.352**	-19.515**	-27.061**
	(5.629)	(7.099)	(7.353)	(10.380)
VAT x lag $log\bar{F}L$ x lag $log\bar{B}L$			14.144	16.138
			(13.336)	(14.171)
Constant	76.618*	221.640***	72.479*	209.916***
	(37.766)	(15.683)	(36.229)	(17.771)
Observations	1942	1942	1942	1942
R-squared	0.572	0.633	0.575	0.635
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry FE				

* p< 0.10 ** p< 0.05 *** p< 0.01

Table A.12: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on total production (**Prod**). The unit of measurement is in 100.000 INR at constant 1960 prices. Standard errors are in parenthesis and are clustered at the state level. **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.12.

	Total production (**Prod**)			
	(1)	(2)	(3)	(4)
VAT dummy	-12.849***		-24.785***	
	(3.977)		(8.259)	
VAT x lag FL	-0.400	-0.683	6.245	7.129
	(0.961)	(1.012)	(4.155)	(4.405)
VAT x lag BL	6.614***	5.566***	11.497***	11.520***
	(1.923)	(1.982)	(3.781)	(3.980)
VAT x lag FL			-2.711	-3.307*
x lag BL			(1.837)	(1.839)
Constant	5.153	42.757***	19.771*	87.662***
	(6.242)	(8.929)	(10.890)	(25.095)
Observations	1942	1942	1942	1942
R-squared	0.573	0.590	0.574	0.591
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry Fixed Effects				

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table A.13: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on total production (**Prod**). The unit of measurement is in 100.000 INR, at constant 1960 prices. Standard errors are in parenthesis and are clustered at the state level. **High FL**: Dummy variable equal to 1 if the forward linkages index of an industry is larger than the median of the FL index at a certain period t. **High BL**: Dummy variable equal to 1 if the backward linkages index of an industry is larger than the median of the FL index at a certain period t. Full description of the model in equation 2.16.

	Total Production (**Prod**)			
	(1)	(2)	(3)	(4)
VAT dummy	0.806		0.619	
	(0.757)		(0.771)	
VAT x High lag FL	-1.646*	-1.786*	-0.613	-0.452
	(0.845)	(0.957)	(0.909)	(1.069)
VAT x High lag BL	8.927***	8.841***	9.724***	10.251***
	(1.841)	(1.873)	(2.154)	(2.217)
VAT x High lag FL x High lag BL			-3.900	-5.376**
			(2.791)	(2.608)
Constant	9.656	23.739***	9.816	23.588***
	(6.442)	(3.022)	(6.069)	(3.118)
Observations	1942	1942	2235	1942
R-squared	0.595	0.613	0.592	0.616
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry and Political Party Fixed Effects				

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table A.14: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on total production (**Prod**). The unit of measurement is in 100.000 INR, at constant 1960 prices. Standard errors are in parenthesis and are clustered at the state level. $\bar{FL}$: Forward linkages index centered around the minimum value of FL. $\bar{BL}$: Backward linkages index centered around the minimum value of BL. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.12.

	Total production (**Prod**)			
	(1)	(2)	(3)	(4)
VAT	-1.366		-4.259***	
	(1.193)		(1.411)	
VAT x $\bar{FL}$	-0.627	-0.683	3.403**	3.332**
	(1.015)	(1.012)	(1.419)	(1.475)
VAT x $\bar{BL}$	6.044***	5.566***	9.774***	9.553***
	(1.728)	(1.982)	(2.352)	(2.774)
VAT x $\bar{FL}$ x $\bar{BL}$			-4.956**	-5.412**
			(2.141)	(2.230)
Constant	10.200*	24.772***	11.533*	33.323***
	(5.618)	(2.795)	(5.850)	(3.899)
Observations	1942	1942	1942	1942
R-squared	0.573	0.590	0.577	0.593
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry FE				
	* p< 0.10	** p< 0.05	*** p< 0.01	

Table A.15: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on total investment (**Inv**). The unit of measurement is in 100.000 INR, at constant 1960 prices. Standard errors are in parenthesis and are clustered at the state level. **High FL**: Dummy variable equal to 1 if the forward linkages index of an industry is larger than the median of the FL index at a certain period t. **High BL**: Dummy variable equal to 1 if the backward linkages index of an industry is larger than the median of the FL index at a certain period t. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.16.

	Total Investment (**Inv**)			
	(1)	(2)	(3)	(4)
VAT dummy	-3.283*		-6.268**	
	(1.629)		(2.926)	
VAT x lag FL	-0.211	-0.235	1.453*	1.572**
	(0.224)	(0.240)	(0.766)	(0.729)
VAT x lag BL	1.884**	1.909**	3.062*	3.203**
	(0.855)	(0.919)	(1.532)	(1.536)
VAT x lag FL x lag BL			-0.652	-0.718*
			(0.432)	(0.402)
Constant	0.433	9.514**	4.858*	21.272***
	(1.368)	(3.581)	(2.482)	(5.831)
Observations	1942	1942	1942	1942
R-squared	0.293	0.321	0.295	0.323
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry and Political Party FE				

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table A.16: Results based on panel differences-in-differences estimations with fixed effects (**FE**) of the VAT reform's impact on total investment (**Inv**). Standard errors are in parenthesis and are clustered at the state level. $\bar{FL}$: Forward linkages index centered around the minimum value of FL. $\bar{BL}$: Backward linkages index centered around the minimum value of BL. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Full description of the model in equation 2.12.

	Total Investment (**Inv**)			
	(1)	(2)	(3)	(4)
VAT	-0.073		-0.793**	
	(0.300)		(0.367)	
VAT x $\bar{FL}$	-0.290	-0.235	0.749***	0.804***
	(0.266)	(0.240)	(0.222)	(0.212)
VAT x $\bar{BL}$	1.610*	1.909**	2.364**	2.676**
	(0.802)	(0.919)	(1.100)	(1.267)
VAT x $\bar{BL}$ x $\bar{FL}$			-0.968*	-1.035*
			(0.547)	(0.545)
Constant	1.489	2.928***	1.900	5.205***
	(1.509)	(0.466)	(1.570)	(0.905)
Observations	1942	1942	1942	1942
R-squared	0.291	0.321	0.298	0.327
State x Year FE	No	Yes	No	Yes
Includes State, Year, Industry and Political Party FE				

* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table A.17: Results based on panel triple differences estimations with fixed effects (**FE**) of the VAT reform impact across industries and above/below the VAT mandatory registration threshold on the proportion of tax registered firms (**Treg**), on the proportion on-registered firms (**Nreg**), and on the proportion of simply registered firms (**Sreg**). Outcomes are measured at the year, state, and industry level for two samples of firms: below and above the VAT mandatory registration threshold. ***Above MRT* is a dummy variable for the sample of firms above the VAT mandatory registration threshold.**; **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Standard errors are in parenthesis and are clustered at the state level. Full description of the model in equation 2.13.

	Sreg	Sreg	Sreg	Sreg	Treg	Treg	Treg	Treg
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VAT dummy	-10.331*		-12.166		-2.621***		0.935	
	(5.295)		(9.196)		(0.743)		(1.009)	
VAT x Above MRT	28.917	59.318**	-39.897	-13.036	10.244	-32.261*	63.565**	41.874
	(23.921)	(28.064)	(32.523)	(34.927)	(18.703)	(18.120)	(28.595)	(24.868)
VAT dummy . . .								
x lag log(FL)	-1.813	-0.237	6.931	13.538	0.454	0.347	-5.563***	-5.733***
	(2.836)	(2.621)	(12.907)	(12.323)	(0.308)	(0.359)	(1.335)	(1.543)
x lag log(BL)	12.424*	14.270**	14.586	20.272	3.173***	2.990***	-1.656	-2.004
	(6.591)	(6.906)	(12.101)	(12.212)	(0.936)	(0.990)	(1.303)	(1.359)
x lag log(FL) x lag log(BL)			-11.629	-18.586			8.460***	8.534***
			(16.275)	(15.230)			(1.731)	(1.916)
VAT x Above MRT . . .								
x lag log(FL)	-18.048**	-17.360**	83.069**	80.116*	-4.434	-0.267	-90.561**	-68.292**
	(6.925)	(7.632)	(37.066)	(40.540)	(7.508)	(7.222)	(36.470)	(29.922)
x lag log(BL)	-36.497	-44.835	45.256	34.235	1.152	-5.222	-61.190*	-50.904*
	(30.139)	(31.367)	(39.347)	(40.331)	(20.548)	(18.625)	(30.314)	(27.478)
x lag log(BL) x lag log(FL)			-116.126***	-111.125**			97.554**	75.825**
			(40.892)	(43.833)			(38.626)	(31.035)
Constant	0.691	0.562	-0.137	261.436	8.137	-47.166***	8.750*	-50.032
	(36.362)	(11.954)	(26.020)	(406.280)	(5.090)	(4.950)	(4.942)	(42.058)
Observations	2567	2567	2567	2567	2567	2567	2567	2567
R-squared	0.538	0.609	0.542	0.614	0.738	0.757	0.742	0.760
State x Year FE	No	Yes	No	Yes	No	Yes	No	Yes
	Includes State, Year, Industry, Political Party FE							

$*p < 0.10$ $** \ p < 0.05$ $*** \ p < 0.01$

Table A.18: Results based on panel triple differences estimations with fixed effects (**FE**) of the VAT reform impact across industries and above/below the VAT mandatory registration threshold on the proportion of simply registered (**Sreg**). Outcomes are measured at the year, state, and industry level for two samples of firms: below and above the VAT mandatory registration threshold. ***Above MRT* is a dummy variable for the sample of firms above the VAT mandatory registration threshold.**; **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Standard errors are in parenthesis and are clustered at the state level. Full description of the model in equation 2.13.

	Proportion of Simply Registered Firms (Sreg)			
	(1)	(2)	(3)	(4)
VAT dummy	-13.835*	-36.336		
	(7.672)	(22.666)		
VAT x lag FL	-1.336	11.216	-0.723	13.927
	(1.361)	(9.917)	(1.336)	(9.690)
VAT x lag BL	6.679*	16.404	7.571**	19.212*
	(3.488)	(10.472)	(3.512)	(10.152)
VAT x lag FL x lag BL		-5.439		-6.515
		(4.332)		(4.189)
VAT x Above MRT	38.719	-15.871	106.310***	26.112
	(30.767)	(77.979)	(37.700)	(62.911)
VAT x Above MRT x lag BL	-15.671	9.441	-19.716	3.101
	(13.972)	(32.402)	(14.643)	(33.199)
VAT x Above MRT x lag FL	-7.649**	13.262	-7.166*	11.158
	(3.467)	(32.408)	(3.848)	(34.137)
VAT x FL x BL x Above MRT		-9.710		-8.457
		(12.606)		(13.389)
Constant	3.364	24.133	9.426	51.760
	(27.902)	(32.260)	(18.310)	(52.372)
Observations	2567	2567	2567	2567
R-squared	0.538	0.540	0.610	0.612
State x Year FE	No	No	Yes	Yes
Includes State, Year, Industry and Political Party FE				

$*p < 0.10$ $** \ p < 0.05$ $*** \ p < 0.01$

Table A.19: Results based on panel triple differences estimations with fixed effects (**FE**) of the VAT reform impact across industries and above/below the VAT mandatory registration threshold on the proportion of tax registered (**Treg**). Outcomes are measured at the year, state, and industry level for two samples of firms: below and above the VAT mandatory registration threshold. ***Above MRT* is a dummy variable for the sample of firms above the VAT mandatory registration threshold.**; **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Standard errors are in parenthesis and are clustered at the state level. Full description of the model in equation 2.13.

	Proportion of Tax Registered Firms (Treg)			
	(1)	(2)	(3)	(4)
VAT dummy	-3.346***	0.078		
	(1.037)	(1.764)		
VAT x lag FL	0.187	-1.764*	0.127	-1.864*
	(0.183)	(0.993)	(0.209)	(1.053)
VAT x lag BL	1.410***	-0.435	1.332**	-0.571
	(0.481)	(0.775)	(0.510)	(0.805)
VAT x lag FL x lag BL		1.071**		1.088**
		(0.432)		(0.446)
VAT x Above MRT	9.540	91.358	-4.671	51.451
	(19.304)	(70.799)	(18.439)	(39.905)
VAT x Above MRT x lag BL	0.944	-34.172	-1.512	-29.966
	(7.961)	(27.938)	(7.208)	(25.935)
VAT x Above MRT x lag FL	-1.600	-40.350	0.772	-31.775
	(3.093)	(32.694)	(2.848)	(29.209)
VAT x FL x BL x Above MRT		16.463		13.642
		(12.732)		(11.223)
			-50.783***	-2.445
Constant	9.647*	13.314**	(5.298)	(5.605)
	(5.156)	(5.147)	2567	2567
Observations	2567	2567	0.757	0.758
R-squared	0.738	0.740	1.000	1.000
State x Year FE	No	No	Yes	Yes
Includes State, Year, Industry and Political Party FE				

$*p < 0.10$ $** \; p < 0.05$ $*** \; p < 0.01$

Table A.20: Results based on panel triple differences estimations with fixed effects (**FE**) of the VAT reform impact across industries and above/below the VAT mandatory registration threshold on the proportion of non registered (**Nreg**). Outcomes are measured at the year, state, and industry level for two samples of firms: below and above the VAT mandatory registration threshold. ***Above MRT* is a dummy variable for the sample of firms above the VAT mandatory registration threshold.**; **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Standard errors are in parenthesis and are clustered at the state level. Full description of the model in equation 2.13

	Proportion of Non Registered Firms (Nreg)			
	(1)	(2)	(3)	(4)
VAT dummy	17.181**	36.258		
	(7.759)	(21.864)		
VAT x lag FL	1.149	-9.453	0.597	-12.063
	(1.311)	(9.514)	(1.268)	(9.214)
VAT x lag BL	-8.088**	-15.970	-8.903**	-18.642*
	(3.444)	(10.106)	(3.465)	(9.727)
VAT x lag FL x lag BL		4.368		5.427
		(4.189)		(4.002)
VAT x Above MRT	-48.259	-75.488	-101.639***	-77.563
	(29.651)	(74.260)	(31.005)	(63.847)
VAT x Above MRT x lag BL	14.727	24.732	21.228*	26.866
	(12.839)	(30.505)	(12.194)	(31.875)
VAT x Above MRT x lag FL	9.249***	27.087	6.394**	20.616
	(3.035)	(33.579)	(3.089)	(34.390)
VAT x FL x BL x Above MRT		-6.753		-5.185
		(13.323)		(13.696)
			141.356***	50.685
Constant	86.989***	62.554*	(20.997)	(53.416)
	(30.119)	(34.479)	2567	2567
Observations	2567	2567	0.730	0.732
R-squared	0.681	0.682	1.000	1.000
State x Year FE	No	No	Yes	Yes
Includes State, Year, Industry and Political Party FE				
$*p < 0.10$	$** \ p < 0.05$	$*** \ p < 0.01$		

Table A.21: Results based on panel triple differences estimations with fixed effects (**FE**) of the VAT reform impact across industries and above/below the VAT mandatory registration threshold on total production (**Prod**: the unit of measurement is in 100.000 INR, at constant 1960 prices). Outcomes are measured at the year, state, and industry level for two samples of firms: below and above the VAT mandatory registration threshold. ***Above MRT* is a dummy variable for the sample of firms above the VAT mandatory registration threshold.**; **FL**: Forward linkages index. **BL**: Backward linkages index. Interquartile range for FL and BL are 1.175 and 0.32 respectively. Standard errors are in parenthesis and are clustered at the state level. Full description of the model in equation 2.13.

	Total production (**Prod**)		
	(1)	(2)	(3)
VAT x Above MRT	24.499***	-7.004	-33.933*
	(4.649)	(9.290)	(16.859)
VAT x Above MRT x lag BL		3.454	14.224
		(3.879)	(10.643)
VAT x Above MRT x lag FL		-2.424**	8.720
		(1.065)	(11.999)
VAT x FL x BL x Above MRT			-4.961
			(4.811)
VAT adoption x lag FL		-0.117*	-0.402
		(0.064)	(0.402)
VAT adoption x lag BL		-0.466	-0.778
		(0.496)	(0.825)
VAT x FL x BL			0.179
			(0.213)
Constant	1.270***	-0.320	-1.424
	(0.442)	(1.505)	(2.522)
Observations	2567	2567	2567
R-squared	0.786	0.791	0.792
State x Year FE	Yes	Yes	Yes
Includes State, Year, Industry, Political Party FE			
$*p < 0.10$	$**\ p < 0.05$	$***\ p < 0.01$	

A.3 Data

This appendix presents additional information on the Annual Survey of Industries and of the Unorganized Manufacturing Enterprises of the National Sample Survey. It also enters into the details of the data preparation process. We extract a subset of variables from the raw data for each wave separately and then stack all waves together before proceeding with the following cleaning process. We create a consistent variable for Indian states that takes into account the creation of Jharkand, Chhatisgarh and Uttarakhand (or Uttaranchal) in 2001 from the states of Bihar, Madhya Pradesh and Uttar Pradesh, respectively. Industries are classified in India using the National Industries Classification (NIC) system that was revised in 1998, 2004, and 2008. First, using the concordance tables provided by MOSPI, we first convert all industry classifications to the 130 sectors classification used in the I-O tables of the Central Statistical Organization (CSO) described in section 2.4.2. Second, following the methodology described in Appendix A.3, we convert these 130 sectors classification to the 60 divisions classification used in Dholakia et al. (2009). All financial amounts are deflated to constant 2004-05 Rupees. Revenue (gross sale) and investments are deflated by a three-digit commodity price deflator contained in the (commodity x commodity) table *Index Numbers of Wholesale Prices in India - By Groups and Sub-Groups (Yearly Averages)* produced by the Office of the Economic Adviser-Ministry of Commerce and Industry.[1]. The sampling rule of the ASI changed over time: all factories with 100 or more workers were included before 2004 (excluded), but only factories with 200 workers or more are included after that. Hence, the sampling rule included one third of factories until 2004 and one fifth afterwards.

Aggregating I-O tables

If the economy is composed of N sectors, then define the following matrices (dimensions in the parentheses):
$Y_{(N\times1)}$ production, $Y_{0(N\times1)}$ final demand, $\mathbf{A}_{(N\times N)}$ Leontief coefficients, Then we have:

$$Y = \mathbf{A}Y + Y_0, \quad Y = (\mathbf{I} - \mathbf{A})^{-1}Y_0$$

To aggregate the sectors to M divisions ($M < N$), assume a linear transformation matrix $\mathbf{W}_{(M\times N)}$ in the sense that $Z_{(M\times1)} = \mathbf{W}Y_{(N\times1)}$ are the aggregated production vector. Then define $\mathbf{W}^{-1}_{(N\times M)}$

[1]Available from the official website of the Office of the Economic Adviser of the Ministry of Commerce and Industry of India

the right inverse of $\mathbf{W}$ such that $\mathbf{W}\mathbf{W}^{-1} = \mathbf{I}_{(M \times M)}$. In this case, we have

$$Z_0 = \mathbf{W}Y_0, \quad Z = (\mathbf{I} - \mathbf{W}\mathbf{A}\mathbf{W}^{-1})^{-1} Z_0$$

The proof comes from the facts that $(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I}+\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3+\ldots$ and $(\mathbf{W}\mathbf{A}\mathbf{W}^{-1})^n = \mathbf{W}\mathbf{A}^n\mathbf{W}^{-1}$. Therefore, if $\mathbf{Z}$ is the diagonalized matrix of vector Z, the linkages indices for new classification can be written as

$$FL = \mathbf{Z}^{-1}(I - \mathbf{W}\mathbf{A}\mathbf{W}^{-1})^{-1} Z, \quad BL = \left(I - (\mathbf{W}\mathbf{A}\mathbf{W}^{-1})'\right)^{-1} J$$

Also, one can show some useful equalities that help computing linkages indices:

$$(\mathbf{I} - \mathbf{W}\mathbf{A}\mathbf{W}^{-1})^{-1} = \mathbf{W}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{W}^{-1}, \qquad \left(\mathbf{I} - (\mathbf{W}\mathbf{A}\mathbf{W}^{-1})'\right)^{-1} = \left(\mathbf{W}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{W}^{-1}\right)'$$

Therefore, we can find LL index as the following

$$LL = \mathbf{W}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{W}^{-1} B$$

Where B is a vector comprising the share of final consumption of the commodity in total final consumption in the economy. The I-O tables published by the CSO consist of use and make tables and 5 associated matrices including commodity $\times$ commodity ($\mathbf{A}$), total value of production (Y), final consumption, and Leontief inverse matrix $(\mathbf{I} - \mathbf{A})^{-1}$. Using them, we compute the linkages indices. Table A.23 presents the estimated FL and BL indices for the industries.

Table A.22: **Underlying data of Figure 2.3.** BL is the backward linkages index computed as (2.2) using I-O transaction tables of Indian economy published by CSO. This index is subtracted by 1 for better illustration. The WPI data is drawn from RBI website (www.rbi.org.in) and its weighted average is computed to match the commodity groups of the I-O table. The earliest year that WPI is available is 2005. The inflation rate is calculated as: INF=log(WPI_{2011})–log(WPI_{2005})

Commodity group	BL	INF	Commodity group	BL	INF
Paddy	3.55	0.49	Silk textiles	0.25	0.20
Wheat	1.68	0.53	Art silk, synthetic fiber textiles	2.40	0.35
Jowar	0.22	0.80	Jute, hemp, mesta textiles	0.98	0.48
Bajra	0.10	0.54	Carpet weaving	0.31	0.32
Maize	0.24	0.57	Readymade garments	3.14	0.08
Gram	0.22	0.52	Furniture and fixtures-wooden	0.89	0.48
Pulses	0.98	0.69	Wood and wood products	0.49	0.40
Sugarcane	0.35	0.50	Paper, paper prods. & newsprint	1.39	0.24
Groundnut	0.31	0.64	Printing and publishing	1.12	0.25
Coconut	0.12	0.29	Leather footwear	0.39	0.28
Other oilseeds	0.67	0.50	Leather and leather products	0.93	0.14
Jute	0.03	0.57	Rubber products	2.22	0.38
Cotton	0.45	1.01	Plastic products	3.62	0.19
Tea	0.03	0.55	Inorganic heavy chemicals	2.04	0.24
Coffee	0.04	0.84	Organic heavy chemicals	1.75	0.25
Rubber	0.08	1.25	Fertilizers	2.34	0.22
Tobacco	0.03	-0.12	Pesticides	0.63	0.11
Fruits	0.22	0.63	Paints, varnishes and lacquers	0.90	0.18
Vegetables	0.29	0.52	Drugs and medicines	2.14	0.16
Other crops	1.62	0.75	Soaps, cosmetics & glycerin	1.60	0.33
Milk and milk products	0.86	0.61	Synthetic fibers, resin	1.42	0.28
Animal services(agricultural)	0.74	0.53	Other chemicals	1.35	0.27
Poultry & Eggs	0.61	0.56	Structural clay products	1.47	0.37
Fishing	1.10	0.79	Cement	2.49	0.40
Coal and lignite	0.57	0.43	Iron and steel casting & forging	2.84	0.23
Crude petroleum	0.67	0.88	Hand tools, hardware	1.62	0.21
Iron ore	0.18	1.63	Tractors and agri. implements	1.36	0.24
Manganese ore	0.01	0.64	Industrial machinery(F & T)	1.07	0.04
Bauxite	0.01	0.49	Industrial machinery(others)	0.87	0.26
Copper ore	0.01	0.57	Machine tools	2.18	0.31
Other metallic minerals	0.05	0.23	Electrical industrial Machinery	3.36	0.22
Lime stone	0.05	0.29	Electrical wires & cables	1.49	0.25
Other non metallic minerals	0.18	0.59	Batteries	0.40	0.22
Sugar	0.85	0.44	Electrical appliances	1.00	0.12
Khandsari, boora	0.33	0.41	Communication equipments	2.04	-0.04
Hydrogenated oil(vanaspati)	0.47	0.26	Electronic equipments(incl.TV)	2.19	-0.14
Edible oils other than vanaspati	2.64	0.35	Ships and boats	3.27	-0.05
Tea and coffee processing	1.74	0.45	Rail equipments	1.89	0.18
Beverages	1.05	0.09	Motor cycles and scooters	1.73	0.25
Tobacco products	0.94	0.53	Bicycles, cycle-rickshaw	0.64	0.38
Khadi, cotton textiles	0.24	0.19	Watches and clocks	0.29	0.11
Cotton textiles	3.29	0.39	Medical, precision&optical instru.s	1.17	0.15
Woolen textiles	0.38	0.30	Miscellaneous manufacturing	2.80	0.19

Table A.23: **Estimated forward and backward linkages overtime.** The estimation are based on original I-O tables of Indian economy for different years. The classification is based on Dholakia et al. (2009)

		Forward Linkages					Backward Linkages				
	Commodity	1989	1993	1998	2003	2007	1989	1993	1998	2003	2007
1	Food Crops	1.26	1.24	1.22	1.48	1.46	1.66	1.65	1.62	2.03	1.90
2	Cash Crops	2.06	1.89	1.89	2.58	2.41	1.40	1.39	1.41	1.62	1.63
3	Plantation Crops	1.49	1.80	2.60	1.61	2.06	1.25	1.26	1.31	1.46	1.52
4	Other crops	1.47	1.80	1.86	1.82	1.83	1.37	1.42	1.35	1.49	1.47
5	Animal Husbandry	1.34	1.36	1.38	1.46	1.78	1.48	1.78	1.73	1.78	1.88
6	Forestry and logging	1.80	1.77	1.78	1.79	2.01	1.20	1.20	1.19	1.20	1.31
7	Fishing	1.10	1.10	1.08	1.09	1.18	1.19	1.32	1.27	1.33	1.33
8	Coal and lignite	3.02	3.91	4.10	4.34	5.42	1.55	1.74	1.54	1.46	1.49
9	Crude petroleum, natural gas	6.55	7.84	8.12	12.90	15.00	1.28	1.30	1.25	1.62	1.65
10	Iron ore	1.60	2.53	3.05	2.12	2.07	1.46	1.62	1.54	1.51	1.33
11	Other Minerals	5.72	6.91	8.68	5.96	4.08	1.26	1.47	1.42	1.59	1.41
12	Sugar	1.18	1.15	1.18	1.32	1.45	2.11	2.15	2.15	2.40	2.51
13	Food Products Excluding Sugar	1.14	1.21	1.21	1.31	1.60	2.00	2.37	2.36	2.52	2.60
14	Beverages	1.12	1.10	1.16	1.25	1.32	1.60	2.08	2.10	2.44	2.64
15	Tobacco products	1.04	1.05	1.08	1.10	1.12	1.73	2.03	2.04	1.85	1.86
16	Cotton textiles	1.53	1.56	1.55	1.46	1.37	2.40	2.18	2.26	2.42	2.30
17	Wool, Silk and synthetic fibre textiles	1.54	1.46	1.47	1.54	1.81	2.39	2.37	2.38	2.56	2.66
18	Jute, hemp, mesta textiles	3.84	2.48	2.52	2.40	2.51	2.84	2.22	2.13	2.11	2.20
19	Textile Products including wearing apparel	1.32	1.31	1.28	1.14	1.39	2.50	2.31	2.29	2.45	2.48
20	Wood and wood products	2.10	2.52	2.56	3.89	2.37	1.43	1.81	1.81	1.84	2.00
21	Furniture and fixtures-wooden	1.40	1.60	1.47	1.22	1.56	1.32	1.81	1.87	2.01	2.09
22	Paper, paper prods. & nesprint	3.11	3.25	3.24	3.29	3.12	2.37	2.51	2.48	2.49	2.56
23	Printing and publishing	1.82	1.66	1.69	2.08	2.12	2.04	2.26	2.20	2.40	2.64
24	Leather and leather products	1.28	1.42	1.49	1.42	1.82	1.96	2.37	2.42	2.28	2.54
25	Plastic and Rubber Products	1.46	2.07	2.03	2.35	2.21	1.83	2.42	2.41	2.62	2.68
26	Petroleum products	2.90	2.88	2.85	2.67	2.05	2.01	2.17	2.09	2.23	2.42
27	Coal tar products	3.39	4.21	4.49	3.43	3.19	2.09	2.60	2.42	2.40	2.16
28	Inorganic heavy chemicals	3.78	4.44	4.28	3.75	4.54	2.15	2.27	2.26	2.49	2.87
29	Organic heavy chemicals	3.68	4.57	4.05	4.14	4.10	2.80	2.39	2.30	2.48	2.75
30	Fertilizers	3.28	2.77	2.94	2.90	4.53	2.67	2.54	2.38	2.56	3.32
31	Paints, varnishes and lacquers	2.35	2.44	2.42	2.95	3.40	2.21	2.41	2.39	2.54	2.48
32	Pesticides, Drugs and other chemicals	1.87	2.45	2.36	2.49	2.67	1.94	2.49	2.34	2.44	2.69
33	Cement	2.15	2.20	2.20	2.37	2.35	2.20	2.32	2.21	2.29	2.28
34	Non-mettalic mineral products	1.87	1.72	1.77	2.37	2.29	2.53	2.10	2.04	2.27	2.34
35	Iron and steel indsutries and foundries	2.88	3.13	3.16	2.59	2.80	2.45	2.78	2.52	2.48	2.60
36	Non-ferrous basic metals	4.45	3.58	5.27	5.91	5.68	2.88	2.54	2.44	2.50	2.68
37	Metal products except Mach and transport equipment	2.77	2.10	2.08	2.41	2.53	2.97	2.46	2.36	2.43	2.69
38	Tractors and agri. implements	2.30	1.41	1.28	1.22	1.28	3.01	2.57	2.52	2.61	2.70
39	Industrial machinery	2.01	2.01	1.87	1.91	2.01	2.77	2.57	2.53	2.58	2.95
40	Other machinery	2.04	1.82	1.86	2.09	2.26	3.01	2.45	2.42	2.51	2.77
41	Electrical, electronic machinery and appliances	1.86	1.76	1.79	1.94	2.34	2.72	2.46	2.46	2.79	2.85
42	Rail equipments	3.53	1.54	1.42	3.77	3.08	2.71	2.55	2.56	2.44	2.74
43	Other transport equipment	1.36	1.65	1.49	1.57	2.21	2.04	2.41	2.45	2.58	3.06
44	Miscellaneous manufacturing industries	1.49	2.03	1.81	2.10	2.05	1.63	2.25	2.46	2.31	2.63
45	Construction	1.19	1.24	1.18	1.29	1.35	1.97	2.20	2.04	2.28	2.32
46	Electricity	2.67	3.13	3.21	3.02	3.14	1.76	2.24	2.29	2.25	2.18
47	Gas and water supply	2.15	1.69	1.67	1.67	1.67	2.18	1.51	1.51	2.56	1.88
48	Railway transport services	2.27	2.64	2.72	2.38	2.31	2.02	1.92	2.08	2.19	1.92
49	Other transport services	1.61	1.94	1.80	1.88	2.07	1.49	2.04	2.06	2.03	2.08
50	Storage and warehousing	2.33	2.88	2.83	2.84	3.38	1.51	1.59	1.81	2.55	1.90
51	Communication	1.93	2.30	2.02	2.51	2.43	1.26	1.32	1.28	1.53	1.59
52	Trade	1.67	1.85	1.84	1.97	2.07	1.37	1.46	1.38	1.35	1.39
53	Hotels and restaurants	1.19	1.33	1.31	1.37	1.72	1.95	2.03	2.13	2.23	2.35
54	Banking	1.97	2.62	2.40	2.60	2.56	1.19	1.23	1.37	1.35	1.30
55	Insurance	3.09	2.92	2.27	2.44	2.47	1.25	1.33	1.53	1.59	1.48
56	Ownership of dwellings	1.00	1.00	1.00	1.00	1.00	1.20	1.12	1.14	1.11	1.10
57	Education and research	1.00	1.01	1.01	1.01	1.12	1.14	1.22	1.23	1.20	1.20
58	Medical and health	1.18	1.23	1.02	1.03	1.05	1.69	2.10	2.37	1.76	1.88
59	Other services	1.83	1.88	2.07	1.67	1.76	1.83	1.93	1.98	1.54	1.55
60	Public administration	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Appendix B

Proofs of the Theoretical Results

Finding the Market Equilibrium

From (2.6), the F.O.C. result

$$x_{mn} = \frac{\alpha_{mn} p_n}{p_m} x_n, \qquad v_n = \alpha_n p_n x_n \tag{B.1}$$

we can find the prices by substituting (B.1) in the production function

$$p_n = (\frac{1}{\alpha_n})^{\alpha_n} \prod_{m=1}^{N} (\frac{p_m}{\alpha_{mn}})^{\alpha_{mn}} \tag{B.2}$$

By taking logarithm we obtain

$$\ln p_n = \sum_{m=1}^{N} \alpha_{mn} \ln p_m + a_{0n} \tag{B.3}$$

where $a_{0n} = -\alpha_n \ln \alpha_n - \sum_m \alpha_{mn} \ln \alpha_{mn}$. The supply-demand identity for each sector becomes

$$x_n = x_{0n} + \sum_{m=1}^{N} x_{nm} \tag{B.4}$$

where x_{0n} is the representative final demand of product n. Therefore, to find the equilibrium production, we need to model how final demand looks like. Assuming a Cobb-Douglas utility and wealth $w' = (1+t)w$ for a representative final consumer, the utility maximization problem becomes

$$\max u(x_1, \ldots, x_N) = \prod_{i=1}^{N} x_i^{\beta_i}, \qquad \text{s.t.} \ \sum_{i=1}^{N} (1+t) p_i x_i = w' \tag{B.5}$$

where $\sum_{i=1}^{N} \beta_i = 1$. This gives final demand for each product as $x_{0n} = w\beta_n/p_n$. By substituting x_{nm} from F.O.C. (B.1) in (B.4), it turns out

$$x_n = w\beta_n/p_n + \sum_{m=1}^{N} \frac{\alpha_{nm}p_m}{p_n} x_m, \quad \Rightarrow \quad p_n x_n = w\beta_n + \sum_{m=1}^{N} \alpha_{nm} p_m x_m \tag{B.6}$$

Define $y_n = p_n x_n$ as the value of total production in sector n. Then, (B.6) is written as

$$y_n = w\beta_n + \sum_{m=1}^{N} \alpha_{nm} y_m \tag{B.7}$$

In order to formulate the expressions in matrix form, define $\ln P$, B, A_0, and Y as $N \times 1$ vectors comprising $\ln(p_n)$, β_n, a_{0n}, and y_n respectively, and matrix $\mathbf{A}$ with dimensionality $N \times N$ in the sense that the element $\mathbf{A}_{(n,m)} = \alpha_{nm}$. Then, we can write (B.3) as

$$\ln P = \mathbf{A}' \ln P + A_0, \quad \Rightarrow \quad \ln P = (\mathbf{I} - \mathbf{A}')^{-1} A_0 \tag{B.8}$$

where $\mathbf{I}$ is the identity matrix, and $\mathbf{A}'$ is the transpose of $\mathbf{A}$. In addition, (B.7) becomes

$$Y = wB + \mathbf{A}Y, \quad \Rightarrow \quad Y = w(\mathbf{I} - \mathbf{A})^{-1} B \tag{B.9}$$

As explained in section 2.3, when the VAT is introduced to the market, the only change in the model is the transformation of w into $w/(1+t)$. Hence, all sectors' production value decline by a homogeneous factor $1/(1+t)$ as long as the consumer preferences are not affected by the tax.

Proof of Proposition 1

Profit maximization in the VAT give the same results as the model without tax. In the ST, the FOCs of (2.10) become

$$x_{mn} = \frac{\alpha_{mn}p_n}{(1+t)p_m} x_n, \qquad v_n = \alpha_n p_n x_n \tag{B.10}$$

and we obtain

$$\ln p_n = \sum_{m=1}^{N} \alpha_{mn} \ln p_m + a_{0n} + (1-\alpha_n)\ln(1+t) \quad \Rightarrow \quad \ln P = (\mathbf{I} - \mathbf{A}')^{-1}\big(A_0 + (J - A_v)\ln(1+t)\big) \tag{B.11}$$

where J and A_v are $N \times 1$ vectors containing 1s (summation vector) and α_n. In summary, we can compare the two systems as follows

Tax system	$\ln P$ (Price vector)
VAT	$(\mathbf{I} - \mathbf{A}')^{-1} A_0$
ST	$(\mathbf{I} - \mathbf{A}')^{-1}\big(A_0 + (J - A_v)\ln(1+t)\big)$

Distortions in prices under a cascading ST are given by

$$\Delta \ln P = (\mathbf{I} - \mathbf{A}')^{-1}(J - A_v)\ln(1+t) \tag{B.12}$$

Given the expansion of $(\mathbf{I} - \mathbf{A}')^{-1} = \mathbf{I} + \mathbf{A}' + \mathbf{A}'^2 + \ldots$, we can write this for each sector in the extensive form as

$$\Delta \ln p_n = \ln(1+t)\Big(1 - \alpha_n + \sum_{m=1}^{N}(1-\alpha_m)\alpha_{mn} + \sum_{m=1}^{N}\sum_{i=1}^{N}(1-\alpha_m)\alpha_{mj}\alpha_{jn} + \ldots\Big) \tag{B.13}$$

Proof of Proposition 2

The production identity for each sector becomes

$$p_n x_n = w\beta_n + \frac{1}{1+t}\sum_{m=1}^{N}\alpha_{nm}p_m x_m \quad \Rightarrow \quad Y = w(\mathbf{I} - \frac{1}{1+t}\mathbf{A})^{-1}B \tag{B.14}$$

In summary, we can compare the two systems as follows

Tax system	Y (Output vector)
VAT	$w(\mathbf{I} - \mathbf{A})^{-1}B$
ST	$w(\mathbf{I} - \frac{1}{1+t}\mathbf{A})^{-1}B$

The only term of output vector that changes in cascading ST is $(\mathbf{I} - \mathbf{A})^{-1}$. Thus,

$$\Delta Y = w\Delta\big[(\mathbf{I} - \mathbf{A})^{-1}\big]B \tag{B.15}$$

To compute the variation in $(\mathbf{I} - \mathbf{A})^{-1}$, when the scalar γ changes, using matrix calculus, one can show that

$$\frac{\Delta(\mathbf{I} - \gamma\mathbf{A})^{-1}}{\Delta\gamma} = (\mathbf{I} - \gamma\mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \gamma\mathbf{A})^{-1} \tag{B.16}$$

Therefore when γ decreases from 1 to $1/(1+t)$, we have

$$\Delta(\mathbf{I} - \mathbf{A})^{-1} = \frac{-t}{1+t}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1} \tag{B.17}$$

$$\Delta Y = \frac{-tw}{1+t}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1}B \tag{B.18}$$

$(\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1}B$ is a measure of length of linkages to final consumers. It is forward-looking similar to FL – due to the common term $(\mathbf{I} - \mathbf{A})^{-1}$ – compared to BL which is backward-looking due to the term $(\mathbf{I} - \mathbf{A}')^{-1}$. If we expand the formula, we have

$$\text{LL} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1}B = (\mathbf{A} + 2\mathbf{A}^2 + 3\mathbf{A}^3 + \cdots + n\mathbf{A}^n + \dots)B \tag{B.19}$$

and in extensive form

$$ll_i = \sum_{k=1}^{N} \alpha_{nk}\beta_k + 2\sum_{k=1}^{N}\sum_{i=1}^{N} \alpha_{nk}\alpha_{ki}\beta_i + 3\sum_{k=1}^{N}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_{nk}\alpha_{ki}\alpha_{ij}\beta_j + \dots \tag{B.20}$$

Chapter 3

The Intensive Margin of Informality: Unreported labour, Inequalities, and Taxation[1]

3.1 Introduction

The intensive margin of informality, defined as the informal employment of workers by formally registered firms, is empirically important even in developed countries. The OECD estimates unreported labour to be responsible for a 30% loss in social security contributions in Hungary, Mexico and South Korea, and for a shortfall larger than 20% in Italy, Turkey, Poland and Spain[2]. This phenomenon is not limited to countries famous for the size of their informal sector. Indeed, the French Central Agency for Social Security (ACOSS)[3] randomly audits every year different sectors of the economy to gather information on the extent of fraud and the profile of dodgers. The hotel, restaurant, and catering industry was for example investigated in 2004. Among interviewed employees, 19% were found not to be declared by their employer. Additionally, sectors with high skill intensity such as the real estate and the legal sectors were found to exhibit fraud rates above average.

Under both the intensive and the extensive margin of informality, employers do not abide by their legal obligations regarding informal employees. On the one hand, firms benefit from increased flexibility and cheaper labour costs. On the other hand, employees do not receive the social security coverage or pension benefits to which they are entitled. Furthermore, both margins also have similar consequences for public finances. Low tax revenues allow only for poor quality of public services and weak rule of law. In turn, the advantages of being part of the formal economy decrease, which depletes further tax revenues. Nevertheless, implications of informality at the intensive margin on

[1]Chapter 3 is based on Briand (2014).

[2]OECD. (2015), *OECD Employment Outlook 2004*, OECD Publishing, Paris.

[3]Fraude sociale et dispositif de répression num. 2007-01, Septembre 2007, Research conducted by the French Central Agency for Social Security (ACOSS)

firms' growth opportunities radically differ from those at the extensive margin. At the difference of unregistered entrepreneurs, non-complying registered employers are not constrained in terms of credit and investment opportunities. In this Chapter, I investigate to which extent inducing greater compliance within formal firms would impact the incentives of firms to register formally. I present empirical evidence on some key patterns of non-compliance by formal entrepreneurs with labour regulations in India and build a theoretical model to explain them. Then, I show that the intensive margin of informality does not necessarily evolve hand-in-hand with formal registration following a given change in the taxation and regulatory environment. Hence, ignoring the impact of a reform on the share of unreported labour within formal firms does not simply lead to an under-estimation of the consequences of informality on the economy. Nevertheless, the intensive margin of informality has been ignored by the literature on taxation and informality at the exception of Kumler et al. (2013) that empirically estimate the importance of payroll taxes evasion in Mexico.

The literature (Helpman et al. (2010), Egger and Kreickemeier (2012)) has underlined that growing inequalities *within* and *between* groups of economic agents are linked to two types of heterogeneity. First, heterogeneity in productivity *across* firms leads to employment policies differing widely from one employer to the other. Indeed, informal labour is used in larger quantities by the least productive formal entrepreneurs.[2] Second, heterogeneity in productivity can also be observed *within* firms across segments of production (Bernard et al. (2011)). It follows that employment policies vary within firms as well, creating incentives for a given firm to hire both formal and informal workers. Empirical studies found that informal workers are employed in segments of production where their employer expertise is lower, or for which demand is only limited or irregular.[2] The first contribution of this Chapter is to take heterogeneity in productivity *within* and *across* firms into account to explain observed patterns of unreported labour in formal firms. Firms produce a good characterized by a continuum of technical and commercial *features*. This industry could be, for instance, the catering industry: restaurants produce a menu characterized by features (starters, main dish...) of which they provide different varieties (tiramisu vs. cheesecake). Crucially, an entrepreneur expertise varies across features. Each feature is produced under increasing return to scale and monopolistic competition as in Melitz (2003). Feature productivity is given by the product of entrepreneurial ability and feature-expertise. Hence, productivity varies within a given firm, creating incentives to hire informal labour in some segments of production and not in others. Indeed, the use of informal labour lowers production costs but entails non-compliance costs that are proportional to revenues.

Consequently, informal labour is more profitable for features for which productivity, revenues, and thus non-compliance costs, remain limited. Finally, larger firms are easier to monitor[4], therefore face higher non-compliance costs and exhibit lower share of informal employment.

The second contribution of this Chapter is to identify for each occupational choice, first the entrepreneurial ability of the individuals having embraced it, and second, the impact of informality on those individuals' welfare. Agents are heterogeneous in entrepreneurial ability and self-select into the occupation allowing them to raise the highest possible income. Imperfect monitoring of compliance enables both types of occupations to be performed formally or informally. The allocation of agents endogenously arising from my model is in accordance with empirical evidence (Banerjee and Duflo (2011); Henrekson and Roine (2005); Perry et al. (2007)). In equilibrium, informal workers have lower ability than formal workers. Besides, two types of informal self-employed arise. "Chosen" self-employment gathers individuals whose level of entrepreneurial ability is high enough for their outside option to be above the wages offered on the market, but not high enough to survive as a formal entrepreneur. By opposition, "forced" self-employment appears when a labour market friction - such as a binding minimum wage - causes labour supply to exceed labour demand. The least skilled agents are then battered into self-employment. As requested by several field studies (Blunch et al. (2001); Perry et al. (2007)), this model reconciles the two strands of literature respectively presenting informal self-employment as chosen or forced.

Finally, I use the model to describe the impact of different tax reforms and labour policies on both margins of informality and on the welfare of each type of agents in the economy.

First, keeping total expenditures on the formally produced good constant, higher formal registration costs decrease firms' formal registration rates but do not impact the intensive margin of informality. Entrepreneurs with the lowest ability level drop out of the formal industry, because their profits are not sufficient to cover the newly increased fixed formal registration costs. Hence, competition eases on the multi-features good market. Surviving formal firms therefore expand, which compensates losses in employment caused by the reduction in the number of formal firms. It is worth noticing that the decrease in welfare caused by higher formal registration costs is attenuated by the presence of within-firm informality. Indeed, unreported labour increases the profitability of all firms and allows them to compensate partially for higher formal costs. Second, following an increase in payroll taxes, formal firms switch the production of some features from formal to informal. Higher payroll

[4] *World Development Report 2005*, Publications of The International Bank for Reconstruction and Development (The World Bank)

taxes cause a decrease in overall profitability in all firms that pushes the least productive formal entrepreneurs out of the market. Again, the decrease in aggregate formal profits is attenuated by the existence of within-firm informality.

Second, fiscal capacity of the government determines whether higher payroll taxes lead to an increase or a decrease in the welfare of informal workers. Fiscal capacity is defined as the aptitude of the government to collect a share of the formal sector's revenues. It increases with the strength of law enforcement and decreases with the level of pre-existent taxes. This Chapter also links fiscal capacity to the distribution of revenues across different segments of production of formal firms: the more dispersed are those revenues, the more sensitive is a formal firm to a decrease in profitability (higher payroll taxes) of its most profitable (formal) segments of production. When fiscal capacity is high, formal firms' attempt to compensate higher payroll taxes by a more widespread use of informal labour is not successful, in the sense that the share of total revenues (formal and informal) allocated to tax payments still rises. On the contrary, when fiscal capacity is low, formal firms successfully compensate for higher formal costs through the intensive margin of informal employment and the share of total revenues that they allocate to tay payments decreases. Therefore, the increase in informal employment is larger when fiscal capacity is low and the real informal wage increases, not only the nominal informal wage as in the high fiscal capacity case.

Section 3.2 discusses the relation of this Chapter to the existing literature.Section 3.3 presents descriptive statistics on the intensive margin of informality in India. Section 3.4 presents a model where the intensive and extensive margin of informality endogenously arise. Section 3.5 solves for the economy equilibrium and section 3.6 analyzes the effects of different tax and policy instruments on welfare and on both margins of informality.

3.2 Literature

The literature on the intensive margin of informality - also referred to as within-firm informality, unreported labour in formally registered firms, or non-compliance with payroll taxes - is limited Nyland et al. (2006) evaluate the role played by different firm characteristics on the outcomes of social security contributions fraud in Shanghai Kleven et al. (2009) develop a theoretical model where tax evasion - and wage under-reporting in particular - arises from the collusion of workers and employers. In their setting, larger firms optimally evade less, because collusion is harder to sustain in a large labour force for two reasons. First, workers are independently subject to random shocks triggering them to "blow the whistle", therefore a larger number of workers translates into a higher probability that one of them becomes a whistle blower. Second, rewards for whistle blowers in their model is proportional to the amount of evaded taxes, in which case workers have a stronger incentive to expose evasion in larger firms. Furthermore, larger firms must comply with more numerous and complex administrative requirements. Concealing fraud by maintaining all formal statements consistent one with another is therefore more difficult Kumler et al. (2013) confirm this result by empirically evaluating the sensitivity of payroll tax compliance in Mexico to changes in employees incentives to monitor their employers (third-party reporting). Tonin (2011) also exploits the idea that tax evasion stems from collusion between employers and firms in a theoretical model used to evaluate the impact of a minimum wage regulation on the welfare of heterogeneous agents in productivity. These authors endogenously derive that enforcement is increasingly efficient when firms grow larger. Ulyssea (2013) and the model presented in this Chapter build on this last result when assuming a reduced form of non-compliance costs that increases with a firm's revenues. Ulyssea (2013) calibrates a dynamic model of within and full informality to Brazilian data and simulate various formalization policies. However, he does not explain why some workers accept to receive a wage lower than formal employees performing the same task within the same firm. Furthermore, these Chapters explain only the overall share of informality in formally registered firms, while the model in this Chapter also explains the allocation of formal and informal workers within firms.

The present Chapter is connected to the literature initiated by Rauch (1991) where informality is an optimal answer to a given economic and institutional environment. In his model, full informality arises due to the presence of labour market frictions such as a minimum wage regulation. The author builds on the work by Lucas (1978), in which agents have heterogeneous managerial ability and choose between being entrepreneurs or employed. In Pavcnik et al. (2004), firms prefer to employ informal workers because they pay them only to their reservation wages, while formal wages are increasing in the probability of being fired. According to Paz et al. (2012), however, these two labour market distortions - minimum wage and protection against layoffs - are not empirically confirmed as causing informality, while it is the case of payroll taxes. His model is inspired by Melitz (2003)'s model of trade and by the efficiency wage model of Shapiro and Stiglitz (1984). Those are only a few papers of an important literature attempting to model the choice of entrepreneurs between formal and informal sectors based on the characteristics of the economy (taxation structure, labour market frictions). Their scope is limited to the extensive margin of informality. The variety of identified possible vectors of informality is a testimony of the phenomenon's complexity Albrecht et al. (2009) and Meghir et al. (2012) break with the traditional view of segmented informal and formal labour markets. They allow firms to locate either in the informal or in the formal labour market, while workers are searching for jobs in both sectors and can therefore switch from one sector to another. However, firms are again assumed to be fully compliant once formally registered: the intensive margin of formality is not considered. Their approach is then complementary to the one taken in this Chapter, where the fixed costs inherent to the use of the production technology with increasing returns to scale are assumed to be too large for firms to be able to fully escape detection by tax authorities. This assumption could be relaxed by the assumption that firms in the increasing return to scale sector could avoid registering if their size remains below a certain threshold. However, the focus of this Chapter is the intensive margin of formality and such an assumption would complicate the model without adding to the analysis.

I build on the optimal taxation literature that analyzes the interactions between tax instruments and full informality. Emran and Stiglitz (2005) introduce an informal sector in the standard model of trade tax reforms. They demonstrate that VAT creates inter-sectoral distortions between the formal and the informal sector. These authors derive the same conditions for welfare enhancing trade reforms as the model of tariffs and VAT reforms of Michael et al. (1993) that assumed no informal sector. Emran and Stiglitz (2005) show that these conditions are implausible in the presence of informality. Keen (2008) highlights an important property of the VAT. Only the formal sector is able to recover the VAT paid on intermediate goods and the chain of crediting and refunding inherent to the use of VAT is broken by informal firms. In these cases, the VAT becomes the equivalent of an input tax for the informal sector: it allows the government to indirectly extract revenues from the informal sector. Most authors investigating the effect of the VAT on informality reduce the welfare analysis to the maximization of the utility of a representative consumer. The informal sector is modeled as a standard sector of the economy, only with the additional property of avoiding some taxes. Its profits are directly integrated in the social welfare function, where they are more or less discounted in function of their social value. Keen (2008) leaves the question of the optimal discount parameter open. However, in the case of a representative consumer, it can be greater or inferior to the unity, but not both. Piggott and Whalley (2001) departed from this approach by offering a more nuanced vision of informality and its consequences. They analyze the welfare consequences of a VAT base broadening in a developed country (Canada). Their model features heterogeneous households: rich and poor. Non-market service goods are provided exclusively by the poor to the rich household. Based on this assumption, Piggott and Whalley (2001) show that such VAT reforms can be welfare improving for the poor, even if welfare worsening at the aggregate level. The "terms of trade" between rich and poor improve, because poor people now informally sell at the gross of the tax price to richer people and, therefore, ripe increased profits. Piggott and Whalley (2001) call it the "distributional effect" of an incomplete VAT coverage.

Therefore, the welfare implications of tax reforms in the presence of informality are still to be understood. The broader the considered definition of informality, the more complex the welfare analysis becomes. Indeed, the definition by Piggott and Whalley (2001) is still restrictive: the informal sector would employ only the poor and produce only for the rich. Furthermore, the size of each group is exogenous in their model. As explained in the introduction of this Chapter, a more recent literature (Helpman et al. (2010), Egger and Kreickemeier (2012)) has studied more thoroughly inequalities between and within-groups of an economy. In particular, labour market frictions have been showed to explain the empirically observed heterogeneity in firms employment policies, and therefore the rising inequalities between and among entrepreneurs and workers. For instance, in Tonin (2011), the share of undeclared labour of each employee is endogenously determined and it is therefore possible to identify the consequences of a minimum wage regulation on the welfare of each range of agents productivity.

Finally, this chapter relates to the literature using heterogeneity of productivity across and within firms to generate their optimal positioning on the industry value chain: diversification in the production of several goods or optimal share of off-shored production. First, Grossman and Rossi-Hansberg (2008) shed light on heterogeneity of expertise in different segments of production within firms. Entrepreneurs decide on which part of their product's global value chain to position themselves and in which country to locate each active segment of production. Second, Bernard et al. (2011) model an economy where firms can produce more than one good and exhibit different level of productivity in each good. This framework is very flexible and can be reinterpreted as the production of a good arising from bundling different features together. At the difference of the paper by Grossman and Rossi-Hansberg (2008), where a continuum of tasks must be performed exactly once for the final good to be consumed, in Bernard et al. (2011), all features need not to be produced for the firm to enter the market.

3.3 Descriptive Statistics: the Intensive Margin of Informality in India

In this section, I quantify the non-compliance of formally registered Indian firms with labour and social security regulations by making an innovative use of three different rounds of a large household survey of the Indian workforce. Apart from the work performed by the French Central Agency for Social Security (ACOSS)[5] in France, there is limited evidence of the magnitude of non-compliance within formal firms. Nevertheless, the sustainability of the social security system and its ability to shield workers from adverse shocks relies on compliance by employers, while enforcement is limited. The 61st, 66th and 68th rounds of the Employment and Unemployment Survey (EUS) of the National Sample Survey (NSS) cover a large variety of topics related to the characteristics of every surveyed household's member (occupation, education and financial situation) for 2004-05, 2008-09 and 2011-12 respectively. Employed workers are defined as regular salaried employees, casual wage labour, and unpaid workers. The EUS contains detailed information on employment conditions (income, type of job contract, availability of social security benefits and method of payment) and the main characteristics of employers (industry, location of workplace, enterprise type, number of workers in the enterprise).

We take advantage of the fact that the main social security and pension laws[6] in India use these same criteria to determine the employers for which it is mandatory to declare and contribute to their workers' coverage. First, the Employee's State Insurance (ESI) is a self-financing social security and health insurance program. An autonomous body created under the ESI Act 1948 manages this fund: the Employee's State Insurance Corporation (ESIC). Workers registered under this scheme are entitled to medical treatment for themselves and for their dependent family members. Additionally, women benefit from unemployment cash benefit and from maternity benefits. The EUS includes information on each one of these benefits. Any business employing more than 10 people is required to register all workers under the ESI scheme. Furthermore, in the case of workers receiving more than Rs. 15,000 (USD 220) per month, the employer and employee respectively contribute 4.75% and 1.75% of the total wage to the scheme. Second, the Employees' Provident Fund Organisation (EPF) is a compulsory contributory pension and insurance fund that is ruled according to three different regulations: the 1952 Employee's Provident Fund Scheme, the 1976 Employee's Deposit

[5] Fraude sociale et dispositif de répression num. 2007-01, Septembre 2007, Research conducted by the French Central Agency for Social Security (ACOSS)

[6] Website of the Ministry of Labour and Employment of the Government of India

Linked Insurance Scheme and the 1995 Employee's Pension Scheme (replacing the 1971 Employee's Family Pension Scheme). It is the largest social security organisation in India and manages assets valued at USD 128 billion in March 2016.[7] The EPF rules regarding the continuity of membership between employment periods and withdrawal before retirement age have been amended in 2016 but remained unchanged between 2004 and 2011. All establishments employing more than 20 workers must be registered with the EPF. For workers earning less than Rs. 6,500 (USD 95) per month (this threshold has been increased to Rs. 15,000 in 2014), both the employee and the employer must make a contribution of 12% of the total wage to the fund. For workers earning more than the EPF mandatory registration threshold, contribution to the fund is optional.

Hence, from this Survey, without the question of illegality ever being raised with enterprises or employees, we can deduce the formal status of 1) the employer and 2) the formal status of the workers employed within formal firms. In order to determine the formal status of an employer, we make the following assumptions:

- Government bodies, public or private limited companies, cooperative, and non-profit institutions are formal employers
- Proprietary partnerships with members of a different household and employing more than 20 workers are formal employers
- Employers providing employees with a written contract are formal employers

Once it is determined that a worker is employed by a formal firm, we make the following assumptions to determine the compliance status of the employer with respect to her employee:

- An employee is eligible for health care and maternity benefits if she is eligible for ESI that is to say if she works in an establishment employing more than 10 people and earns less than Rs. 15,000
- An employee should be registered in a pension fund if she is eligible for registration within EPF that is to say if she works in an establishment employing more than 20 people and earns less than Rs. 6,500

[7] *EPFO to invest more in government bonds amid corporate loan defaults.* The Economic Times. Retrieved 2016-03-18

3.4 Model

3.4.1 Overview

The economy is characterized by two sectors of production and by a measure M of individuals that are heterogeneous in entrepreneurial ability. Agents select among four options the occupation that promises them the highest possible income. They choose whether to be workers or entrepreneurs. Both occupations can be performed formally or informally.

In the first sector, firms use an increasing return to scale (IRS) technology, with fixed production costs and decreasing marginal costs in productivity as in Melitz (2003). There is monopolistic competition in the production of a good characterized by a finite continuum of features. Entrepreneurs choose whether to produce each of those features. In the affirmative case, an entrepreneur produces a unique variety of each feature. Productivity is given by the product of the entrepreneur's ability and feature-expertise. Therefore, it varies *across* and *within* firms. Based on his feature-productivity, an entrepreneur chooses whether to employ workers formally or informally. The use of informal labour is less costly than the use of formal labour, but it implies non-compliance costs proportional to revenues.

In the second sector, a basic good is produced using a constant return to scale (CRS) technology. The productivity of a firm depends on the ability of its employees. On the contrary, in the IRS sector, entrepreneurial ability and expertise determine the productivity of the firm independently of workers ability. This captures the difference between a craftsman using a hand-weaving loom and the assembly line textile industry. An employer's decision to hire a given employee informally or formally does not depend on the employee's characteristics in this model, only on the employer's. However, entrepreneurial ability is positively correlated with the ability to find employment. Therefore, it plays a role in determining the type of workers willing to accept an informal job.

The government collects taxes and designs labour regulations. Formal registration with the government entails some fixed administrative costs F for firms. Due to limited public resources, compliance with the law is only imperfectly monitored. Formally registered firms can then use informal labour. The resulting economy is illustrated in figure 3.1. To simplify computations, the following assumption is made:[8]

[8]The second part of `assumption 1` simply makes sure that all firms in the IRS sector register formally, and do not give up on producing some - profitable - features using formal labour in order to avoid fixed formal registration costs. If `assumption 1` were not verified, the fully informal sector would also include firms in the IRS sector that employ only informal labour, because the additional profits generated by formally expanding above the detection threshold of the government would not be sufficient to cover the formal registration costs F. Employment in the fully informal sector is therefore under-estimated in the present model. `Assumption 1` could be relaxed to give rise to a more realistic

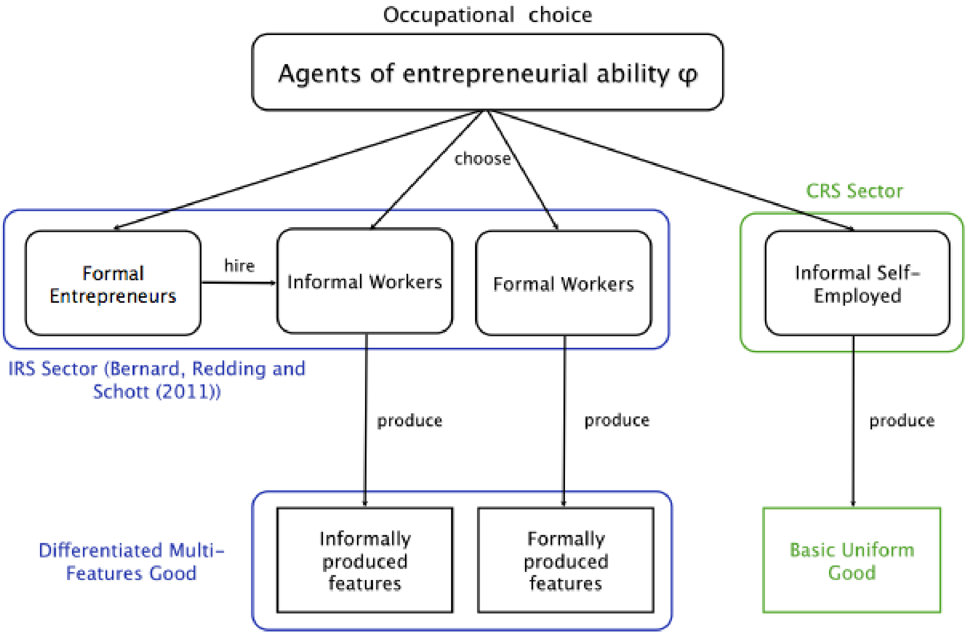


Figure 3.1: Overview of the Economy

Assumption 1: *Single-person firms can fully escape detection and therefore never register formally with tax authorities. On the contrary, fixed production costs for a single feature in the IRS sector are assumed to be large enough for the existence of a firm producing this feature to be detected by the government.*

3.4.2 Preferences

Consumers derive utility from the consumption of the basic good and of the different features produced in the IRS sector. The segment of existing features is normalized to $[0, 1]$. There is a constant elasticity of substitution across features so that a consumer with income Y maximizes the following utility function:

$$U = Rln\left(\int_0^1 C_j^{\nu} dj\right)^{\frac{1}{\nu}} + A \qquad 0 < \nu < 1,\ R > 0 \tag{3.1}$$

where R is the constant aggregate expenditure on the IRS multi-features good, $A = Y - R$ is the expenditure on the basic good, $\kappa = \frac{1}{1-\nu} > 1$ is the elasticity of substitution between different features

setting, but this would complicate further the model without adding to the analysis. Indeed, the present results would not be affected.

of the multi-features good, and C_j is an index summarizing the consumption of all varieties $\omega \in \Omega_j$ of feature j:

$$C_j = \left(\int_{\omega \epsilon \Omega} x_j \left(\omega \right)^{\rho} dj \right)^{\frac{1}{\rho}} \qquad 0 < \rho < 1 \tag{3.2}$$

where $\sigma = \frac{1}{1-\rho}$ is the elasticity of substitution between the different varieties of one feature and $x_j \left(\omega \right)$ the consumption of variety ω of feature j.

Given these quasi-linear preferences, aggregate demand for the basic good X_{CRS} and the multi-features good X_{IRS} are:

$$X_{CRS} = \frac{A}{p_0} = \frac{Y-R}{p_0} \quad X_{IRS} = \frac{R}{P} \tag{3.3}$$

where P is the price index in the multi-features good sector. The price of the basic good p_0 is determined on the international market, treated as exogenous in this Chapter for simplicity and normalized to 1 in the remaining of this Chapter.

Demand for variety ω of feature j is:

$$x_j(\omega) = \left(\frac{p_j(\omega)}{P_j} \right)^{-\sigma} \left(\frac{P_j}{P} \right)^{-\kappa} X_{IRS} \tag{3.4}$$

where $p_j(\omega)$ is the price of feature j variety ω, P_j is the price index for feature j in the economy. Demand for a feature's variety depends *only on the price charged for this variety* by a firm relative to prices charged by its competitors. Demand for a firm's multi-features good is the sum of demands for all the feature's varieties produced by this firm.[9] Since there is a continuum of identical features in the IRS sector, $P_j = P$ and $x_j \left(\omega \right) = x \left(\omega \right)$ for all features j. Demand for each feature's variety reduces to:

$$x \left(\omega \right) = RP^{\sigma-1} p(\omega)^{-\sigma} \tag{3.5}$$

Welfare W is given by $W = Y + CS$, where $CS = Rln\left(\frac{R}{P}\right) - R$ denotes the consumer surplus.

[9]Eventual synergies between features are therefore not taken into account to preserve the tractability of the model.

3.4.3 Occupational Choice and Income

Distribution of entrepreneurial ability

Individuals entrepreneurial ability φ is drawn from a Pareto distribution G of density g with scaling parameter k' over $[\varphi_0, +\infty)$

$$G(\varphi) = 1 - \left(\frac{\varphi_0}{\varphi}\right)^{k'} \tag{3.6}$$

with $\varphi_0 > 0$. Based on their ability, individuals select the available occupation that promises the highest income.

Entrepreneurs

In both sectors, entrepreneurs retain the profits of their firms as income. Section 3.4.4 elabourates further on these profit functions.

Self-employment in the CRS sector is the outside option of individuals in the economy. The role of informal entrepreneurship as the outside option of the poor has been documented in particular by Banerjee and Duflo (2011). They underline that high entrepreneurial skills and the availability of resources to cover fixed production costs are determinant in explaining successful entrepreneurship. Therefore, the most disadvantaged agents in terms of skills and endowments would rather be employees receiving a fixed wage than entrepreneurs generating a small income.[10]

Workers

When formally or informally employed in the IRS sector, individuals receive a fixed wage independent of their entrepreneurial ability. Informal workers are not declared as employees to the government. The informal wage w_I - where I stands for informal - is determined by clearing the informal labour market. In addition to their wage, formally declared workers benefit from all the advantages $\delta > 0$ laid down by labour regulations.[11] Formal workers also receive the wage gap $\mu \geq 0$ between formal and informal wages that is determined by clearing the formal labour market. The formal wage

[10] In this model, modeling unemployment benefits would simply increase the outside option of all agents in the economy: such benefits can be combined with informal self-employment or employment. By increasing the informal and formal wages in the economy, it would impact the overall level of employment and informality in the economy. However, the results of this Chapter will not be affected.

[11] In reality, formal advantages for a worker also consist in access to insurance against income shocks (contracts limiting abusive lay-offs for example or entitlement to unemployment benefits) or against health shocks (access to a public health insurance). As there is no uncertainty in this model, advantages of formality are monetized and included in δ.

received by the worker w_F - where F stands for formal - is then given by $w_F = w_I(1+\delta+\mu)$. It differs from the formal wage paid by the employer $w_F^T = w_I(1+\delta+\mu+t)$ that includes payroll taxes.

The informal and formal labour markets clear for two reasons. First, profits of the self-employed in the CRS sector - the outside option - are strictly increasing in individuals' entrepreneurial ability. Therefore, firms adjust wages until they exactly equal the outside option of the last worker they wish to hire. Second, entrepreneurial skills are positively correlated with the ability to find employment. Informal employment arises, because agents with lower level of ability are crowded out of formal employment by agents of higher ability. Workers never voluntarily choose to be informally rather than formally employed. But, at low levels of entrepreneurial ability, formal employment is not available, while informal employment brings a higher income than self-employment.[12]

3.4.4 Production Technologies

The Constant Return to Scale Sector

Firms operate under perfect competition in the production of a basic good using a technology exhibiting CRS in labour. Profits for a worker of ability φ are given by a continuous and increasing function $\pi_{CRS}(\varphi)$ verifying:

$$\frac{\partial \pi_{CRS}}{\partial \varphi} > 0 \quad \frac{\partial^2 \pi_{CRS}}{\partial \varphi \partial \varphi} < 0 \quad \lim_{\varphi \to \infty} \pi_{CRS} = \overline{\pi_{CRS}} \quad \pi_{CRS}(\varphi) > 0 \quad \forall \varphi \tag{3.7}$$

There is a physical limit to the output quantity a single person can produce however talented he is. Therefore $\pi_{CRS}(\varphi)$ is bounded above. Even agents at the lowest end of the entrepreneurial skill distribution can generate a positive income in the CRS sector.

To convince an individual to work for him, an entrepreneur in the CRS sector would need to offer a wage at least equal to profits generated by this individual as a self-employed in the CRS sector. But an additional employee is profitable to the entrepreneur, if her wage is inferior to the additional revenues she creates, in other words, inferior to the minimum wage she would accept. Therefore, all firms in the CRS sector are single person firms. Given `assumption 1` on imperfect law enforcement, it is never profitable for them to register formally.

I denote by φ_F the level of skills of the individual whose outside option - profits as self-employed - is equal to the formal wage $w_F = w_I(1+\delta+\mu)$. It is the highest level of skills among agents willing

[12] If formal advantages δ were also partially paid by workers, and not only by employers as in this model, informal employment could be preferred to formal employment by some workers.

to be formal workers. φ_I denotes the entrepreneurial ability of the individual whose outside option is equal to the informal wage w_I. It is the highest level of skills among agents willing to be informal workers. Therefore, φ_F and φ_I are determined by:

$$\pi_{CRS}(\varphi_F) = w_I(1+\delta+\mu) \tag{3.8}$$

$$\pi_{CRS}(\varphi_I) = w_I \tag{3.9}$$

where w_I and μ are determined in general equilibrium. The mass of agents M_I whose outside option is inferior to the informal wage is:

$$M_I = M\int_{\varphi_0}^{\varphi_I} g(\varphi)d\varphi = M\varphi_0\left[\varphi_0^{-k'} - \left(\pi_{CRS}^{-1}(w_I)\right)^{-k'}\right] \tag{3.10}$$

The mass of agents M_F whose outside option is in between the formal and informal wage is:

$$M_F = M\int_{\varphi_I}^{\varphi_F} g(\varphi)d\varphi = M\varphi_0\left[\left(\pi_{CRS}^{-1}(w_I)\right)^{-k'} - \left(\pi_{CRS}^{-1}\left((1+\delta+\mu)\,w_I\right)\right)^{-k'}\right] \tag{3.11}$$

where π_{CRS}^{-1} is the inverse function of π_{CRS}.

The Increasing Return to Scale Sector

In the IRS sector, as in Bernard et al. (2011), firm productivity θ_j in producing feature j is the product of two components: entrepreneurial ability φ - common to all features - and *feature-specific* expertise λ_j. Expertise λ_j is drawn from a Pareto distribution Z of density z with scaling parameter $\sigma - 1 < k < k'$ over $[\lambda_0, +\infty)$ with $\lambda_0 > 0$. Distributions G and Z are independent, and expertise λ_j is independently and identically distributed across features. A higher level of entrepreneurial skills φ raises the productivity θ_j across all features j.

Entrepreneurs decide whether to enter the IRS sector, which features to produce and - for each produced feature - whether to hire formal or informal labour. Fixed and variable costs are lower when using informal workers on a given segment of production, as the firm avoids abiding by labour regulations and payroll taxes $t \in [0,1]$. Production fixed costs are covered with the basic good. As described in equation (3.4), demand for variety $\omega \in \Omega_j$ of feature j depends on the variety's price and on the price index P. Because each firm produces one of a continuum of varieties, it has a negligible impact on the price index P and wages in the IRS sector. Section 3.5 details how the price index and wages are determined in general equilibrium. Therefore, profit maximization in the IRS sector reduces for an entrepreneur to pricing its variety ω of feature j by maximizing feature-profits π_j given his ability φ and feature-expertise λ_j and the prices in the economy, all taken as given. The firm produces quantity $q(p_j)$ to satisfy demand arising from the profit maximizing price p_j:

$$\underset{p_j}{Max} \quad \pi_j(p_j, q(p_j)) = \underset{p_j}{Max} \quad n(\varphi) p_j q(p_j) - \frac{wq(p_j)}{\varphi \lambda_j} \tag{3.12}$$

where $w \in \{w_I, w_F^T\}$ is the wage that depends on whether the firm employs formally or informally on feature j's segment of production, and $0 < n(\varphi) < 1$ is the fraction of revenues retained by an entrepreneur of ability φ when employing informal labour. There is $n(\varphi) = 1$ when the firm employs formal labour. The fraction $(1 - n(\varphi))$ of informal revenues lost by the entrepreneur can be interpreted as non-compliance costs, or alternatively, as the strength of law enforcement in the economy. The specific form taken by $n(\varphi)$ is described in details in section 3.4.5. I already introduce here the following restriction on non-compliance and regulation costs:

Assumption 2: *Non-compliance costs $n(\varphi)$ in the economy satisfy for all φ:*

$$n(\varphi) < (1 + \delta + t)^{1-\sigma} \tag{3.13}$$

Formal fixed costs E_F and informal fixed costs E_I satisfy:

$$E_F > \frac{E_I}{n(\varphi)(1 + \delta + t)^{\sigma - 1}} \tag{3.14}$$

`Assumption 2` is equivalent to the restriction on foreign and domestic production fixed costs imposed in all models adapted from Melitz (2003). Namely, fixed export costs must be larger than a lower bound that depends on fixed domestic costs and some iceberg costs melting away a fraction of export revenues. The tax wedge and non-compliance costs in my model are equivalent to iceberg costs: they are simply the share of revenues retained by the entrepreneur after paying taxes or non-compliance costs. Instead of two destinations for the produced good (foreign and domestic) with one characterized by iceberg transportation costs, my model features two modes of production (formal and informal), each characterized by a fraction of lost revenues due to the presence of taxes and non-compliance costs. The restriction in Melitz (2003) insures that no firm finds it profitable to produce only for the foreign market, and that only the most productive firms self-select into exporting. By parallel reasoning, `Assumption 2` makes sure first that there exists for each firm a level of expertise for which informal labour is optimal, and second that formal labour is more profitable than informal labour at the highest level of expertise.

Solving equation (3.12) leads to the standard result of the equilibrium price for each feature variety $p_j(\omega)$ being a constant mark-up over the marginal cost:

$$p_j(\omega) = p_j(\varphi, \lambda_j) = \frac{\sigma}{\sigma - 1} \frac{w}{n(\varphi)\varphi\lambda_j} \tag{3.15}$$

Revenues derived from a feature j by an entrepreneur of ability φ and expertise λ_j *before* paying the non-compliance costs (in the case of informal employees) or the formal tax wedge (in the case of formal employees) are obtained from combining the pricing rule in equation (3.15) to the demand function for variety ω in equation (3.5):

$$r_j(\varphi, \lambda_j) = R\left(\frac{\sigma - 1}{\sigma} \frac{w_I}{\varphi\lambda_j} \frac{1}{P}\right)^{1-\sigma} \tag{3.16}$$

Features revenues of an entrepreneur are therefore increasing in ability φ and expertise λ_j. Total expenditures on the IRS sector R is fixed by consumers' preferences (see utility maximization problem in equation (3.1)) and constant across features. Crucially, equation (3.16) implies that the ratio of revenues for two varieties produced using the same type of labour - formal or informal - only depends on their relative productivity:

$$r_j(\varphi'', \lambda''_j) = \left(\frac{\varphi''}{\varphi'}\right)^{\sigma-1} \left(\frac{\lambda''_j}{\lambda'_j}\right)^{\sigma-1} r_j(\varphi', \lambda'_j) \tag{3.17}$$

Informal and formal zero-profit expertise cutoffs

Informal profits - indexed by I - from feature j when employing informal workers for the production of feature j are given by variable profits minus informal fixed production costs E_I:

$$\pi_{jI}(\varphi, \lambda_j) = \frac{n(\varphi) r_j(\varphi, \lambda_j)}{\sigma} - E_I \tag{3.18}$$

Formal profits - indexed by F - from feature j when employing formal workers for the production of feature j are:

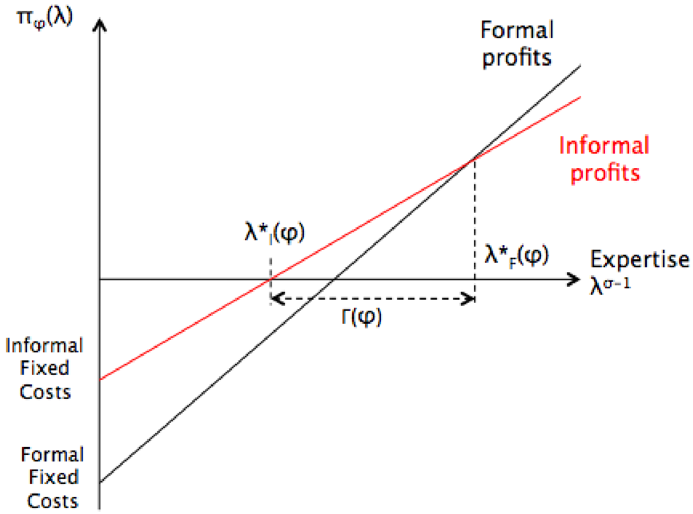


Figure 3.2: Formal and Informal Profits for a given φ

$$\pi_{jF}(\varphi, \lambda_j) = \frac{(1+\delta+t+\mu)^{1-\sigma} r_j(\varphi, \lambda_j)}{\sigma} - E_F \tag{3.19}$$

Since expertise λ is identically and independently distributed across features, the zero-profit informal condition and the equal-profit formal condition are identical across features: $\pi_{jF}(\varphi, \lambda) = \pi_F(\varphi, \lambda)$ and $r_j(\varphi, \lambda) = r(\varphi, \lambda)$ for all j. I will therefore drop the notation j from this point on. For each entrepreneurial ability level φ, there is an *informal zero-profit expertise cutoff* $\lambda_I^*(\varphi)$ and a *larger-than-informal-profits expertise cutoff* $\lambda_F^*(\varphi)$. A firm makes positive profits using informal labour in the production of features for which expertise λ is larger than $\lambda_I^*(\varphi)$. For feature-expertise levels λ larger than $\lambda_F^*(\varphi)$, the use of formal labour is more profitable than the use of informal labour. Figure 3.2 shows the formal and informal profit functions for a feature j and an entrepreneur of ability φ across all possible levels of expertise $\lambda(\varphi)$ under `Assumption 2`.

For each level of entrepreneurial skills, the value of $\lambda_I^*(\varphi)$ is defined by the informal zero-profit condition:

$$\pi_I\left(\varphi, \lambda_{jI}^*(\varphi)\right) = 0 \quad \Leftrightarrow \quad n(\varphi)\, r(\varphi, \lambda_I^*(\varphi)) = \sigma E_I \tag{3.20}$$

Using equation (3.16) for feature revenues, the informal zero-profit expertise cutoff can be rewritten as an explicit function $\lambda_I^*(\varphi, P)$ of entrepreneurial ability φ and price index P:

$$\lambda_I^*(\varphi, P) = \frac{\sigma-1}{\sigma}\frac{w_I}{\varphi P}\left(\frac{E_I}{n(\varphi)}\frac{\sigma}{R}\right)^{\frac{1}{\sigma-1}} \tag{3.21}$$

The larger-than-informal-profits cutoff for feature expertise using formal labour $\lambda_F^*(\varphi)$ is defined by the equality condition for formal and informal profits:

$$\pi_I(\varphi, \lambda_F^*(\varphi)) = \pi_F(\varphi, \lambda_F^*(\varphi)) \tag{3.22}$$

$$\Leftrightarrow \tag{3.23}$$

$$n(\varphi)\, r(\varphi, \lambda_F^*(\varphi)) - \sigma E_I = (1+\delta+t+\mu)^{1-\sigma} r(\varphi, \lambda_F^*(\varphi)) - \sigma E_F$$

The informal zero-profit expertise cutoff can be rewritten as an explicit function $\lambda_F^*(\varphi, P)$ of entrepreneurial ability φ and price index P:

$$\lambda_F^*(\varphi, P) = \frac{\sigma-1}{\sigma}\frac{w_I}{\varphi P}\left(\frac{E_F - E_I}{(1+t+\delta+\mu)^{1-\sigma} - n(\varphi)}\frac{\sigma}{R}\right)^{\frac{1}{\sigma-1}} \tag{3.24}$$

As a benchmark case, I can also derive the *formal zero-profit expertise cutoff* $\lambda^*(\varphi)$. An entrepreneur of ability φ makes positive profits using formal labour in the production of feature j when his expertise is larger than $\lambda^*(\varphi)$. The formal zero-profit condition is:

$$\pi_F\left(\varphi, \lambda_j^*(\varphi)\right) = 0 \tag{3.25}$$

Fraction of produced features

By the law of large numbers, the fraction of features produced using formal labour $\alpha_F(\varphi)$ by a entrepreneur of type φ equals the probability of this entrepreneur drawing a feature expertise above $\lambda_F^*(\varphi)$:

$$\alpha_F(\varphi) = 1 - Z(\lambda_F^*(\varphi)) \tag{3.26}$$

and the fraction of informally produced features $\alpha_I(\varphi)$ is given by the probability of this entrepreneur drawing a feature expertise between $\lambda_I^*(\varphi)$ and $\lambda_F^*(\varphi)$:

$$\alpha_I(\varphi) = Z(\lambda_F^*(\varphi)) - Z(\lambda_I^*(\varphi)) \tag{3.27}$$

The total fraction $\alpha(\varphi)$ of features produced by an entrepreneur of ability φ is then given by:

$$\alpha(\varphi) = 1 - Z(\lambda_I^*(\varphi)) \tag{3.28}$$

As in Bernard et al. (2011), the lower the ability φ of an entrepreneur, the higher the cutoff of informal and formal expertise $\lambda_I^*(\varphi)$ and $\lambda_F^*(\varphi)$, and the lower the probability of drawing a feature expertise sufficiently high to profitably produce a given feature j. Entrepreneurs with lower ability have lower expected profits on each feature market and manufacture a smaller fraction of features.

Formal and informal labour demand functions

Informal and formal labour demand by a firm across all produced features are respectively given by $l_I(\varphi)$ and $l_F(\varphi)$:

$$l_I(\varphi) = \int_0^1 I_j \left[E_I + \frac{q_I(\varphi, \lambda_j)}{\varphi\lambda_j}\right] dj\,; \qquad l_F(\varphi) = \int_0^1 F_j \left[E_F + \frac{q_F(\varphi, \lambda_j)}{\varphi\lambda_j}\right] dj \tag{3.29}$$

where I_j (resp. F_j) is an indicator variable equal to one if feature j is produced informally (resp. formally) and to zero otherwise. $q_I(\varphi, \lambda_j)$ (resp. $q_F(\varphi, \lambda_j)$) represents the quantity of feature j informally (resp. formally) produced by an entrepreneur with ability φ and expertise λ_j.

$$q_I(\varphi, \lambda_j) = \frac{n(\varphi)\, r(\varphi, \lambda)}{p(\varphi, \lambda)} = \frac{n(\varphi)\, R}{P^{\sigma-1}} \left(\frac{\sigma-1}{\sigma}\frac{w_I}{\varphi\lambda_j}\right)^{-\sigma} \tag{3.30}$$

$$q_F(\varphi, \lambda_j) = \frac{(1+t+\delta+\mu)^{1-\sigma} r(\varphi, \lambda)}{p(\varphi, \lambda)} = \frac{R}{P^{\sigma-1}} \left(\frac{\sigma-1}{\sigma}\frac{(1+t+\delta+\mu)\, w_I}{\varphi\lambda_j}\right)^{-\sigma} \tag{3.31}$$

Formal registration costs

Firms incur fixed formal registration costs F that are independent of the number of produced features. These costs are paid to the government and are covered by the entrepreneur's own labour. Therefore, an entrepreneur would not be undertaking these administrative activities if he did not expect his future profits to entirely cover them. Formal registration costs represent a minimum level of profits to be generated across features to successfully enter the multi-features good market and a fixed component of entrepreneurial income in the IRS sector.

Assumption 3: *Fixed administrative costs in the IRS sector are superior to the upper bound of* *profits in the CRS sector.*

$$F > \overline{\pi_{CRS}} \tag{3.32}$$

This assumption maintains the tractability of the model. It corresponds to the case where formal registration costs are so high that they exceed the revenues generated by any individual in the CRS sector. It implies that agents will always prefer entrepreneurship in the IRS sector to self-employment in the CRS sector if they can cover formal registration costs F. As `assumption 1`, `assumption 3` leads to an under-estimation of full informality in the economy, but results are not affected.

Zero-profit ability cutoffs

An entrepreneur of ability φ decides whether or not to enter the multi-features good market by comparing fixed administrative costs F and total profits across all produced features. With a continuum of identical features, by the law of large numbers, a firm's expected revenues across all features $r(\varphi)$ equal its expected revenues from an individual feature.

$$\begin{aligned} r(\varphi) &= r_I(\varphi) + r_F(\varphi) \\ &= \int_{\lambda_I^*(\varphi)}^{\lambda_F^*(\varphi)} n(\varphi)\, r(\varphi,\lambda) z(\lambda) d\lambda + \int_{\lambda_F^*(\varphi)}^{\infty} (1+t+\delta+\mu)^{1-\sigma} r(\varphi,\lambda)\, z(\lambda) d\lambda \end{aligned} \tag{3.33}$$

where $r_I(\varphi)$ and $r_F(\varphi)$ are total firm revenues respectively derived from all informally and formally produced features.

Hence, total profits across all features equal expected profits for an individual feature minus the fixed administrative costs:

$$\begin{aligned} \pi(\varphi) &= \pi_I(\varphi) + \pi_F(\varphi) - F \\ &= \int_{\lambda_I^*(\varphi)}^{\lambda_F^*(\varphi)} \left(\frac{n(\varphi)\, r(\varphi,\lambda)}{\sigma} - E_I \right) z(\lambda)\, d\lambda + \int_{\lambda_F^*(\varphi)}^{\infty} \left(\frac{(1+t+\delta+\mu)^{1-\sigma} r(\varphi,\lambda)}{\sigma} - E_F \right) z(\lambda)\, d\lambda - F \end{aligned} \tag{3.34}$$

where $\pi_I(\varphi)$ and $\pi_F(\varphi)$ are total firm profits respectively derived from all informally and formally produced features.

3.4.5 Non-compliance costs - Strength of law enforcement

Formally registered firms in the IRS sector face non-compliance costs $(1 - n(\varphi))$ when employing informal labour. Non-compliance costs can be interpreted as the outcome of the monitoring process implemented by the government on formal firms. Entrepreneurial ability is not observable, but is known to be positively correlated with the (observable) fraction of produced features. All factors contributing to the efficiency of law enforcement are embedded in $n(\varphi)$: the government auditing strategy, the vigilance of civil society "watchdogs", the power of labour unions, the extent to which tax inspectors are corruptible, and finally, entrepreneurs' own ability to circumvent regulations or persuade tax inspectors to accept bribes. To maintain the tractability of the model, I assume that entrepreneurs are myopic. This is a realistic assumption, considering the number and complexity of factors to be taken into account in the estimation of these costs.

Exogenous monitoring function $(1 - m(\lambda))$

Tax authorities audit firms based on the exogenous monitoring function $(1 - m(\lambda))$, according to which an entrepreneur observed to produce a fraction $(1 - Z(\lambda))$ of features retains a fraction $m(\lambda)$ of its informal revenues. This monitoring function can be interpreted as a certain level of expenditures made by the government on law enforcement as in Kopczuk and Slemrod (2006).
Consistently with empirical evidence.[3], larger firms face higher non-compliance costs: $m(\lambda)$ is decreasing in the fraction of produced features. Indeed, labour force tends to be more efficiently organized in large corporations, which renders collusion between workers and employers more difficult to sustain as in Kleven et al. (2009). Wrongdoings of large firms also attract more public attention. First, because they are often considered invested of a social responsibility. Second, because most profitable firms can pay higher bribes and fines.
For tractability, I assume that $m(\lambda)$ is a step-function characterized by N steps as shown in figure 3.3. Every step $s \in [1, N]$ has boundaries $\overline{\lambda_s}$ and $\overline{\lambda_{s+1}}$ with $\overline{\lambda_s} \in (\lambda_0, +\infty)$ and a level of efficiency $0 < n_s < 1$. All these parameters are exogenous. Each step s corresponds to a level of non-compliance costs - the share n_s of retained informal revenues - and the interval $\left[\overline{\lambda_s}, \overline{\lambda_{s+1}}\right]$ of firm size to which this level of non-compliance costs applies. Since fraud is easier to detect in larger firms, there is:

$$\overline{\lambda_2} > ... > \overline{\lambda_s} > \overline{\lambda_{s+1}} > ... > \overline{\lambda_N} > \lambda_0 \tag{3.35}$$

$$1 < n_N < n_{N-1} < ... < n_s < n_{s-1} < ... < n_2 < n_1 < 0 \tag{3.36}$$

Myopic entrepreneurs produce the largest possible - while maintaining positive profits - fraction of features given their ability φ. Tax authorities monitor a firm following step s of $m(\lambda)$ when the

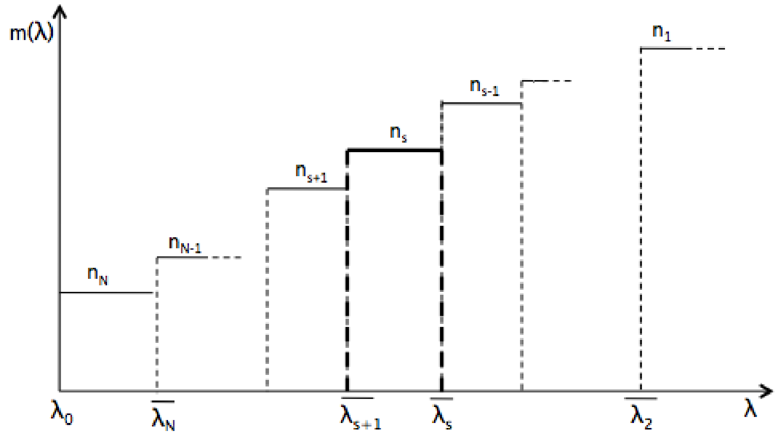


Figure 3.3: Non-compliance costs as a function of firms size

following conditions are satisfied:

1. the firm makes positive profits across all features given non-compliance costs n_s
2. the firm is one of the largest making positive profits given non-compliance costs n_s
3. the firm would not satisfy conditions (1) and (2) for a stricter step of $m(\lambda)$

When an entrepreneur of ability φ is allocated on step s of the monitoring function $m(\lambda)$ by the government, he faces non-compliance costs:

$$n(\varphi) = n_s \tag{3.37}$$

Allocation of entrepreneurs on the steps of the monitoring function $m(\lambda)$

From equations (3.21) and (3.27), firm size $(1 - Z(\lambda_I^*(\varphi)))$ depends on the efficiency of law enforcement $n(\varphi)$, which in turn depends on firm size. It is then not straightforward to determine the non-compliance costs experienced by an entrepreneur of ability φ: $n(\varphi)$ is only implicitly defined by equation (3.37). I introduce additional notations to write an explicit expression of $n(\varphi)$.

Step-specific zero-profit ability and expertise cutoffs For each step $s \in [1, N]$, there is a *zero-profit ability cutoff* φ_s^* such that entrepreneurs of ability $\varphi \geq \varphi_s^*$ make positive profits:

$$\pi(\varphi_s^*) \mid_{n(\varphi)=n_s} = 0 \tag{3.38}$$

and a *minimum expertise cutoff* $\lambda_{sI}^* = \lambda_{sI}^*(\varphi_s^*)$ such that the firm of lowest ability φ_s^* produces a given feature when $\lambda \geq \lambda_{sI}^*$:

$$\pi_I(\varphi_s^*, \lambda_{sI}^*)\ |_{n(\varphi)=n_s} \quad = \quad 0 \tag{3.39}$$

From equation (3.17), the minimum level of productivity of active firms $\theta_s^* = \varphi_s^* \lambda_{sI}^*$ on each step $s \in [1, N]$ can be rewritten as a function of the minimum level of productivity $\theta_1^* = \varphi_1^* \lambda_{1I}^*$ on the most lenient step of the law enforcement technology:

$$\varphi_s^* \lambda_{sI}^* \quad = \quad \varphi_1^* \lambda_{1I}^* \left(\frac{n_1}{n_s} \right)^{\frac{1}{\sigma - 1}} \tag{3.40}$$

The stricter is law enforcement on step s, the higher is the level of productivity $\varphi_s^* \lambda_{sI}^*$ required to survive when confronted with those non-compliance costs. The ratio $\frac{\theta_s^*}{\theta_1^*}$ is lower for higher values of elasticity of substitution σ. Indeed, a lesser love of variety translates into a lower incentive to produce a larger number of features. The use of informal labour - allowing the production of additional features at a lower cost - is therefore less determinant for firms survival on the market, along with being subject to a stricter law enforcement. The jump in productivity necessary for a firm to survive a higher level of non-compliance costs is thus reduced.

From equation (3.17), (3.20) and (3.21), the informal zero-profit expertise cutoff $\lambda_{sI}^*(\varphi)$ of an entrepreneur of ability $\varphi \geq \varphi_s^*$ when monitored according to step $s \in [0, N]$ depends only on his ability φ and on the zero-profit ability cutoff φ_s^* and minimum expertise cutoffs λ_{sI}^* for step s.

$$\lambda_{sI}^*(\varphi) \quad = \quad \frac{\varphi_s^* \lambda_{sI}^*}{\varphi} \tag{3.41}$$

Intuitively, as in Bernard et al. (2011), an increase in ability φ reduces the informal zero-profit expertise cutoff $\lambda_{sI}^*(\varphi)$. For a given level of non-compliance costs, higher entrepreneurial ability raises productivity for all features, and fixed production costs E_I can be covered at a lower expertise level. On the contrary, an increase in $\varphi_s^* \lambda_{sI}^*$ implies a higher level of aggregated productivity in the IRS sector and thus a tougher competitive environment. Prices charged by competitors are lower, which decreases demand for a firm's variety and increases the level of expertise required to make positive profits.

Monitoring cutoffs I define $\overline{\varphi_s}$ the level of entrepreneurial ability for which $\overline{\lambda}_s = \lambda_s^*(\overline{\varphi_s})$. $\overline{\varphi_s}$ is the lowest level of entrepreneurial ability for which an entrepreneur produces a fraction larger than $\left[1 - Z\left(\overline{\lambda}_s\right)\right]$ of total features when subject to step $s \in [1, N]$ of the law enforcement technology. It is given by:

$$\overline{\varphi_s} = \frac{\varphi_s^* \lambda_{sI}^*}{\overline{\lambda}_s} \tag{3.42}$$

Finally, the indirect informality cost function can be rewritten as a function of entrepreneurial ability:

$$n(\varphi) = n_s \quad if \quad \overline{\varphi_{s+1}} \geq \varphi \geq \overline{\varphi_s} \tag{3.43}$$

An entrepreneur loses a fraction $(1 - n_s)$ of revenues generated using informal labour, when (1) its firm makes positive profits: $\varphi \geq \overline{\varphi_s} \geq \varphi_s^*$ when subject to step s, (2) its firm is one of the largest firms making profits: $\lambda_{Is}^*(\varphi) \leq \overline{\lambda_s}$ when subject to step s, and (3) its firm is not subject to a stricter step of the monitoring function: $\overline{\varphi_{s+1}} \geq \varphi$.

Discussion: Implications of a monitoring step-function

The definition of the law enforcement technology as a step function can appear cumbersome, because it requires the introduction of additional notations. However, it allows for the computation of a closed form solution of the equilibrium. Moreover, the ability and expertise cutoffs on each step s of the function - respectively φ_s^* and λ_{sI}^* - as well as each step boundaries $\overline{\varphi_s}$ and $\overline{\varphi_{s+1}}$ can be rewritten as simple functions of the ability and expertise cutoffs φ_1^* and λ_{1I}^* of the most lenient step of $n(\varphi)$. The lowest level of entrepreneurial ability present in the IRS sector φ^* is therefore given by:

$$\varphi^* = \varphi_1^*$$

and the expertise cutoff for the entrepreneur of lowest ability λ_I^* is simply:

$$\lambda_I^* = \lambda_{1I}^*$$

With a step function $m(\lambda)$, firms subject to the same step of $m(\lambda)$ face the same non-compliance costs. Therefore, non-compliance costs do not play a role in determining those firms relative revenues. On the contrary, in the case of a continuously increasing monitoring function $m(\lambda)$, for every level of entrepreneurial ability φ, the competitive advantage in terms of productivity of the most able entrepreneurs with respect to less able entrepreneurs is always compensated to a certain extent by a non-compliance disadvantage. Hence, using a step-function rather than a continuous monitoring function $m(\lambda)$ leads to an *under-estimation* of the impact of within-firm informality on the intensity of competition in the formal sector.

Note that $\overline{\varphi_s}$ and $\overline{\varphi_{s+1}}$, the boundaries of the law enforcement function $n(\varphi)$, are endogenous. Even though the criteria $\overline{\lambda_s}$ and $\overline{\lambda_{s+1}}$ are exogenous and constant, non-compliance costs faced by an entrepreneur of ability φ can vary following a change in taxation or labour market policies in the economy. Nevertheless, the extent of the range of entrepreneurial ability of firms located on a step of $n(\varphi)$[13] - facing the same level of non-compliance costs - is determined by the exogenous monitoring function $m(\lambda)$:

$$\frac{\overline{\varphi_{s+1}}}{\overline{\varphi_s}} = \frac{\overline{\lambda_s}}{\overline{\lambda_{s+1}}}\left(\frac{n_s}{n_{s+1}}\right)^{\frac{1}{\sigma-1}} \tag{3.44}$$

The under-estimation of the effects of within-firm informality caused by the adoption of a step-function $m(\varphi)$ can be reduced at will by increasing the number of steps of the monitoring function $m(\lambda)$. It does not drive the results as I will show later on when deriving the general equilibrium in subsection 3.5.

3.5 General Equilibrium

3.5.1 Derivation of the Equilibrium

The equilibrium is defined by the set:

$$\{\varphi^*, \lambda_I^*, P, w_I^*, \mu^*\} \tag{3.45}$$

The zero-profit ability cutoff φ^* is the lowest level of skills among formal entrepreneurs. By equation (3.41), together with the minimum expertise cutoff λ_I^*, it pins down the zero-profit expertise cutoff $\lambda_I^*(\varphi)$ of any entrepreneur of ability $\varphi \geq \varphi^*$ active on the IRS market. In equilibrium, minimum firm productivity $\theta^* = \varphi^*\lambda_I^*$ is consistent with the price index P, the labour market clearing informal wage w_I^*, and the formal wage gap μ^*, which are sufficient to determine profits in the IRS sector

[13] The extent of the range of entrepreneurial ability facing the same level of non-compliance costs can be interpreted as the degree of sophistication of the monitoring strategy of the government. It seems plausible to assume that more sophisticated monitoring strategies are more costly to implement. The issue of an optimal degree of sophistication of the government auditing strategy is left for future research.

in equilibrium. Once the informal and formal wages are known, conditions (3.8) and (3.9) give the highest level of skills among formal workers φ_F and the highest level of skills among informal workers φ_I . Agents whose level of skills is between φ^* and φ_F are self-employed, as described in figure 3.4. I briefly describe in this subsection 3.5 how I solve for each component of the equilibrium. More details can be found in the mathematical appendix in subsection C.

Minimum Formal and Informal Expertise Cutoffs

To determine λ_I^*, I combine the zero-profit condition for a firm with ability φ^* in equation (3.38) with the expression for firm profits in (3.34). I substitute for formal and informal revenues using the relationship between relative variety revenues in equation (3.17) and the informal zero-profit condition for expertise in equation (3.20) and the equality condition for formal and informal profits in equation (3.22). I obtain:

$$\lambda_F^* = \left(\frac{\sigma - 1}{k - \sigma + 1} \frac{E_F + E_I \left(\Gamma_1^k - 1 \right)}{F} \right)^{\frac{1}{k}} \tag{3.46}$$

$$\lambda_I^* = \left(\frac{\sigma - 1}{k - \sigma + 1} \frac{E_I + \Gamma_1^{-k} \triangle FC}{F} \right)^{\frac{1}{k}} \tag{3.47}$$

As a benchmark case, I also derive the zero-profit formal expertise cutoff using the formal zero-profit condition for expertise in equation (3.25):

$$\lambda^* = \lambda^*(\varphi^*) = \left(\frac{\sigma - 1}{k - \sigma + 1} \frac{E_F}{F} \right)^{\frac{1}{k}} \tag{3.48}$$

Therefore, it is always verified that:

$$\lambda_I^* < \lambda^* < \lambda_F^* \tag{3.49}$$

Proposition 3. *When monitoring of formal firms is imperfect, an entrepreneur of ability φ ...*

- *produces a larger number of features*
- *employs a smaller number of formal employees*

than in the perfect monitoring case (full compliance i.e no within-firm informality).

Zero-Profit Ability Cutoff and Price Index

From equation (3.38) and the equilibrium informal zero-profit expertise cutoff defined in equation (3.47), I derive the zero-profit ability cutoff φ^* as a function of the price index on the multi-features good market P:

$$\varphi^*(P) = \frac{1}{\lambda_I^* P}\left(\frac{\sigma}{\sigma-1}w_I^*\right)\left[\frac{\sigma E_I}{n_1 R}\right]^{\frac{1}{\sigma-1}} \tag{3.50}$$

The price index on the multi-features good market can in turn be derived as a function of the zero-profit cutoff for ability using the following expression:

$$P = P(\varphi^*) = \left[\sum_{s\epsilon[1,N]} P_{sI}(\varphi^*)^{1-\sigma} + \sum_{s\epsilon[1,N]} P_{sF}(\varphi^*)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{3.51}$$

where P_{sI} (respectively P_{sF}) is the price index of features informally (respectively formally) produced by firms facing step $s \in [1, N]$ of the monitoring function $m(\lambda)$:

$$P_{sI}(\varphi^*)^{1-\sigma} = \int_{\overline{\varphi_s}}^{\overline{\varphi_{s+1}}}\left[\int_{\lambda_{sI}^*(\varphi)}^{\lambda_{sF}^*(\varphi)} p_I(w)^{1-\sigma} z(\lambda)d\lambda\right] g(\varphi)d\varphi \tag{3.52}$$

$$P_{sF}(\varphi^*)^{1-\sigma} = \int_{\overline{\varphi_s}}^{\overline{\varphi_{s+1}}}\left[\int_{\lambda_{sF}^*(\varphi)}^{\infty} p_F(w)^{1-\sigma} z(\lambda)d\lambda\right] g(\varphi)d\varphi \tag{3.53}$$

Combining $P(\varphi^*)$ and $\varphi^*(P)$, I get the following expressions for the price index and the minimum ability cut-off in equilibrium:

$$P = \left(\frac{\sigma}{\sigma-1}w_I^*\right)\left[\frac{k'}{k'-k}\frac{k\varphi_0}{k-\sigma+1}\sum_s (n_s)^{\frac{k-\sigma+1}{\sigma-1}} S_s(\mu^*) C_s\right]^{-\frac{1}{k'}}\left[\frac{\sigma E_I}{R}\right]^{\frac{1}{k'}\frac{k'-\sigma+1}{\sigma-1}} \tag{3.54}$$

$$\varphi^* = \varphi^*(P^*) = \frac{1}{\lambda_I^*}\left[\frac{\sigma E_I}{R}\frac{k'}{k'-k}\frac{k\varphi_0}{k-\sigma+1}\sum_s (n_s)^{\frac{k-\sigma+1}{\sigma-1}} S_s(\mu^*) C_s\right]^{\frac{1}{k'}} (n_1)^{\frac{1}{1-\sigma}} \tag{3.55}$$

where, for all steps $1 < s < N$, the term C_s is a measure of the - exogenous - sophistication of the monitoring strategy implemented by the government in the economy as discussed in subsection 3.4.5.

$$C_s = \left(\overline{\lambda}_s (n_s)^{\frac{1}{\sigma-1}}\right)^{k'-k} - \left(\overline{\lambda_{s+1}} (n_{s+1})^{\frac{1}{\sigma-1}}\right)^{k'-k} \tag{3.56}$$

The larger is C_s , the larger is the range of entrepreneurial ability of the firms located on the steps of $n(\varphi)$ and the more intense is competition in the IRS sector. Indeed, from equation (3.44), the mass of firms located on a same step of $n(\varphi)$ increases with C_s. For each given entrepreneur, the number of competitors of higher ability with which they share the same non-compliance level is then larger. The extent to which non-compliance can mitigate a lower productivity is reduced, and it is more difficult to enter the market $\left(\frac{\partial \varphi^*}{\partial C_s} > 0\right)$. As active firms are more productive on average, they charge lower prices $\left(\frac{\partial P}{\partial C_s} < 0\right)$.

For all $1 < s < N$, the endogenous term $\Gamma_s^{\sigma-k-1}\left(1 - (1+t+\delta+\mu^*)^{1-\sigma}\right)$ is the share of total revenues (formal and informal) allocated to tax payments. More productive firms, located on stricter steps of the law enforcement function, dedicate a larger fraction of their total revenues to taxes. S_s represents the share of total revenues retained by an entrepreneur facing step s of the monitoring function after paying taxes:

$$S_s(\mu^*) \quad = \quad 1 - \Gamma_s^{\sigma-k-1}\left(1 - (1+t+\delta+\mu^*)^{1-\sigma}\right) \tag{3.57}$$

All firms facing a given step s of the monitoring function retain the same share of total revenues after taxes. When this share increases, larger firms with larger scale of production benefit more than smaller firms due to increasing returns to scale. Competition becomes more intense $\left(\frac{\partial \varphi^*}{\partial S_s} > 0\right)$ and prices decrease $\left(\frac{\partial P}{\partial S_s} < 0\right)$.

S_s and C_s embody the two opposite effects of within-firm informality on equilibrium ability cutoff φ^* and price index P. On the one hand, within-firm informality increases prices and decreases aggregate productivity by providing the least able entrepreneurs with an advantage in non-complying that partially compensates for their productivity disadvantage. Indeed, within-firm informality plays a larger role in determining firms relative revenues when the monitoring function of the government is sophisticated (C_s is small). On the other hand, within-firm informality increases S_s the share of retained revenues after taxes - even if not uniformly across productivity levels - and therefore decreases prices and increases aggregate productivity.

Clearing the informal labour market

The informal wage w_I is determined so that aggregate informal labour demand $L_I(\varphi^*, w_I^*, \mu^*)$ can be satisfied. Informal labour supply is equal to $M_I(\varphi^*, w_I^*, \mu^*)$, the mass of agents in the economy whose outside option is inferior to the informal wage, given by equation (3.10). The informal labour market clearing condition is:

$$L_I\left(\varphi^*, w_I^*, \mu^*\right) = M_I\left(\varphi^*, w_I^*, \mu^*\right) \tag{3.58}$$

$$\sum_{1<s<N} l_I^s\left(\varphi^*, w_I^*, n_s, \mu^*\right) = M_I\left(\varphi^*, w_I^*, \mu^*\right) \tag{3.59}$$

where $l_I^s(\varphi^*, w_i, n_s, \mu)$ is total informal labour demand from firms located on step s of the monitoring function:

$$l_I^s\left(\varphi^*, w_I^*, n_s, \mu^*\right) = \int_{\overline{\varphi_s}}^{\overline{\varphi_{s+1}}} l_I(\varphi) g(\varphi) d\varphi \tag{3.60}$$

Informal labour demand of a firm facing step s of the monitoring function is given by equation (3.29).

Clearing the formal labour market

The formal wage gap μ is determined so that aggregate formal labour demand $L_F\left(\varphi^*, w_I^*, \mu^*\right)$ can be satisfied. Formal labour supply is equal to $M_I\left(\varphi^*, w_I^*, \mu^*\right) + M_F\left(\varphi^*, w_I^*, \mu^*\right)$, the mass of agents whose outside option is inferior to the formal wage, given by equations (3.10) and (3.11). In equilibrium, formal positions are filled in by workers of higher entrepreneurial ability, because it is correlated with the ability to find a job. Informal workers in $M_I\left(\varphi^*, w_I, \mu\right)$ are therefore pushed out of formal employment into informal employment. The formal labour market clearing condition is:

$$L_F\left(\varphi^*, w_I^*, \mu^*\right) = M_F\left(\varphi^*, w_I^*, \mu^*\right) \tag{3.61}$$

$$\sum_{1<s<N} l_F^s\left(\varphi^*, w_I^*, n_s, \mu^*\right) = M_F\left(\varphi^*, w_I^*, \mu^*\right) \tag{3.62}$$

where $l_F^s(\varphi^*, w_I^*, n_s, \mu^*)$ is total formal labour demand from firms located on step s of the monitoring function:

$$l_F^s\left(\varphi^*, w_I^*, n_s, \mu^*\right) = \int_{\overline{\varphi_s}}^{\overline{\varphi_{s+1}}} l_F(\varphi) g(\varphi) d\varphi \tag{3.63}$$

Formal labour demand of a firm facing step s of the monitoring function is given by equation (3.29). Without giving an explicit form to the profit function π_{CRS} of self-employed in the CRS sector, the equilibrium conditions for the formal and informal labour markets are only implicitly defined. The full expression is included in the mathematical appendix C.

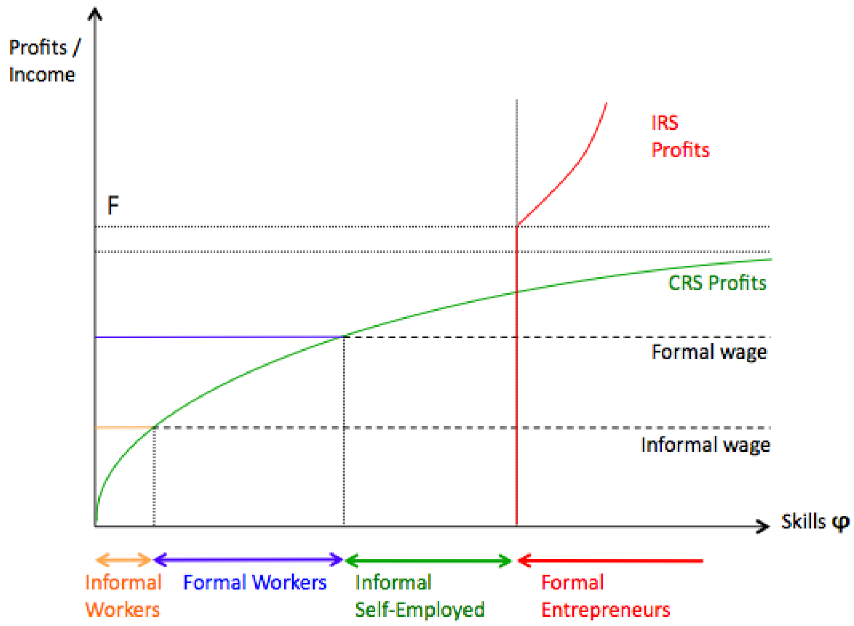


Figure 3.4: Allocation of Agents in Equilibrium

3.5.2 Allocation of Agents in the Economy

`Assumption 1, 2, 3` and `4` give rise to the allocation of agents in the occupational choices of the economy as described in figure 3.4.

Share of Within-Firm Informality The attractiveness of informal labour relative to formal labour - on a given step of the monitoring function - is summarized by Γ_s :

$$\Gamma_s = \left(\frac{n_s}{E_I} \frac{\triangle FC}{\triangle RR_s} \right)^{\frac{1}{\sigma - 1}} \quad (3.64)$$

where $\triangle FC = E_F - E_I$ and $\triangle RR_s = (1 + \delta + t + \mu)^{1-\sigma} - n_s$ respectively represent the difference between formal and informal fixed costs and the difference between formal and informal retained revenues. For all φ, the relationship between the informal zero-profit expertise cutoff $\lambda^*_{sI}(\varphi)$ - defined by equation (3.21) - and the formal zero-profit expertise cutoff $\lambda^*_{sF}(\varphi)$ - defined by equation (3.24) - is given by:

$$\lambda^*_{sF}(\varphi) = \Gamma_s \lambda^*_{sI}(\varphi) \quad (3.65)$$

The larger is Γ_s, the larger is the share of features generated by the possibility of using informal labour. It is verified that:

$$1 < \Gamma_N < ... < \Gamma_2 < \Gamma_1 \quad (3.66)$$

Firms subject to lower non-compliance costs - larger Γ_s - enjoy an advantage in concealing revenues generated through the use of informal labour. The share of within-firm informality $\alpha_{WFI}(\varphi)$ in the firm of an entrepreneur of ability φ facing step s of the law enforcement technology:

$$\alpha_{WFI}(\varphi) \quad = \quad 1 - \frac{1}{\Gamma_s^k} \tag{3.67}$$

Proposition 4. *Firm productivity and formal employment are negatively correlated with the share of informal employees (within-firm informality). Within a given firm, informal workers are employed in the segments of production where firm expertise is the lowest.*

The most able entrepreneurs run the most productive firms and are producing a larger number of features. Therefore, they are easier to monitor by the government and face higher non-compliance costs than their smaller and less productive competitors. The use of informal labour is thus less attractive for them and the share of informal workers in their total labour force is lower.

3.6 Properties of the General Equilibrium

Bernard et al. (2011) derive the properties of an IRS sector with heterogeneous firms when the extensive margin of production is endogenous. Properties of the equilibrium of Bernard et al. (2011) are therefore also present in this model. There is a positive correlation between intensive and extensive margins of production. Entrepreneurial ability enters in the productivity of all features, so that more able entrepreneurs manufacture a larger share of features and a greater quantity of each feature. The percentage difference in revenues between two entrepreneurs of different ability is therefore greater when the range of produced features is endogenous. On the contrary, the percentage difference between formal firms weighted average productivity is lower than if formal firms were producing a single feature. Indeed, entrepreneurs of higher ability are able to produce features for which they have a lower expertise. Within-firm informality in this model is simply an opportunity for formal firms to expand along the extensive margin of production at a lower cost. Nevertheless, the most able entrepreneurs suffer from a disadvantage in non-complying. Within-firm informality then causes a stronger decrease in firms weighted average productivity for the least able entrepreneurs. The dispersion of average weighted productivity is thus larger than in the full compliance case.

3.6.1 Effects of an Increase in Formal Registration Costs

Proposition 5.

- *Higher formal registration costs increase the zero-profit ability cutoff φ^* in the formal sector. The zero-profit productivity cutoff $\theta^* = \varphi^* \lambda_I^*$ in the formal sector - and thus the price index, formal and informal employment - is unchanged.*
- *Full informality increases: the least able formal entrepreneurs become informally self-employed in the CRS sector. Welfare of these agents decreases.*
- *This expansion in full informality is lower than if formal firms were fully compliant.*

Higher formal registration costs F increase the minimum level of profits that an entrepreneur must realize across all features in order to remain active on the market. Therefore, the ability cut-off φ^* increases (direct positive effect). The number of sold varieties drops, which eases competitive pressure on all features markets and triggers entrepreneurs to start producing previously non-profitable features. Hence, formal and informal expertise cut-off λ_F^* and λ_I^*. This expansion along the extensive margin of production increases overall profits in active firms and φ^* decreases (indirect negative effect). As in Bernard et al. (2011), the possibility for firms to respond more easily along the extensive margin of trade to changes in the economy dampens the reaction of the zero-profit ability cut-off. Hence, within-firm informality decreases the welfare loss caused by an increase in formal registration costs.

$$\frac{d\varphi^*}{dF} = \frac{1}{k}\frac{\varphi^*}{F} > 0 \tag{3.68}$$

Higher formal registration F costs have a "floating boat" effect on employment and prices: resources and market shares are reallocated across firms, but average productivity $\theta^* = \varphi^* \lambda_I^*$ in the formal sector is unchanged. Indeed, the increase in the zero-profit ability cutoff φ^* is exactly compensated by the decrease in the zero-profit expertise cutoff λ_I^*. Entrepreneurs active in the formal sector are more able on average, but they also manufacture features at a lower level of expertise.

$$\frac{d\lambda_I^*}{dF} = -\frac{1}{k}\frac{\lambda_I^*}{F} < 0 \tag{3.69}$$

$$\frac{d\theta^*}{dF} = 0 \tag{3.70}$$

Employment levels and the price index are therefore unchanged $\frac{dP}{dF} = \frac{dw_I}{dF} = \frac{d\mu}{dF} = 0$. This result must be taken with caution, because it is derived under the assumption that changes in formal registration costs F have no impact on aggregate expenditures on the IRS sector. Nonetheless, it provides a good benchmark case and it is a reasonable assumption if one considers an economy with several industries and a change in the formal registration costs of only one of those industries.

3.6.2 Effect of an Increase in Payroll Taxes

Assumption 4: ***The function $\pi_{CRS}(\varphi)$ describing the profits as self-employed of a agent of ability φ verifies the following condition:***

$$\frac{d}{dw}\left[\frac{\left(\pi_{CRS}^{-1}(w)\right)^{k'+1}}{\left(\pi_{CRS}^{-1}(w)\right)'}\right] \leq 0 \tag{3.71}$$

Assuming a simple functional form for $\pi_{CRS}(\varphi)$ verifying `Assumption 3` in section (3.4.4) and the conditions described in equation (3.7):

$$\pi_{CRS}(\varphi) = \overline{\pi_{CRS}}\left(1 - \frac{\varphi_0}{\varphi^x}\right)$$

(where $x > 0$), `Assumption 4` is equivalent to:

$$x > k'$$

Namely, the concavity of the self-employed profit function dominates the convexity of the Pareto distribution of entrepreneurial ability. This guarantees that informal labour demand is always satisfied by an increase in the informal wage. In other words, an increase in the informal wage leads to an excess supply of informal labour, holding the formal wage gap μ constant:

$$\frac{\partial M_F}{\partial w_I} \geq \frac{\partial M_I}{\partial w_I}$$

On one hand, the concavity of $\pi_{CRS}(\varphi)$ causes φ_F to increase by more than φ_I for a given increase in the informal wage w_I. On the other hand, the convexity of the Pareto distribution of φ implies that a lower mass of agents can be found at higher levels of entrepreneurial ability. Since $\varphi_F > \varphi_I$, the overall effect on informal labour supply is ambiguous. `Assumption 4` guarantees that the mass of agents pushed out of formal employment by an increase in M_F is sufficient to satisfy informal labour demand.

The profit function of self-employed is concave in ability and bounded above. It implies that one can distinguish between two parts of the curve. On the first "steep" part, marginal returns of ability are significant for the lower end of the ability distribution, while on the second "flat" part, marginal returns of ability are negligible for higher levels of ability. The more concave is $\pi_{CRS}(\varphi)$ and the smaller is the range of ability for which marginal return of entrepreneurial ability are significant. When k' increases, dispersion of ability decreases and a larger mass of agents falls into the range where the marginal return of ability is significant. `Assumption 4` states that there is a significant mass of agents for which marginal return of ability are negligible. An intuition for this assumption could be that, because informal self-employed produce a basic good, consumers are not willing to differentiate too much between different informal producers. Entrepreneurs at the very bottom of the ability distribution may have difficulty to produce at all, or to market their production, but the competitive advantage of the most able entrepreneurs remains limited. Another intuition could also simply be that law enforcement in the economy is sufficiently strong to prevent the most able entrepreneurs of the informal sector to fully take advantage of their competitive advantage, by limiting their ability to advertise themselves for example.

Proposition 6. *Higher payroll taxes...*

- *increase the informal wage and total informal employment in the economy:*

$$\frac{dw_I}{dt} > 0 \tag{3.72}$$

- *decrease the formal wage gap:*

$$\frac{d\mu}{dt} < 0 \tag{3.73}$$

- *increase in the entrepreneurial ability cut-off of the formal increasing return to scale sector:*

$$\frac{d\varphi^*}{dt} > 0 \tag{3.74}$$

- *increase the range of produced – and informally produced – features within each formal firm:*

$$\frac{d\lambda_I^*}{dt} < 0\,; \quad \frac{d\lambda_F^*}{dt} > 0 \tag{3.75}$$

- *increase the price index:*

$$\frac{dP}{dt} > 0 \tag{3.76}$$

Profitability of formally produced features decreases. Fixed formal production costs E_F and fixed formal registration costs F are more difficult to cover for the least able entrepreneurs in the IRS sector and they must exit the market. The zero-profit ability cut-off φ^* increases. By a general equilibrium effect already highlighted in (Bernard et al., 2011), this increase in the entrepreneurial ability cut-off is lower than if firms could not expand along the extensive margin of production. Indeed, as some firms exit the market, the intensity of competition eases. First, some varieties disappear and second, demand for labour decreases, which lowers the upward pressure on wages. As a result, entrepreneurs remaining on the market start producing previously non-profitable features: they expand along the extensive margin of production $\frac{d\lambda_I^*}{dt} < 0$. Furthermore, features that were previously more profitable when using formal labour are now produced using informal labour $\frac{d\lambda_F^*}{dt} > 0$. The price index is driven up both by the decrease in the lower level of expertise for which firms can produce and by the increase in production costs (higher informal wage and larger payroll taxes).

Proposition 7. *If, for each step $s \in [1, N]$ of the monitoring function, the sufficient condition - henceforth referred to as the* high fiscal capacity[14] condition - *is verified (μ_{pre} is the formal wage gap previous to the change in the tax rate):*

$$1 - n_s > \frac{k}{\sigma - 1}\left[1 - (1 + t + \delta + \mu_{pre})^{1-\sigma}\right] \tag{3.77}$$

Then, higher payroll taxes...

- *decrease the share of total revenues retained by formal entrepreneurs after paying taxes (see equation (3.57)):*

$$\frac{dS_s}{dt} < 0 \tag{3.78}$$

- *decrease the welfare of informal and formal workers in the economy.*

[14] The term "fiscal capacity" usually refers to the ability of a government to raise taxes, both in terms of enforcement and level of taxes. Here, only payroll taxes are raised by the government, so that the term fiscal capacity refers to the government ability to raise payroll taxes only.

Proposition 8. *If, for each step $s \in [1, N]$ of the monitoring function, - henceforth referred to as low fiscal capacity condition - is verified:*

$$1 - n_s \quad < \quad \frac{k}{\sigma - 1}\left[1 - (1 + t + \delta)^{1-\sigma}\right] \tag{3.79}$$

Then, higher payroll taxes...

- *increase the share of total revenues retained by formal entrepreneurs after paying taxes:*

$$\frac{dS_s}{dt} \quad > \quad 0 \tag{3.80}$$

- *increase the welfare of agents that remained informal workers following the tax increase.*

Equation (3.77) is referred to as high fiscal capacity condition, because it is verified when either law enforcement is strict (large non-compliance costs $(1 - n_s)$), or taxes are low (low value of $(\delta + t)$), or both. It implies that the government can easily collect taxes through strict enforcement or raise existent low tax rates. On the contrary, when fiscal capacity is low, enforcement is low, or taxes are high, or both, which prevent further tax increase.

For the same level of enforcement and taxes, a larger dispersion of feature revenues within a formal firm $\frac{k}{\sigma-1}$ increases the likelihood of the low fiscal condition (3.79) to be verified. Indeed, a large value of $\frac{k}{\sigma-1}$ means a low dispersion of firms internal distribution of revenues across features. In other words, formal firms are homogeneously deriving revenues from their set of manufactured features. I refer to this situation as an *efficiently diversified* industry. On the contrary, a low value of $\frac{k}{\sigma-1}$ implies a large dispersion of features revenues within firms. The bulk of firm profits in the IRS sector thus comes from a few features, in which entrepreneurs have high levels of expertise. I refer to this situation as an *inefficiently diversified* industry. In both types of industries, firms expand as much as possible along the extensive margin of production, but features of lower expertise generate a larger share of total revenues in an efficiently diversified industry. The government fiscal capacity is therefore lower in an efficiently diversified industry, because formal firms derive a larger share of their revenues from informally manufactured features. Total revenues are thus less sensitive to changes in payroll taxes.

On one hand, larger payroll taxes increase the amount of formal revenues that is remitted to the government, but on the other hand, the share of formal revenues in total revenues decreases as the production of some features is switched from formal to informal. Furthermore, informal revenues increase, because firms start producing new features at the extensive margin of production. The final effect on the share of total revenues that is retained after taxes is therefore ambiguous. The lower is fiscal capacity, the more attractive is informal labour and the more efficient is the tax evasion strategy of firms in the formal sector: the share of retained revenues after taxes increases $\left(\frac{\partial S}{\partial t} > 0\right)$. The effect on total revenues of the increase in informal revenues dominates the effect of the increase in payroll taxes. When fiscal capacity is high in the economy and informal labour attractiveness is limited, the share of total revenues remaining after taxes decreases $\left(\frac{\partial S}{\partial t} < 0\right)$. In this case, the increase in informal revenues is dominated and informal employment does not rise as much as in the low fiscal capacity case. Therefore, the informal wage rises more in the low fiscal capacity case than in the high fiscal capacity.[15] In fact, in the low fiscal capacity case, the real informal wage rises, while it decreases in the high fiscal capacity case.

[15]The high and low fiscal capacity conditions in equations (3.77) and (3.79) apply to specific steps of the monitoring function. For simplicity, I will limit the analysis to the case where all steps (or none) of the monitoring function satisfy them. One could imagine that these conditions are satisfied only for some steps of the monitoring function, and that it would be sufficient for the real informal wage to evolve in one direction or another. However, this complicates the model without adding to the analysis.

3.7 Conclusion

Although within-firm informality is empirically important both in developing and developed countries, it has been ignored by the theoretical literature of informality and optimal taxation. This Chapter introduces a general equilibrium model where agents in the economy are heterogeneous in entrepreneurial ability and self-select into the different occupational choices: formal or informal, workers or entrepreneurs. The key result of the Chapter is that within-firm informality does not evolve in the same way as full informality to different tax instruments and regulations. Furthermore, the impact of within-firm informality on workers welfare depends both on the fiscal capacity and on the degree of diversification of the relevant industry. Hence, the results of this chapter underline that enforcement vary significantly from one industry to another and can be used as a stepping stone for further research on the optimal strategy of a government to monitor formal firms and improve the enforcement of labour regulations.

Appendices

Appendix C

Proofs of the Theoretical Results

Derivation of the equilibrium

For each step of the monitoring function, the derivation of the equilibrium follows the strategy detailed in Bernard et al. (2011). In this appendix, I explain the computations that are not explicitly derived in the main body of this chapter.

Price index P and Ability cutoff φ^*

The price index of informally produced features P_{sI} under LET_s is given by:

$$\begin{aligned} P_{sI}\left(\varphi^*\right)^{1-\sigma} &= \int_{\overline{\varphi_s}}^{\overline{\varphi_{s+1}}} \left[\int_{\lambda^*_{si}(\varphi)}^{\lambda^*_{sf}(\varphi)} p_i(w)^{1-\sigma} z(\lambda) d\lambda \right] g(\varphi) d\varphi \\ &= \frac{k'}{k'-k} \frac{k\varphi_0}{k-\sigma+1} \left(\frac{\sigma}{\sigma-1} w_i \right)^{1-\sigma} \left(\varphi^* \lambda^*_I\right)^{\sigma-k'-1} \left[1 - \Gamma_s^{\sigma-k-1} \right] C_s \end{aligned}$$

where C_s is an exogenous measure of the sophistication of the law enforcement technology:

$$C_s = (n_1)^{\frac{\sigma-k'-1}{\sigma-1}} \left(\overline{\lambda}_{(s-1)} \left(n_{s-1}\right)^{\frac{1}{\sigma-1}} \right)^{k'-k} - \left(\overline{\lambda}_s \left(n_s\right)^{\frac{1}{\sigma-1}} \right)^{k'-k}$$

Similarly, the price index of formally produced features P_{sf} under LET_s is given by:

$$\begin{aligned} {P_{sf}}^{1-\sigma} &= \int_{\overline{\varphi_s}}^{\overline{\varphi_{s+1}}} \left[\int_{\lambda^*_{sf}(\varphi)}^{\infty} p_F(w)^{1-\sigma} z(\lambda) d\lambda \right] g(\varphi) d\varphi \\ &= \frac{(1+t+\delta+\mu)^{1-\sigma}}{\Gamma_s^{k-\sigma+1} - 1} P_{si}(\omega)^{1-\sigma} \end{aligned}$$

The price index P can be written as a function of the ability cut-off φ^*:

$$P(\varphi^*) = \left(\frac{\sigma}{\sigma-1}w_i\right)(\varphi_1^*\lambda_{i1}^*)^{\frac{\sigma-k'-1}{1-\sigma}}(n_1)^\beta\left[\frac{k'}{k'-k}\frac{k\varphi_0}{k-\sigma+1}\sum_s S_s(\mu,\Gamma_s)(n_s)^{\frac{\sigma-k-1}{1-\sigma}}C_s\right]^{\frac{1}{1-\sigma}}$$

with $\beta = \frac{k'-\sigma+1}{(\sigma-1)^2}$ and $S_s(\mu,\Gamma_s) = 1-\Gamma_s^{\sigma-k-1}\left(1-(1+t+\delta+\mu)^{1-\sigma}\right)$ for all $1 < s < S$. In addition, there is $\overline{\varphi_1} = \varphi^*$, $\overline{\lambda_0} = \lambda_{i1}^*$, $\overline{\varphi_S} = \infty$, $\frac{1}{\overline{\lambda_S}} = 0$, $n_0 = 1$.

Comparative statics with respect to Payroll Taxes

Sketch of the proof - Part 1: The application of the implicit function theorem on the equilibrium conditions of the informal and formal labour market leads to:

$$\frac{k'w_iM(1+\mu)\Omega}{\underbrace{\frac{\partial g}{\partial\mu}}_{>0}-\underbrace{\frac{\partial g}{\partial w_i}\left(\frac{\partial f}{\partial w_i}\right)^{-1}}_{>0}\underbrace{\frac{\partial f}{\partial t}}_{<0}} = 1+\frac{d\mu}{dt}$$

where $\Omega > 0$ is the change in the mass of agents whose outside option is below the informal wage when the informal wage rises. From the following:

$$-\underbrace{\frac{\partial f}{\partial t}}_{<0}\left(\underbrace{1+\frac{d\mu}{dt}}_{>0}\right)\left(\underbrace{\frac{\partial f}{\partial w_i}}_{>0}\right)^{-1} = \frac{dw_i}{dt}$$

I conclude:

$$\frac{dw_i}{dt} > 0$$

Sketch of the proof - Part 2: The sign of $\left(1+\frac{d\mu}{dt}\right)$ determines the sign of the derivatives of the expertise cut-off in the formal increasing return to scale sector:

$$\frac{d\lambda_{si}^*}{dt} = -\frac{\lambda_{si}^*}{(1+t+\mu(t))^\sigma}\left(1+\frac{d\mu}{dt}\right)\frac{1}{\Delta RR_s}\left(\frac{\Delta FC}{E_i{\Gamma_s}^k+\Delta FC}\right)$$

$$\frac{d\lambda_f^*}{dt} = \frac{\lambda_f^*}{(1+t+\mu(t))^\sigma}\left(1+\frac{d\mu}{dt}\right)\frac{1}{\Delta RR_s}\left(\frac{E_i}{E_i{\Gamma_s}^k+\Delta FC}\right)$$

and:

$$\frac{d\varphi^*}{dt} = \frac{\varphi^*}{(1+t+\mu)^\sigma}\left(1+\frac{d\mu}{dt}\right)\left[\frac{\Delta FC}{\Delta RR_1}\left(\frac{\Delta FC}{E_i\Gamma_1^k+\Delta FC}\right)+\frac{k-\sigma+1}{k'}\left(\frac{\sum_s Fisc\Gamma_{s'}^{\sigma-k-1}C_s}{\sum_s S_s Cs}\right)\right]$$

with

$$Fisc = \frac{k}{\sigma-1}\left[1-(1+t+\delta+\mu)^{1-\sigma}\right]-1+n_s$$

Sketch of the proof - Part 3: If `condition 1a` is respected, the equilibrium condition for the formal wage gap can also be rewritten as:

$$\begin{aligned}
\int_{\varphi^*}^{\infty}\left[\sum_s\int_{\lambda^*_{fs}(\varphi)}^{\infty}\left(E_f+\frac{q_f(\varphi,\lambda_j,t)}{\varphi\lambda_j}\right)\frac{k}{\lambda^{k+1}}d\lambda\right]-M\varphi_0\left[\left(\pi_{CRS}^{-1}(w_i)\right)^{-k'}-\left(\pi_{CRS}^{-1}\left((1+\delta+\mu)\,w_i\right)\right)^{-k'}\right] &= 0\\
g\left(w_i,t,\mu,\lambda^*_f,\varphi^*\right) &= 0\\
h\left(w_i,t,\mu,\lambda^*_f,\varphi^*\right)-M\varphi_0\left[\left(\pi_{CRS}^{-1}(w_i)\right)^{-k'}-\left(\pi_{CRS}^{-1}\left((1+\delta+\mu)\,w_i\right)\right)^{-k'}\right] &= 0
\end{aligned}$$

Total differentiation:

$$\underbrace{\frac{\partial h}{\partial\mu}}_{<0}\frac{d\mu}{dt} = \underbrace{-\underbrace{\frac{\partial h}{\partial t}}_{<0}-\underbrace{\frac{\partial h}{\partial w_i}}_{<0}\underbrace{\frac{dw_i}{dt}}_{>0}-\underbrace{\frac{\partial h}{\partial\lambda^*_f}}_{<0}\underbrace{\frac{d\lambda^*_f}{dt}}_{>0}-\underbrace{\frac{\partial h}{\partial\varphi^*}}_{<0}\underbrace{\frac{d\varphi^*}{dt}}_{>0}}_{>0}$$

so that finally:

$$\frac{d\mu}{dt}<0$$

While, if `condition 1b` is respected, $0>\frac{d\mu}{dt}+1$ implies $\frac{d\mu}{dt}<0$.

Sketch of the proof - Part 4 Note that:

$$\frac{dP}{dt} = \frac{P}{w_i}\frac{dw_i}{dt}-\frac{1}{k'}\frac{P}{A}\sum_s C_s\frac{dS_s}{dt}$$

Appendix D

Descriptive Statistics: Non-Compliance of Formal Employers with Labour Regulations in India

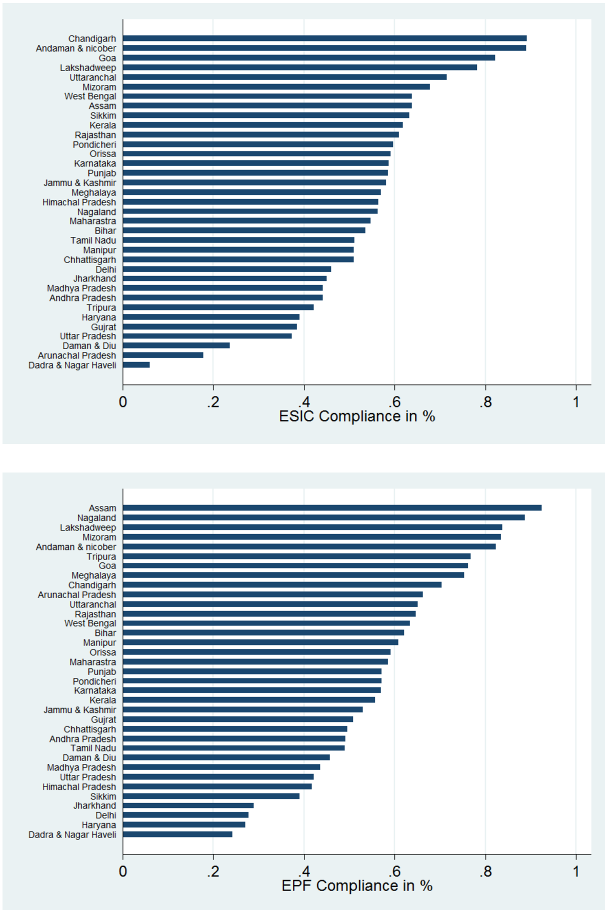


Figure D.1: **Compliance of Formal Employers with the ESI and EPF regulations**
State-wise 2004-05 (% formal employment)

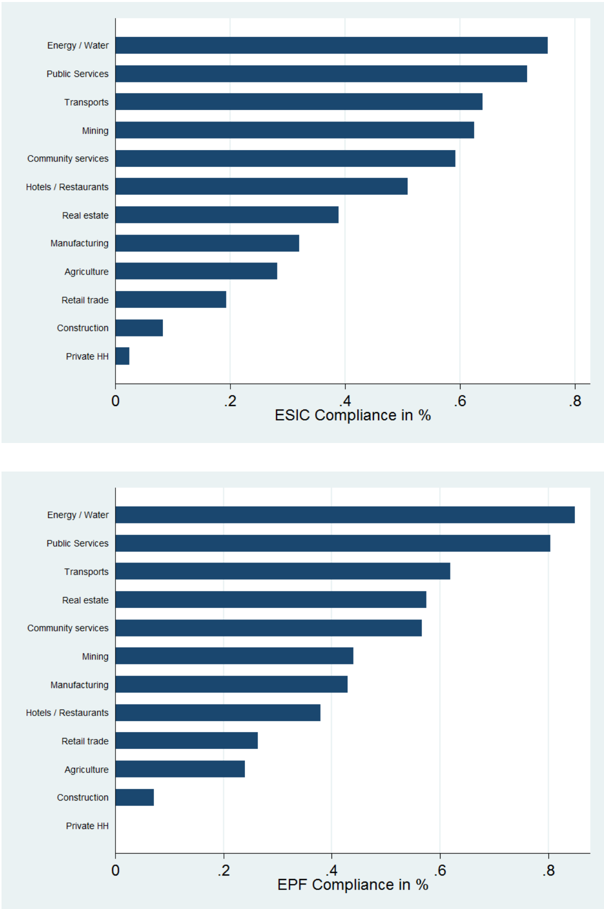


Figure D.2: **Compliance of Formal Employers with the ESI and EPF regulations**
Industry-wise 2004-05 (% formal employment)

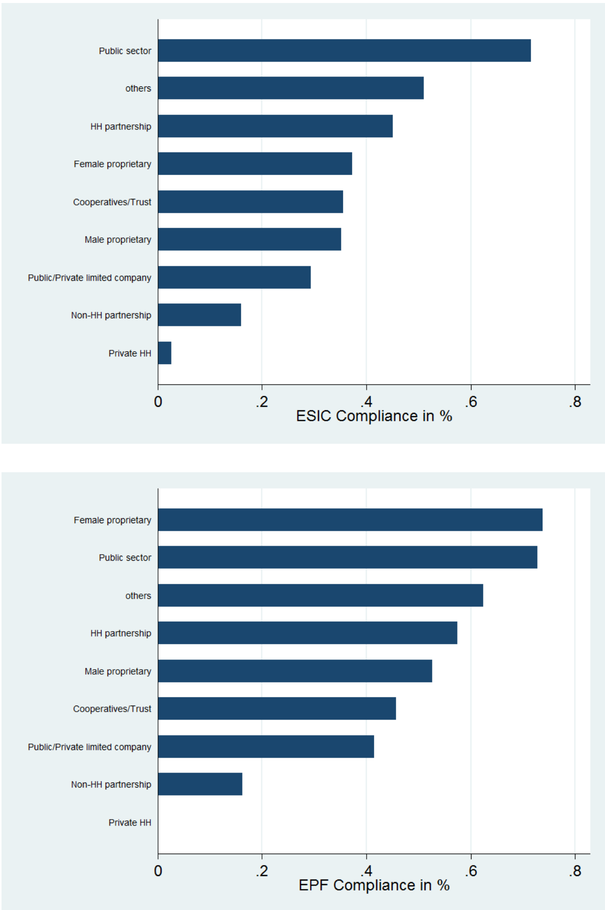


Figure D.3: **Compliance of Formal Employers with the ESI and EPF regulations** across Enterprise Type 2004-05 (% formal employment) (HH denotes household)

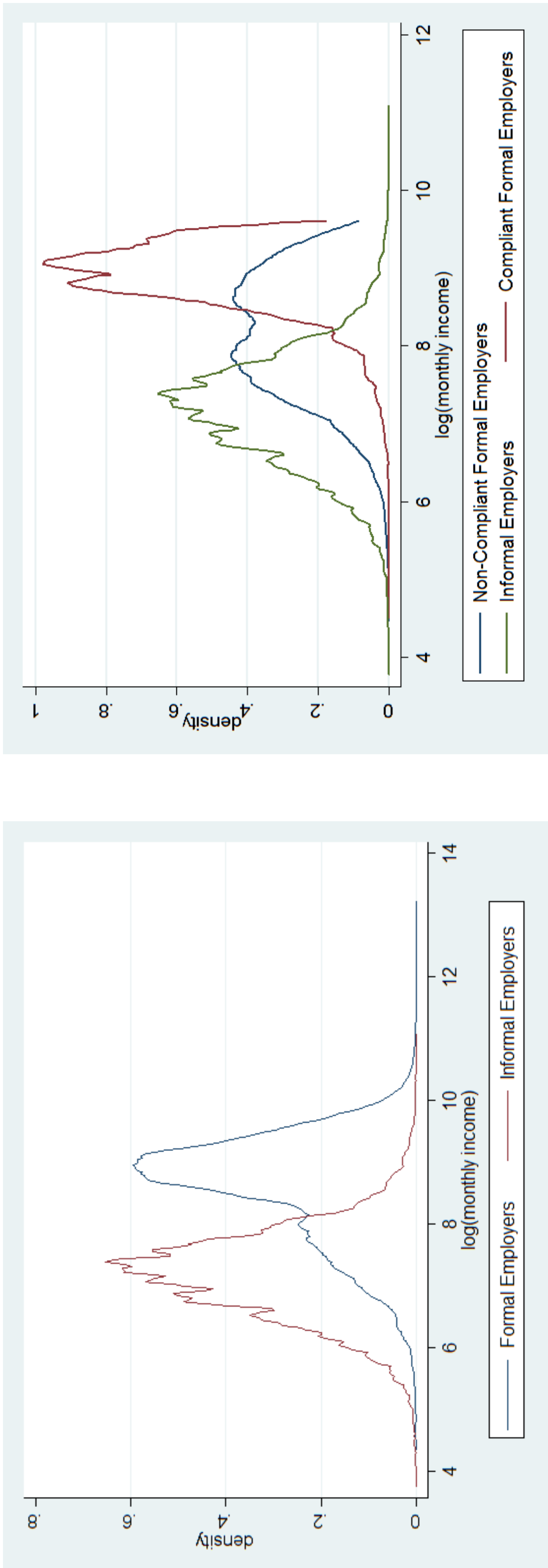


Figure D.4: Estimated Density of log earnings per Employer Status:Informal employers pay lower wages than formal employers. Wages paid by non-compliant formal employers are in-between the wages paid by compliant formal employers.

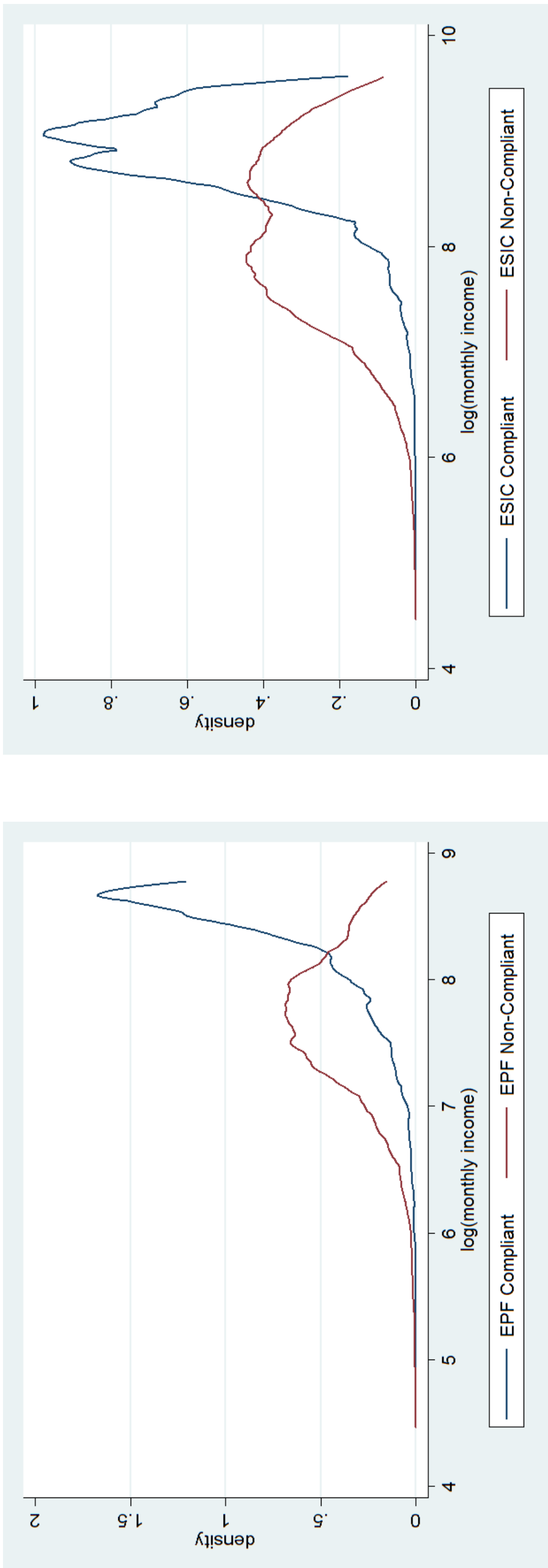


Figure D.5: Estimated Density of log earnings per Employee Status:Employees that are fraudulently non-registered with the ESI and EPF schemes are also systematically less paid than employees for which employers are compliant.

Chapter 4

Crowding-out of Energy Efficient Investments by Self-Generation: Evidence from India[1]

[1]Chapter 4 is based on Baccianti and Briand (2016).

4.1 Introduction

Most energy policies aim to support economic growth while reducing its impact on the environment. Indeed, the global and local externalities associated with the use of fossil fuels exacerbate climate change and generate significant health and economic costs from air, water, and land pollution (Greenstone and Hanna (2011), Hanna and Oliva (2015)). Rapid growth in energy demand requires constant and massive investment in power production and related energy infrastructures. For countries that do not keep a satisfactory level of investment, the resulting unbalance between supply and demand of energy disrupts firms productivity (Allcott et al. (2016)) and lowers household well-being (Sagar (2005), Nussbaumer et al. (2012)). Moreover, dependency on specific energy sources and ecosystems vulnerable to climate change exacerbates geopolitical conflicts (IPCC (2014)) and increases the vulnerability of the poorest (Majra et al. (2009)).

Investments in energy efficiency and in self-generation of renewable energy are often referred to as the twin pillars of energy policy (Prindle et al. (2007)). On the one hand, demand-side policies reduce the growing pace of energy demand and address environmental externalities. On the other hand, supply-side policies boost electricity supply and support economic growth. The India's Electricity Act 2003 aims to remediate to the chronic power deficit in India by simultaneously reducing aggregate demand through the promotion of energy efficient technologies and boosting electricity supply through the liberalization of the energy market, which includes lower costs of self-generation for manufacturers. Furthermore, all power producers, including captive power generators, are obligated to use renewable energy sources for a certain share of their total capacity (renewable purchase obligations). This policy is based on the idea, shared by other contemporary energy policies like the Energy Roadmap of the European Union,[2] that investments in energy efficiency and self-generation are complementary in addressing both environmental concerns and economic needs. Energy efficiency reduces the costs of self-generation by decreasing the required total generation capacity, while self-generation from renewable energy sources[3] decreases the emission intensity of the economy (i.e it reduces the average emission rate of pollutants of producers in the economy). Crucially, this reasoning is valid only if economic agents self-generate electricity from renewable energy sources like solar or wind.

This combination of supply and demand-side policies raises two concerns. First, the adoption of

[2] The Energy Roadmap of the European Union aims at the transformation of national energy systems through wider adoption of clean technologies and the reduction in the amount of energy required to produce one unit of Gross Domestic Product (GDP). See the official website of the Energy Roadmap of the European Union.

[3] While many renewable energy projects are conducted on a large scale, renewable energies are also suited to small scale production in particular in rural or remote areas. In 2011, 44 million Indian households waste digesters to produce biogas used in lighting and cooking and 166 million Indian households rely on biomass cookstoves (source: www.powerliners.com).

energy efficient technologies remained low in both developing and developed countries in spite of the provided fiscal incentives and of evidence of the positive impact of these investments on profitability. Second, fossil fuel-based captive plants[4] are proliferating alongside self-generation from renewable energy sources. Indeed, these technologies are more readily available,[5] better understood, and cheaper than renewable energy sources. When self-generation occurs through dirty energy sources like kerosene, it is more environmentally harmful than publicly operated larger utility power plants (Ishii and Stevens (1998) and Nag (2010)).[6] Under such conditions, the promotion of fossil fuel-based self-generation undermines the effectiveness of supply-side policies targeting greater use of energy efficient technologies. Indeed, firms do not internalize the negative externalities caused on the environment of self-generating electricity from fossil fuel. When self-generating electricity from dirty energy sources is cheaper than purchasing it from the public grid, it decreases the share of energy in total costs, which decreases in turn their incentives to invest in energy efficiency. Furthermore, in the presence of credit constraints, firms are not necessarily able to adopt both technologies even if their revenues would increase following the adoption of energy efficient technologies. Hence, in the presence of power outages, independently of the relative costs of self-generation and electricity grid, firms give the priority to investments in self-generation because it allows them not to interrupt production during power outages.

The contribution of this paper is to investigate theoretically and empirically how short-term incentives given to firms by supply-side policies can undermine in the long-run the success of policies targeting sustainable economic development through lower energy demand and higher energy savings. First, supply-side policies are one of the main factors (along with restricted access to credit) explaining the low take-up of investments in energy efficient technologies. Second, the promotion of fossil fuel-based self-generation counteracts the effects of environmental policies and hinder efforts to end poverty.[7]

Our theoretical model is based on Chaney (2008) and Bustos (2011): firms are heterogenous in pro-

[4]Captive power plants are power plants set up by private agents to generate electricity primarily for their own use. It includes generating plants operated by cooperative societies or association of private agents. The definition of "primarily" varies across countries. Furthermore, the definition of captive units implies that electricity producers may want to sell excess power.

[5]For example, a market based instrument, the Renewable Energy Certificate (REC) was subsequently created to address the mismatch between availability of renewable sources and the legal requirements of the Electricity Act 2003. See the official website of the Renewable Energy Certificate Registry of India.

[6]First, the power grid is also supplied by renewable sources of energy like hydro-electricity. The installed hydro-electric capacity in India represents 14.35% of total demand in 2016 (Source: Official website of the Ministry of Power of India.). Second, a public power provider may enjoy returns to scale.

[7]See the report: World Bank, last modified 6 Feb. 2015. *Climate Change Complicates Efforts to End Poverty.* Washington, DC: World Bank.

ductivity and can adopt either one, or the other, or the combination of two technologies. An energy efficient technology allows to produce a higher level of output with the same amount of energy and a self-generation technology makes possible the production of electricity from fuel. Firms choose the technology or combination of technology that provides them with the lowest possible variable costs, conditional on being productive enough to cover the fixed adoption costs. Productivity, relative prices of electricity and fuel, the technological frontier determine the most profitable production mode for each firm. To meet the empirical evidence presented in section 4.10.3 of this chapter, we assume that the fixed adoption costs of self-generation are higher than those of energy efficiency. Consequently, generators and energy efficient firms are present in the economy only when fuel is relatively cheaper than electricity. In this case, the most productive firms self-select into self-generation. Additionally, only the most productive generators adopt energy-efficient technologies in spite of the decrease in variable costs that it represents. Other generators are not able to afford the fixed adoption costs of both self-generation and energy efficiency: this is the *direct* crowding-out effect of energy efficiency by self-generation. Among non-generators, the most productive firms can afford the fixed adoption costs of the energy efficient technology but face a more competitive market because generators benefit from lower variable prices and charge a lower price. Hence, market shares and profits are reallocated from non-generators to generators, which decreases the share of output that is produced in an energy efficient way: this is the *indirect* crowding-out effect of energy efficiency by self-generation. Arguably, the economic significance of the direct crowding out of investments in energy efficient technologies by supply-side policies would be reduced if firms benefited from a greater access to credit markets. This chapter does not consider this dimension and simply acknowledge that access to credit is limited for most firms, in particular in India.[8]

Following the results of Allcott et al. (2016) on the significant economic costs of power shortages in India, we introduce power outages in our theoretical model during which non-generators lose a fraction of their revenues. The intensity of power outages becomes another factor that affects the relative profitability of production technologies. Both direct and indirect crowding-out effects are magnified in the presence of power outages because it enhances the profitability of self-generation. Our theoretical model shows that both the indirect and direct crowding-out effects of investments in energy efficiency by self-generation take place in an economy in which the price of fuel is larger than the price of electricity and the use of fuel to self-generate electricity. Comparing the relative costs

[8]The total credit gap in 2011 for formal SMEs worldwide was estimated between 1.5 and 1.6 trillion USD. Between 42 and 51% comes from medium enterprises. In developing countries, this figure is between 0.9 and 1.1 trillion USD. For micro-enterprises in developing countries, the credit gap is between 0.4 and 0.5 trillions USD. Source: Internatinal Finance Corporation (IFC) Enterprise Finance Gap Database.

and benefits of both technologies, self-generation should be a dominated choice in such an economy. Nevertheless, in the presence of severe electricity shortages, self-generators benefit from lower average variable costs than energy-efficient firms because they do not interrupt production during power outages.

To investigate empirically the relationship between self-generation and energy efficiency and test the main result of our model on the crowding out of investments in energy efficient technologies by self-generation, we use a repeated cross-section representative of the Indian manufacturing sector for the period 2000-10. The Indian manufacturing industry is strongly relying on self-generation of electricity. Captive power units of capacity superior to 1 megawatt (MW) represented 25 MW in 2008, or 17% of the total Indian capacity. It increased to 31 MW in 2010, while its share in total capacity remained constant. In the United States, less than 5% of total energy consumption was produced on site. These high numbers are even more surprising when one considers that the relative cost of grid-purchased and self-generated electricity for manufacturing firms in India is equal to $7/4.5 \approx 1.56$[9]. However, our theoretical model captures and explains these facts by arguing that self-generation is a way for manufacturing firms to overcome a systemic failure in balancing electricity demand and supply. Based on our theoretical results, we then expect a direct crowding-out of investments in energy efficient technologies by self-generation in India.

Data on investment in energy-saving technologies is not available to our knowledge and electricity intensity is not an accurate measure of the level of energy efficiency of a plant. Therefore, following Filippini and Hunt (2015), we rely on a stochastic frontier model to estimate the technological frontier in India and rank plants by comparing their observed electricity consumption with the most energy efficient plants in the economy. We use this measure of efficiency in electricity at the plant level as a proxy for investments in energy efficient technologies by single firms. Our theoretical model predicts a negative impact of self-generation activities of a firm on its energy efficiency ranking.

Two sources of endogeneity arise when studying the causal relationship between energy efficiency and self-generation. First, we face a reverse causality problem: firms with higher energy cost shares face higher productivity losses from electricity shortages and they have a higher incentive to invest in self-generation. Second, we face an omitted variable bias problem: the most productive firms self-select into investments in captive power generator and in the adoption of energy efficient tech-

[9] These numbers are extracted from Allcott et al. (2016) that use that use the 2005 World Bank Enterprise Survey medians to compute the relative cost of grid-purchased and self-generated electricity. Section 4.10.1 provides more details on the data sources.

nologies. To address endogeneity, we adopt an instrumental variable approach based on the cost of self-generation for Indian manufacturing firms. In India, almost all electricity generation units are run with fossil fuels like coal, furnace oil or diesel. We exploit the change in fuel taxation from the 2003-2008 VAT reform to get an exogenous variation on the variable cost of self-generation. The tax reform was intended to replace the Retail Sales Tax system with a Value Added tax system, and it has not been implemented by Indian states simultaneously. The intensity and the direction of change in fuel taxation also differed across states. Finally, electricity prices charged by the public power grid are regulated and were unaffected by changes in the sales tax reform. This results to be a powerful instrument for the IV regression. Our results suggest that self-generation crowds out investment in energy-saving technologies, in line with the predictions of the theoretical model.

Two different strands of the literature separately investigated the determinants of self-generation and energy efficiency. First, Kenneth Gillingham and Palmer (2009) give a comprehensive review of the literature on the determinants of energy efficiency. Robust evidence was found on the importance of energy prices with respect to technology adoption decisions (Henri L.F. De Groot and Nijkamp (2001), Linn (2008)) but also with respect to innovation by inducing progress on the technological frontier of energy efficient technologies (Adam B. Jaffe and Stavins (1999),Aghion et al. (2016), and Popp (2002)). The possibility to invest in energy efficient technologies implies that energy demand is characterized by low-price elasticity in the short run but by high-price elasticity in the long-run (Atkeson and Kehoe (1999)). Eventually, the literature on the determinants of the adoption of energy efficient technologies focused on market failures, market barriers, and on behavioral aspects (high discounting of future cash flows, incomplete information) that can explain the observed modest pace of diffusion of energy efficiency across households and firms (Kenneth Gillingham and Palmer (2009), Gerarden et al. (2015)). This chapter suggests that inconsistencies between energy policies may also be an explanation. Second, Allcott et al. (2016) show that power outages have a significant and positive impact on self-generation because it allows firms to insure[10] themselves against the revenue losses that they cause. Rud (2012) develops a model of heterogeneous firms in productivity with technology adoption based on Chaney (2008) and Bustos (2011) to explain the impact of low reliability of infrastructures on the adoption rates of self-generation among Indian firms. In his model,

[10]Installed self-generation capacity may not be sufficient to cover the firm's entire energy requirements and thus allow the firm to operate without any publicly-provided electricity. Thus, firms might not be fully insure against power outages, but self-generation still decreases losses in profits caused by outages. Allcott et al. (2016) model explicitly firms' self-generating self-capacity decision. In this chapter, we model only the extensive margin of self-generation and assume that generators are able to produce enough electricity to cover their energy needs. We discuss the robustness of our results to this assumption in section 4.5.2.

self-generation requires a fixed production cost but reduces variable energy costs. This decrease in variable costs is proportional to each firm's idiosyncratic productivity but fixed adoption costs are constant across firms. Hence, the most productive firms self-select into self-generation. The author argues that this mechanism mitigates the negative effects on output of the cross-subsidies of electricity prices from industries to households in India. As expected, the share of generators increases with the intensity of electricity shortages, with the price of electricity from the public grid, and with barriers to competition. Rud (2012) tests his model's predictions and his results are consistent with the indirect crowding-out effect highlighted in this paper. The introduction of an alternative technology in the economy reallocates market shares towards the most productive adopters, which raises competition and the minimum level of productivity required to adopt it. We adopt a more comprehensive approach of firms' energy-related decisions and investigate the impact of electricity self-generation on firms' incentives to invest in another technology enhancing energy efficiency.

Section 4.2 gives an overview of the power sector and manufacturing industry in India. It also highlights the significant economic costs of pollution and climate change for the Indian economy and gives an overview of the environment policies in the country. In section 4.4, we lay out a basic model with heterogeneous firms producing different varieties of a good using electricity and labour. In section 4.5, we introduce the possibility for firms to invest in two different technologies: an energy efficient technology that reduces variable costs, and an electricity generating technology that allows the firm to substitute electricity purchased from the public power grid by electricity it produces itself from fuel. In section 4.8, we introduce electricity outages in the economy: electricity supply from the power grid is limited and firms that are not able to generate electricity from fuel lose a share of their revenues during these interruptions. Section 4.9 enters into the details of our empirical strategy to test the main prediction of the theoretical model. A description of the data is given in section 4.10. Section 4.10.2 describes our strategy to measure firms' energy efficiency. Section 4.11 presents the empirical results.

4.2 Institutional Framework In India

4.2.1 Energy Consumption in the Indian Manufacturing Sector

By 2030–35, India is forecasted to be the country with the highest energy demand according to the 2014 energy outlook report of the oil company British Petroleum.[11] Per capita consumption of energy crossed the 1,000 kilowatts (KW) in 2015 but is still far below the average global consumption. In the same year, generation capacity in India was of 276 gigawatts (GW) from which 69% come from thermal energy (coal and fuel), 15% hydroelectricity, 13% from renewable, and 2% from nuclear energy. Within the industry, captive power plants of capacity superior to 1 megawatt (MW) represented 25 MW in 2008, that is to say, 17% of the economy's total capacity. It increased to 31 MW in 2010 but its share in total capacity remained constant.

The manufacturing sector in India suffers from several issues when it comes to its access to electricity. First, electricity prices for industrial consumers are higher than the marginal cost of supply. On the contrary, agricultural consumers and households are responsible for around half of electricity consumption but benefit from a reduced price. Thanks to these cross-subsidies from the industry, farmers pay only 12% of the average cost of power supply[12], while households pay only 54%. Second, the Indian power grid suffers from systemic shortages as described in Allcott et al. (2016). These authors show that electricity shortages are one of the determinants of the rapid growth of self-generation by Indian industrials. Reasons for these shortages are multiple. Electricity prices never differentiate between peak and low demand periods. Hence, users do not adjust their consumption during peak time, which worsens existing electricity waste caused by the absence of proper consumption metering and low prices for non-industrials users. Allcott et al. (2016) note that there is no correlation between shortages and the median electricity price paid by manufacturing industries, once state and year fixed effects are controlled for. It demonstrates that prices do not adjust to supply and demand in India, largely because the State Electricity Boards do not regularly update electricity tariffs. Power utilities lack the financial capacity to increase production even though their financial losses are covered by government subsidies. Given the importance of the necessary investment to expand and increase the reliability and capacity of public electric supply in India, the situation of the Indian power grid is only expected to improve slowly. Only 51.5% of the capacity

[11] The report can be found on the BP corporate website.

[12] The average price per kilowatt-hour is computed dividing the revenue collected by the quantity sold to a given category of consumers. Official statistics probably under-estimate the average price paid by agriculture because only some of the consumers are actually paying their bills, or even registered as official consumers. Nevertheless, the same is true for the average price by industrials.

addition targeted in the 10^{th} five-year plan period (2002-2007) was reached, and only 66% of the capacity targeted in the 11^{th} plan (2007-2012).

Two important features of the Indian power sector could potentially impact our results. First, 26% of electricity generated in India in 2010 was lost due to illegal connections to the power grid. In 2005, these losses represented as much as 34% of total production (Allcott et al. (2016)). Power theft are linked to the political inability to crack down on the responsible because they fear a voter backlash.[13] Power theft also implies that some consumers have a free access to electricity: there is evidence that industrials register their power connection as private.[14] Power theft decreases the costs of electricity for firms but may also worsens power shortages. However, it does not impact the costs of self-generation or energy efficient technology and thus does not undermine the results of our analysis. The second problem is due to the limitations of our data: the Annual Survey of Industry (ASI)[15] includes only formally registered firms. The informal manufacturing sector, as shown in Chapter 2, is responsible for a large share of employment and represents the majority of firms. Furthermore, informal firms may be responsible for a large share of power theft. Unfortunately, the Unregistered Manufacturing Enterprise (UME) survey used in Chapter 2 does not contain any information on the energy consumption profile of informal firms. Nevertheless, we can argue that the construction of a captive power plant requires a significant financial investment that credit constrained informal firms would find difficult to cover.

4.2.2 Energy Policy in India

The Indian policy agenda has attempted to tackle the electricity-related problems faced by manufacturing firms with policies targeting both the supply and demand side of the electricity sector. On the one hand, the central government sought to open the electricity market to new players and to introduce performance-based competition, instead of guaranteed arrangements between public companies and the private sector. The Electricity Act was enacted in 2003. It covers major issues involving generation, distribution, transmission and trading in power. Electricity generation was delicensed and private firms no longer require clearance from the Central Electricity Authority (CEA) to produce and sell electricity. In some states, private producers are even able to sell directly to other private consumers. Licenses for electricity distribution have been abolished in rural areas, and the licensing

[13]See McRae (2014) for a discussion of the infrastructure quality and subsidy trap: the vicious circle between poor quality of public services, for which voters refuse to pay for. Consequently, governments must subsidize power utility companies to limit their losses and support their voters.

[14]See the article: Ashish Royl, "MSEDCL launches drive against Illegal connections", *The Times of India*, 23 June 2016, accessed 13 Aug. 2016

[15]See section 4.10.1 for a detailed description of the data

regime has been simplified in urban areas. A renewable purchase obligation (RPO) scheme has also been implemented to drive demand for solar and wind energy. Under this scheme, electricity distributors can either generate a minimum amount of renewable power, or purchase renewable energy certificates (REC). On the other hand, the central government took action to decrease the energy intensity of the Indian economy with the Energy Conservation Act 2001. One of the objectives of this piece of legislation was to increase the energy efficiency of Small and Medium Enterprises (SMEs), in particular through the purchase of energy efficient technologies. SME financing had already been defined as a priority sector lending in India, which means that commercial banks must meet some requirements in terms of lending to this type of firms. For example, targets set for commercial banks include a 20% annual growth in the growth of their SME credit portfolio.

4.2.3 Environmental Risks and Policy in India

Vulnerability to climate change is a function of exposure, sensitivity, and adaptive capability (IPCC (2014)). Arid areas, high mountain zones, and densely populated coastal areas are considered to be particularly vulnerable to small changes in global temperature. As discussed in a World Bank's report,[16] India is then particularly at risk with the 7500 km long Great Himalayas and the densely populated coast line in the south of the country. Indeed, Majra et al. (2009) underline that 700 million Indian people depend on climate-sensitive sectors (agriculture and forestry for instance) and fragile ecosystems (biodiversity, mangroves, coastal zones). Furthermore, air pollution has already become a sanitary concern: the World Health Organization (WHO)[17] finds that the five most heavily polluted cities in the world in 2016 are in India. WHO's statistics also indicate that India is the country with the highest death rate from chronic respiratory diseases and the largest number of deaths from asthma, which adds pressure to the overloaded health infrastructures and to the sprouting welfare system.

In spite of the seriousness of the situation, Indian representatives blocked negotiations during the 2015 United Nations Climate Change Conference by refusing to commit towards specific actions. Indian politicians emphasized at this occasion the priority of the fight against poverty on environmental concerns.[18] This apparent paradox is explained by the fact that Indian politicians hold developed countries responsible of climate change through their own earlier economic growth. However, this

[16]World Bank. 2013. *Turn Down the Heat: Climate Extremes, Regional Impacts, and the Case for Resilience.* A report for the World Bank by the Potsdam Institute for Climate Impact Research and Climate Analytics. Washington, DC:World Bank.

[17]See the website of the Global Health Observatory (GHO) that includes statistics on 1,600 cities in 91 countries.

[18]See the article: Ellen Barry, "For Indians, smog and poverty are higher priority than talks in Paris", *New York Times*, 9 Dec. 2015, accessed 13 Aug. 2016

neither implies that the Indian population is insensitive to the issue of climate change and air pollution nor that there are no environmental policies in India.[19] On the contrary, even the poorest have been found to be aware of the economic and health benefits of renewable energy: Urpelainen (2016) conducted a survey in Uttar Pradesh and shows that most of the 1597 surveyed households simultaneously exhibit dissatisfaction with the use of kerosene lighting and high awareness of the potential benefits of replacing it by using solar power.[20]

In the world's largest democracy, politicians have already picked up[21] on this national sentiment but environmental policies have been mostly imbedded into broader social, economic and urban policies rather than established as stand-alone regulations.[22] Indeed, the fight against poverty overlaps because climate change "hits the poorest people the hardest".[23] Hence, several policies are already implemented to mitigate the effects of climate change and improve the adaptability of the most vulnerable population. For example, the Smart Cities Mission[24] aims to improve the sustainability of 100 selected cities. This includes to improve the treatment of water sewage, limit traffic congestion, garbage disposal, and other aspects that will significantly limit air, water, and land pollution in metropolitan areas. A further example is the Biological Diversity Act that was implemented in 2002 after India ratified the Convention on Biological Diversity of the United Nations Environment Programme (UNEP). In accordance, a National Biodiversity Action Plan (NBAP)[25] is prepared and implemented every year. Crucially, it does not only aim to preserve wildlife but is designed to preserve the livelihood of population residing in fragile ecosystems and to reassert the ownership by these populations of natural resources (medicinal plants, precious woods, and profits from tourism...).

In this chapter, we argue that self-generation from dirty sources of energy can not only crowd investments in energy efficient technologies out, but it also undermines the effects of other environmental policies led by the Indian government.

[19] See the article: The Editorial Board, "Cutting Through India's Smog", *New York Times*, 23 Feb. 2015, accessed 13 Aug. 2016

[20] Nevertheless, Urpelainen (2016) also shows that households appear reluctant to purchase this service from a private market provider and would rather trust the government to offer this service. Nevertheless, Aklin et al. (2015) show that the relationship between the electorate's satisfaction with implemented energy policies and their propensity to take seriously a politician's electorate promises on this topic is complex. The authors conduct a survey in two villages in Uttar Pradesh and show that voters that experience low level of services tend to discard as non-credible the promises made by politicians when it comes to improve these services and do not take them into account when choosing their preferred candidates during an election.

[21] See the article: Gardiner Harris, "Plan to Verify Delhi Pollution Data Raises Suspicions", *New York Times*, 11 March 2015, accessed 13 Aug. 2016

[22] See the following blog article funded by the European Climate Foundation's International Policy and Politics Initiative (IPPI): Sunita Dubey and Srinivas Krishnaswamy, "India and Elections: Will the environment and climate ever be a priority for the world's largest democracy?", *Nivela Organization Blog*, 6 March 2014, accessed 14 Aug. 2016

[23] See the report: World Bank, last modified 6 Feb. 2015. *Climate Change Complicates Efforts to End Poverty.* Washington, DC: World Bank.

[24] See the official website of the Smart Cities Mission of the Ministry of India of Urban Development.

[25] See the official website of the Ministry of India for Environment, Forest and Climate Change.

4.3 Introductory Note to the Theoretical Model

Throughout this Chapter, we use the term *efficient use of electricity* as a qualification of the production technology used by the firm: we refer to the use of electricity as *efficient* when a smaller amount of electricity is required to produce one unit of output, not when it increases profitability. This Chapter considers two different types of crowding out of investments in energy efficient technologies by self-generation: the *direct* and *indirect* crowding out effects as defined below:

Definition 1. *We refer to the* ***direct crowding-out*** *of investments in energy efficiency by investments in self-generation, when profits from adopting the energy efficient technology are positive and superior to those obtained from using the standard technology but firms do not invest in energy intensive technology because investments in self-generation are more profitable.*

Definition 2. *We refer to the* ***indirect crowding-out*** *of investments in energy efficiency by investments in self-generation, when firms at the top of the productivity distribution are adopting both the energy efficient and the self-generation technologies: Market shares are reallocated towards those firms that benefit from a larger decrease of their variable costs and consequently of their prices. First, the minimum productivity level that is required for investments in the energy efficient technology to be profitable increases: the share of energy efficient firms decreases. Second, the relative share of output produced by energy efficient firms in the economy also decreases.*

First, our theoretical model aims to show the mechanisms through which these two different types of crowding-out are taking place in India where there are power outages and where electricity from the grid is cheaper than fuel used in captive power plants. Second, we test empirically the predictions of this model using Indian data in section 4.9. Hence, our theoretical model focuses on the **empirically relevant case for India of electricity from the public grid being cheaper than fuel**. Indeed, Allcott et al. (2016) use the 2005 World Bank Enterprise Survey medians to compute the relative cost of grid-purchased and self-generated electricity in India. They find that it is equal to $7/4.5 \approx 1.56$. Hence, the relative price of electricity to the relative price of fuel is larger than 1. **Case 1 (cheap electricity case)** describes an economy in which the use of fuel is more expensive than the use of electricity from the grid. Other possible cases on the relative prices of fuel and electricity from the public grid are described in table 4.1. The equilibria corresponding to these cases will be solved in section 4.7, in which we refer to an economy in which the use of fuel is cheaper than the use of electricity from the grid. The different possible equilibria when fuel is cheaper than electricity from the grid are not empirically relevant for our chapter on India but we will use these results to simplify the analysis of the equilibrium in Case 1 in the presence of power outages as shown in table 4.2. Indeed, the Indian Central Authority estimated power outages in India to be 7.1% on average across all plants in 2005.[26] Thus, the **empirically relevant case for theoretical predictions** on the impact of self-generation of electricity on the adoption of energy efficient technologies is **when electricity from the grid is cheaper than fuel (Case 1) and there are power outages**.

Table 4.1: Summary of the possible equilibria **without power outages** based on relative energy prices and technological parameters. (**Use of Electricity** refers to the unit cost of energy when using electricity from the grid and standard technologies (not energy efficient), **Efficient Use of Electricity** refers to the unit cost of energy when using electricity from the grid and energy efficient technologies, **Fuel** refers to the unit cost of energy when self-generating electricity from fuel.

Without Power Outages

Electricity price lower than Fuel price	Fuel price lower than Electricity price	
Standard use of Electricity cheaper than Fuel	Efficient use of Electricity cheaper than Fuel	Fuel cheaper than Efficient use of Electricity
Case 1	Case 2	Case 3
No Crowding-Out	Indirect Crowding-Out	Direct and Indirect Crowding-Out

Table 4.2: Summary of the possible equilibria **with power outages** based on technological parameters in Case 1 when electricity from the grid is cheaper than fuel. (**Use of Electricity** refers to the unit cost of energy when using electricity from the grid and standard technologies (not energy efficient), **Efficient Use of Electricity** refers to the unit cost of energy when using electricity from the grid and energy efficient technologies, **Fuel** refers to the unit cost of energy when self-generating electricity from fuel.

With Power Outages

Electricity price lower than Fuel price		
Rare outages and/or Low technological progress	Frequent Outages and/or Medium technological progress	Very frequent Outages and/or High technological progress
Standard use of Electricity cheaper than Fuel	Efficient use of Electricity cheaper than Fuel	Fuel cheaper than Efficient use of Electricity
Case 1	Case 2	Case 3
No Crowding-Out	Indirect Crowding-Out	Direct and Indirect Crowding-Out

We derive the equilibrium for case 1 when there are no power outages in the economy in section 4.6. In section 4.8, we solve the equilibrium in case 1 in the presence of power outages and build the

[26]Section 4.10.1 provides more details on the data sources. Allcott et al. (2016) mention that average electricity outage for the year 2005 is very close to the average over the 1999-2000 sample in which the national shortage average ranged from 6.4 to 11.1%.

theoretical predictions that we will test empirically in section 4.9.

4.4 Baseline Model

4.4.1 Production

There are two sectors in the economy: The first sector produces a homogenous numéraire good X with constant returns to scale, and the second sector produces a differentiated good. It is characterized by a continuity of firms that are heterogenous in productivity. Each firm produces a different variety $\omega \in \Omega$ with increasing returns to scale. The production of one unit of X requires one unit of labour so that the wage is equal to one in the economy, and total labour income is fixed to one as well[27].

4.4.2 Preferences

The representative consumer maximizes a quasi-linear utility function of the quantities of the homogenous good X and of a composite good Q of the different varieties $\omega \in \Omega$ produced in the economy:

$$U(X,Q) \quad = \quad Rln(Q) + X \tag{4.1}$$

with $R > 0$ and Q the CES composite good with $\sigma > 1$ the elasticity of substitution between different varieties:

$$Q \quad = \quad \left(\int_{\Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \tag{4.2}$$

The utility maximization problem results in constant aggregate expenditure on the differentiated sector $R = PQ$ with $P = P(\omega)$ the price index in the economy. The sum of aggregate profits and labor income defines total income $Y = \Pi + 1$, which is spent on the composite and homogenous good $Y = R + X$. Thus, the representative consumer's demand for each variety is:

$$q(\omega) \quad = \quad \left(\frac{p(\omega)}{P} \right)^{-\sigma} \frac{R}{P} \tag{4.3}$$

and the price index in the economy is:

[27]For the differentiated good to be produced in equilibrium, aggregate labour demand in this sector must be assumed inferior to one. All results hold when the mass of labour is greater than one.

$$P(\omega) = \left(\int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}} \tag{4.4}$$

Hence, total spendings on variety ω depends on its price relative to other prices in the economy. Firms charging a lower price will obtain a larger share of the market.

4.4.3 Technology and Firm Maximization Problem

Firms purchase electricity from the public grid at a regulated (thus exogenous) price p_E and use it as an input to produce a unique variety ω_i that is sold to the representative consumer on a market characterized by monopolistic competition. At no extra cost, a firm draws its productivity φ from a Pareto distribution with shape parameter[28] $k > \sigma$. The cumulative distribution function is given by $V(\varphi) = 1 - \left(\frac{1}{\varphi}\right)^k$, whose probability density function is $v(\varphi) = \frac{k}{\varphi^{k+1}}$ for $\varphi \geq k$. Explicitly, the productivity parameter φ is the inverse of the firm's variable per-unit electricity production requirements.[29]After a firm knows its productivity draw, firms has to pay a fixed entry cost consisting of f units of electricity to start producing. Thus, the cost function of a firm of productivity φ producing quantity q is given by:[30]

$$C(\varphi) = p_E \left[\frac{q(\varphi)}{\varphi} \right] + f \tag{4.5}$$

A firm sets the price p_φ of its variety to maximize total profits $\pi(\varphi) = p_\varphi q(p_\varphi) - p_E \left(\frac{q(p_\varphi)}{\varphi}\right) - f$. Due to CES demand functions, profit maximization leads to the standard constant markup pricing

[28]For the equilibrium quantity produced to be well-defined, we assume $k > \sigma$.

[29]Given its productivity draw, each firm produces a single variety but different firms can have the same productivity. Since all firms of productivity φ will charge the same price, and sell the same quantities, we can relabel equations 4.3 and 4.4 and study the distribution of the productivity parameter φ rather than varieties ω.

[30]This production function implies that the share of energy expenses in total input costs evolves hand in hand with energy prices, and decreases with productivity. The elasticity of substitution between energy and value-added (understood here as the composite of all other intermediate inputs) is generally assumed to be lower than one: energy and value-added are thus complements in production. When the price of energy rises, the share of energy costs in total production costs rises as well. With a Cobb-Douglas function of the type $C(\varphi) = (p_E)^\beta w^{1-\beta} \left[\frac{q(\varphi)}{\varphi}\right] + f$, variations in input prices have no impact on the respective cost shares of these inputs. We observe that in all countries the aggregate energy cost share has declined over time as a consequence of energy-saving technical change, and therefore the use of a Cobb-Douglas would not meet empirical evidence gathered by the literature on the complementarity of energy and value-added (Koesler and Schymura (2012) and Baccianti (2013)).

rule: $p(\varphi) = \frac{\sigma}{\sigma-1}\frac{p_E}{\varphi}$ and to the following profit functions:

$$\pi(\varphi) = \frac{R}{\sigma}\left(\frac{\sigma-1}{\sigma}\frac{\varphi}{p_E}P\right)^{\sigma-1} - f \tag{4.6}$$

with firms revenue functions:

$$r(\varphi) = R\left(\frac{\sigma-1}{\sigma}\frac{\varphi}{p_E}P\right)^{\sigma-1} \tag{4.7}$$

4.4.4 Equilibrium

Following Chaney (2008), the set of possible firm entrants J is fixed to one. Only a subset of firms will be active in equilibrium. The economy is in steady state such that firms entry equals firm exit. All active firms make non-negative profits. The productivity level of the marginal firm making zero profits is the minimum level of productivity φ^* necessary to produce in the economy, and verifies $\pi(\varphi^*) = 0$. Using equation 4.6, we solve for the *production cut-off* φ^* :

$$\varphi^*(P) = \frac{\sigma}{\sigma-1}\frac{p_E}{P}\left(\frac{\sigma f}{R}\right)^{\frac{1}{\sigma-1}} \tag{4.8}$$

Therefore, the mass of active firms is $J_A = (1 - V(\varphi^*))$, and the price index can be written as a function of the production cut-off:

$$P(\varphi^*) = p_E\frac{\sigma}{\sigma-1}\frac{k}{k-\sigma+1}(\varphi^*)^{\sigma-k-1} \tag{4.9}$$

Combining equations 4.8 and 4.9, we get the equilibrium cut-off and the equilibrium price index:

$$\varphi^* = \left(p_E\frac{\sigma}{\sigma-1}\right)^{\frac{\sigma}{k}}\left(\frac{k}{k-\sigma+1}\right)^{k}\left(\frac{\sigma f}{R}\right)^{1-\frac{k-\sigma+1}{k(\sigma-1)}} \tag{4.10}$$

$$P^* = \left(p_E\frac{\sigma}{\sigma-1}\right)^{\frac{k-\sigma}{k}}\left(\frac{k-\sigma+1}{k}\right)^{k}\left(\frac{\sigma f}{R}\right)^{\frac{k-\sigma+1}{k(\sigma-1)}} \tag{4.11}$$

Aggregate productivity of active firms is conveniently given by:

$$\overline{\varphi}(\varphi^*) = \left[\frac{1}{J_A}\int_{\varphi^*}^{\infty}\varphi^{\sigma-1}v(\varphi)\,d\varphi\right]^{\frac{1}{\sigma-1}} = \varphi^*\left(\frac{k}{k-\sigma+1}\right)^{\frac{1}{\sigma-1}} \tag{4.12}$$

Firms of higher productivity φ benefit from lower variable costs, and are thus able to charge a lower price. They capture a larger share of the market and generate larger output, revenues and profits. The price index is taken as exogenous by every single firm but it decreases[31] with the average productivity in the economy, which is itself increasing with the minimum production cut-off as seen in equation 4.12. Indeed, as the productivity in the pool of active firms increases, those firms charge lower prices. Competition on the market intensifies, which makes it more difficult for firms at the lower end of the production distribution to enter the market.

4.4.5 Welfare and Pollution

Based on the quasilinear preferences of the representative consumer, total expenditure on the differentiated good sector is fixed and aggregate profits in the economy are given by:

$$\Pi = \int_{\omega} \pi(\omega)\, d\omega = \frac{\sigma-1}{\sigma}\frac{R}{k} \tag{4.13}$$

Thus, aggregated profits do not play a role in determining welfare because they are fixed by the constant aggregated expenditure on the differentiated sector. Any change in this sector results in the reallocation of profits across its firms, not in a change in aggregate profits. Welfare is $W = Y + CS - D$ where total income in the economy $Y = \Pi + 1$ equals total profits Π added to labour income fixed to 1 by assumption, CS is the consumer surplus, and D is an increasing function of the total level of electricity used in production. It accounts for the negative externality caused by the use of energy and thus increases with the level of electricity consumed for production $\left(\frac{\partial D}{\partial I} > 0\right)$. The consumer surplus is a decreasing function of the price index in the economy: $CS = Rln(Q) - PQ = Rln\left(\frac{R}{P}\right) - R$. Hence, welfare can be written as a function of the equilibrium price index:

$$W(P^*) = \frac{\sigma-1}{\sigma}\frac{R}{k} + 1 + R(ln(R) - 1) - Rln(P^*) - D(I^*) \tag{4.14}$$

First, tighter competition and higher production cut-off φ^* decrease the price index P^* and increase the consumer surplus. Second, total electricity consumption is given by $I = \frac{Q}{p_E\overline{\varphi}} = \frac{R}{p_E P\overline{\varphi}(P)}$. Equations 4.9 and 4.12 imply:

$$I^* = \frac{R}{p_E^2}\frac{\sigma-1}{\sigma}\left(\frac{R}{\sigma f}\frac{k-\sigma+1}{k}\right)^{\frac{1}{\sigma-1}}, \qquad \frac{\partial I(P)}{\partial P} = 0 \tag{4.15}$$

With constant expenditures on the differentiated good sector, a lower price index P^* goes hand in hand with a larger level of output ($R = P^*Q$) and, from equation 4.8, a higher aggregate productivity

[31] From equations 4.9 and 4.12, we have $P^* = p_E\frac{\sigma}{\sigma-1}\left(\frac{k-\sigma+1}{k}\right)^{\frac{-k}{\sigma-1}}\overline{\varphi}^{1-\frac{k}{\sigma-1}}$ with $1 - \frac{k}{\sigma-1} < 0$ since $k > \sigma$.

in the economy. This implies in turn a more efficient use of energy. The increase in required energy caused by the increase in output in an economy with a lower price index will be exactly compensated for by a more efficient use of energy by the firms that remain on the market and are more productive. Hence, in the baseline model presented in this section, the assumption that expenditures on the differentiated sector are constant implies that variations in output do not impact the level of negative externality caused by the use of energy[32]. Hence, welfare is impacted only by changes in the consumer surplus through the price index in the baseline model.

4.5 Economy with Technology Adoption

The model presented in section 4.4 is a simple autarky version of the Chaney (2008) 's heterogenous firms framework. In this section, we introduce two different types of production technologies that firms can use after paying some fixed adoption costs: and energy efficient technology and a self-generation technology. In this chapter, "energy efficiency" or "efficient use of electricity" means that the technology requires a smaller amount of electricity to produce one unit of output. It **does not necessarily** mean that it is always efficient for a firm to use this technology in the economic sense that it always increases profitability. The energy efficient technology allows for a more efficient use of every unit of electricity consumed and thus reduces variable costs. The self-generating technology allows firms to generate electricity from fuel for their own consumption and thus frees them from relying on the public power grid for production. Similarly to fixed production costs, fixed adoption costs are covered by labour. Firms that do not adopt an alternative production technology will be referred to as *standard firms* in the remaining sections of this chapter.

The introduction of two production technologies in the economy gives rise to four possible production modes characterized by i) the use of electricity grid, ii) the efficient use of electricity grid, iii) the use of electricity self-generated from fuel (*the use of fuel thereafter*), and iv) the efficient use of electricity self-generated from fuel (*the efficient use of fuel thereafter*: combination of an electricity generator requiring fuel and of the efficient use of fuel produced by this generator.). Each production mode exhibits different fixed and variable costs. Intuitively, firms choose the technology or combination of technologies that offers the lowest variable costs conditional on the fact that they are productive

[32] With a single production technology, the level of energy consumed in the economy results from preferences and not from the intensity of competition on the market. Hence, the regulator could modify it only through policies affecting the level of expenditures spent by the representative consumer on the differentiated sector. Our baseline model refers then to a situation where consumption patterns are a society's choice and are not constrained by paternalistic regulations. In section 4.5, we introduce alternative production technologies. The regulator can then impact the energy input mix used by the differentiated sector. The average unit price of energy paid by firms in the differentiated sector is then no longer equal to p_E and I^* evolves hand in hand with this average unit price.

enough to cover the fixed costs induced by the adoption of this technology. We can already note that the efficient use of fuel will dominate the efficient use of electricity if the use of fuel is more profitable than the use of electricity (and vice-versa). However, the efficient use of fuel may dominate the efficient use of electricity, even if the use of fuel does not dominate the use of electricity (and vice-versa). Finally, the efficient use of fuel can be ruled out only in the case where the use of fuel is dominated by the use of electricity. The remaining of this section presents each alternative production technology, its associated costs, revenue, and profit functions, and the consequences of its use on welfare, as well as other useful notations.

4.5.1 Energy Efficient Technology

Following Bustos (2011), firms can upgrade to a more energy efficient (EE) technology and reduce variables costs by paying an additional fixed cost f_{EE}. The cost function of a firm of productivity φ using the energy efficient technology to produce quantity q becomes (with $\gamma > 1$):

$$C_{EE}(\varphi) = f + f_{EE} + p_E \left[\frac{q(\varphi)}{\gamma\varphi}\right] \tag{4.16}$$

Firms producing with the energy efficient technology will be referred to as *energy efficient firms*. Total profits are given by $\pi_{EE}(\varphi) = p_\varphi q(p_\varphi) - p_E\left(\frac{q(p_\varphi)}{\gamma\varphi}\right) - f - f_{EE}$, which implies that energy efficient firms will charge a price $p_{EE}(\varphi) = \frac{\sigma}{\sigma-\mathbf{1}}\frac{\mathbf{p_E}}{\gamma\varphi}$. The revenue function of an energy efficient firm of productivity φ is given by $r_{EE}(\varphi) = \gamma^{\sigma-1} r(\varphi)$.

Minimum Profitable Productivity cut-off

There exists an energy efficient production cut-off $\varphi_{EE} = \left(\frac{f+f_{EE}}{f}\right)^{\frac{1}{\sigma-1}}\varphi^*$ such that all firms of productivity $\varphi \geq \varphi_{EE}$ will generate positive profits when adopting the energy efficient technology $(\pi_{EE}(\varphi) \geq 0)$. Conditional on their productivity, firms face a trade-off between higher fixed costs and lower variable costs following the adoption of EE technology. Assuming that $f_{EE} \geq f\left(\gamma^{\sigma-1} - 1\right)$, [33] profits maximization induced by equations 4.5 and 4.16 implies that it will be profitable for the most productive only to invest into an EE technology. The key mechanism at play is that the decrease in variable costs caused by the more technologically efficient use of electricity is proportional to intrinsic firm productivity, while fixed adoption costs are constant across firms. Hence, there exists

[33] This restriction on the model's parameters insures that, consistently with the empirical evidence presented in section 4.10.3, the fixed adoption costs of the energy efficient technology are sufficiently high for the most productive active firms only to afford it ($\varphi'_{EE} > \varphi^*$).

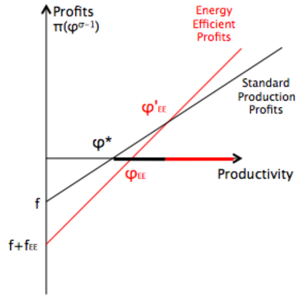


Figure 4.1: Profit functions of Standard and Energy Efficient firms

a productivity cut-off $\varphi'_{EE} > \varphi_{EE} > \varphi^*$ above which firms find profitable to pay the fixed adoption costs in order to enjoy lower variable costs. As illustrated in figure 4.1, the EE cut-off is defined by:

$$\pi\left(\varphi'_{EE}\right) = \pi_{EE}\left(\varphi'_{EE}\right) \tag{4.17}$$

Solving equation 4.17, we obtain:

$$\varphi'_{EE} = \varphi^*\left(\frac{f_{EE}}{\left(\gamma^{\sigma-1}-1\right)f}\right)^{\frac{1}{\sigma-1}} = \chi\varphi^* \tag{4.18}$$

where $\frac{1}{\chi} = \left(\frac{f_{EE}}{(\gamma^{\sigma-1}-1)f}\right)^{\frac{-1}{\sigma-1}} < 1$ represents the affordability of an upgrade from the use of electricity to the technologically efficient use of electricity: It decreases in the extra adoption fixed costs but increases in the relative gains in variable costs caused by the efficient use of electricity. The less affordable the efficient technology, the larger the increase in productivity necessary to realize positive profits after adopting it.

Efficiency Gain

Energy efficient firms produce a larger level of output $q_{EE}(\varphi) = RP^{\sigma-1}\left(\frac{\gamma\varphi}{p_E}\frac{\sigma-1}{\sigma}\right)^{\sigma} > q(\varphi)$ after adopting the technologically efficient use of electricity. The efficiency gain $E(\varphi) = \left(\frac{q_{EE}(\varphi)}{\varphi} - \frac{q_{EE}(\varphi)}{\gamma\varphi}\right)p_E$ made by a firm of productivity φ is the difference between the true energy efficient variable costs and what they would have been if $q_{EE}(\varphi)$ had been produced using the standard technology:

$$E_{EE}(\varphi) = \frac{\sigma-1}{\sigma}\frac{r_{EE}(\varphi)}{p_E}(\gamma-1) \tag{4.19}$$

Note that total variable costs and input quantities used by energy efficient firms also increase after the adoption of the energy efficient technology:

$$\frac{q(\varphi)}{\varphi} < \frac{q_{EE}(\varphi)}{\gamma\varphi} < \frac{q_{EE}(\varphi)}{\varphi} \tag{4.20}$$

This is consistent with the fact that this paper understands energy efficiency as the ability to produce a larger amount of output using a given level of energy and causing a certain level of pollution.

4.5.2 Self-Generation of Electricity

While the energy efficient technology reduces the variable costs of firms by increasing the amount of output they can produce using a single unit of electricity, self-generation allows firms to use fuel as an alternative source of energy. Firms can purchase fuel at price p_{SG} and self-generate electricity from it by paying an additional fixed cost f_{SG}. The price of fuel is determined on the international market and therefore considered exogenous. Generators retain the possibility to purchase electricity from the public power grid and can decide of the share ν^* of production they produce with electricity purchased from the grid or with self-generated electricity.The new energy input price p_{SG} is a function of the share of total energy consumed that the firm decides to self-generate ν^*. It is set by the firm to minimize its energy unit cost:

$$\underset{1\geq\nu^*\geq0}{Min} \quad \nu\,(p_F) + (1-\nu)\,p_E \tag{4.21}$$

The unit price of energy for a generator is then given by:

$$p_{SG} = \begin{cases} \nu^*\,(p_F) + (1-\nu^*)\,p_E & \text{if} \quad 1 \geq \nu^* > 0 \\ \varnothing \quad \text{if} \quad \nu^* = 0 \end{cases} \tag{4.22}$$

where $s \in \{s_L, s_H\}$ is the intrinsic ability of the industry to self-generate industry[34] with $s_H \geq 1 \geq s_L > 0$.

Without power outages in the economy, firms consider adopting an alternative production technology only if it lowers production costs and the solution to the minimization problem in equation 4.21 is straightforward. There is no incentive to self-generate when electricity from the grid is cheaper than

[34] Chaney (2008)'s model of heterogenous firms can also be interpreted as a model of heterogenous industries. Following this interpretation, the parameter s reflects the fact that different industries exhibit different abilities to generate electricity. For example, the sugar industry uses leftovers from sugar canes as fuel for their captive power plants, which reduces the variable costs of producing electricity.

fuel because self-generation would increase both fixed and variable costs. Then, if $p_F \geq p_E$ (cheap electricity), firms do not self-generate: $\nu^* = 0$, which implies $p_{SG} = \emptyset$. On the contrary, when fuel is relatively cheaper than electricity, self-generated electricity becomes cheaper than electricity provided by the grid. In this case, generators have lower variable costs than standard firms and self-generation represents an alternative technology as discussed in Bustos (2011) and similar to the energy efficient technology introduced in section 4.5.1. Thus, if $p_F < p_E$ (cheap fuel), generators produce electricity at full capacity[35]: $\nu^* = 1$, which implies $p_{SG} = p_F$.

Based on the energy input price p_{SG} arising from equations 4.21 and 4.22, the cost function of a firm of productivity φ that self-generates electricity (a *generator* thereafter) becomes:

$$C_{SG}(\varphi) = f + f_{SG} + p_{SG}\left[\frac{q(\varphi)}{\varphi}\right] \tag{4.23}$$

Total profits of self-generating firms are given by $\pi_{SG}(\varphi) = p_\varphi q(p_\varphi) - p_{SG}\left(\frac{q(p_\varphi)}{\varphi}\right) - f - f_{SG}$, which implies that generators charge a price $p_G(\varphi) = \frac{\sigma}{\sigma-1}\frac{p_{SG}}{\varphi}$ to the representative customer. The revenue function of a generator of productivity φ is given by:

$$r_{SG}(\varphi) = \left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} r(\varphi) = \left(\frac{1}{\gamma}\frac{p_E}{p_{SG}}\right)^{\sigma-1} r_{EE}(\varphi) \tag{4.24}$$

Minimum Profitable Productivity cut-off

There exists a self-generation production cut-off $\varphi_{SG} = \frac{p_{SG}}{p_E}\left(\frac{f+f_{SG}}{f}\right)^{\frac{1}{\sigma-1}}\varphi^*$ such that all firms of productivity $\varphi \geq \varphi_{SG}$ will generate positive profits when self-generating ($\pi_{SG}(\varphi) \geq 0$).

If electricity is more expensive than fuel ($p_E > p_{SG}$) and $f_{SG} \geq f\left(\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} - 1\right)$,[36] there exists a productivity level φ'_{SG} above which self-generation is more profitable than the use of electricity ($\pi(\varphi) < \pi_{SG}(\varphi)$) for all $\varphi > \varphi'_{SG}$:

$$\pi(\varphi'_{SG}) = \pi_{SG}(\varphi'_{SG}) \tag{4.25}$$

$$\varphi'_{SG} = \chi_4\varphi^* \tag{4.26}$$

[35] A more realistic way to model generators' choice of their capacity level would be to introduce an infinity of generation technology $t \in (0, +\infty)$ whose respective fixed costs $f_t = f_t(\nu)$ are increasing functions of generation capacity ν. Self-generation of a small amount of electricity would induce lower additional fixed costs and allow a greater control of firms on their variable costs $p_F < p_{SG} < p_E$.

[36] This restriction on the model's parameters insures that, consistently with empirical evidence presented in section 4.10.3, fixed adoption costs of self-generation are sufficiently high for the most productive active firms only to afford it ($\varphi'_{SG} > \varphi^*$).

where $\frac{1}{\chi_4} = \left(\frac{f_{SG}}{\left(\left(\frac{p_E}{p_{SG}} \right)^{\sigma-1} - 1 \right) f} \right)^{\frac{-1}{\sigma-1}} < 1$ represents the affordability of an upgrade from the standard use of electricity to the use of fuel (self-generation): It decreases in the fixed adoption costs of self-generation but decreases in the relative gains of variable costs caused by this technology change. Similarly, if the technologically efficient use of electricity is more expensive than fuel ($\gamma p_{SG} > p_E$), there exists a productivity level φ'' above which self-generation is more profitable than the technologically efficient use of electricity ($\pi_{EE}(\varphi) < \pi_{SG}(\varphi)$) for all $\varphi > \varphi''$:

$$\pi_{SG}(\varphi'') = \pi_{EE}(\varphi'') \tag{4.27}$$

Self-Generation and Welfare

The generation of electricity from the consumption of fuel has a higher environmental cost than if firms were directly purchasing electricity from the public grid.[37] Negative externalities caused by electricity self-generation are taken into account in the following welfare function:

$$D = D(I_{Elec}, I_{Fuel}) \qquad \frac{\partial D(I_{Elec})}{\partial I_{Elec}}(I_{Elec}, I_{Fuel}) < \frac{\partial D(I_{Fuel})}{\partial I_{Fuel}}(I_{Elec}, I_{Fuel})$$

where I_{Elec} is the total consumption of electricity purchased from the grid in the economy by standard and energy efficient firms, I_{Fuel} is the total consumption of fuel in the economy by generators, and total energy use is $I = I_{Elec} + I_{Fuel}$.

Efficiency Gain versus Pollution Intensity If self-generation offers lower marginal costs than either the standard or the energy efficient technology, its use represents an efficiency gain to the extent that it allows to produce a higher level of output from a given quantity of energy. Nevertheless, pollution intensity of fuel used by self-generation is higher than pollution intensity of electricity.[38] Hence, at equal quantity of energy used, the level of negative externality in the economy rises in the presence of self-generation.

$$\triangle D(I) > 0 \tag{4.28}$$

[37] First, the power grid is also supplied by renewable sources of energy like hydro-electricity. The installed hydro-electric capacity in India represents 14.35% of total demand in 2016 (Source: Official website of the Ministry of Power of India.). Second, a public power provider may enjoy returns to scale.

[38] See footnote 37.

The social planner faces a trade-off between the potential gains in consumer surplus arising from a lower price index and the decrease in welfare caused by higher levels of pollution. Section 4.5.3 takes into account that, when profitable, the technologically efficient use of fuel represents an additional efficiency gain that mitigates the increase in pollution caused by electricity self-generation from fuel.

4.5.3 Multiple Technology Adoption: Self-Generation and Efficient Use of Energy

Firms also consider whether the efficient use of self-generated electricity is more profitable than the efficient use of electricity from the grid. The fixed cost of energy efficiency for a generator are $f_{ESG} > f_{EE}$. Indeed, while a standard firm purchasing electricity from the grid needs to improve its industrial processes to produce the same level of output with a lower quantity of energy, a generator additionally needs to improve the efficiency of its generation process and its links with the production process to benefit from the same decrease γ in variable costs. The cost function of a firm of productivity φ that is self-generating electricity *and* using the energy-efficient technology (an *energy efficient generator* thereafter) becomes:

$$C_{ESG}(\varphi) \quad = \quad f + f_{ESG} + f_{SG} + p_{SG}\left[\frac{q(\varphi)}{\gamma\varphi}\right] \tag{4.29}$$

Total profits of an energy efficient generator are given by $\pi_{ESG}(\varphi) = p_\varphi q(p_\varphi) - p_{SG}\left(\frac{q(p_\varphi)}{\gamma\varphi}\right) - f - f_{ESG} - f_{SG}$, which implies that these firms charge a price $p_{ESG}(\varphi) = \frac{\sigma}{\sigma-1}\frac{p_{SG}}{\gamma\varphi}$ to the representative customer. The revenue function of an energy efficient generator of productivity φ is given by:

$$r_{ESG}(\varphi) \quad = \quad \left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1} r(\varphi) \quad = \quad \left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} r_{EE}(\varphi) \quad = \quad \gamma^{\sigma-1} r_{SG}(\varphi) \tag{4.30}$$

Minimum Profitable Productivity cut-offs

First, there exists an energy efficient self-generation production cut-off $\varphi_{ESG} = \frac{\gamma p_{SG}}{p_E}\left(\frac{f+f_{ESG}+f_{SG}}{f}\right)^{\frac{1}{\sigma-1}}\varphi^*$ verifying $\pi_{ESG}(\varphi_{ESG}) = 0$ such that all firms of productivity $\varphi \geq \varphi_{ESG}$ will generate positive profits when adopting the energy efficient technology ($\pi_{ESG}(\varphi) \geq 0$).

Second, assuming that $f_{ESG} > (f + f_{SG})\left(\gamma^{\sigma-1} - 1\right)$,[39] there exists a productivity level $\varphi_{SG/ESG}$ for

[39] This restriction on the model's parameters insures that, consistently with empirical evidence presented in section 4.10.3, the combination of fixed adoption costs of the alternative technologies are sufficiently high for the most productive firms only to afford them $\left(\varphi_{SG/ESG} > \varphi_{SG}\right)$.

which $\pi_{SG}(\varphi) < \pi_{ESG}(\varphi)$ for all $\varphi > \varphi_{SG/ESG}$:

$$\pi_{SG}\left(\varphi_{SG/ESG}\right) = \pi_{ESG}\left(\varphi_{SG/ESG}\right) \tag{4.31}$$

$$\varphi_{SG/ESG} = \varphi^{*}\chi_{5} \tag{4.32}$$

where $\frac{1}{\chi_5} = \left(\frac{f_{ESG}}{f\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1}\left(\gamma^{\sigma-1}-1\right)}\right)^{\frac{-1}{\sigma-1}} < 1$ represents the affordability of an upgrade from the use of fuel to the technologically efficient use of fuel: It decreases in the fixed costs of this technology change but increases in the gain in variables costs it causes.

Third, if the efficient use of electricity is more expensive than fuel ($\gamma p_{SG} > p_E$), there exists a productivity level φ'_{ESG} for which the efficient use of fuel is more profitable than the use of electricity ($\pi(\varphi) < \pi_{ESG}(\varphi)$) for all $\varphi > \varphi'_{ESG}$:

$$\pi\left(\varphi'_{ESG}\right) = \pi_{ESG}\left(\varphi'_{ESG}\right) \tag{4.33}$$

$$\varphi'_{ESG} = \varphi^{*}\chi_{2b} = \left(\frac{\chi_{2b}}{\chi}\right)\varphi_{SG/ESG} \tag{4.34}$$

where $\frac{1}{\chi_{2b}} = \left[\frac{f_{ESG}+f_{SG}}{\left(\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1}-1\right)f}\right]^{\frac{-1}{\sigma-1}} < 1$ represents the affordability of an upgrade from the use of electricity to the efficient use of fuel: It decreases in the extra adoption fixed costs but increases in the gains in variable costs caused by the efficient use of fuel.

Finally, if the use of electricity is more expensive than fuel ($p_{SG} < p_E$), there exists a productivity level $\varphi_{EE/ESG}$ for which the efficient use of fuel is more profitable than the efficient use of electricity $\pi_{EE}(\varphi) < \pi_{ESG}(\varphi)$ for all $\varphi > \varphi_{EE/ESG}$:

$$\pi_{EE}\left(\varphi_{EE/ESG}\right) = \pi_{ESG}\left(\varphi_{EE/ESG}\right) \tag{4.35}$$

$$\varphi_{EE/ESG} = \chi_{2a}\varphi^{*} = \left(\frac{\chi_{2a}}{\chi}\right)\varphi'_{EE} \tag{4.36}$$

where $\frac{1}{\chi_{2a}} = \left(\frac{\left(f_{SG}+f_{ESG}-f_{EE}\right)}{\left(\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1}-\gamma^{\sigma-1}\right)f}\right)^{\frac{-1}{\sigma-1}} < 1$ represents the affordability of an upgrade from the efficient use of electricity to the efficient use of fuel: It decreases in the extra adoption fixed costs but increases in the gains in variable costs caused by the efficient use of fuel.

Efficiency Gains

Energy efficient generators produce a larger level of output $q_{ESG}(\varphi) = \gamma^{\sigma} q_{SG}(\varphi) = \left(\frac{p_E}{p_{SG}}\right)^{\sigma} q_{EE}(\varphi)$ after adopting the efficient use of fuel but the relative prices of electricity and fuel determine whether it produces more or less than if it were using electricity in a technologically efficient way. Hence, the efficiency gain with respect to self-generation $ESG_{SG}(\varphi) = \left(\frac{q_{ESG}(\varphi)}{\varphi} - \frac{q_{ESG}(\varphi)}{\gamma\varphi}\right) p_{SG}$ made by a firm of productivity φ is the difference between the true technologically efficient variable costs using fuel and what they would be if $q_{ESG}(\varphi)$ was produced using fuel:

$$ESG_{SG}(\varphi) = \frac{\sigma - 1}{\sigma} \frac{r_{ESG}(\varphi)}{p_{SG}} (\gamma - 1) \tag{4.37}$$

Similarly, the efficiency gain $ESG_{EE}(\varphi) = \frac{q_{ESG}(\varphi)}{\gamma\varphi} p_E - \frac{q_{ESG}(\varphi)}{\gamma\varphi} p_{SG}$ with respect to the efficient use of electricity is given by the difference between the true variable costs of a given firm and what these variable costs would be with the efficient use of electricity:

$$ESG_{EE}(\varphi) = \frac{\sigma - 1}{\sigma} \left(\frac{p_E - p_{SG}}{p_E}\right) \left(\frac{p_E}{p_{SG}}\right)^{\sigma} r_{EE}(\varphi) \tag{4.38}$$

4.5.4 Electricity Outages

Power outages can occur with probability $0 < 1 - \delta < 1$ in the economy. Firms that do not self-generate cannot produce[40] during those outages: they earn only a fraction δ of their potential revenues $r(\varphi)$.[41] In the presence of power outages, the profits of a standard firm of productivity φ are given by $\mathring{\pi}(\varphi) = \delta p_{\varphi} q(p_{\varphi}) - \frac{p_E q(p_{\varphi})}{\varphi} - f$, while profits of an energy efficient firm are $\mathring{\pi}_{EE}(\varphi) = \delta p_{\varphi} q(p_{\varphi}) - \frac{p_E q(p_{\varphi})}{\gamma\varphi} - f$. Standard and energy efficient firms respectively charge prices $\mathring{p}(\varphi) = \frac{p(\varphi)}{\delta}$ and $\mathring{p}_{EE}(\varphi) = \frac{p_{EE}(\varphi)}{\delta} = \frac{p(\varphi)}{\gamma\delta}$ to the representative customer. Firms that self-generate electricity from fuel keep producing, and their profit, revenue and price functions are identical to those explicated in section 4.5.2. Similar to the case without power outages, if $f_{EE} > f\left(\gamma^{\sigma-1} - 1\right)$, there exists an energy efficient productivity cut-off $\mathring{\varphi}'_{EE} > \mathring{\varphi}^*$ such that $\mathring{\pi}'_{EE}(\varphi) \geq \mathring{\pi}(\varphi)$ for all firms of productivity

[40] Another incentive to self-generate electricity, which we do not discuss in this chapter, because it would be modeled in the same way and would not affect results, is the low quality of power provided by the grid. Differences in voltages can damage machines and interrupt production in the same way as a power outage. Empirically, the reliability of the power grid should then optimally be captured not only by the frequency of power outages but also by the variations in the characteristics of supplied power.

[41] Outages are modeled as iceberg costs: $(1 + \delta)$ units must be produced so that 1 unit can be sold. For example, the goods of an ice cream producer melt during a power outage and a larger quantity must be produced to generate a given level of revenues. When considering non-perishable goods, the model would capture the fact that firms must produce for a larger number of periods to achieve a given level of output.

$\varphi \geq \mathring{\varphi}'_{EE}$. These productivity cut-offs verify:

$$\mathring{\varphi}^* = \delta^{\frac{1}{1-\sigma}} \varphi^* > \varphi^* \qquad \mathring{\varphi}'_{EE} = \delta^{\frac{1}{1-\sigma}} \varphi'_{EE} > \varphi'_{EE} \tag{4.39}$$

Revenue functions of firms in the economy are given by:

$$\mathring{r}(\varphi) = \delta r(\varphi) \qquad \mathring{r}_{EE}(\varphi) = \delta r_{EE}(\varphi)$$

Note that we have:

$$\frac{\varphi_{\mathring{EE}}^*}{\mathring{\varphi}^*} = \frac{\varphi_{EE}^*}{\varphi^*} \tag{4.40}$$

Variable Costs of Self-Generation with Power Outages

As in section 4.5.2, energy input price p_{SG} is a function of the share $1 \geq \nu^* \geq 0$ of total energy consumed that the firm decides to self-generate:

$$p_{SG} = \begin{cases} \nu^* (p_F) + (1-\nu^*) p_E & \text{if} \quad 1 \geq \nu^* > 0 \\ \varnothing \quad \text{if} \quad \nu^* = 0 \end{cases} \tag{4.41}$$

As in section 4.5.2, the energy mix $1 \geq \nu^* \geq 0$ is the solution to the following minimization problem:

$$\underset{1 \geq \nu^* \geq 0}{Min} \quad \nu(p_F) + (1-\nu) p_E \tag{4.42}$$

When fuel is cheaper than electricity from the grid, generators produce at full capacity as in section 4.5.2: $\nu^* = 1$ and $p_{SG} = p_F$. Crucially, when fuel is more expensive than electricity directly purchased from the grid (cheap electricity), self-generation induces higher variable costs but allows to retain the entire generated revenues. In this case, firms choose their self-generation capacity such that $\nu^* = 1 - \delta$ and $p_{SG} = (1-\delta) p_F + \delta p_E \in [p_E, p_F]$ if $p_F \geq p_E$.

4.6 Cheap Electricity Equilibrium (Case 1) Without Power Outages

In this section, we focus on the **empirically relevant case of electricity from the public grid being cheaper than fuel** (see section 4.3). **Case 1 (cheap electricity case)** describes an economy in which the use of fuel is more expensive than the use of electricity from the grid, that is to say, when fuel and electricity prices verify:

$$p_{SG} \geq p_E \tag{4.43}$$

Other possible cases on the relative prices of fuel and electricity from the public grid are described in table 4.4. The equilibria corresponding to these cases will be solved in section 4.7, in which we refer to an economy in which the use of fuel is cheaper than the use of electricity from the grid.[42] Again, the different possible equilibria when fuel is cheaper than electricity from the grid are not empirically relevant but we will use these results to simplify the analysis of the equilibrium in Case 1 in the presence of power outages as shown in table 4.5. Indeed, as explained in section 4.3, the **empirically relevant case for theoretical predictions** on the impact of self-generation of electricity on the adoption of energy efficient technologies is **when electricity from the grid is cheaper than fuel (Case 1) and there are power outages**.

We derive the equilibrium for case 1 when there are no power outages in the economy in section 4.6. In section 4.8, we solve the equilibrium in case 1 in the presence of power outages and build the theoretical predictions that we will test empirically in section 4.9.

Definitions, Notations and Introductory Results

This section introduces some preliminary results necessary to derive the equilibrium and recalls the concepts of *direct* and *indirect crowding out* of investments in energy efficient technologies by self-generation. A summary of the main notations used in this chapter is provided in table 4.3.

The definition of the equilibrium in our model of heterogenous firms with technology adoption starts with the vector of productivity cut-offs $(\varphi^*, \varphi^*_{EE}, \varphi^*_{SG}, \varphi^*_{ESG})$ respectively defining the minimum productivity levels of standard firms, energy efficient firms, generators, and energy efficient generators. All other variables (aggregate profits, revenues, consumption levels of both the homogenous and the composite good) will be derived from these key variables and from the exogenous parameters in the model. Energy prices (p_E and p_{SG}) and technological parameters (the adoption fixed costs f_{EE}, f_{SG}, and f_{ESG}, as well as the level of technological progress γ) are exogenous: it is therefore important to differentiate between the different cases in terms of "true" range of available production technologies

[42] Recall that the *efficient* use of electricity is a qualification of the production technology used by the firm: we refer to the use of electricity as *efficient* when a smaller amount of electricity is required to produce one unit of output, not when it increases profitability.

that they imply. Indeed, exogenous prices and technological parameters can immediately imply that some production modes are dominated for all firms. For example, the use of fuel is always dominated in an economy in which fuel is more expensive than electricity and there are no power outages.

As previously explained, three cases can be defined based on the relative prices of electricity from the grid and fuel. Additionally, we derive in Appendix E the implications of the combination of exogenous energy prices with exogenous technological parameters for the choice set of production technologies available to firms of different productivity levels. Proposition 25 describes the technological conditions for the efficient use of electricity grid to dominate the standard production technology for at least some firms when the use of self-generation is dominated by all other production modes at all productivity levels. Proposition 26 describes the technological conditions for the use of fuel (self-generation) to dominate the standard production technology for at least some firms when the efficient use of electricity grid is dominated by all other production modes at all productivity levels. Finally, proposition 27 presents the conditions on exogenous parameters for all production technologies to be present in equilibrium, that is to say for every production mode to offer the highest level of profitability to at least some firms in the economy: neither the use of fuel nor the use of electricity is dominated. Table 4.4 explicits the role played by these propositions in defining the different possible equilibria.

We now recall the definition of the two key concepts of this chapter. We refer to the **direct crowding-out** of investments in energy efficiency by investments in self-generation, when profits from adopting the energy efficient technology are positive and superior to those obtained from using the standard technology but firms do not invest in energy intensive technology because investments in self-generation are more profitable. We refer to the **indirect crowding-out** of investments in energy efficiency by investments in self-generation, when firms at the top of the productivity distribution are adopting both the energy efficient and the self-generation technologies: Market shares are reallocated towards those firms that benefit from a larger decrease of their variable costs and consequently of their prices. First, the minimum productivity level that is required for investments in the energy efficient technology to be profitable increases: the share of energy efficient firms decreases. Second, the relative share of output produced by energy efficient firms in the economy also decreases.

Henceforth, we use P^*_{Casei} (respectively φ^*_{Casei}) to refer to the price index (resp. the production cut-off) in the economy characterized by the energy prices corresponding to the definition of case

Table 4.3: Summary of Notations: **SG**=Self-Generation, **ESG**=Energy Efficient Self-Generation, **EE**=Energy Efficient Electricity Use, and **S**=Standard production mode.

Notation	Definition
φ^*	$\pi(\varphi^*) \geq 0$ for all $\varphi \geq \varphi^*$
φ_{EE}	$\pi_{EE}(\varphi_{EE}) \geq 0$ for all $\varphi \geq \varphi_{EE}$
φ_{SG}	$\pi_{SG}(\varphi_{SG}) \geq 0$ for all $\varphi \geq \varphi_{SG}$
φ_{ESG}	$\pi_{ESG}(\varphi_{ESG}) \geq 0$ for all $\varphi \geq \varphi_{ESG}$
φ'_{EE}	$\pi_{EE}(\varphi'_{EE}) \geq \pi(\varphi'_{EE})$ for all $\varphi \geq \varphi'_{EE}$
φ'_{SG}	$\pi_{SG}(\varphi'_{SG}) \geq \pi(\varphi'_{SG})$ for all $\varphi \geq \varphi'_{SG}$
φ'_{ESG}	$\pi_{ESG}(\varphi'_{ESG}) \geq \pi(\varphi'_{ESG})$ for all $\varphi \geq \varphi'_{ESG}$
φ''	$\pi_{EE}(\varphi'') \geq \pi_{SG}(\varphi'')$ for all $\varphi \geq \varphi''$
$\varphi_{EE/ESG}$	$\pi_{ESG}(\varphi_{EE/ESG}) \geq \pi_{EE}(\varphi_{EE/ESG})$ for all $\varphi \geq \varphi_{EE/ESG}$
$\varphi_{SG/ESG}$	$\pi_{ESG}(\varphi_{SG/ESG}) \geq \pi_{SG}(\varphi_{SG/ESG})$ for all $\varphi \geq \varphi_{SG/ESG}$
	Minimum productivity level of ...
φ^*	Standard firms
φ^*_{EE}	Energy Efficient firms
φ^*_{SG}	Generators
φ^*_{ESG}	Energy Efficient Gener

"i". We compare the level of welfare under the technology adoption choice of firms induced by this set of electricity and fuel prices with the level of welfare in two other hypothetical states of the economy: i) when only the standard production mode is available, and ii) when the energy efficient technology is available but not self-generation. In the first case, P^* (respectively φ^*) would be the price index (resp. the production cut-off) in an economy in which firms do not have the possibility to adopt energy efficient technologies as described in the baseline model in section 4.4.4. In the second case, P^{**} (respectively φ^{**}) would be the price index (resp. the production cut-off) in an economy in which firms have the possibility to invest neither in energy efficiency nor in self-generation.

Cheap Electricity Equilibrium Without Power Outages

Equilibrium Prices and Productivity Cut-Offs

If the fuel necessary to self-generate electricity is more expensive than electricity from the grid, self-generation induces both higher fixed and variable costs and is always a dominated technology choice. G1 in figure 4.2 represents the profits $\pi_{SG}\left(\varphi^{\sigma-1}\right)$ of self-generating firms as a function of productivity $\varphi^{\sigma-1}$ when energy prices and technological progresses are such that the variable costs of a generator

Table 4.4: Summary of the possible equilibria **without power outages** based on relative energy prices and technological parameters. (**SG**=Self-Generation, **ESG**=Energy Efficient Self-Generation, **EE**=Energy Efficient Electricity Use, and **S**=Standard production mode. **(*)** indicates the proposition describing the conditions on exogenous technological parameters corresponding to each case or sub-case. **(**) Dominated Technology** refers to the case when a particular production mode is never profitable because it implies both higher variable costs and higher fixed costs. See table 4.3 for a summary of the notations used in this table.)

	Case 1	Case 3		Case 2	
Energy Prices	$p_F \geq p_E$	$p_F < p_E$			
Unit Costs	$p_F > p_E > \frac{p_E}{\gamma}$	$\frac{p_E}{\gamma} > p_F$		$p_F > \frac{p_E}{\gamma}$	
Dominated Technology**	SG ESG	EE	None	SG	
		Proposition 24*		Proposition 25*	
		Case 3a	Case 3b	Case 2a	Case 2b
		$\varphi_{SG} < \varphi_{EE}$	$\varphi_{SG} > \varphi_{EE}$	$\varphi'_{ESG} > \varphi'_{EE}$	$\varphi'_{ESG} \leq \varphi'_{EE}$
Technology in Equilibrium	S EE	See below.		S, EE ESG	S ESG
Crowding-Out	No	See below.		Indirect	Indirect

	Case 3a		Case 3b	
	$\varphi_{SG} \leq \varphi_{EE}$		$\varphi_{SG} > \varphi_{EE}$	
	Proposition 26*		Proposition 27*	
	Case 3a1	Case 3a2	Case 3b1	Case 3b2
	$\varphi_{SG/ESG} \leq \varphi'_{SG}$	$\varphi_{SG/ESG} > \varphi'_{SG}$	$\varphi_{SG/ESG} > \varphi''$	$\varphi_{SG/ESG} \leq \varphi''$
Technology in Equilibrium	S ESG	S, SG ESG	S, SG ESG	See below.
Crowding-Out	Indirect	Direct Indirect	Direct Indirect	See below.

	Case 3b2	
	$\varphi_{SG/ESG} \leq \varphi''$	
	Proposition 25*	
	Case 3b2'	Case 3b2"
	$\varphi'_{ESG} > \varphi'_{EE}$	$\varphi'_{ESG} \leq \varphi'_{EE}$
Technology in Equilibrium	S, EE ESG	S ESG
Crowding-Out	Indirect	Indirect

Table 4.5: Summary of the possible equilibria **with power outages when electricity from the grid is cheaper than fuel (Case 1)** based on relative energy prices and technological parameters. (*) **Similar** equilibria refer to equilibria discussed in sections 4.7.1 and 4.7.2 when there are no power outages but where energy input prices lead to similar allocation of firms across different production modes. Recall that $c_\epsilon(\delta,\gamma)\,p_F > g\epsilon(\delta)$ for all ϵ, δ and γ. See table 4.3 for a summary of the notations used in this table.

Energy Prices	$p_F > p_E$		
Outages Intensity / Technological Progress	$g\epsilon(\delta) > p_E$	$p_E \in [c_\epsilon(\delta,\gamma)\,p_F, g\epsilon(\delta)]$	$p_E > c_\epsilon(\delta,\gamma)\,p_F$
Crowding-Out	No	Indirect	Direct and Indirect
Similar* Equilibria	Case 1 Without Outages	Case 2 Without Outages	Case 3 Without Outages

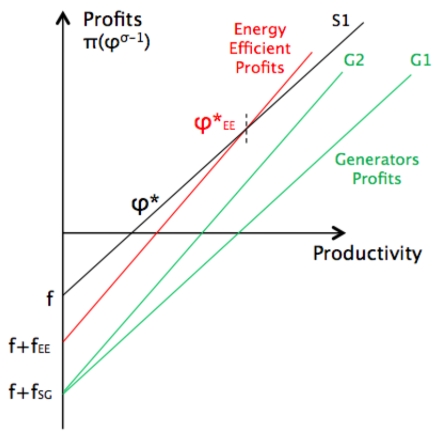


Figure 4.2: Self-Generation as a Strictly Dominated Technology Choice (Case 1)

of productivity φ equal its standard variable costs ($p_F = p_E$). In this state of the economy, G1 is parallel to S1 that represents standard profits $\pi\left(\varphi^{\sigma-1}\right)$. Hence, there is no productivity level φ verifying the equation $\pi(\varphi) \leq \pi_{SG}(\varphi)$. When the price of fuel increases (resp. decreases) with respect to electricity, the profit function of self-generation $\pi_{SG}\left(\varphi^{\sigma-1}\right)$ rotates towards the right (resp. left). **Case 1** refers to the state of the economy when *electricity is cheaper than fuel*, that is to say, when fuel and electricity prices verify:

$$p_{SG} \quad \geq \quad p_E \tag{4.44}$$

G2 in figure 4.2 represents the profits of self-generating firms when $\gamma p_F = p_E$, that is to say, when the variable costs of a generating firm are equal to the variable costs of an energy efficient firm. Case 1 implies that there is no productivity level φ verifying the equation $\pi_{EE}(\varphi) \leq \pi_{SG}(\varphi)$. Firms self-select into the use of electricity and the efficient use of electricity following:

$$\begin{aligned} \varphi^*_{SG} &= \varphi'_{SG} = \varphi'' = \emptyset \\ \varphi^*_{ESG} &= \emptyset \\ \varphi^*_{EE} &= \varphi'_{EE} \end{aligned} \tag{4.45}$$

Proposition 9. *In case 1 when fuel is relatively more expensive than electricity and in the absence of power outages, there are no generators or efficient generators in the economy. There is neither direct nor indirect crowding-out of investments in energy efficient technology by investments in self-generation.*

Proposition 9 is intuitively straightforward: this chapter shows theoretically and empirically that it does not hold in the presence of power outages in sections 4.8 and 4.11 respectively.

The equilibrium price index is given by $P^*_{Case1} = \left[\int\limits_{\omega\in\Omega_S} p(\omega)^{1-\sigma}d\omega + \int\limits_{\omega\in\Omega_{EE}} p(\omega)^{1-\sigma}d\omega\right]^{\frac{1}{1-\sigma}}$ where Ω_S is the subset of varieties produced by standard firms and Ω_{EE} is the subset of varieties produced by energy efficient firms. There is $\Omega = \Omega_S \cup \Omega_{EE}$ and:

$$P^*_{Case1} \quad = \quad P^*\left(1 + \frac{f_{EE}}{f}\left(\frac{1}{\chi}\right)^k\right)^{\frac{-1}{\sigma-1}} < P^* \tag{4.46}$$

where $\frac{1}{\chi} < 1$ represents the affordability of an upgrade from the use of electricity to the efficient use of electricity as described in section 4.5.1. The more affordable the energy efficient technology, the larger is its effect on the price index because a larger share of firms can adopt it and decrease their price. Since the minimum productivity cut-off in the economy is a decreasing function of the price index, we have:

$$\varphi^*_{Case1} = \varphi^* \left(1 + \frac{f_{EE}}{f}\left(\frac{1}{\chi}\right)^k\right)^{\frac{1}{\sigma-1}} \quad > \quad \varphi^* \tag{4.47}$$

It follows from equation 4.12 that average productivity of active firms $\overline{\varphi}$ is also larger than in an economy in which energy conservation investments do not exist.[43] The most productive firms can offer a lower price to the representative customer after adopting the energy efficient technologies, which decreases the price index and increases the competition in the economy. Market shares are reallocated towards the energy efficient firms that generate larger profits and benefit relatively more than their standard competitors through an increase in demand for their goods. The least productive firms exit the market because they are not able to generate enough revenues to cover fixed standard production costs. Average productivity of energy efficient firms $\overline{\varphi_{EE}}$ is given by:

$$\overline{\varphi EE}_{Case1} \quad = \quad \left[\frac{1}{J_{EE}}\int_{\varphi^*_{EE}}^{\infty} \varphi^{\sigma-1} v\left(\varphi\right) d\varphi\right]^{\frac{1}{\sigma-1}} = \chi\overline{\varphi_{Case1}} > \overline{\varphi} \tag{4.48}$$

where $J_{EE} = (1 - V(\varphi^*_{EE}))$ is the mass of energy efficient firms. The less affordable the energy efficient technology, the larger χ and the larger the difference in productivity between the pool of standard firms and energy efficient firms. Aggregate revenues and aggregate profits can be split up into revenues and profits of energy efficient and standard firms: $R = R_{EE} + R_S$ and $\Pi = \frac{\sigma-1}{\sigma}\frac{R}{k} = \Pi_{EE} + \Pi_S$. Combined with equation 4.48, aggregate revenues of energy efficient firms can be written as $R_{EE} = J_{EE} r_{EE}\left(\overline{\varphi_{EE}}\right)$, which implies that aggregate profits of energy efficient firms are $\prod_{EE} = J_{EE}\pi_{EE}\left(\varphi^*_{EE}\right) = \frac{R_{EE}}{\sigma} - J_{EE}\left(f + f_{EE}\right)$. Furthermore, the share of energy efficient firms depends only on the ratio of the production and the EE cut-off:

$$\frac{J_{EE}}{J_A} \quad = \quad \left(\frac{\varphi^*_{Case1}}{\varphi'_{EE}}\right)^k = \left(\frac{1}{\chi}\right)^k \tag{4.49}$$

The share of energy efficient firms increases when the energy efficient technology is more affordable and χ decreases. From equation 4.19, aggregate efficiency gains for firms having adopted the efficient

[43] In case 1, the equilibrium in an economy where self-generation does not exist is similar to the equilibrium in an economy where self-generation does exist but is never a profitable choice. In particular, we would have $\varphi^*_{Case1} = \varphi^{**}$ and $P^*_{Case1} = P^{**}$.

use of electricity are:

$$E_{Case1} = \frac{\sigma - 1}{\sigma}\frac{R_{EE}}{p_E}(\gamma - 1) > 0 \quad (4.50)$$

Welfare with Energy Efficient Technology

The most productive firms self-select into the technologically efficient use of electricity and capture a larger share of the constant expenditure on the differentiated sector by charging a lower price. It decreases the price index $P^*_{Case1} < P^*$ and increases the consumer surplus:

$$\begin{aligned} \triangle CS &= Rln\left(\frac{P^*}{P^*_{Case1}}\right) > 0 \quad (4.51) \\ &= \frac{1}{\sigma - 1} Rln\left(1 + \frac{f_{EE}}{f}\left(\frac{1}{\chi}\right)^k\right) \end{aligned}$$

Energy efficient firms produce a larger quantity of output to satisfy the larger demand arising from lower prices. They are also more technologically efficient as shown by equation 4.50 that represents the increase in additional input that would have been necessary to get the positive change in welfare described by equation 4.51. The variation $\triangle CS$ is the *shadow price of the pollution intensity* $D(E_{Case1})$: the damage that it would have been necessary to inflict to the environment in order to reach this higher level of consumer surplus without energy efficient technology. As explained in section 4.4.5, there will be no effect on aggregate profits and energy consumption, only a reallocation of profits and production towards energy efficient firms. From equations 4.46 and 4.51, we see that the positive effect of the energy efficient technology on the consumer surplus is increasing in the affordability of this technology. Indeed, its effect on the price index will increase with the share of firms that can afford it.

Proposition 10. *In case 1 when fuel is relatively more expensive than electricity and in the absence of power outages, welfare is higher in an economy in which energy efficient technologies are available. The more affordable and performant this technology, the larger is its positive impact on welfare.*

4.7 Cheap Fuel and Intermediate Prices Equilibria (Case 2 and Case 3) Without Power Outages

Case 2 and Case 3 refer to an economy in which the use of fuel is cheaper than the use of electricity from the grid, but fuel is still more expensive than the efficient use of electricity from the grid in case 2[44]. Hence, in **case 2 (intermediate prices case)**, fuel and electricity prices verify:

$$\gamma p_F \quad \geq \quad p_E \quad > \quad p_F \tag{4.52}$$

and in **case 3 (cheap fuel case)**, energy prices verify:

$$p_E \quad \geq \quad \gamma p_F \tag{4.53}$$

This section presents the equilibrium in these two different cases and summarizes the aspects that will be used in section 4.8 as a reference point for the effect of power outages on prices and the allocation of firms across the different production modes. Full solutions and intuition are provided in appendix E.

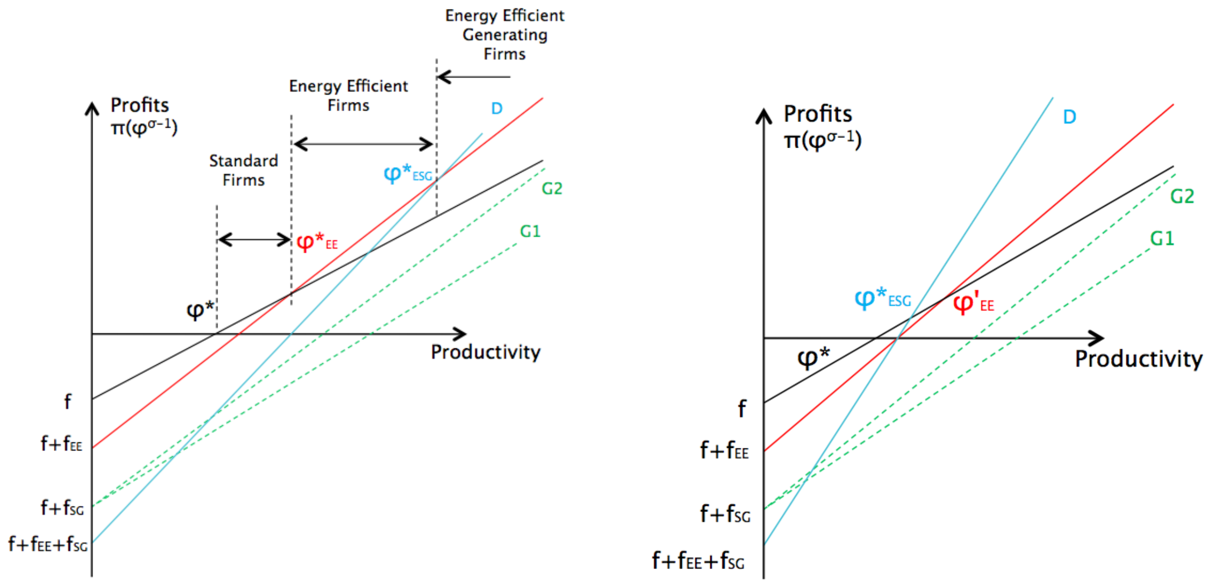
4.7.1 Technology Adoption without Power Outages: Intermediate Prices Equilibria (Case 2)

Even when the use of fuel generates lower variable costs than the use of electricity, it may be less profitable than the efficient use of electricity. **Case 2** refers to the state of the economy when *fuel is cheaper than electricity but more expensive than the efficient use of electricity.* In Case 2, the efficient use of fuel is less expensive than the efficient use of electricity. Generators' profits in case 2 are represented by lines located between G1 and G2 in figure 4.2 or figure 4.3a. Thus, even if generating electricity is more profitable than purchasing electricity for at least some levels of productivity ($sp_{SG} < p_E$), it will never be more profitable than adopting the EE technology when energy prices and the technological parameter γ verify:

$$\gamma p_F \quad \geq \quad p_E \quad > \quad p_F \tag{4.54}$$

Figure 4.7.1 presents the two possible cases of production technology adopted by firms in case 2 when firms never adopt the use of fuel. D represents the profits $\pi_{ESG}\left(\varphi^{\sigma-1}\right)$ of an energy efficient generator as a function of its productivity φ. Firms self-select into the use of electricity, or the efficient use of electricity or the efficient use of fuel according to the following:

[44]Recall that the *efficient* use of electricity is a qualification of the production technology used by the firm: we refer to the use of electricity as *efficient* when a smaller amount of electricity is required to produce one unit of output, not when it increases profitability.



(a) Equilibrium with standard and energy efficient firms, and energy efficient generators (Case 2a)

(b) Equilibrium with standard firms and energy efficient generators (Case 2b)

Figure 4.3: Self-Generation as a Dominated Technology Choice (Case 2)

$$\begin{aligned} \varphi^*_{SG} &= \emptyset \\ \varphi^*_{EE} &= \varphi'_{EE} \end{aligned} \tag{4.55}$$

Furthermore, profits derived from the efficient use of fuel will be represented by profit curves located above G2. Hence, there exists a minimum productivity level for which the efficient use of fuel will be the most profitable production mode in the economy:

$$\varphi^*_{ESG} \neq \emptyset \tag{4.56}$$

Equilibria in Case 2a and Case 2b and the proof of the following propositions 11, 12, and 13 are fully derived in appendices E.

Proposition 11. *In case 2, in the absence of power outages, firms self-select into the use of electricity, the efficient use of electricity, or the efficient use of fuel. The most productive firms in the economy find it profitable to adopt the energy efficient technology and to self-generate electricity. There is no direct crowding-out of investments in energy efficient technologies by investments in self-generation.*

Proposition 12. *In case 2, in the absence of power outages, if the technological condition in equation E.3 is verified, then there are standard firms, energy efficient firms and energy efficient generators* *in the economy (Case 2a):*

$$\varphi^*_{ESG} = \varphi_{EE/ESG} \tag{4.57}$$

If the technological condition in equation E.3 is not verified, then there are only standard firms and energy efficient generators in the economy (Case 2b):

$$\varphi^*_{ESG} = \varphi'_{ESG} \tag{4.58}$$

Welfare with Energy Efficient Generators: Indirect crowding-out

The most productive firms self-select into the energy efficient use of fuel and capture a larger share of the constant expenditure on the differentiated sector by charging a lower price. First, it decreases the price index and increases the consumer surplus both with respect to a situation $(**)$ when the energy efficient technology only is available or a situation $(*)$ when no technology upgrade is possible:

$$\triangle CS^* = Rln\left(\frac{P^*}{P^*_{Case2}}\right) > 0 \tag{4.59}$$

$$\triangle CS^{**} = Rln\left(\frac{P^{**}}{P^*_{Case2}}\right) > 0 \tag{4.60}$$

Second, it increases the share of fuel in total energy consumption I. In this model, where total energy consumption does not depend on the price index but on energy prices and technological constraints as described by equation 4.15, it implies that the pollution intensity of output increases.

Proposition 13. *In case 2, in the absence of power outages, there is an indirect crowding-out of investments in energy efficient technology by investments in self-generation, when the most productive firms adopt the efficient use of fuel. Market shares are reallocated towards energy efficient generators, which decreases revenues and profits in the rest of the industry. The use of dirty energy (fuel) crowds out the use use of cleaner energy (electricity). Furthermore, self-generation decreases (Case 2a) or fully erases (Case 2b) the efficiency gains made by energy efficient firms using the cleaner source of energy.*

4.7.2 Technology Adoption without Power Outages: Cheap Fuel Equilibria (Case 3)

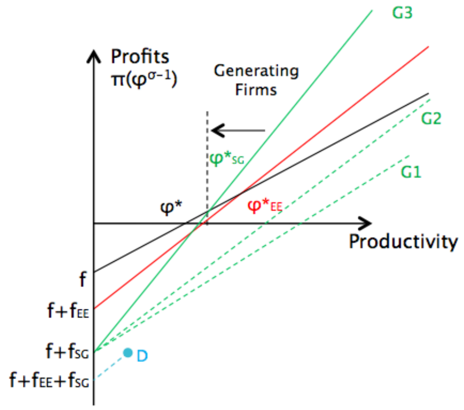
Self-generation reduces variable costs with respect to both the standard and the energy efficient production technology, when fuel is cheaper than the efficient use of electricity from the grid:

$$p_E \geq \gamma p_F \tag{4.61}$$

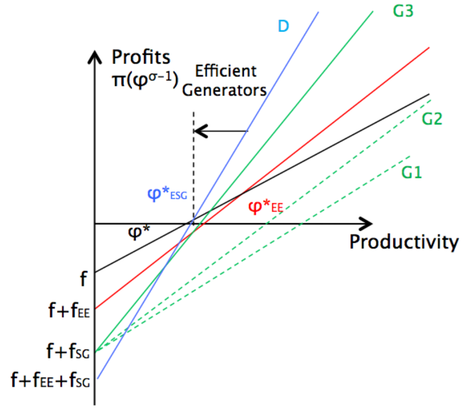
Figures 4.4 and 4.5 describe the possible cases based on this set of energy prices and technological constraints. As in figure 4.2, G1 represents the profits of generators when self-generation variable costs are equal to standard variable costs. G2 represents the profits of generators when self-generation variable costs are lower than standard variable costs but equal to variable costs when using the EE technology. **Case 3** refers to the state of the economy when *fuel is cheaper than both electricity and the efficient use of electricity*. Hence, the efficient use of fuel is the production technology that offers the lowest variables costs in the economy. In this type of economy, self-generation profits are represented by lines located above G2: G3 is an example of such a state of the economy. There exists a *self-generation productivity cut-off* φ_{SG} such that self-generation is the production choice offering the highest profits to firms of productivity $\varphi \geq \varphi^*_{SG}$. In figure 4.4, the use of fuel is much cheaper than the use of electricity or the fixed adoption costs of self-generation are only marginally superior to those of the energy efficient technology. Hence, the minimum level of productivity required to make positive profits is lower for generators than for energy efficient firms ($\varphi_{SG} < \varphi_{EE}$). The use of fuel always dominates the efficient use of electricity. This case is referred to as Case 3a. On the contrary, in figure 4.5, the minimum profitable productivity cut-off is lower for energy efficient firms than for generators ($\varphi_{SG} > \varphi_{EE}$). The efficient use of electricity is thus the most profitable production mode for firms in the middle of the productivity distribution. This case is referred to as Case 3b.

Case 3a - Dominated Use of Electricity: Equilibria without Energy Efficient Firms

This state of the economy corresponds to the allocation of firms across the different production technologies described in figure 4.4. The self-generation technology requires a lower level of productivity than the energy efficient technology to be profitable: $\varphi_{SG} < \varphi_{EE}$. The efficient use of electricity is always dominated by the use of fuel. Firms self-select into production, the use of fuel, and the



(a) Equilibrium with standard firms, generators, and energy efficient generators



(b) Equilibrium with standard firms and energy efficient generators

Figure 4.4: Energy Efficiency as a Dominated Technology Choice (Case 3a)

efficient use of fuel according to the following productivity cut-offs:

$$\varphi^*_{SG} = \varphi'_{SG} \quad (4.62)$$

$$\varphi^*_{EE} = \emptyset \quad (4.63)$$

Furthermore, the efficient use of fuel is the production mode offering the lowest variable costs. Thus, there exists a minimum productivity level for which the efficient use of fuel will be the most profitable production mode in the economy:

$$\varphi^*_{ESG} \neq \emptyset \quad (4.64)$$

Proposition 14. *In case 3a, in the absence of power outages, if the technological condition in equation E.5 is verified, then there are standard firms, generators, and energy efficient generators in the economy:*

$$\varphi^*_{ESG} = \varphi_{SG/ESG} \quad (4.65)$$

If the technological condition in equation E.5 is not verified, then there are only standard firms and energy efficient generators in the economy:

$$\varphi^*_{ESG} = \varphi'_{ESG} \quad (4.66)$$

Proposition 15. *In Case 3a and in the absence of power outages, there is an indirect crowding-out of investments in energy efficiency by investments in the self-generation technology. Furthermore, if*

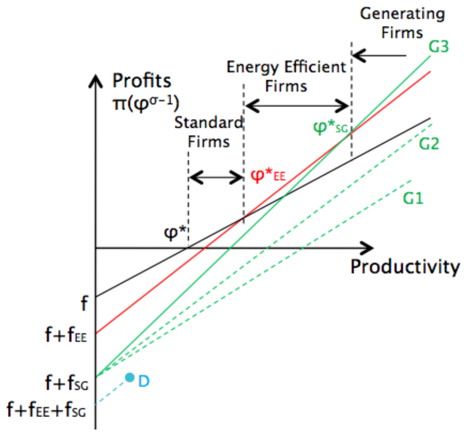


Figure 4.5: Equilibrium with all Production Technologies (Case 3b)

the inequality in equation E.5 is verified, there is also a direct crowding-out of investments in energy efficiency by investments in the self-generation technology.

Case 3b - No dominated technology: Equilibria with Energy Efficient Firms

In this state of the economy, the efficient use of fuel is the production technology offering the lowest variable costs followed by the use of fuel and the efficient use of electricity. However, the fixed adoption costs of the use of fuel or the efficient use of fuel are too costly for some firms in the middle of the productivity distribution. Hence, firms self-select into all four production modes according to the following productivity cut-offs:

$$\varphi^*_{SG} = \varphi'' \tag{4.67}$$

$$\varphi^*_{EE} = \varphi'_{EE} \tag{4.68}$$

Furthermore, the efficient use of fuel is the production mode offering the lowest variable costs. Thus, there exists a minimum productivity level for which the efficient use of fuel will be the most profitable production mode in the economy:

$$\varphi^*_{ESG} \neq \emptyset \tag{4.69}$$

Proposition 16. *In case 3b, in the absence of power outages, if the technological condition in equation E.7 is verified, then all four production modes are present in the economy:*

$$\varphi^*_{ESG} = \varphi_{SG/ESG} \tag{4.70}$$

If the technological condition in equation E.3 is not verified, but the technological condition in equation

E.7 is verified, then there are standard firms, energy efficient firms and energy efficient generators in the economy:

$$\varphi^*_{ESG} \quad = \quad \varphi_{EE/ESG} \tag{4.71}$$

If neither the technological condition in equation E.3 nor technological condition in equation E.7 is verified, then there are only standard firms and energy efficient generators in the economy:

$$\varphi^*_{ESG} \quad = \quad \varphi'_{ESG} \tag{4.72}$$

Proposition 17. *In Case 3b and in the absence of power outages, there is an indirect crowding-out of investments in energy efficiency by investments in the self-generation technology. Furthermore, if the inequality in equation E.7 is verified, there is also a direct crowding-out of investments in energy efficiency by investments in the self-generation technology. Pollution intensity and consumer surplus increase.*

4.8 Cheap Electricity Equilibrium (Case 1) With Power Outages

In this section, we show that the presence of generators in India in spite of electricity from the grid being cheaper than costs of electricity self-generation can be explained by power outages.

Proposition 18. *There exists an indirect crowding-out function* $g_\epsilon(\delta) = \frac{\delta^{\epsilon-1}(1-\delta)}{1-\delta^\epsilon} \in \left[0, \frac{1}{\epsilon}\right]$ *with* $\epsilon = \frac{\sigma}{\sigma-1}$ *that decreases with the intensity of outages in the economy* ($g'_\epsilon(\delta) > 0$).

Proposition 19. *When energy prices verify:*

$$g_\epsilon(\delta)\, p_F \quad < \quad p_E < p_F \tag{4.73}$$

there are energy efficient generators in the economy in spite of the price of fuel being higher than the price of electricity.

There exists a productivity cut-off $\mathring{\varphi}'_{SG}$ such that firms of productivity $\varphi \geq \mathring{\varphi}'_{SG}$ make larger profits when self-generating electricity than when purchasing it from the grid: $\mathring{\pi}_{SG}(\varphi) \geq \mathring{\pi}(\varphi)$ for all firms of productivity $\varphi \geq \mathring{\varphi}_{SG}$. Hence, there is:

$$\mathring{\varphi}^*{}_{SG} \quad \neq \quad \emptyset \tag{4.74}$$

There also exists a productivity cut-off $\mathring{\varphi}_{EE/ESG}$ such that firms of productivity $\varphi \geq \mathring{\varphi}'_{SG}$ find the energy efficient use of fuel more profitable than the efficient use of electricity: $\mathring{\pi}_{EE/ESG}(\varphi) \geq \pi_{\mathring{E}E}(\varphi)$ for all firms of productivity $\varphi \geq \mathring{\varphi}_{EE/ESG}$. Hence, there is:

$$\mathring{\varphi}^{*}{}_{ESG} \neq \emptyset \tag{4.75}$$

Proposition 20. *There exists a direct crowding-out function* $c_{\epsilon}(\delta,\gamma) = \frac{\gamma\left(\delta^{\epsilon-1}-\delta^{\epsilon}\right)}{\left(1-\gamma\delta^{\epsilon}\right)} > g_{\epsilon}(\delta)$ *that decreases with the intensity of outages in the economy* $\left(\frac{\partial c_{\epsilon}}{\partial \delta} > 0\right)$*, and increases with technological progress* $\left(\frac{\partial c_{\epsilon}}{\partial \gamma} > 0\right)$ *if and only if shortages in the economy are high or the level of technological progress too limited:*

$$\gamma < \frac{1}{\delta^{\epsilon}} \tag{4.76}$$

Proposition 21. *When energy prices are such that:*

$$c_{\epsilon}(\delta,\gamma)\, p_F < p_E < p_F \tag{4.77}$$

there are generators and energy efficient generators in the economy in spite of the price of fuel being higher than the price of electricity. However, if shortages in the economy are low or the energy efficient technology is very performant $\left(\gamma > \frac{1}{\delta^{\epsilon}}\right)$*, no firm in the economy will self-generate without adopting the energy efficient technology.*

Proposition 22. *In case 1 and in the presence of power outages, self-generation remains a dominated technology choice and the energy efficient productivity cut-off verifies* $\mathring{\varphi}^{*}_{EE} = \mathring{\varphi}'_{EE}$*, as described in section 4.6, when energy prices verify:*

$$p_E < g_{\epsilon}(\delta)\, p_F < c_{\epsilon}(\delta,\gamma)\, p_F < p_F \tag{4.78}$$

4.8.1 No Crowding Out: Equilibrium without Generators

By proposition 22, if $c_{\epsilon}(\delta,\gamma)\, p_F > g_{\epsilon}(\delta)\, p_F > p_E$, the use of fuel and the efficient use of fuel are dominated by the use of electricity. There are no generators and the economy behaves in the same way as described in case 1 in section 4.6. Nevertheless, from equation 4.39, we see that the minimum levels of productivity necessary to produce and to adopt the energy efficient technology are larger in the presence of outages. Firms must produce a larger level of output and charge higher prices in order to generate the same revenues. Hence, the price index in the economy is higher in the presence of power outages, while aggregate productivity of active and energy efficient firms decreases. Consequently, the increase in the consumer surplus caused by the energy efficient technology is also lower. However, the share of energy efficient firms in the economy is unaltered because power outages affect standard and energy efficient firms in the same way: they lose an equal share of revenues.

4.8.2 Indirect Crowding-Out: Equilibria with Efficient Generators

By propositions 18 and 22, if $c_\epsilon(\delta,\gamma)\,p_F > p_E > g_\epsilon(\delta)\,p_F$, the use of fuel dominates the use of electricity but not the efficient use of electricity, in spite of the loss of revenues caused by power outages. Nevertheless, the efficient use of fuel dominates the efficient use of electricity in the presence of power outages. We have:

$$\varphi_{EE/ESG}^{\circ} = \varphi_{EE}^{'}\left(\frac{\left(\gamma^{\sigma-1}-1\right)}{\left(\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1}-\delta\gamma^{\sigma-1}\right)}\frac{\left(f_{SG}+f_{ESG}-f_{EE}\right)}{f_{EE}}\right)^{\frac{1}{\sigma-1}} \tag{4.79}$$

Similarly to equation E.3, we have that $\varphi_{EE/ESG}^{\circ} > \varphi'_{EE}$ if the following equation is verified:

$$\left(\gamma^{\sigma-1}-1\right)\left(f_{SG}+f_{ESG}-f_{EE}\right) > \left(\left(\frac{\gamma p_E}{p_{SG}}\right)^{\sigma-1}-\delta\gamma^{\sigma-1}\right)f_{EE} \tag{4.80}$$

The interpretation of equation 4.80 is similar to the interpretation of equation E.3 but reflects the fact that the profitability of the use of electricity and efficient use of electricity is reduced by the presence of shortages in the economy: condition E.3 is more restrictive than condition 4.80.

Equilibrium with Energy Efficient Firms and Energy Efficient Generators Firms self-select into the use of electricity, into the efficient use of electricity, and into the efficient use of fuel following:

$$\begin{aligned}\mathring{\varphi}_{ESG}^{*} &= \mathring{\varphi}_{EE/ESG}\\ \mathring{\varphi}_{EE}^{*} &= \mathring{\varphi}_{EE}^{'}\\ \mathring{\varphi}_{SG}^{*} &= \emptyset\end{aligned} \tag{4.81}$$

The economy behaves in the same way as Case 2a in section 4.7.1. The details of all derivations are presented in Appendix E. We denote to $\mathring{\chi}_{2a}$ the inverse of the affordability of an upgrade from the efficient use of electricity to the efficient use of fuel. It verifies:

$$\mathring{\chi}_{2a} = \chi_{2a} \left(\frac{\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} - \delta}{\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}} > \chi_{2a} \tag{4.82}$$

where χ_{2a} is the inverse of the affordability of an upgrade from the efficient use of electricity to the efficient use of fuel in an economy where prices are defined according to $c_\epsilon(\delta,\gamma) > p_E > g_\epsilon(\delta)$ but in which there are no power outages. Nevertheless, the relative price of electricity with respect to fuel is lower than in case 2a.

Equilibrium with Standard Firms and Energy Efficient Generators If the condition in equation 4.80 is not verified, then we have $\mathring{\varphi}'_{EE} < \mathring{\varphi}_{EE/ESG}$ and the efficient use of fuel always dominates the efficient use of electricity. Firms self-select into the use of electricity and into the efficient use of fuel following:

$$\begin{aligned} \mathring{\varphi}^*_{ESG} &= \mathring{\varphi}'_{ESG} \\ \mathring{\varphi}^*_{EE} &= \emptyset \\ \mathring{\varphi}^*_{SG} &= \emptyset \end{aligned} \tag{4.83}$$

The economy behaves in the same way as Case 2b in in section 4.7.1 (details of the computations in section E). We denote $\mathring{\chi}_{2b}$ the inverse of the affordability of an upgrade from the use of electricity to the efficient use of fuel. It verifies:

$$\mathring{\chi}_{2b} = \chi_{2b} \left(\frac{\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1} - \delta}{\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}} > \chi_{2b} \tag{4.84}$$

where χ_{2b} is the inverse of the affordability of an upgrade from the efficient use of electricity to the efficient use of fuel in an economy where prices are defined according to $c_\epsilon(\delta,\gamma) > p_E > g_\epsilon(\delta)$ but in which there are no power outages. Nevertheless, the relative price of electricity with respect to fuel is lower than in case 2b.

4.8.3 Direct and Indirect Crowding-Out: Equilibrium with Generators and Energy Efficient Generators

If $p_E > c_\epsilon(\delta, \gamma) p_F > g_\epsilon(\delta) p_F$, the use of fuel dominates the efficient use of electricity. The economy behaves in a similar way as described in section 4.7.2.

4.9 Summary of Theoretical Predictions and Empirical Model

Proposition 23. *In the presence of power outages, self-generation of electricity from fuel, either alone or in combination with energy efficient technologies, is no longer a strictly dominated technology choice when fuel is more expensive than electricity from the grid if the conditions described in propositions 19 and 21 in section 4.8 are verified (high intensity of shortages or limited technological progress). Hence, self-generation of electricity from fuel affects the level of energy efficiency through direct and indirect crowding-out of investments in energy efficient technologies. First, the size of self-generation capacity is the result of a sequence of past investment that proposition 15 and proposition 17 predict to have crowded out energy-saving investment choices (direct crowding out defined in section 4.6). Second, energy conservation decisions can alter the break-even period of past self-generation investment as shown in proposition 13, proposition 15, and proposition 17 if the installed captive power capacity is sufficiently high (indirect crowding out defined in section 4.6).*

Therefore, in India, by proposition 23, where power outages are significant and electricity from the grid is cheaper than fuel, we expect a decrease in the costs of self-generation through lower taxes on fuel to crowd-out investments in energy efficient technologies and therefore to decrease the energy efficiency of firms. While specifying an empirical model to test this theoretical prediction, we account for the lack of data on annual investment in captive power plants and in energy-saving technologies and we transform our variables of interest from flows to stocks. In the empirical model, we replace plant-level energy-saving investment with a measure of the level of energy efficiency of the plant and we include the observed status of self-generation in each plant instead of the yearly investment in generation capacity. In this way, the analysis is based on the observed outcome of past investment in both electricity self-generation and energy conservation.

The measure of energy efficiency at plant i and time t, EE_{it}, is our variable of interest. Our aim is to estimate the impact of electricity self-generation on EE_{it} to test for the theoretical prediction of a negative effect. We include two self-generation variables. The first one is the status of self-generator, D_{it}^{SG}, which measures the extensive margin of self-generation. It takes value of one if the plant generates a positive amount of electricity and zero otherwise. On the aggregate, it captures the fraction of self-generating plants over the year, but not counting for the actual amount of electricity produced. The second one is the variable SSG_{it}, the ratio of self-generated electricity in total electricity consumption. This variable combines both the extensive and intensive margin of self-generation because it has non-zero values for all plants that use a own power plant. We define it as an indicator of "electricity autarky". The variable SSG_{it} not only captures the degree of reliance of the plant from the grid to satisfy its own electricity demand, but also the intensity of investment in self-generation, relative to the plant size. A larger share is associated with a higher capacity and, therefore, higher fixed costs incurred by the plant.

Furthermore, sections 4.5.1, 4.5.2 and 4.5.3 of our theoretical model underline how both energy conservation and self-generation decisions are determined by the firm productivity level φ: only the most productive firms can afford the fixed costs associated with the investment in energy-saving technologies and in self-generation capacity. Moreover, total factor productivity should be directly related to the energy efficiency index EE, as energy is one of the production inputs. A measure of firm-level productivity need to be included in the econometric model to avoid an omitted variable bias. Note that the expansion of power generation capacity is not changing the efficiency of the production process and therefore it has no direct impact on productivity.

The following econometric model with repeated cross-section data:

$$EE_{it} = \alpha + \beta_1 D_{it}^{SG} + \beta_2 SSG_{it} + \gamma \mathbf{X}_{it} + \delta_1 \phi_{it} + \epsilon_{it} \tag{4.85}$$

summarizes the discussion we presented up to this point. The energy efficiency level EE_{it} in firm i depends on the status of self-generator D_{it}^{SG}, the share of self-generated electricity on total electricity consumption SSG_{it}, the firm-level productivity level ϕ_{it} and a vector of firm characteristics, $\mathbf{X}_{it}$, which includes state, year and industry fixed effects and a dummy-variable to indicate the status of seller of self-generated electricity on the grid. The error term is $\epsilon_{it} \sim N(0, \sigma_\epsilon)$.

4.9.1 Empirical Strategy

Coefficients resulting from the estimation of model (4.85) with Ordinary Least Squares (OLS) are likely to be biased because of endogeneity issues. On the extensive margin, more energy efficient firms have lower electricity bills and higher profits, especially in electricity intensive sectors. They are more capable of financing self-generation investment and own a captive power plant, which introduces a problem of reverse causality. Similarly, on the intensive margin, firms that are more energy efficient require a smaller scale of self-generation capacity to satisfy internal demand and therefore face lower a fixed cost of setting up a back-up generation system. Fixed costs are in fact proportional to the electricity generator capacity. As a result, the reverse causality between EE_{it} and the self-generation variables, D_{it}^{SG} and SSG_{it} might bias the OLS estimates of equation (4.85).

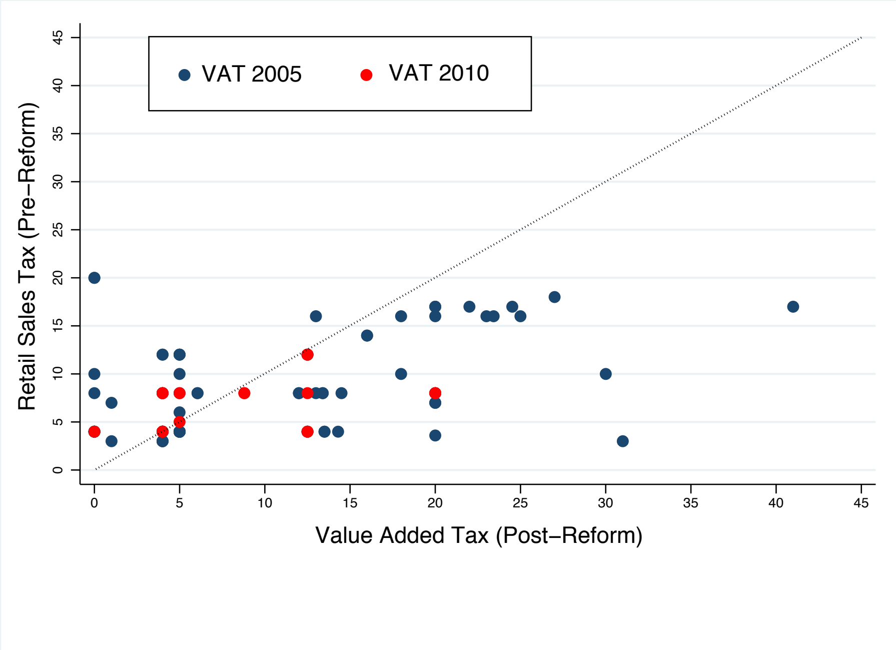
We exploit an exogenous variation in the cost of self-generation across time and geographical location (i.e. Indian states) to identify the causal impact of an increase in the incentives to self-generate electricity on energy efficiency. As described in section 2.2 of Chapter 2, a fiscal reform initiated by the Indian federal government mandated the states to replace their retail sales tax (RST) system with a value-added tax (VAT) system. The reform implementation changed the tax rate applied to all fuels used in most captive power plants[45], but did not affect the electricity price that is regulated by the central government. This exogenous change in fuel prices can be regarded as a shock to the cost of self-generation. Indeed, most of the captive power generation in India is based on fossil fuels like naphta, diesel, coal and gas. In 2005, renewable energy covered only 1% of total installed captive power capacity. Steam turbines (mostly coal-fired), diesel generators and gas turbines respectively accounted for 46.6%, 37.4% and 15% of the total (CEA, 2006). Furthermore, states did not uniformly change the tax rate on fuels. Figure 4.6 displays the pre- and post-reform tax rates for fuel commodities that are also used with electricity generators, such as diesels, coal, kerosene and furnace oil. Whereas the tax rate increased in some states - as shown by the points below the 45 degree line - fuels also experienced a drop in taxation in other states.

According to our model, the change in the price of fuel, p_F, affects marginal costs of self-generation and the productivity cut-offs φ_{SG} and φ_{ESG}. Energy conservation investment instead has not been affected by the reform, because the price of electricity has not been modified during tax system overhaul. Our dependent variable is defined over electricity only and therefore does not depend directly on the price of fuels.

In the case of a fuel tax rise, plants require higher productivity levels to afford the self-generation option. Self-generation becomes relatively less attractive than purchasing electricity from the grid. Yet, in the presence of electricity outages, the hike in fuel taxes lowers profits of plants facing power cuts. The response depends on the fuel mix of self-generators and differences in changes of tax rates across fuels. We have limited information on the plant-level fuel-mix, as the ASI has too many missing values for variables relative to the oil and coal costs. We construct sector j and state s

[45]Section 4.2 gives an overview of the power sector and captive power plants in India.

Figure 4.6: Pre- and Post-reform Tax rates on Fuels: Blue dots are fuel taxes in states that reformed the tax system before 2005. Red dots refer instead to those states that reformed between 2005 and 2010. Data below the 45 degrees line show an increase in the tax applied to a specific fuel good after the VAT reform. Fuel types include: Coal and coke, kerosene, light diesel oil, high-speed diesel oil, furnace oil. Note: dots might overlap when fuels have equal tax rates.



specific tax rates, $T_{js,t}$, for the pre-reform and post-reform time periods $t \in \{PRE, POST\}$, that are used as indicators of the tax rate faced by plants. Fuel-specific tax rates are aggregated into 2-digits NIC indicators. The aggregation of tax rates $\tau_{sh,t}$ over fuel types, h, is based on weights calculated from sectoral fuel-mixes of captive power plant capacity data, released by the Central Electricity Authority's *General Review* (CEA, 2006). Fuel weights, ω_{jh}, are India-wide and sector-specific. They are also time invariant, with 2005 as a benchmark year. The use of constant India-wide weights avoids issues related to endogeneity, as the self-generation energy-mix is endogenous and is likely to respond to a variation in fuel tax rates at the state-level. The weights should be interpreted as pre-reform, because they refer to March 2005 and only one state implemented the reform before that year. Even if more than half of Indian states reformed their VAT system exactly in 2005, we expect these multiple interventions not to affect the fuel mix in electricity self-generation at the beginning of the same year. First, as explained in section 2.2.2 of Chapter 2, there was a lot of uncertainty regarding the timing and the terms of the VAT reform's implementation. Therefore, it seems unlikely that firms would undergo the significant investments required to change their self-generation capacity based on speculations on the outcome of states legislative elections. Second, self-generation capacity is part of fixed capital. Hence, it takes time to change the fuel mix used by captive power plants.

The sector-state **fuel tax** index is finally computed as:

$$T_{js,t} = \sum_h \tau_{sh,t}\omega_{jh}. \tag{4.86}$$

which is the instrument we use in the 2SLS estimation of equation 4.85. We exploit the variations in tax rates across states and sectors as well as across time, to instrument for the self-generation dummy SG_{it} and the intensity of self-generation SSG_{it}. The first-stage regression based on the instrument $T_{js,t}$, reads:

$$vsg_{it} = \alpha + \beta_1 T_{js,t} + \gamma\mathbf{X}_{it} + \delta_1\phi_{it} + \varepsilon_{it} \tag{4.87}$$

for each one of the self-generation variables, $vsg_{it} \in \{SG_{it}, SSG_{it}\}$. Together with a set of plant-level controls, we include alternative sets of dummy variables and time trends: sector, state and year dummies, as well as their interactions, and industry-specific time trends. The VAT reform was broad and affected taxation level of the majority of consumption and input goods - except the price of electricity.[46] Standard errors are corrected for the presence of heteroscedasticity.

The second-stage equation is composed of the non-endogenous regressors, also included in the model specified in equation 4.87 and the predicted value of the self-generation variable obtained in the first-stage estimation, $\widehat{vsg}$:

$$EE_{it} = \alpha^{'} + \beta_1^{'}\widehat{vsg}_{it} + \gamma^{'}\mathbf{X}_{it} + \delta_2^{'}\phi_{it} + \eta^{'}D_{it}^{SG} + \epsilon_{it} \tag{4.88}$$

where notations are similar to those used in equations 4.85 and 4.87.

In the remainder of this chapter, we describe the implementation of the empirical strategy and the results we have obtained. In section 4.10.1, we present the dataset. Section 4.10.2 shows how we construct our firm-level indicator EE_{it} of energy efficiency and section 4.10.3 presents descriptive statistics on energy efficiency and self-generation in the Indian manufacturing sector and underlines how they are consistent with the assumptions and results of our theoretical model. Finally, section 4.11 presents our results.

[46] As a robustness check, we introduced in equations 4.85, 4.87, and 4.88 an additional dummy variable, $VAT_{si,t}$, to capture the potential effects the reform might have had on energy efficiency through other channels. Additionally, the indirect effect of the VAT reform will be captured by the addition of a state x year fixed effect as reported in the appendix.

4.10 Data and Measurement of Energy Efficiency

4.10.1 Data

Plant-level Data The main data sources are three rounds of the Annual Survey of Industries (ASI), a nation-wide enterprise survey in India, over the period 2000-2010, published by the Ministry of Statistics and Programme Implementation (MOSPI). The ASI consists of manufacturing establishments registered under the Factories Act, a social legislation that has been enacted for occupational safety, health and welfare of workers at work places. Each year, all establishments with more than 100 employees, and at least 12% of the rest, are surveyed with a representative sample at the state and 4-digit industry code level. Section 2.4 and Appendix A.3 of Chapter 2 provide a detailed description of the ASI survey.

In this chapter, we use information on the energy consumption and generation of formal firms contained in the ASI for the years 2000, 2005 and 2010. The three rounds of the ASI between 2000 and 2010 include information on the quantity and purchase value of electricity and coal consumed by each firm, as well as on the quantity of generated electricity. The value of electricity generated and sold is included, as well as the purchase value of other sources of energy such as different types of fuels (petrol, diesel) and chemicals. The 2000 and 2010 ASI rounds include the quantity and purchased value of gas consumed, while the 2000 round contains the quantity and purchased value of a detailed list of energy sources including biomass, wind and solar power. However, the large number of missing values in these variables (from 91% for coal to 100% for solar energy costs and quantities) prevents their use in our empirical analysis.

We must note that the ASI dataset only provides information on the amount of electricity generated, not on the installed capacity. In line with Allcott et al. (2016), we expect firms with captive power plants to keep production running during power outages and use the generators at full capacity. Therefore, we assume that the utilization rate of power generators is similar across plants and not correlated with plant-level characteristics. In this way, we are able to relate generation capacity to the observed electricity generated over a year period.

We use the gross value of new addition to other types of fixed assets (plants, transports, building, computer equipment, and land) as the value of new investments in this type of assets. Unfortunately, we cannot distinguish from investments in land and building for the purpose of generating electricity from other purposes: we discuss our strategy to identify an exogenous source of variations in the intensive and the extensive margin of self-generation in section 4.9.1. Similarly, any new investment in plants, building or transports can represent an investment in energy efficiency. For example, a new building could be better insulated, or a new piece of machinery could be more performant simply because it is of a more recent technology vintage. Section 4.10.2 describes our strategy to measure energy efficient firms in each industry and/or state. Finally, the questionnaire of the ASI in 2010 reports whether the firm is certified by any ISO Certification of the 14000 series,[47] and the gross value of existing and new addition to the stock of pollution control equipments. We interpret the latter as the gross value of investments in pollution control equipments. Pollution control differs from energy efficiency but the willingness of a firm to reduce its impact on the environment can also have a positive impact on its energy efficiency. Higher investments in pollution control can also be correlated with a stricter local enforcement of environmental regulations. Tables 4.6 and 4.9 present general descriptive statistics from the ASI on the variables of interest in this paper. Already in the original dataset, some states and industries show a very low number of observations and this raises potential issues in the estimation phase. The state Sikkim does not have observations in year 2000 and other four minor states have less than one hundred observation per year, namely Manipur, Meghalaya, Nagaland and the Andaman and Nicobar islands. As there are several missing values in the variables we use to estimate the energy efficiency index, the sample size for this states turn to be an issue for the estimation of some interaction between dummy variables later on. The number of remaining Indian states is 25. For the same reason, we exclude 4 out of 63 industries classified according to the NIC 3-digits code. Such data cleaning step alleviates the disparity in the number of observations between state and industry groups.

Deflators We encounter the common issue of lacking quantity data for some variables of interest like output. The use of quantities should always be preferred to the use of monetary value for variables capturing levels inputs and output because it captures only the variations in the production function of the firm and not variations in the nominal price through inflation or in the real price through changes in supply and demand. As a second-best solution, we derive real term values for variables of interest, such as revenues and input costs, by deflating the original values with constant 2004 price indexes. All nominal values are deflated to constant 2004 prices. This adjustment is done with different deflators. For fuel expenditures, we use the corresponding India-wide wholesale price index by commodity. Electricity quantities are retrieved from total costs, when necessary,

[47] ISO 14000 is a family of standards and practical tools for environmental management that support firms in minimizing their impact on the environment (in particular through water, land and air pollution) and facilitate their compliance with existing laws (www.iso.org/iso/iso14000).

using the electricity prices explained below. For capital, materials and output, we use the price deflators provided by Allcott et al. (2016). Deflators for materials are constructed combining the official commodity price indexes with India's input-output table. The industrial classification is reconverted from NIC-1987 into NIC-2004 codes. Capital price indexes are the implied national deflators calculated from the nominal and real quantities provided in the "Sector-wise Gross Capital Formation" Table from the Reserve Bank of India's Handbook of Statistics on the Indian Economy.

Temperature Data and Electricity Prices State-level temperature data have been collected through the Wolfram WeatherData database. We use minimum and maximum average temperatures over the year, expressed in degrees Celsius. State-level electricity prices are available from the Annual Report of the Central Electricity Authority. As the earliest Report accessible through the CEA website refers to year 2009, we have limited information on electricity prices for year 2000. Yet, the document reports tariffs that might have been effective from earlier dates, so we are able to collect values that apply to 2005 but only rarely to the year 2000. Missing values are replaced by the 2005 tariff deflated with the India commodity price index for electricity. Data are average values that apply to the single states, disaggregated by consumer category. In some states the average rates differ between urban and rural areas. Electricity rates are differentiated across industrial firms, depending on their monthly electricity consumption. Heavy industries consuming more than 8.7 million kWh per month benefit from a discount on marginal rates compared to less energy intensive firms.

Sales Tax Rates on Fuel VAT tax rates are retrieved from official documents. The pre-reform Retail Sales Tax schedules for single fuel commodities are obtained[48] from NIPFP (1994). Fuel-specific tax rates are aggregated into 2-digits NIC indicators. The aggregation of tax rates τ_{sh} over fuel types, h, is based on weights calculated from sectoral fuel-mixes of captive power plant capacity data, released by the Central Electricity Authority's *General Review* (CEA, 2006).

4.10.2 Measurement of Energy Efficiency

Energy intensity, the ratio of electricity consumed over output, is an imperfect measure of a firm's level of energy efficiency. Indeed, energy intensity is the final outcome of many other factors that are not directly related to the technical and organizational characteristics underlying energy efficiency.

[48]RST rates for some states could not be found, namely Arunachal Pradesh, Assam, Delhi, Himachal Pradesh, Jammu and Kashmir, Meghalaya, Mizoram Nagaland, Sikkim, Tripura. We attribute an average RST rate to these states for the period before the VAT reform.

	Obs.	Mean	St. Dev.	Min	Max
Production					
Net Sales Value (Million Rs.)	92270	4497	6820	0.0014	1.44x10^7
Capital Stock (Million Rs.)	87097	156	2890	0	631000
Total Wage Payments (Million Rs)	113926	18.8	141	0	16300
Total Workers (units)	113926	168.36	697.72	0	44195
Electricity Consumption					
Electricity Purchased (GWH)	109580	2.41	3.80	0	6540
Electricity Generation					
Electricity Generated (GWH)	39259	6.83	279	0	4210
Self-Generator (Yes/No)	142383	.275	.447	0	1
Share Self-generation on Elec. Consumption	38000	.206	.236	0.0001	1
Selling on Grid (Yes/No)	142383	.00534	.0729	0	1
Value Electricity Self-generated and Sold (Rs.)	737	7976.642	29775.57	.0008	437419
Other Variables					
High Temperature	142383	32.26	3.070	7	25
Low Temperature	142383	20.39	1.94	20	35
State-level Shortages	141245	.081	.06624	0	.288
Total Observations	142383				

Table 4.6: Descriptive Statistics: Presentation of the Sample

For instance, production processes that are more capital intensive also tend to be more energy intensive but, at the same time, they can employ the most advanced and energy efficient technologies. Moreover, the use of energy intensity introduces sources of endogeneity when jointly analysed with self-generation. Energy intensive sectors like the chemical and steel production industries have a high energy consumption per unit of output. Hence, they face higher costs in the event of power cuts. In this case, following the predictions of our model in section 4.6, higher energy intensity is associated with a higher probability of owning a captive power plant.

We adopt a stochastic frontier approach[49] to estimate the pure energy efficiency component from energy consumption data. The idea is to spot inefficiencies by comparing the energy use of plants that share similarities with respect to output level, input mix and location. For a given set of input prices and a technological frontier, inefficiencies in energy use arise as the input mix does not

[49]The stochastic frontier model is used in a large literature studying the conditions under which targets in production, cost, revenue, or profit can be attained. The methodology was first developed by Aigner et al. (1977). An alternative to the parametric stochastic frontier (SFA) approach is the non-parametric data envelopment analysis (DEA). The main difference between the SFA and the DEA approach is that the DEA generates a deterministic frontier from the observed firms and some firms will then be energy efficient by construction. On the contrary, the SFA generates a frontier that takes into account potential stochastic elements. Many varieties of the SFA have appeared in the literature and we follow the specifications recommended when using a cross-sectional dataset in the literature review of Filippini and Hunt (2015).

minimize production costs. The stochastic frontier analysis is a parametric method that measures the distance between the optimal and the actual energy use of a specific observation, disentangling factor intensity from the pure technical efficiency. The analysis of the energy efficiency through a stochastic frontier approach is done in two steps. In the first step, the frontier of energy efficiency is estimated. In the second step, the estimated frontier is used to construct a firm-specific measure of inefficiency: the amount by which each firms fails to reach the optimum level of electricity, in other words the distance to the frontier. This measure of inefficiency is used to estimate the model presented in section 4.9. It allows us to build a relative ranking of all firms in our sample with respect to energy efficiency.

We follow the literature (Filippini and Hunt (2015)) and specify the econometric model disaggregates energy consumption of plant i as composed of two parts. The first one is the technology frontier, specified as a stochastic exponential function:

$$f(\mathbf{X}_i, u_i) = e^{\beta \mathbf{X}_i} e^{u_i}, \tag{4.89}$$

where $u_i \sim N(0, \sigma_u^2)$ is an idiosincratic disturbance term. The frontier is determined by production-specific and location-specific characteristics included in $\mathbf{X}_i$ and it is common to groups of observations. The frontier could be interpreted as "best practices", predicting an optimal energy consumption level equal to $e_i^* = e_i - \nu_i - u_i$. The second component of the model, e^{ν_i}, measures the distance of plant i to the technology frontier, with a non-negative random variable ν_i, which is half-normally distributed, $\nu_i \sim HN(0, \sigma_\nu^2)$. We apply a logarithmic transformation to the general stochastic frontier model (4.89):

$$e_i = \underbrace{\beta \mathbf{X}_i + u_i}_{\text{Stochastic frontier}} + \underbrace{\nu_i}_{\text{Inefficiency}} .$$

The assumption on ν_i to be non-negative and half-normally distributed allows to have a specific term dedicated to capture inefficiencies, that is, how much the actual input mix differs from the optimal one suggested by the technological frontier. For the specification of $\mathbf{X}_i$, we follow the input demand frontier function of Filippini et al. (2011), in which the frontier is determined by the scale of production, the input mix, the price of electricity and location-specific controls:

$$e_i = \alpha + \alpha_y y_i + \alpha_k k l_i + \alpha_p p_i^E + \beta_1 T_i^{max} + \beta_2 T_i^{min} + \delta D_i^s + \delta_t D_i^t + \nu_i + u_i, \tag{4.90}$$

where lower-case indicates the natural logarithm of each variable. Electricity consumption e is regressed on a set of production variables, output y, the capital-labour ratio kl_i (the fixed capital stock over the number of workers), as well as two temperature variables measuring the minimum and maximum average temperature measured by year-state combination, T_i^{max} and T_i^{min}. Sector and time dummies are introduced to remove sector-specific and year-specific factors that affect electricity consumption. We allow for state-level clustering of standard errors for u_i.

Figure 4.7 illustrates the stochastic frontier approach for the one dimensional case with log-output y_i. The technology frontier is a fitted curve, similar to the one we would obtain with an OLS estimator. However, the introduction of a non-negative error term, ν_i, constrains the curve to lie on the left of the observed plant-level observations. Net of a pure random term u_i, to produce a level of log-output y_i all observations must employ not less than the optimal amount of energy, $e^* = e_i - \nu_i - u_i$. Plants that have $\hat{\nu} = 0$ lie on the technological frontier.

We estimate an India-wide technological frontier for each sector, allowing for shifts over time. Sector and year dummies enter the estimated models and, as illustrated in Figure 4.7, they determine the non-stochastic part of the frontier. Production technologies in different sectors are so different, that energy efficiency is better measured when comparing plants within a quite narrow category of economic activity. Therefore, the frontier curve is sector-specific and varying over time. Plants within each state might face correlated idiosyncratic disturbances, which calls for the clustering of error terms u_i at the state-level. Yet, we correct standard errors for heteroskedasticity using the Huber-White sandwich estimator of the variance-covariance matrix but we do not specify a structure for potential state-level clustering. The reason is related to the structure of data clusters. The estimation of within-state correlation is hindered by the quite small number of state groups and the high unbalance in group sizes. Whereas some states have less than 500 observations, others have more than 6,000. Moreover, the small number of clusters, which is 25, does not satisfy the asymptotic assumption of the estimator.

We estimate equation 4.90 and results are displayed in Table F.1, together with alternative specifications. Results from the baseline econometric model presented in equation 4.90 are presented in column (1). In this case the production scale effect is completely captured by output - which estimated coefficient rises - and the capital/labour ratio instead represents the input mix in production. Firms with higher capital used per worker will tend to demand more electricity. Due to the use of air conditioning, an increase in temperature leads to higher consumption of electricity, while a decrease in temperature decreases electricity consumption. Heating systems are usually not using electricity and are not common in India, which explain why we find only a limited negative effect

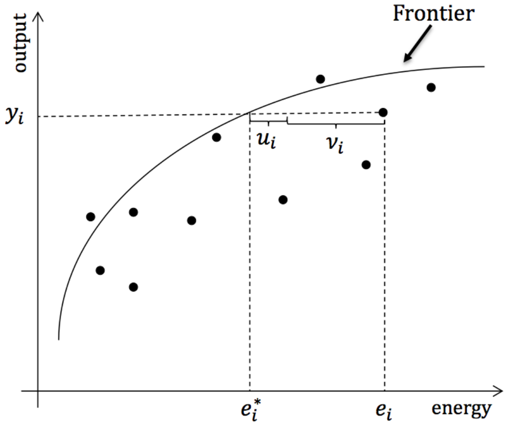


Figure 4.7: Intuition of the Stochastic Frontier: The frontier represents the optimal input mix, given the set of technologies adopted by the plant included in the dataset. Two plants might share very similar characteristics but have different levels of energy use because they employ different production technologies. Plants that produce the maximum amount of output with the minimum use of energy determine the frontier. The fitted line is the deterministic frontier, which allows for idiosyncratic error terms u_i. The model separates noise u_i, from the purely technical inefficiency component: the amount by which the observed firm fail to reach the optimum, that is to say the frontier. Hence, the inefficiency level, the distance-to-frontier indicator, is measured by the positive random term v_i.

of minimum temperature on electricity consumption. In columns (2) and (3), the model includes labour and capital as separate inputs instead of the capital-labour ratio. In (3), labour is measured with total salaries, to control for the quality of labour, i.e. human capital. In column (4), the model additionally controls for two indicators of the environmental management of the firm, the ISO 14000 certification and the capital stock of pollution abatement devices. Because of data availability, the last case is limited to a small subsample but both variables appear to be statistically significant. It is interesting to note that investing in pollution control or in the ISO certification of environmental standards has a positive impact on electricity consumption. It could be either that pollution control and environmental management require more sophisticated and electricity intensive production systems or that the least energy efficient firms invest in pollution control.

The estimated inefficiency term, $\hat{\nu}_i$, is used to compute the energy efficiency index as $EE_i = exp(-\hat{\nu}_i)$. We obtain a distance-to-frontier indicator EE_i, ranking all plants according to their energy efficiency level. Filippini and Hunt (2015) highlight that model (4.90) controls for both technical and allocation efficiency, as electricity prices are included. Notice that, throughout the analysis, we use the term "energy efficiency" meaning the technical efficiency in electricity use, which does not include the efficiency of captive power plants.

The estimated EE_i is robust to alternative model specifications. The correlation matrix between estimated EE_i obtained from alternative models is show in table F.1. High correlation between the results of alternative model specification suggests that the use of the stochastic frontier model gives robust energy efficiency indexes.

Finally, we acknowledge for measurement error in our dependent variable, energy efficiency, as we estimate it using the methodology described above. As shown in the literature, measurement error in the dependent variable that is not correlated with other independent variables in the model does not affect the consistency of the estimated coefficient using OLS or 2SLS. To prevent this, we re-estimate the energy efficiency indicator adding all the regressors included in equations 4.85 and 4.88. Nevertheless, measurement error inflates standard errors, which implies that our empirical results will under-estimate the true effect of self-generation investments on energy efficiency.

Energy Efficient Indices				
	(1)	(2)	(3)	(4)
(1)	1			
(2)	0.879***	1		
(3)	0.899***	0.993***	1	
(4)	0.940***	0.832***	0.853***	1

Table 4.7: **Robustness of the Energy Efficient Index**. Correlation matrix between the energy efficiency indices resulting from alternative specifications of the stochastic frontier model in equation 4.90 respectively displayed in columns (1), (2), (3), and (4) of table F.1.

4.10.3 Descriptive Statistics: Energy Efficiency and Self-generation

There exists variations in energy efficiency and self-generation across Indian States (Figure F.1). By assumption in the model presented in equation 4.90, the technological frontier is the same across states every year. Thus, cross-states differences should not be due to industry composition but rather to lags in innovation and technology adoption. With respect to self-generation, the right panel of Figure F.1 shows the extensive margin: the fraction of firms owning and using an electricity generator.

Consistently with our theoretical results summarized in section 4.9, figure F.2 shows that energy efficiency tends to decline when the level of self-generated electricity increases. The year-specific fitted lines indicate that the correlation is stable over time. Table 4.10 displays mean differences in energy efficiency levels between non-generators and generators. Self-generators have on average a lower energy efficiency index compared to non-generators. Furthermore, table 4.8 shows that labour productivity is higher in energy efficient firms than in non-energy-efficient non-generators and higher in energy efficient generators than in non-energy-efficient generators. This last piece of empirical evidence is also visible in tables 4.9 and F.3.

	(φ vs φ_{EE})	(φ_{EE} vs φ_{SG})	(φ_{SG} vs φ_{ESG})
Labour Productivity	-0.345***	-0.0253	-0.707***
	(0.000)	(0.104)	(0.000)
N	42399	40646	33591

P-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.8: **Differences in Mean for Labor Productivity between different groups of firms** φ includes non-energy-efficient non-generators, φ_{EE} non-energy-efficient non-generators, φ_{SG} non-energy-efficient generators, and φ_{ESG} energy efficient generators (*Efficient* designates a firm ranked in the top 20% of the sample according to our energy efficiency indicator i.e a firm on the technological frontier.)

	obs.	% obs.	En. Eff.	Labour Prod. (*)	Coal Costs (*)	Gas Costs (*)	Biomass Costs (*)
Generators	39242	0.725	.640	17543.91	39697.1	4231.9	18618.7
Sellers	761	98.06	.617	42176.34	805372.7	28451.4	321929.4
Non Sellers	38481	1.94	.64	17263.33	23634.8	3649.5	14961.1
Non Generators	103141	.275	.650	12920.6	4801.6	292.8	8230.7
Mean Differences							
Non Gen. vs Gen.			.0171 ***	-4970.842***	-7.62x10^7***	-1.5210^7**	-5109148***
			(.000436)	(438.14)	(7538717)	(9251740)	(1693334)
Non Sellers vs Sellers			.0298***	-29534.27***	-8.14x10^7***	-3.44x10^8***	-2.3710^7***
			(.00201)	(2094.48)	(4.85x10^7)	(4.49x10^7)	(6627846)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.9: Descriptive Statistics on Electricity Generation (*) Thousand Rupies

Year	2000	2005	2010
Energy Efficiency	0.0195***	0.0180***	0.0165***
	(9.44)	(28.35)	(25.94)
N	3512	32430	31506

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.10: Mean Differences for Energy efficiency between Non-Generators and Generators: N denotes the total sample size

4.11 Empirical Results

We estimate the model presented in equation (4.85) with OLS and 2SLS, including firm-level characteristics such as the plant's years of operation and the status of rural or urban plant. Standard errors are corrected for the presence of heteroskedasticity using the Huber-White sandwich estimator. Table 4.11 shows the estimated impact of self-generation on energy efficiency with the two estimators, under alternative model specifications. We compare two versions of model 4.85. Whereas both include the interaction between state and year fixed effects, the first model has sector-specific time trends and the second one has the interaction terms between year and sector dummies. The complete set of results is reported in the Appendix in Table F.2, for OLS, and in Table F.3 (first stage results) and Table F.4 (second stage results), for the 2SLS estimation.

Results point at a negative effect of self-generation on plant-level energy efficiency. The OLS estimation already shows that plants with own generation of electricity tend to be more distant from the sectoral energy efficiency frontier, as measured by the index estimated with the stochastic frontier model of Section 4.10.2. The 2SLS regressions confirm not only the negative effect, but also the presence of an upward bias in the OLS estimates.

The coefficients of the other regressors are also statistically significant (see the Appendix). Energy efficiency increases in labour productivity, measured as the plant output per worker, and decreases with the amount of capital and labour used in production, possibly meaning that efficiency decreases with the plant size. There is also a negative correlation of the plants' age with energy efficiency. The result is in line with Linn (2008), which finds entrants to have a lower energy intensity than existing firms: newly built plants benefit first from the fact that technologies available at their time of construction are more energy efficient due to technological progress. Plants in rural areas have a higher energy efficiency level. This effect is weakly significant, but the sign is puzzling, as we would expect plants in rural areas to use less advanced technologies and suffer from higher transportation costs. Finally, we point out that the VAT dummy of equation (4.85) is not directly included in the regression equation, as the interaction between state and year fixed effects would be collinear with such indicator. Therefore, the indirect effect of the tax reform on energy efficiency is captured by the interaction terms.

The first-stage estimations indicate that the instrument has a negative effect on self-generation. We perform diagnostic tests to evaluate the IV approach. As errors are non-identically distributed, we apply the Kleibergen-Paap Wald (KPW) statistic to test for the presence of weak instruments. The

Table 4.11: Overview of Estimation Results: The table reports the estimated coefficient for the two self-generation variables using OLS and 2SLS. The columns indicate alternative model specifications. Both include a base set of controls: labour productivity, age of the firm, rural or urban status, electricity price, capital stock and total salaries, as well as state, industry and year fixed effects. All models include year and state fixed effects **(FE)** along with year x state fixed effects. In all regressions, observations are weighted with ASI sample weights. Robust standard errors in parenthesis and F-statistics in square brackets. The Weak IV statistic is the Kleibergen-Paap Wald statistic.

	Energy Efficiency			
Self-Generation Indicator	*Status (Yes/No)*		*Share of Electricity Use*	
	(1)	(2)	(1)	(2)
OLS	-0.014***	-0.014***	-0.021***	-0.021***
	(0.001)	(0.001)	(0.002)	(0.002)
	[53.589]	[43.471]	[50.488]	[40.352]
R Square	0.225	0.228	0.218	0.221
2SLS	-0.137**	-0.136**	-0.246***	-0.230***
	(0.059)	(0.060)	(0.087)	(0.083)
	[30.528]	[24.792]	[38.774]	[32.112]
R Square	-0.636	-0.608	-0.123	-0.073
Weak IV Statistic	8.498	8.125	25.947	27.851
Observations	63202	63203	63202	63202
Industry Trends	Yes	No	No	No
Year x State FE	No	Yes	No	Yes
Year x Industry FE	No	No	Yes	Yes

role of the test is to detect the potential bias introduced by the use of an instrument that is poorly correlated with the endogenous regressor. According to Stock and Yogo (2005), the critical value of 8.96 ensures a maximal 15 percent bias. The instrument is strong in the 2SLS estimation based on the share of self-generation - the last two columns in Table 4.11 - but the KPW statistic falls slightly below the critical value in the other model. When self-generation is only captured by the binary variable, the fuel tax instrument is expected to have a bias that is a bit higher than 15 percent.

Tables F.3 and F.4 display the results. The negative effect of self-generation on the level of energy efficiency is confirmed by the IV approach and survives alternative specifications of control variables. We see that the coefficient for the share of SG on electricity consumption is larger than the one of the self-generation dummy, indicating that the combination of the intensive margin of self-generation, namely the total installed capacity relative to the firm size, has an additional effect on energy efficiency. Not only self-generating plants have a lower level of energy efficiency, but inefficiencies in energy use tend to rise with a larger self-generation capacity.

The OLS estimation was expected to be upward biased because higher energy efficiency favors the conditions for investing in self-generation, by improving the plant profitability and reducing the size of the required generator capacity. This is consistent with our IV estimates. The results in the IV estimations are characterized by a lower - more negative - coefficient for self-generation variables D_i^{SG} and SSG_i compared to the OLS case. The magnitude of the negative effect of self-generation on energy efficiency is much larger when accounting for the endogeneity of SSG_i and D_i^{SG} in (4.85). Here, we deal with these variables separately and we find the share of self-generation on total electricity consumption, SSG_i, to have a stronger negative effect on energy efficiency compared to the self-generation status. As the former variable captures both the intensive and extensive margin, we interpret such difference in the estimated coefficients like a stronger negative impact of the intensive margin on energy efficiency.

Robustness The empirical models already have a broad set of control variables but we perform additional robustness checks to validate our results. First, we make use of an alternative instrument that only captures the post-reform variation in fuel tax rates. This instrument is the interaction of the VAT dummy with the difference in pre- and after-reform tax rates, $T_{js,t}$. Indeed, it could be that the differences in fuel taxes between industries and states are affected by differences in the reliance of firms on self-generation through lobbying. Variations across states in the average dependence on self-generation from fuel could be due to differences in the industry mix or in reliability of the electricity grid. To control for this potential source of endogeneity, we use an instrument that takes into account only the change in taxes caused by the VAT reform under which the fuel taxes had to be more or less aligned on the rates recommended by the state government. Hence, our instrument captures the extent of the change from the fuel sales tax to the fuel VAT tax. More precisely, the new variable is constructed as follows:

$$T_{js}^D = VAT_{st}(T_{js,POST} - T_{js,PRE}), \tag{4.91}$$

where VAT$_{st}$ is a reform dummy taking the value 1 for those observations belonging to a post-reform year in state s. The first and second stage estimation results are reported in Tables (F.5) and (F.6) in the Appendix. The point estimates share the same signs of the previous ones and the effects of self-generation variables are less negative, yet still statistically significant at 5 percent level. The test for weak instruments suggests that this tax reform indicator explains better the variation in self-generation levels. The KPW statistic is higher than in the previous specification and the model including D_i^{SG}, the self-generation dummy, has now a lower bias due to the lower of correlation between the instrument and the endogenous regressor.

The second robustness check deals with the estimation of the variance-covariance matrix. We have so far controlled for heteroskedasticity and not for within-group correlation, i.e. clustering. As we discussed in Section 4.10.2, the use of cluster-robust standard errors is, in our case, potentially affected by small-sample bias. The presence of 25 states is not a sufficiently high number of groups and, moreover, the number of within-group observations varies widely across clusters. This is why we place the estimation with clustered-robust standard errors among the robustness checks. We re-run the estimations of the stochastic frontier model, as well as the OLS and IV models under the assumption of state-level clustering. As a result, we find clustering to inflate the standard errors of the self-generation estimates in the 2SLS regression. In the IV model, both self-generation variables are not statistically significant at 10 percent level.

4.12 Conclusion

This paper is an investigation of the potential conflicts between demand-side and supply-side energy policies. Worldwide, governments intervene to dampen energy consumption and promote distributed electricity generation not only for the purpose of climate change policies, but also to improve the availability of electricity and lower the frequency of power cuts. We argue that these two policies might be conflictual and, using Indian data, we offer evidence showing that fostering small-scale self-generation of electricity has a negative feedback on plant-level energy efficiency. Our theoretical framework highlights the presence of conflicts between self-generation of electricity and investment in energy efficiency and find empirical evidence for a negative effect of the former on the latter variable. We allow for the presence of electricity outages. Both the installation of self-generation capacity and investment in energy-saving technologies incur in fixed costs and have little degree of complementary. On the one hand, self-generation does not affect the level of energy efficiency of the production unit and, on the other hand, higher energy efficiency does not help in the presence of power cuts. In a model of heterogenous firms in productivity, we show that crowding out might arise between these two types of investment through two different channels. First, self-generation directly crowds out investments in energy conservation when firms are not profitable enough to adopt both types of technologies. Hence, they choose the most profitable one: in the presence of power outages or low fuel prices, self-generation is a more profitable investment than energy efficiency for firms in the middle of the productivity distribution. Second, self-generation indirectly crowds out investments in energy conservation, because it strengthen the comparative advantage of the most profitable adopters. Thus, it intensifies competition on the market and increases the minimum level of productivity required to invest in energy efficiency. The empirical evidence from Indian data points at the existence of a crowding out effect of self-generation on the energy efficiency level, measured as the distance to the technological frontier. This result is supported by the use of an IV strategy overcoming endogeneity problems. Given the limited information available on the type of self-generation technology used by plants, we could not differentiate between renewable energy and fossil fuel generation units. This distinction is important for environmental policy, as the crowding out effect might differ between these two classes of energy technologies.

Appendices

Appendix E

Theoretical Results: Proofs

The derivation of the baseline economy equilibrium follows the strategy described in Chaney (2008) and the derivation of the equilibrium with different technologies is very similar to the strategy followed in Bustos (2011). In this appendix, I enter into the details of computations that are not explicitly derived in the main body of this chapter.

Introductory Results on Technological Parameters

Proposition 25 describes the technological conditions for the efficient use of electricity grid to dominate the standard production technology for at least some firms when the use of self-generation is dominated by all other production modes at all productivity levels. Proposition 26 describes the technological conditions for the use of fuel (self-generation) to dominate the standard production technology for at least some firms when the efficient use of electricity grid is dominated by all other production modes at all productivity levels. Finally, proposition 27 presents the conditions on exogenous parameters for all production technologies to be present in equilibrium, that is to say for every production mode to offer the highest level of profitability to at least some firms in the economy: neither the use of fuel nor the use of electricity is dominated. Table 4.4 presents a summary of the different possible cases in the absence of power outages.

Proposition 24. *The following technological condition:*

$$(f + f_{SG})\left(\left(\frac{p_{SG}}{p_E}\right)^{\frac{1}{\sigma-1}} \quad > \quad (f + f_{EE}) \tag{E.1}$$

is equivalent to:

$$\varphi_{EE} < \varphi_{SG} \tag{E.2}$$

Proposition 25. ***Firms' Technology Choice when Self-Generation is a Dominated choice:***

The following technological condition:

$$(f_{SG} + f_{ESG} - f_{EE})\left(\gamma^{\sigma-1} - 1\right) \quad > \quad f_{EE}\left(\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1} - \gamma^{\sigma-1}\right) \tag{E.3}$$

is equivalent to:

$$\varphi_{EE/ESG} > \varphi'_{ESG} \tag{E.4}$$

The left-hand side of the inequality E.3 is increasing in the extra fixed costs for a standard firm of becoming an energy efficient generator but the right-hand side is increasing in the revenue gains of doing so. Similarly, the right-hand side of the inequality is increasing in the fixed adoption costs of the energy-efficient technology but the left-hand side is increasing in the revenue gains from electricity efficiency. The inequality in equation E.3 is depicted in figure 4.3a and is equivalent to $\chi_{2a} > \chi_{2b}$: the upgrade from the efficient use of electricity to the efficient use of fuel is less affordable than the upgrade from the use of electricity to the efficient use of fuel. Thus, we have $\varphi_{EE/ESG} > \varphi'_{ESG}$, which is relevant only if the use of fuel is a dominated choice in the economy as in Case 2 in section 4.3a when electricity and fuel prices verify $\gamma p_F \geq p_E > p_F$. Hence, the condition in equation E.3 implies that the adoption of the energy efficient efficient technology is more profitable than the standard production technology for firms that can afford the fixed adoption costs of the efficient use of electricity but not the combined fixed costs of self-generation and the energy efficient technology ($\varphi'_{EE} < \varphi < \varphi_{EE/ESG}$). On the contrary, if the inequality in equation E.3 is not verified, as depicted in figure 4.3a, then we have the opposite situation and firms of productivity $\varphi'_{EE} < \varphi < \varphi_{EE/ESG}$ are more profitable when using the standard production technology.

Proposition 26. ***Firms' Technology Choice Set when the Efficient Use of Electricity Grid is a dominated Choice***

The following technological condition:

$$f_{ESG}\left(\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1}-1\right) \quad > \quad f_{SG}\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1}\left(\gamma^{\sigma-1}-1\right) \tag{E.5}$$

is equivalent to:

$$\varphi_{SG/ESG} > \varphi'_{SG} \tag{E.6}$$

The condition in equation E.5 means that the change of production technology from the use of fuel (self-generation) to the efficient use of fuel (energy efficient self-generation) generates a lower increase in profitability than the change from the use of electricity to the use of fuel. Hence, the incentive to start using fuel in a more energy efficient way is not as large as the incentive to switch from the use of electricity to the use of fuel. The left-hand side of the inequality E.5 is increasing in the difference in fixed costs between self-generation and energy efficient self-generation but the right-hand side is increasing in the revenue gains of doing so. Similarly, the right-hand side of the inequality is increasing in the extra fixed costs of self-generation but the left-hand side is increasing in the revenue gains it allows for. Thus, the incentive to start using fuel in an energy efficient way is not as large as the incentive to switch from using electricity to using fuel and we have $\varphi_{SG/ESG} > \varphi'_{SG}$. This is relevant only when the adoption of the efficient use of electricity grid is a dominated choice as in Case 3a depicted in figure 4.4 in section sub: Case 3a. Hence, the condition in equation E.5 implies that self-generation is more profitable than the standard production technology for firms that can afford the fixed costs of self-generation but not the combined fixed costs of self-generation and the energy efficient technology ($\varphi'_{SG} < \varphi < \varphi_{SG/ESG}$).
On the contrary, if the inequality in equation E.3 is not verified, as depicted in figure 4.4, then we have the opposite situation and firms of productivity $\varphi'_{SG} < \varphi < \varphi_{SG/ESG}$ are more profitable when using the standard production technology.

Proposition 27. ***Firms' Technology Choice Set when No Technology is Dominated***
The following technological condition:

$$f_{ESG}\left(\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1}-\gamma^{\sigma-1}\right) \quad > \quad (f_{SG}-f_{EE})\left(\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1}-\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1}\right) \tag{E.7}$$

is equivalent to:

$$\varphi_{SG/ESG} > \varphi'' \tag{E.8}$$

The condition in equation E.7 means that the change of production technology from the use of fuel (self-generation) to the efficient use of fuel (efficient self-generation) is less affordable than the change from the energy efficient use of electricity to the use of fuel. Hence, the incentive to start using fuel in a more energy efficient way is not as large as the incentive to switch from the efficient use of electricity to the standard use of fuel. The left-hand side of the inequality E.7 is increasing in the extra fixed costs of becoming energy efficient for a generator but the right-hand side is increasing in the revenue gains of doing so. Similarly, the right-hand side of the inequality is increasing in the difference in adoption fixed costs between the self-generating and the energy-efficient technology but the left-hand side is increasing in the difference in revenue gains.

Cheap Fuel Equilibria (Case 2 and 3)

Intermediary Prices Equilibria (Case 2) without Power Outages

By proposition 25, the inequality in equation E.3 implies that $\varphi_{EE/ESG} > \varphi'_{EE}$. Hence, there exist energy efficient firms for which the fixed adoption costs of self-generation are too high: they adopt only the technologically efficient use of electricity from the grid. As depicted in figure 4.3a, there are *three types of producers* in the differentiated good sector: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{EE}]$, ii) energy efficient firms of productivity $\varphi \in [\varphi^*_{EE}, \varphi^*_{ESG}]$, and iii) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$. The equilibrium prices and productivity cutoffs in this case are derived in section E. As depicted in figure 4.3b, if $\varphi_{EE/ESG} < \varphi'_{EE}$, the efficient use of fuel always dominates the efficient use of electricity: there are only *two types of producers* in the differentiated good sector: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{ESG}]$ and iii) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$. The equilibrium prices and productivity cutoffs in this case are derived in section E.

Case 2a - Non-Dominated Energy Efficiency: Equilibrium with standard firms, energy efficient firms and energy efficient generators

If the inequality in equation E.3 is verified, then $\varphi_{EE/ESG} > \varphi'_{EE}$. This is the case depicted in figure 4.3a. The efficient use of electricity dominates the efficient use of fuel for firms in the middle of the productivity distribution and the double technology adoption cut-off is defined as $\varphi^*_{ESG} = \varphi_{EE/ESG}$. The equilibrium price index is given by $P^*_{Case2} = [P + P_{EE} + P_{ESG}]^{\frac{1}{1-\sigma}}$ where the price indices of subsets of varieties respectively offered by standard firms, generators and energy efficient generators are given by: $P = \int\limits_{\omega\in\Omega_S} p(\omega)^{1-\sigma}d\omega$, $P_{EE} = \int\limits_{\omega\in\Omega_{EE}} p(\omega)^{1-\sigma}d\omega$, and $P_{ESG} = \int\limits_{\omega\in\Omega_{ESG}} p(\omega)^{1-\sigma}d\omega$. Ω_S and Ω_{EE} are defined as before. Ω_{ESG} is the subset of varieties produced by energy efficient generators. There is $\Omega = \Omega_S \cup \Omega_{EE} \cup \Omega_{ESG}$ and P^*_{Case2a} becomes:

$$P^*_{Case2a} = P^* \left(1 + \frac{f_{EE}}{f}\left(\frac{1}{\chi}\right)^k + \left(\frac{f_{SG} + f_{ESG} - f_{EE}}{f}\right)\left(\frac{1}{\chi_{2a}}\right)^k\right)^{\frac{-1}{\sigma-1}} \tag{E.9}$$

$$< P^{**} < P^* \tag{E.10}$$

where $\frac{1}{\chi_{2a}} < 1$ represents the affordability of an upgrade from the efficient use of electricity to the efficient use of fuel as described in section 4.5.3. Self-generation, as an alternative production mode in the economy, reduces variable costs and prices, which intensifies competition and increases the minimum level of productivity required to be profitable on the market:

$$\varphi^*_{Case2a} > \varphi^{**} > \varphi^* \tag{E.11}$$

Aggregate productivity of energy efficient firms $\overline{\varphi EE}_{Case2a}$ is given by equation 4.48 and the share of energy efficient firms by equation 4.49. Note that the share of energy efficient firms - including energy efficient generators - is unchanged with respect to an economy where only the adoption of energy efficient technologies is possible but the mass of energy efficient firms must be smaller from equation E.11 and $\varphi'_{EE} = \chi\varphi^*$. Aggregate productivity of energy efficient generators equals the aggregate productivity of generators in case 2a:[1]

$$\overline{\varphi ESG}_{Case2a} = \overline{\varphi SG}_{Case2a} = \left[\frac{1}{1 - V\left(\varphi^*_{ESG}\right)}\int_{\varphi^*_{ESG}}^{\infty} \varphi^{\sigma-1} v\left(\varphi\right) d\varphi\right]^{\frac{1}{\sigma-1}} \tag{E.12}$$

$$= \chi_{2a}\overline{\varphi}_{Case2a} = \left(\frac{\chi_{2a}}{\chi}\right)\overline{\varphi EE}_{Case2a} > \overline{\varphi EE}_{Case2a} \tag{E.13}$$

The less affordable the upgrade from the use of electricity to the efficient use of fuel, the larger χ_{2a} and the larger the difference in average productivity between energy efficient generators and other

[1] Note that the inequality in equation E.3 that defines case 2a tells us that the incentive to switch from the efficient use of electricity to the efficient use of fuel is smaller (less affordable) than the incentive to switch from the use of electricity to the efficient use of electricity. It implies that $\chi_{2a} > \chi$ which is why there are energy efficient firms in the economy in spite of the variable costs of the efficient use of fuel being lower than those of the efficient use of electricity.

firms. The less affordable the upgrade from the use of electricity to the efficient use of electricity, the larger χ and the smaller the difference in average productivity between energy efficient generators and energy efficient firms: A smaller range of (higher) productivity levels can afford the technologically efficient use of electricity.

Aggregate revenues and aggregate profits can be split up into revenues and profits of standard, energy efficient firms, and energy efficient generators: $R = R_S + R_{EE} + R_{ESG}$ and $\Pi = \Pi_S + \Pi_{EE} + \Pi_{ESG}$. Aggregate revenues of energy efficient generators are $R_{ESG} = J_{ESG} r_{ESG} (\overline{\varphi_{ESG}})$. The share of energy efficient generators equals the share of firms having adopted the energy efficient technology:

$$\frac{J_{ESG}}{J_A} = \frac{J_{SG}}{J_A} \quad = \quad \left(\frac{\varphi^*_{Case2a}}{\varphi^*_{ESG}}\right)^k = \left(\frac{1}{\chi_{2a}}\right)^k \tag{E.14}$$

The share of energy efficient firms increases when upgrading from the use of electricity to the efficient use of fuel is more affordable and χ_{2a} decreases. The share of fuel in total energy consumption is given by:

$$\frac{I_{ESG}}{I} \quad = \quad 1 - \frac{1}{\gamma}\left(\frac{p_E}{\gamma p_{SG}}\right)^{\sigma} (\chi_{2a})^{\sigma-k-1} \tag{E.15}$$

The more affordable the upgrade to the efficient use of fuel and the larger the share of fuel in aggregate input. This share decreases in the performance of the energy efficient technology γ and when the relative cost of using electricity with respect to the costs of using fuel efficiently decreases. The aggregate efficiency gain in the economy in comparison to a case when only the energy efficient technology is available is given by:

$$E_{Case2a} \quad = \quad E_{EE} + E_{ESG} \tag{E.16}$$

$$E_{Case2a} \quad = \quad \left(\frac{P^*_{Case2a}}{P^{**}}\right)^{k-\sigma+1} \left(E^{**}_{EE} + \left(\frac{p_E}{p_{SG}}\right)^{\sigma} (p_E - p_{SG})\, E^{**}_{ESG}\right) \tag{E.17}$$

where E_{EE} is the efficient gain made by energy efficient non-generators and E_{EE} the gain of energy efficient generators. E^*_{EE} (respectively E^*_{ESG}) is the efficiency gains that energy efficient firms (respectively energy efficient generators) in the range of productivity $\varphi \in [\varphi^*_{EE}, \varphi^*_{ESG}]$ (respectively $\varphi \geq \varphi^*_{ESG}$) would realize if the self-generation technology was not available as given in equation 4.50. Hence, the gains made by energy efficient firms decreases when self-generation is introduced in the economy because it decreases the market shares of energy efficient firms through the decline in the price index:

$$E_{EE} = \left(1 + \frac{f_{SG} + f_{ESG} - f_{EE}}{f_{EE}}\left(\frac{\chi}{\chi_{2a}}\right)\right)^{1-\frac{k}{\sigma-1}} E_{EE}^{**} < E_{EE}^{**} \tag{E.18}$$

Case 2b - Dominated Energy Efficiency: Equilibrium with standard firms and energy efficient generators

If the inequality in equation E.3 is not verified, then $\varphi_{EE/ESG} < \varphi'_{EE}$. The energy efficient use of electricity is always dominated by the efficient use of fuel and the double technology adoption cut-off is defined as $\varphi^*_{ESG} = \varphi'_{ESG}$. This is the case depicted in figure figure 4.3b. The equilibrium price index is given by $P^*_{Case2b} = [P + P_{ESG}]^{\frac{1}{1-\sigma}}$ where the price indices of subsets of varieties respectively offered by standard firms and energy efficient generators are defined as before. There is $\Omega = \Omega_S \cup \Omega_{ESG}$ and P^*_{Case2b} becomes:

$$\begin{aligned} P^*_{Case2b} &= P^*\left(1 + \left(\frac{f_{ESG} + f_{SG}}{f}\right)\left(\frac{1}{\chi_{2b}}\right)^k\right)^{\frac{-1}{1-\sigma}} \\ &< P^{**} < P^* \end{aligned} \tag{E.19}$$

where $\frac{1}{\chi_{2b}} < 1$ represents the affordability of an upgrade from the use of electricity to the efficient use of fuel as described in section 4.5.3.[2] Hence, we have:

$$\varphi^*_{Case2b} > \varphi^{**} > \varphi^* \tag{E.20}$$

Aggregate productivity of energy efficient generators equals the aggregate productivity of generators and the aggregate productivity of energy efficient firms in case 2b, is given by:

$$\begin{aligned} \overline{\varphi ESG}_{Case2b} = \overline{\varphi SG}_{Case2b} = \overline{\varphi EE}_{Case2b} &= \left[\frac{1}{1 - V\left(\varphi^*_{ESG}\right)}\int_{\varphi^*_{ESG}}^{\infty} \varphi^{\sigma-1} v\left(\varphi\right) d\varphi\right]^{\frac{1}{\sigma-1}} \\ &= \chi_{2b}\overline{\varphi}_{Case2b} > \overline{\varphi}_{Case2b} \end{aligned} \tag{E.21}$$

The share of energy efficient generators is equal to the share of energy efficient firms and to the share of generators:

$$\frac{J_{ESG}}{J_A} = \frac{J_{SG}}{J_A} = \frac{J_{EE}}{J_A} = \left(\frac{\varphi^*}{\varphi^*_{ESG}}\right)^k = \left(\frac{1}{\chi_{2b}}\right)^k \tag{E.22}$$

[2]The reverse inequality defining Case 2b in equation E.3 implies that $\chi_{2b} < \chi$.

As expected, the share of energy efficient firms increases when the technologically efficient use of fuel becomes more affordable as embodied by the term χ_{2b}. The share of energy efficient firms in case 2b is larger than it would be in an economy where self-generation is not possible: the additional decrease in variable costs allowed by the use of fuel decreases the price index and the minimum level of productivity necessary to adopt the efficient use of fuel. The share of fuel in total energy consumption is given by:

$$\frac{I_{ESG}}{I} = 1 - \frac{1}{\gamma}\left(\frac{p_E}{\gamma p_{SG}}\right)^{\sigma}(\chi_{2b})^{\sigma-k-1} \tag{E.23}$$

Aggregate efficiency gains with respect to an economy where only the energy efficient technology is available are given by:

$$E_{Case2b} = \left(\frac{P^*_{Case2b}}{P^{**}}\right)^{k-\sigma+1}\left(\frac{p_E}{p_{SG}}\right)^{\sigma}(p_E - p_{SG})\,E^{**}_{ESG} \tag{E.24}$$

Cheap Fuel Equilibria (Case 3) without Power Outages

Case 3a

By proposition 26, the inequality in equation E.5 implies that $\varphi_{SG/ESG} > \varphi'_{SG}$ and means that the change of production technology from the use of fuel (self-generation) to the efficient use of fuel (energy efficient self-generation) generates a lower increase in profitability than the change from the use of electricity to the use of fuel. As depicted in figure 4.4a, if $\varphi_{SG/ESG} > \varphi'_{SG}$, there are *three types of producers* if $\varphi^*_{ESG} = \varphi_{SG/ESG}$: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{ESG}]$, ii) generators of productivity $\varphi \in [\varphi^*_{SG}, \varphi^*_{ESG}]$, and iii) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$.

On the contrary, as depicted in figure 4.4b, if the inequality in equation E.5 is not verified, then $\varphi_{SG/ESG} < \varphi'_{SG}$, there are only *two types of producers* when $\varphi^*_{ESG} = \varphi'_{ESG}$: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{ESG}]$, and ii) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$. We now derive the equilibrium prices indices and productivity cutoffs corresponding to each case. In both cases, the use of fuel is relatively cheaper than the efficient use of electricity and/or the fixed adoption costs of self-generation are only marginally superior to those of the EE technology. Therefore, self-generation of electricity dominates the efficient use of fuel at all productivity levels: no firm will adopt energy efficient technologies in the economy.

Equilibrium with Standard Firms, Generators and Energy Efficient Generators As depicted in figure 4.4a, if $\varphi_{SG/ESG} > \varphi'_{SG}$, the fixed adoption costs of the efficient use of fuel are too high for firms in the middle of the productivity distribution to afford. They adopt only the use of fuel and the double technology adoption cut-off is defined as $\varphi^*_{ESG} = \varphi_{SG/ESG}$. The equilibrium price index is $P^*_{Case3a} = [P + P_{SG} + P_{ESG}]^{\frac{1}{1-\sigma}}$ where P and P_{ESG} are defined as before, and $P_{SG} = \int\limits_{\omega \in \Omega_{SG}} p(\omega)^{1-\sigma} d\omega$ is the price index of the subset of varieties offered by generators. There is $\Omega = \Omega_S \cup \Omega_{SG} \cup \Omega_{ESG}$ and P^*_{Case3a} becomes:

$$\begin{aligned} P^*_{Case3a} &= P^* \left(1 + \left(\frac{1}{\chi_4}\right)^k \left(\frac{f_{SG}}{f}\right) + \left(\frac{1}{\chi_5}\right)^k \left(\frac{f_{ESG}}{f}\right)\right)^{\frac{1}{1-\sigma}} \\ &< P^{**} < P^* \end{aligned} \tag{E.25}$$

where $\frac{1}{\chi_4} < 1$ represents the affordability of an upgrade from the use of electricity to the efficient use of fuel as described in subsection 4.5.2 and $\frac{1}{\chi_5} < 1$ represents the affordability of an upgrade from the use of fuel to the efficient use of fuel as described in section 4.5.3. Hence, we have:

$$\varphi^*_{Case3a} \;>\; \varphi^{**} > \varphi^* \tag{E.26}$$

Average productivity of energy efficient generators $\overline{\varphi_{ESG}}$ is equal to the average productivity of energy efficient firms:

$$\overline{\varphi ESG}_{Case3a} = \overline{\varphi EE}_{Case3a} = \varphi_{SG/ESG} \left(\frac{k}{k-\sigma+1}\right)^{\frac{1}{\sigma-1}} = \chi_5 \overline{\varphi} \tag{E.27}$$

Average productivity of generators $\overline{\varphi_{SG}}$ is[3]:

$$\overline{\varphi SG}_{Case3a} = \chi_4 \overline{\varphi} < \overline{\varphi ESG}_{Case3a} \tag{E.28}$$

The share of fuel in total energy consumption is given by:

$$\frac{I_{SG} + I_{ESG}}{I} = 1 - \left(\frac{p_E}{p_{SG}}\right)^{\sigma} \left((\chi_4)^{\sigma-k-1} + \gamma^{\sigma-1} (\chi_5)^{\sigma-k-1}\right) \tag{E.29}$$

[3]Note that the inequality in equation E.5 that defines this case guarantees that $\chi_4 < \chi_5$.

The share $\frac{J_{ESG}}{J_A}$ of active energy efficient self-generators is equal to the share of energy efficient firms $\frac{J_{EE}}{J_A}$ but inferior to the share of active generators $\frac{J_{SG}}{J_A}$:

$$\frac{J_{SG}}{J_A} = \left(\frac{\varphi^*}{\varphi^*_{SG}}\right)^k = \left(\frac{1}{\chi_4}\right)^k \qquad \frac{J_{ESG}}{J_A} = \frac{J_{EE}}{J_A} = \left(\frac{\varphi^*}{\varphi^*_{ESG}}\right)^k = \left(\frac{1}{\chi_5}\right)^k \tag{E.30}$$

The share of non-energy-efficient generators is given by:

$$\frac{J_{NSG}}{J_A} = \left(\frac{1}{\chi_4}\right)^k - \left(\frac{1}{\chi_5}\right)^k > 0 \tag{E.31}$$

Equilibrium with Standard Firms and Energy Efficient Generators As depicted in figure 4.4b, if $\varphi_{SG/ESG} < \varphi'_{SG}$, the efficient use of fuel always dominates the use of fuel and all generators will be energy efficient generators. $\varphi^*_{ESG} = \varphi'_{ESG}$. The equilibrium price index is $P^*_{Case3a} = [P + P_{ESG}]^{\frac{1}{1-\sigma}}$ where the price indices of the subsets of varieties produced by standard firms and energy efficient generators are defined as before. The economy behaves in a similar way as in Case 2b described in section E. Nevertheless, the relative price of electricity with respect to fuel is higher in Case 3a than in Case 2b. Hence, we have:

$$\chi_{2b}\left(p_E^{Case2b}\right) > \chi_{2b}\left(p_E^{Case3a}\right) \tag{E.32}$$

The upgrade from the use of electricity to the efficient use of fuel is less affordable when the price of electricity is lower. It implies that the decrease in the price index and the increase in aggregate productivity caused by the introduction of self-generation in the economy are larger in Case 3a than in Case 2b. Furthermore, equation E.23 implies that the share of fuel in total energy consumption is larger in Case 3a than in Case 2b.

Case 3b

First, by proposition 27, the inequality in equation E.7 implies that $\varphi_{SG/ESG} > \varphi''$ and means that the change of production technology from the use of fuel (self-generation) to the efficient use of fuel (energy efficient self-generation) generates a lower increase in profitability than the change from the efficient use of electricity to the use of fuel. There are *four types of producers* in the differentiated good sector: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{EE}]$, ii) energy efficient firms of productivity $\varphi \in [\varphi^*_{EE}, \varphi^*_{SG}]$, iii) generators of productivity $\varphi \in [\varphi^*_{SG}, \varphi^*_{ESG}]$, and iv) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$.

Second, by proposition 27 and proposition 25, the inequalities in equations E.7 and E.3 imply that $\varphi_{SG/ESG} < \varphi''$ and $\varphi_{EE/ESG} > \varphi'_{EE}$. There are *three types of producers* in the differentiated good sector: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{EE}]$, ii) energy efficient firms of productivity $\varphi \in [\varphi^*_{EE}, \varphi^*_{ESG}]$, iii) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$. On the contrary, if $\varphi_{SG/ESG} < \varphi''$ and $\varphi_{EE/ESG} < \varphi'_{EE}$, there are *two types of producers* in the differentiated good sector: i) standard firms of productivity $\varphi \in [\varphi^*, \varphi^*_{ESG}]$, ii) energy efficient generators of productivity $\varphi \geq \varphi^*_{ESG}$.

Equilibrium with Standard Firms, Energy Efficient Firms, Generators and Energy Efficient Generators As depicted in figure 4.5, if the inequality described in equation E.7 is verified, we have $\varphi_{SG/ESG} > \varphi''$ and all four production modes are present in the economy. The equilibrium price index is $P^*_{Case3b} = [P + P_{EE} + P_{SG} + P_{ESG}]^{\frac{1}{1-\sigma}}$ where notations are identical to previous sections. There is $\Omega = \Omega_S \cup \Omega_{EE} \cup \Omega_{SG} \cup \Omega_{ESG}$. P^*_{Case3b} becomes:

$$P^*_{Case3b} = P^* \left(1 + \frac{f_{EE}}{f}\left(\frac{1}{\chi}\right)^k + \left(\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} - \gamma^{\sigma-1}\right)\chi_4^{\sigma-k-1} + \left(\frac{1}{\chi_5}\right)^k \left(\frac{f_{ESG}}{f}\right)\right)^{\frac{1}{1-\sigma}} \quad \text{(E.33)}$$

where $\chi_i,\ i \in [1,5]$ and P^* are defined as before. Hence, we have:

$$\varphi^*_{Case3b} > \varphi^{**} > \varphi^* \quad \text{(E.34)}$$

Average productivity of generators and efficient generators are respectively given by $\overline{\varphi_{SG}}$ and $\overline{\varphi_{ESG}}$ are as described in equations E.28 and E.12. The share of firms of productivity $\varphi \in [\varphi^*_{EE}, \varphi^*_{SG}] \cup [\varphi^*_{ESG}, \infty)$ that invest in energy efficient technology includes energy efficient firms and energy efficient generators but not generators. Thus, average productivity of energy efficient firms is:

$$\overline{\varphi EE}_{Case3b} = \overline{\varphi}_{Case3b}\left[\left(\frac{\chi_4 - \chi}{\chi_4^{\sigma-1} - \chi^{\sigma-1}}\right) + \chi_5\right]^{\frac{1}{1-\sigma}} \quad \text{(E.35)}$$

Equilibrium with Standard Firms, Energy Efficient Firms and Energy Efficient Generators If the inequality in equation E.7 is not verified, then $\varphi_{SG/ESG} < \varphi''$. Additionally, if the inequality in equation E.3 is verified, then $\varphi_{EE/ESG} > \varphi'_{EE}$. The economy behaves as in Case 2a described in section E. Nevertheless, the price of electricity is higher in Case 3b than in Case 2a.

Hence, we have:

$$\chi_{2a}\left(p_E^{Case2a}\right) \quad > \quad \chi_{2a}\left(p_E^{Case3b}\right) \tag{E.36}$$

The upgrade from the use of electricity to the efficient use of fuel is less affordable when the price of electricity is lower. It implies that the decrease in the price index and the increase in aggregate productivity caused by the introduction of self-generation in the economy is larger in Case 3b than in Case 2a. Furthermore, equation E.15 implies that the share of fuel in total energy consumption is larger in Case 3b than in Case 2a.

Equilibrium with Standard Firms and Energy Efficient Generators If neither the inequality in equation E.7 nor the inequality in equation E.3 are verified, then we have $\varphi_{SG/ESG} < \varphi''$ and $\varphi_{EE/ESG} < \varphi'_{EE}$. The economy behaves as in Case 2b described in section E. Nevertheless, the price of electricity is higher in Case 3a than in Case 2b. Hence, we have:

$$\chi_{2b}\left(p_E^{Case2b}\right) \quad > \quad \chi_{2b}\left(p_E^{Case3b}\right) \tag{E.37}$$

The upgrade from the use of electricity to the technologically efficient use of fuel is less affordable when the price of electricity is lower. It implies that the decrease in the price index and the increase in aggregate productivity caused by the introduction of self-generation in the economy is larger in Case 3b than in Case 2b. Furthermore, equation E.23 implies that the share of fuel in total energy consumption is larger in Case 3b than in Case 2b.

Aggregate Efficiency Gains of Energy Efficient Firms

In case 2a, the aggregate efficiency gains with respect to an economy without alternative technologies achieved by firms having adopted the energy efficient technology are given by integrating the efficiency gains of all energy efficient firms as given by equation 4.19:

$$E \quad = \quad \frac{\sigma-1}{\sigma}\frac{(\gamma-1)}{p_E}\frac{1}{J_A}\int_{\varphi^*_{EE}}^{\infty} r_{EE}\left(\varphi\right)v\left(\varphi\right)d\varphi = \frac{\sigma-1}{\sigma}\frac{R_{EE}}{p_E}\left(\gamma-1\right) \tag{E.38}$$

We exploit the properties of the model developed by Chaney (2008) and use $R_{EE} = J_{EE} r_{EE} \left(\overline{\varphi_{EE}}\right)$ and $\overline{\varphi_{EE}} = \chi\overline{\varphi}$. From equation 4.12 and $r(\varphi^*) = \sigma f$, we obtain:

$$E = \frac{J_{EE}(\chi\gamma)^{\sigma-1}}{p_E}\frac{k(\sigma-1)}{k-\sigma+1} f(\gamma-1) \tag{E.39}$$

We replace the mass of energy efficient firms in the economy $J_{EE} = \chi^{-k} J_A$ from equation 4.49 and finally use the properties of the Pareto distribution to obtain $J_{EE} = J(\varphi^*)^{-k}$. We note here that it implies that the ratio of the energy efficiency gains realized in these three cases depends only on the ratio of the associated minimum production cut-offs or, alternatively, of the associated equilibrium price indices. Using equations 4.10, 4.47 and E.11, we plug-in the minimum production cutoff corresponding to case 1, case 2a or case 3b and obtain the equilibrium aggregate efficiency gains of each case.

In case 2b, energy efficient firms are also generators. We can define the aggregate efficiency gains realized by these firms as the difference between the true variable costs and what variable costs would have been if those firms had produced the same quantity using either the standard technology, or the self-generation technology (aggregating the efficiency gains in equation 4.37), or the energy efficient technology (aggregating the efficiency gains in equation 4.38). We proceed following the same strategy as in the paragraph above, but integrate the efficiency gains from φ^*_{ESG} to ∞ instead of φ^*_{EE} to ∞.

In case 3b with all four production technologies present in the economy, we integrate energy efficient firms' efficiency gains from φ^*_{EE} to φ^*_{SG}. Using $\varphi^*_{SG} = \varphi'_{SG} = \varphi^*\chi_4$ and $\varphi^*_{EE} = \varphi'_{EE} = \varphi^*\chi$.

Proof of Proposition 13

In both cases 2a and 2b, it is straightforward to prove that $\triangle CS^* = Rln\left(\frac{P^*}{P^*_{Case2}}\right) > 0$. To simplify the notations, let's call $A = \left(\frac{1}{\chi}\right)^k \left(\frac{f_{EE}}{f}\right) > 0$, $B = \left(\frac{1}{\chi_{2a}}\right)^k \left(\frac{f_{ESG}+f_{SG}-f_{EE}}{f}\right) > 0$ and $C = \left(\frac{f_{ESG}+f_{SG}}{f}\right)\left(\frac{1}{\chi_{2b}}\right)^k > 0$ such that we have

$$P^{**} = P^*(1+A)^{\frac{-1}{\sigma-1}} \tag{E.40}$$

$$P^*_{Case2a} = P^*(1+A+B)^{\frac{-1}{\sigma-1}} \tag{E.41}$$

$$P^*_{Case2b} = P^*(1+C)^{\frac{-1}{\sigma-1}} \tag{E.42}$$

The definition of case 2a in equation E.3 implies that we have that $\chi_{2a} < \chi$. Hence, we have:

$$\frac{P^{**}}{P^{*}_{Case2a}} = \left(1 + \frac{B}{A}\right)^{\frac{1}{\sigma-1}} = \left(1 + \frac{f_{SG} + f_{ESG} - f_{EE}}{f_{EE}}\left(\frac{\chi}{\chi_{2a}}\right)\right)^{\frac{1}{\sigma-1}} > 1 \quad \text{(E.43)}$$

On the contrary, the reverse inequality defining Case 2b in equation E.3 implies that $\chi_{2b} < \chi$. Furthermore, we assumed $f_{EE} < f_{ESG}$. We have $C > A$

$$\frac{P^{**}}{P^{*}_{Case2b}} = \left(\frac{1 + C}{1 + A}\right)^{\frac{1}{\sigma-1}} = \left(\frac{1 + \left(\frac{f_{ESG}+f_{SG}}{f}\right)\left(\frac{1}{\chi_{2b}}\right)^{k}}{1 + \left(\frac{1}{\chi}\right)^{k}\left(\frac{f_{EE}}{f}\right)}\right)^{\frac{1}{\sigma-1}} > 1 \quad \text{(E.44)}$$

Proof of Proposition 15

It is straightforward to show that $\triangle CS^{*} = Rln\left(\frac{P^{*}}{P^{*}_{Case3a}}\right) > 0$. Additionally, the definition of case 3a implies that $\varphi'_{SG} < \varphi'_{EE} \Leftrightarrow f_{EE}\left(\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} - 1\right) < f_{SG}\left(\gamma^{\sigma-1} - 1\right) \Leftrightarrow \chi_4 < \chi$. Furthermore, by assumption on the technology parameters ($f_{SG} > f_{EE}$) and on the definition of case 3 ($p_E > p_{SG}$), we have $\chi_5 < \chi$. It follows that $\left(\frac{1}{\chi_4}\right)^{k}\frac{f_{SG}}{f} > \frac{f_{EE}}{f}\left(\frac{1}{\chi}\right)^{k}$, $P^{*}_{Case3a} < P^{**}$, and $\triangle CS^{**} = Rln\left(\frac{P^{**}}{P^{*}_{Case3a}}\right) > 0$.

Case 1 and Power Outages: Indirect Crowding Out Function

In case 1, energy prices verify $p_F \geq p_E$ and in the presence of power outages, the price of self-generation becomes: $p_{SG} = (1-\delta)\,p_F + \delta p_E$. First, self-generation as a production technology dominates standard production if and only if $\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} > \delta$. Second, energy efficient self-generation dominates the energy efficient technology if and only if $\left(\gamma\frac{p_E}{p_{SG}}\right)^{\sigma-1} > \gamma^{\sigma-1}\delta$. Both conditions are equivalent to:

$$p_E > \frac{(1-\delta)\,\delta^{\epsilon-1}}{1-\delta^{\epsilon}}p_F \quad \text{(E.45)}$$

with $\epsilon = \frac{\sigma}{\sigma-1} > 1$. Note that since $\delta^{\epsilon-1} < 1$, it is always verified that $\frac{(1-\delta)\delta^{\epsilon-1}}{1-\delta^{\epsilon}} < 1$. We study the function $g_{\epsilon}(\delta) = \frac{\delta^{\epsilon-1}-\delta^{\epsilon}}{1-\delta^{\epsilon}} = \frac{1-\delta}{\delta}\frac{\delta^{\epsilon}}{1-\delta^{\epsilon}}$. We find $g'_{\epsilon}(\delta) = -\left(\frac{1}{\delta}\right)\frac{\delta^{\epsilon}}{1-\delta^{\epsilon}} + \frac{1-\delta}{\delta}\left(\frac{\epsilon\delta^{\epsilon-1}}{(1-\delta^{\epsilon})^{2}}\right)$ and solving for the inequality $g'_{\epsilon}(\delta) > 0$ leads to the following inequality:

$$\left(\frac{1-\delta}{\delta}\right)\left(\frac{\epsilon}{1-\delta^{\epsilon}}\right) > 1 \quad \text{(E.46)}$$

which is always true for $0 > \delta > 1$ and $\epsilon > 1$. Finally, we have $g(0) = \frac{(1-0)0^{\epsilon-1}}{1-0} = 0$ and the limit of $g_\epsilon(\delta)$ when $\delta \to 1$ is equal to $\frac{1}{\epsilon} = \frac{\sigma-1}{\sigma} < 1$.

Power Outages and Direct Crowding Out

In case 1, energy prices verify $p_F \geq p_E$ and in the presence of power outages, the price of self-generation becomes: $p_{SG} = (1-\delta)p_F + \delta p_E$. Self-generation as a production technology dominates the energy efficient technology if and only if $\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1} > \gamma^{\sigma-1}\delta$. If the following inequality is verified,

$$\gamma < \frac{1}{\delta^\epsilon} \tag{E.47}$$

self-generation brings lower marginal costs than the energy efficient technology if and only if:

$$p_E > \frac{\gamma(1-\delta)\delta^{\epsilon-1}}{(1-\gamma\delta^\epsilon)} p_F \tag{E.48}$$

with $\epsilon = \frac{\sigma}{\sigma-1} > 1$. Note that the inequality in equation E.47 implies high levels of shortages or poor performance of the energy efficient technology in terms of revenue gains **(High shortage case)**. If it is not verified, shortages are limited or the energy efficient technology is very performant **(Low shortage case)**. In the low shortage case, self-generation brings lower marginal costs than the energy efficient technology if and only if:

$$p_E > \frac{\gamma(1-\delta)\delta^{\epsilon-1}}{(1-\gamma\delta^\epsilon)} p_F \tag{E.49}$$

We study the function $c_\epsilon(\delta) = \frac{\gamma(1-\delta)\delta^{\epsilon-1}}{(1-\gamma\delta^\epsilon)}$. First, we differentiate with respect to γ the parameter representing the performance of the energy efficient technology:

$$\frac{\partial c_\epsilon}{\partial \gamma}(\gamma,\delta) = \frac{g(\delta)(1-\delta^\epsilon)}{(1-\gamma\delta^\epsilon)^2} > 0 \tag{E.50}$$

Second, we differentiate with respect to δ the parameter negatively related to the intensity of power outages in the economy:

$$\frac{\partial c_\epsilon}{\partial \delta}(\gamma,\delta) = \gamma g(\delta)\epsilon\delta^{\epsilon-1}\frac{(\gamma-1)}{(1-\gamma\delta^\epsilon)^2} + g'(\delta)(\gamma(1-\delta^\epsilon)) > 0 \tag{E.51}$$

High Shortage Case We have $c_\epsilon\left(\gamma,\delta\right)=\frac{\gamma(1-\delta^\epsilon)}{(1-\gamma\delta^\epsilon)}g_\epsilon\left(\delta\right)>g_\epsilon\left(\delta\right)$.

Low Shortage Case It is never verified that $\left(\frac{p_E}{p_{SG}}\right)^{\sigma-1}>\gamma^{\sigma-1}\delta$: the self-generation technology never dominates the energy efficient technology.

Appendix F

Empirical Results

Measurement of Energy Efficiency: Stochastic Frontier Approach

	Electricity consumption (Log)			
	(1)	(2)	(3)	(4)
Output	0.771***	0.319***	0.370***	0.815***
	(0.003)	(0.005)	(0.005)	(0.021)
Electricity Price	0.539***	0.287***	0.364***	0.425**
	(0.034)	(0.030)	(0.030)	(0.187)
Capital/Labor Ratio	0.233***			0.245***
	(0.004)			(0.028)
Temperature min	-0.018***	-0.031***	-0.034***	0.021
	(0.003)	(0.003)	(0.003)	(0.018)
Temperature max	0.051***	0.064***	0.067***	0.020
	(0.004)	(0.004)	(0.004)	(0.027)
Capital Stock		0.343***	0.364***	
		(0.005)	(0.005)	
Total Salaries		0.257***		
		(0.006)		
Total Workers			0.218***	
			(0.007)	
ISO14000 Certification				0.196***
				(0.068)
Pollution Control				0.059***
Equipment - Stock				(0.012)
Constant	1.533***	-0.448***	1.531***	0.001
	(0.188)	(0.110)	(0.109)	(0.844)
Inefficiency term v_i				
	-1.270**	-12.863***	-12.813***	-0.903
	(0.630)	(0.096)	(0.101)	(1.353)
Idiosincratic term u_i				
	0.431***	0.283***	0.299***	0.281*
	(0.041)	(0.007)	(0.007)	(0.159)
Observations	66563	66563	66489	2176

Robust Standard errors in parentheses
* $p < 0.10$ ** $p < 0.05$ ***$p < 0.01$

Table F.1: Results based on the stochastic frontier model estimating the effect of various firm's characteristics on the level of energy consumption with year, state and industry fixed effects. Full description of the model in equation 4.90.

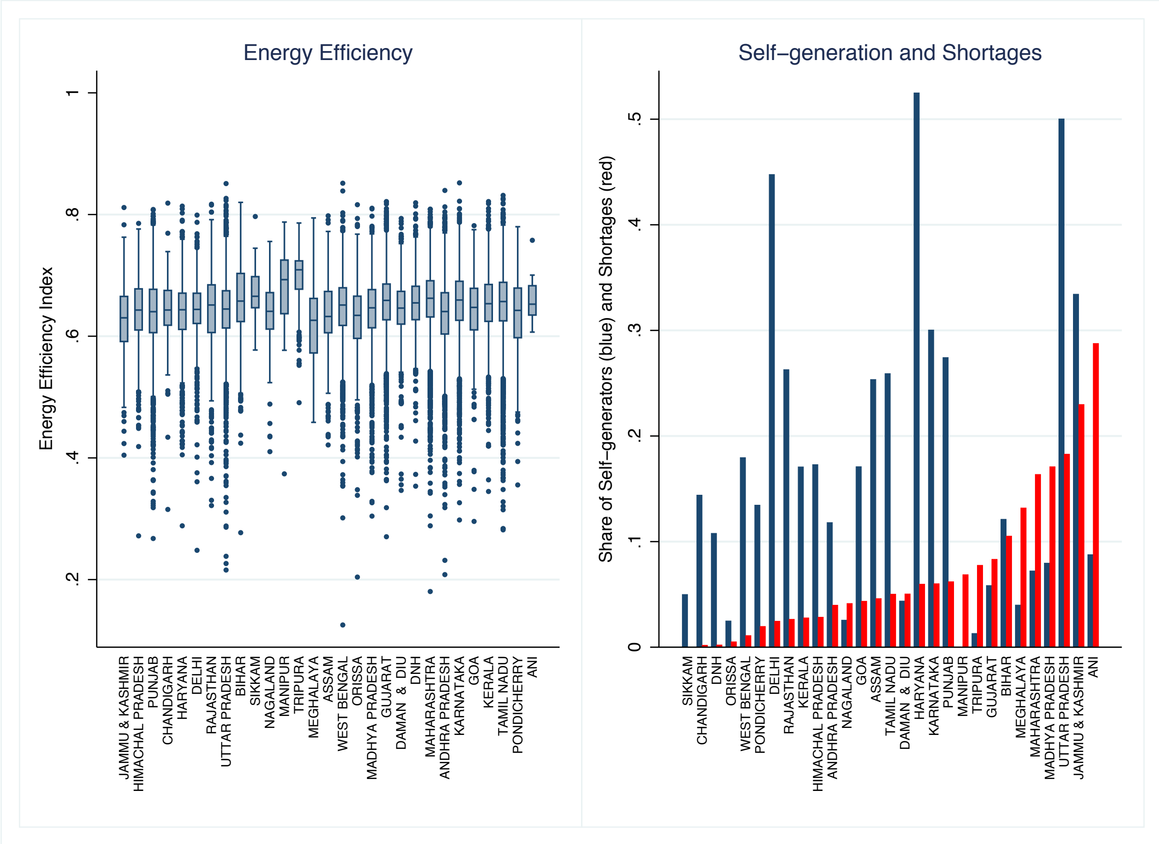


Figure F.1: **Energy efficiency distributions, share of self-generators and shortages, by state.** The blue box in the boxplot indicates the median value and the first and third quartiles. The lines above and below each box indicate the values of the distribution that are within 1.5 times the interquartile range. Data on electricity shortages by state and year are obtained from the India Energy Data Repository prepared by Hunt Allcott, Allan Collard-Wexler and Stephen D. O'Connell for their Allcott et al. (2016) study. Their online Appendix gives a thorough description of the data.

Energy Efficiency and Self-Generation: Descriptive Statistics

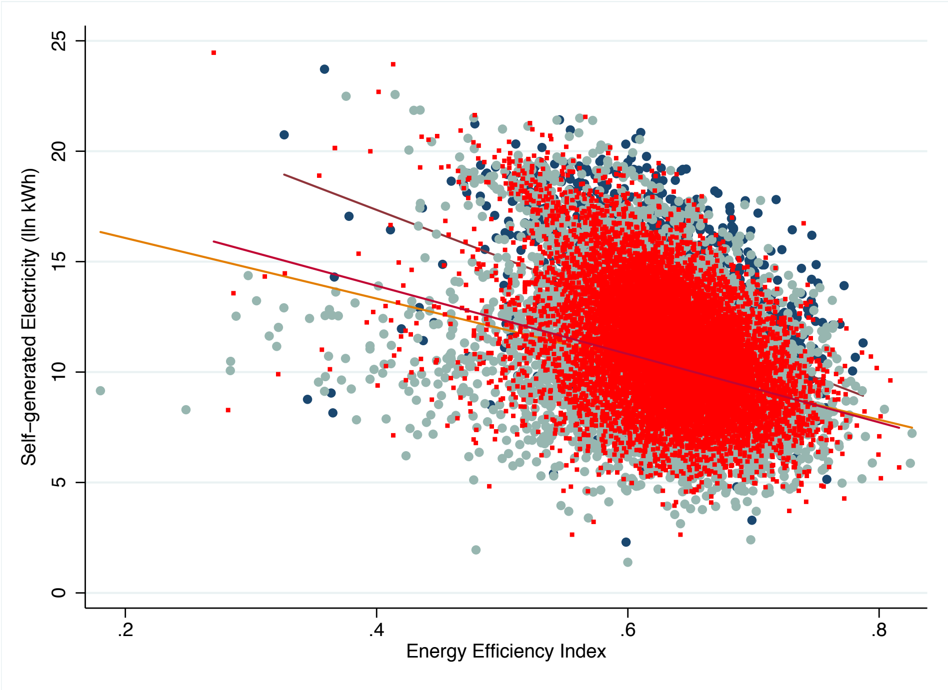


Figure F.2: **Correlation between self-generation and energy efficiency for self-generators.** Fitted line by year.

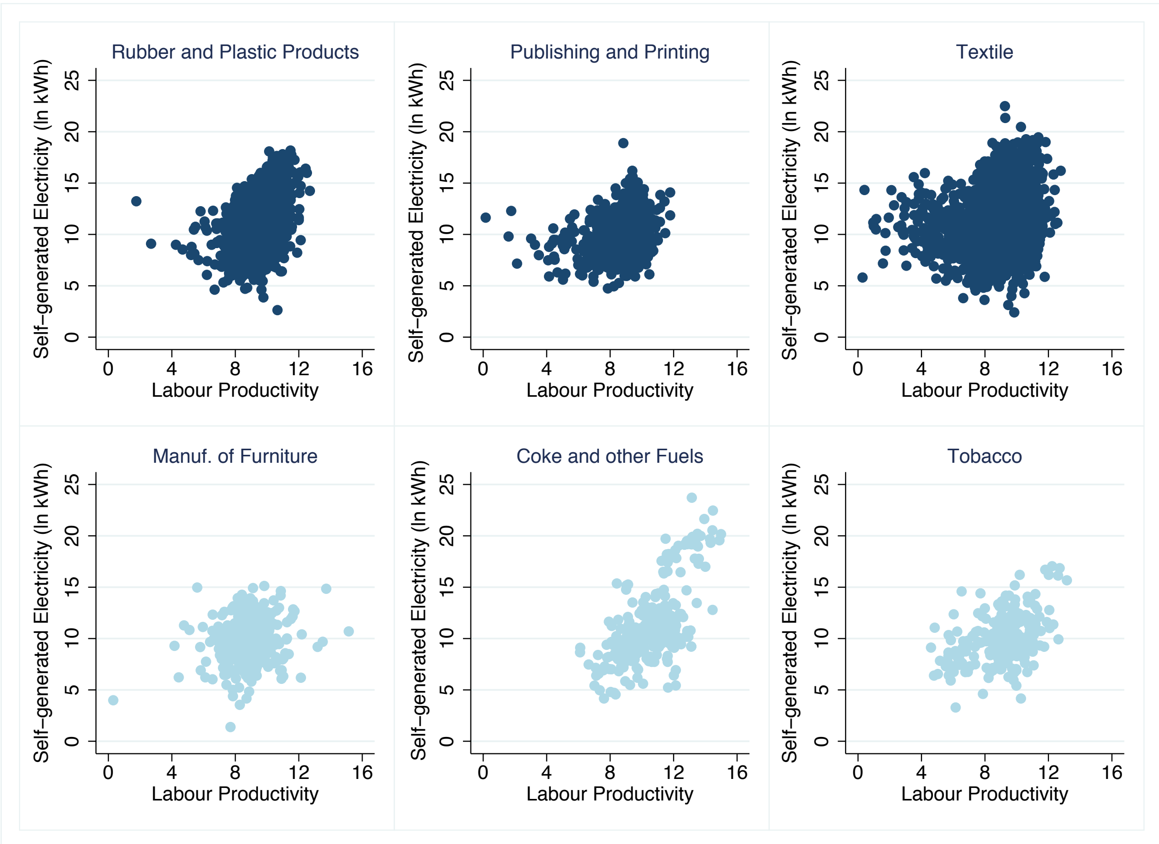


Figure F.3: **Labour productivity of self-generators** for the most electricity intensive sectors (first row) and the least electricity intensive sectors (second row).

Energy Efficiency and Self-Generation: Results

OLS Estimation Results

Table F.2: Results based on the OLS model estimating the impact of self-generation captured by the the dummy *Self-Generator (Yes/No)* equal to 1 when the firm is self-generating electricity (columns (1) to (3)) and the share of self-generated electricity in the firm's total energy consumption *Share SG* (columns (4) to (6)) on energy efficiency with state and industry fixed effects (**FE**). Full description of the model in equation 4.85.

	Energy Efficiency Index					
	(1)	(2)	(3)	(4)	(5)	(6)
Share SG	-0.021***	-0.021***	-0.021***			
	(0.002)	(0.002)	(0.002)			
Electricity Price	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Labour Productivity	0.017***	0.017***	0.017***	0.017***	0.017***	0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Capital Stock	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total Salaries	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rural (Yes/No)	0.001*	0.001	0.001	0.001*	0.001*	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Self-Generator (Yes/No)				-0.014***	-0.014***	-0.014***
				(0.001)	(0.001)	(0.001)
Constant	0.695***	0.583***	0.583***	0.687***	0.572***	0.572***
	(0.028)	(0.025)	(0.025)	(0.027)	(0.025)	(0.025)
Observations	63202	63202	63202	63202	63202	63202
R-squared	0.218	0.221	0.221	0.225	0.228	0.228
Adjusted R-squared	0.216	0.218	0.218	0.223	0.225	0.225
Time Trend	Yes	No	No	Yes	No	No
Industry Trend	Yes	No	No	Yes	No	No
Ind x Year FE	No	Yes	Yes	No	Yes	Yes
F Statistic	50.488	40.352	40.352	53.589	43.471	43.471

State, Year, Industry, Industry x Year and State x Year fixed effects included
Robust Standard errors in parentheses
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Instrumental Variable Estimation Results

Table F.3: First Stage Results for IV Regression with the Fuel Tax Instrument for the dummy *Self-Generator (Yes/No)* equal to 1 when the firm is self-generating electricity and the share of self-generated electricity in the firm's total energy consumption *Share SG* as in equation 4.87.

	Self-Generator (Yes/No)		Share SG	
	(1)	(2)	(3)	(4)
Fuel Tax Instrument	-0.004***	-0.004***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.000)	(0.000)
Electricity Price	-0.063***	-0.063***	-0.019***	-0.018***
	(0.017)	(0.017)	(0.007)	(0.007)
Age	0.014***	0.014***	0.003***	0.003***
	(0.002)	(0.002)	(0.001)	(0.001)
Labour Productivity	0.023***	0.023***	0.004***	0.004***
	(0.002)	(0.002)	(0.000)	(0.000)
Capital Stock	0.024***	0.024***	0.004***	0.004***
	(0.001)	(0.001)	(0.000)	(0.000)
Total Salaries	0.049***	0.049***	0.013***	0.013***
	(0.002)	(0.002)	(0.000)	(0.000)
Rural (Yes/No)	0.015***	0.015***	0.006***	0.006***
	(0.004)	(0.004)	(0.001)	(0.001)
Constant	-0.498***	-0.514***	-0.080	-0.103
	(0.146)	(0.173)	(0.066)	(0.075)
Observations	63202	63202	63202	63202

State, Year, Industry, Industry x Year and State x Year fixed effects included
Robust Standard errors in parentheses
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table F.4: Second Stage Results for IV Regression with the Fuel Tax Instrument for the dummy *Self-Generator (Yes/No)* equal to 1 when the firm is self-generating electricity and the share of self-generated electricity in the firm's total energy consumption *Share SG* as in equation 4.88. The weak instrument test is based on Stock and Yogo (2005): the critical value of 8.96 ensures a maximal 15 percent bias.

	Energy Efficiency Index			
	(1)	(2)	(3)	(4)
Electricity Price	-0.041***	-0.041***	-0.037***	-0.036***
	(0.005)	(0.005)	(0.003)	(0.003)
Age	-0.001	-0.001	-0.003***	-0.003***
	(0.001)	(0.001)	(0.000)	(0.000)
Labour Productivity	0.020***	0.020***	0.018***	0.018***
	(0.001)	(0.001)	(0.000)	(0.000)
Capital Stock	-0.001	-0.001	-0.003***	-0.003***
	(0.001)	(0.001)	(0.000)	(0.000)
Total Salaries	0.002	0.002	-0.001	-0.001
	(0.003)	(0.003)	(0.001)	(0.001)
Rural (Yes/No)	0.003**	0.003**	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Self-Generator (Yes/No)	-0.137**	-0.136**		
	(0.059)	(0.060)		
Share SG			-0.246***	-0.230***
			(0.087)	(0.083)
Constant	0.547***	0.556***	0.596***	0.602***
	(0.036)	(0.038)	(0.018)	(0.021)
Observations	63202	63202	63202	63202
R-squared	-0.636	-0.608	-0.123	-0.073
Industry Trend	Yes	No	Yes	No
Industry x Year	No	Yes	No	Yes
F Statistic	30.528	24.792	38.774	32.112
Weak Instrument Test	8.498	8.125	25.947	27.851

State, Year, Industry, and State x Year fixed effects included
Robust Standard errors in parentheses
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table F.5: First Stage Results for IV Regression with the VAT Reform instrument for the dummy *Self-Generator (Yes/No)* equal to 1 when the firm is self-generating electricity and the share of self-generated electricity in the firm's total energy consumption *Share SG* as in equation 4.87.

	Self-Generator (Yes/No)		Share SG	
	(1)	(2)	(3)	(4)
VAT Reform Instrument	-0.005***	-0.005***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.000)	(0.000)
Electricity Price	-0.063***	-0.063***	-0.019***	-0.018***
	(0.017)	(0.017)	(0.007)	(0.007)
Age	0.014***	0.014***	0.003***	0.003***
	(0.002)	(0.002)	(0.001)	(0.001)
Labour Productivity	0.023***	0.023***	0.004***	0.004***
	(0.002)	(0.002)	(0.000)	(0.000)
Capital Stock	0.024***	0.024***	0.004***	0.004***
	(0.001)	(0.001)	(0.000)	(0.000)
Total Salaries	0.050***	0.049***	0.013***	0.013***
	(0.002)	(0.002)	(0.000)	(0.000)
Rural (Yes/No)	0.015***	0.016***	0.006***	0.006***
	(0.004)	(0.004)	(0.001)	(0.001)
Constant	-0.538***	-0.552***	-0.102	-0.126*
	(0.145)	(0.172)	(0.065)	(0.075)
Observations	63202	63202	63202	63202

State, Year, Industry, Industry x Year and State x Year fixed effects included
Robust Standard errors in parentheses
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table F.6: Second Stage Results for IV Regression with the VAT Reform instrument for the dummy *Self-Generator (Yes/No)* equal to 1 when the firm is self-generating electricity and the share of self-generated electricity in the firm's total energy consumption *Share SG* as in equation 4.88. The weak instrument test is based on Stock and Yogo (2005): the critical value of 8.96 ensures a maximal 15 percent bias.

	Energy Efficiency Index			
	(1)	(2)	(3)	(4)
Electricity Price	-0.038***	-0.038***	-0.035***	-0.035***
	(0.004)	(0.004)	(0.003)	(0.003)
Age	-0.002***	-0.002***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.000)	(0.000)
Labour Productivity	0.019***	0.019***	0.017***	0.017***
	(0.001)	(0.001)	(0.000)	(0.000)
Capital Stock	-0.002**	-0.002**	-0.003***	-0.004***
	(0.001)	(0.001)	(0.000)	(0.000)
Total Salaries	-0.000	-0.000	-0.002**	-0.002**
	(0.002)	(0.002)	(0.001)	(0.001)
Rural (Yes/No)	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Share SG			-0.171**	-0.163**
			(0.073)	(0.071)
Self-Generator (Yes/No)	-0.084**	-0.083**		
	(0.039)	(0.039)		
Constant	0.576***	0.585***	0.604***	0.610***
	(0.024)	(0.025)	(0.015)	(0.018)
Observations	63202	63202	63202	63202
R-squared	-0.053	-0.043	0.068	0.085
Industry Trend	Yes	No	Yes	No
Industry x Year	No	Yes	No	Yes
F Statistic	40.981	33.249	43.435	35.624
Weak Instrument Test	13.095	12.861	30.438	32.118

State, Year, Industry, and State x Year fixed effects included
Robust Standard errors in parentheses
* $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Bibliography

Adam B. Jaffe, R. G. N. and Stavins, R. N. (1999). *Energy-efficient technologies and climate change policies: issues and evidence.* Climate Issues Working Paper.

Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., and Reenen, J. V. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, 124(1):1–51.

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *journal of Econometrics*, 6(1):21–37.

Aklin, M., Bayer, P., Harish, S., and Urpelainen, J. (2015). The political economy of energy access: Survey evidence from india on state intervention and public opinion. *Energy Research & Social Science*, 10:250–258.

Albrecht, J., Navarro, L., and Vroman, S. (2009). The effects of labour market policies in an economy with an informal sector. *The Economic Journal*, 119(539):1105 – 1129.

Allcott, H., Collard-Wexler, A., and O'Connell, S. D. (2016). How do electricity shortages affect industry? evidence from india. *American Economic Review*, 106(3):587–624.

Atkeson, A. and Kehoe, P. (1999). Models of energy use: Putty-putty vs. putty-clay. *American Economic Review*, 89(1028-1043).

Baccianti, C. (2013). Estimation of sectoral elasticities of substitution along the international technology frontier. *ZEW-Centre for European Economic Research Discussion Paper*, (13-092).

Baccianti, C. and Briand, O. (2016). Crowding-out of energy efficient investments by self-generation: Evidence from india. *Working Paper.*

Bagchi (1994). Reform of domestic trade taxes in india: Issues and options. Study report, National Institute of Public Finance and Policy New Delhi.

Banerjee, A. V. and Duflo, E. (2011). Poor economics – a radical rethinking of the way to fight global poverty. *Book.*

Bernard, A. B., Redding, S. J., and Schott, P. K. (2011). Multi-product firms and trade liberalization. *The Quarterly Journal of Economics*, 126(3):1271 – 1318.

Blunch, N.-H., Canagarajah, S., and Raju, D. (2001). The informal sector revisited: A synthesis across space and time. *World Bank Working Paper no. 119.*

Boadway, R. and Sato, M. (2009). Optimal tax design and enforcement with an informal sector. *American Economic Journal: Economic Policy*, 1(1):1–27.

Briand, O. (2014). Unreported labor, inequalities and taxation. *Working Paper.*

Briand, O. and Hoseini, M. (2016). Production efficiency and self-enforcement under the vat: Empirical evidence from india. *mimeo.*

Brockmeyer, A., Kleven, H., Spinnewijn, J., and Waseem, M. (2015). Production vs revenue efficiency with limited tax capacity: Theory and evidence from pakistan. *Journal of Political Economy*, 123(5):000.

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *American Economic Review*, 101(1):304–40.

CEA (2006). General review 2006 - all india electricity statistics. Technical report, Central Electricity Authority - Government of India, Ministry of Power.

Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. 98.4 (2008): 1707-1721.. *American Economic Review*, 98(4):1707–1721.

Davies, R. and Paz, L. (2011). Tariffs versus VAT in the presence of heterogenous firms and an informal sector. *International Tax and Public Finance*, 18(5):533–554.

De Paula, A. and Scheinkman, J. A. (2011). The informal sector: An equilibrium model and some empirical evidence from brazil. *Review of Income and Wealth*, 57(s1):S8–S26.

Dholakia, R. H., Agarwalla, A., Bazaz, A. B., and Agarwal, P. (2009). Trends in technical progress in india analysis of input-output tables from 1968 to 2003. *Indian Institute of Management Ahmedabad (IIMA) Working Paper*, pages 2009–11–02.

Diamond, P. A. and Mirrlees, J. A. (1971). Optimal taxation and public production i: Production efficiency. *The American Economic Review*, 61(1):8–27.

Egger, H. and Kreickemeier, U. (2012). Fairness, trade, and inequality. *Journal of International Economics*, 86(2):184–196.

Emran, M. S. and Stiglitz, J. E. (2005). On selective indirect tax reform in developing countries. *Journal of Public Economics*, 89(4):599 – 623.

Filippini, M. and Hunt, L. C. (2015). Measurement of energy efficiency based on economic foundations. *Energy Economics*, 52:S5–S16.

Filippini, M., Hunt, L. C., et al. (2011). Energy demand and energy efficiency in the oecd countries: a stochastic demand frontier approach. *Energy Journal*, 32(2):59–80.

Friedlaender, A. F. (1967). Indirect taxes and relative prices. *The Quarterly Journal of Economics*, pages 125–139.

Gerarden, T., Newell, R. G., and Stavins, R. N. (2015). Deconstructing the energy-efficiency gap: Conceptual frameworks and evidence. *American Economic Review*, 105(5):183–86.

Greenstone, M. and Hanna, R. (2011). Environmental regulations, air and water pollution, and infant mortality in india. Technical report, National Bureau of Economic Research.

Grossman, G. and Rossi-Hansberg, E. (2008). Trading tasks: a simple theory of offshoring. *American Economic Review*, 98(5):1978 – 1997.

Hanna, R. and Oliva, P. (2015). The effect of pollution on labor supply: Evidence from a natural experiment in mexico city. *Journal of Public Economics*, 122(C):68–79.

Helpman, E., Itskhoki, O., and Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica*, 78(4):1239–1283.

Henrekson, M. and Roine, J. (2005). Promoting entrepreneurship in the welfare state. *Available at SSRN 996851.*

Henri L.F. De Groot, E. T. V. and Nijkamp, P. (2001). Energy saving by firms: decision-making, barriers and policies. *Energy Economics*, 23(6):717–740.

Hoseini, M. (2015). Value-added tax and shadow economy: the role of inter-sectoral linkages. *CentER Discussion Paper Series*, pages 2015–058.

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.

IPCC (2014). Climate change 2014: Mitigation of climate change. contribution of working group iii to the fifth assessment report of the intergovernmental panel on climate change. Technical report, IPCC.

Ishii, R. K. and Stevens, G. (1998). Evaluation of self-generation for energy efficiency programs. Technical report, California Board for Energy Efficiency.

Kanbur, R. and Keen, M. (2014). Thresholds, informality, and partition of compliances. *International Tax and Public Finance*, 21(4):536–559.

Keen, M. (2008). Vat, tariffs and withholding: Border taxes and informality in developing countries. *Journal of Public Economics*, 92(10 - 11):1892 – 1906.

Keen, M. and Lockwood, B. (2010). The value added tax: Its causes and consequences. *Journal of Development Economics*, 92(2):138–151.

Kenneth Gillingham, R. G. N. and Palmer, K. (2009). Energy efficiency economics and policy. Technical Report w15031, National Bureau of Economic Research.

Kleven, H. J., Kreiner, C. T., and Saez, E. (2009). Why can modern governments tax so much? an agency model of firms as fiscal intermediaries. Technical report, National Bureau of Economic Research.

Koesler, S. and Schymura, M. (2012). Substitution elasticities in a ces production framework-an empirical analysis on the basis of non-linear least squares estimations. *ZEW-Centre for European Economic Research Discussion Paper*, (12-007).

Kopczuk and Slemrod (2006). Putting firms into optimal tax theory. *American Economic Review Papers and Proceedings*, 92(2):131–134.

Kumler, T., Verhoogen, E., and Frías, J. A. (2013). Enlisting employees in improving payroll-tax compliance: Evidence from mexico. Technical report, National Bureau of Economic Research.

Linn, J. (2008). Energy prices and the adoption of energy-saving technology. *The Economic Journal*, 118(533):1986–2012.

Lucas (1978). On the size distribution of business firms. *The Bell Journal of Economics*, 9(2):508 – 523.

Majra, J., Gur, A., et al. (2009). Climate change and health: Why should india be concerned? *Indian journal of occupational and environmental medicine*, 13(1):11.

McRae, S. (2014). Infrastructure quality and the subsidy trap. *The American Economic Review*, 105(1):35–66.

Meghir, C., Narita, R., and Robin, J.-M. (2012). Wages and informality in developing countries. Technical report, National Bureau of Economic Research.

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695 – 1725.

Mukhopadhyay, S. (2005). Vat in haryana. *Economic and Political Weekly*, 40(8):731–732.

Nag, T. (2010). Captive generation in india. the dilemma of dualism. *India infrastructure report*, pages 198–207.

NIPFP (1994). *Reform of domestic trade taxes in India: issues and options : report of a study team.* National Institute of Public Finance and Policy.

Nussbaumer, P., Bazilian, M., and Modi, V. (2012). Measuring energy poverty: Focusing on what matters. *Renewable and Sustainable Energy Reviews*, 16(1):231–243.

Nyland, C., Smyth, R., and Zhu, C. (2006). What determines the extent to which employers will comply with their social security obligations? evidence from chinese firm-level data. *Social Policy and Administration*, 40(2):196 – 214.

Pavcnik, N., Blom, A., Goldberg, P., and Schady, N. (2004). Trade liberalization and industry wage structure: Evidence from brazil. *The World Bank Economic Review*, 18(3):319–344.

Paz, L. et al. (2012). The impacts of trade liberalization on informal labor markets: an evaluation of the brazilian case. *MPRA paper*, 38858.

Perry, G., Arias, O., Fajnzylber, P., Maloney, W., Mason, A., and Saavedra, J. (2007). Informality: Exit and exclusion. *Washington DC: World Bank.*

Piggott, J. and Whalley, J. (2001). VAT base broadening, self supply, and the informal sector. *American Economic Review*, 91(4):1084–1094.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105(8):2539–69.

Popp, D. (2002). Induced innovation and energy prices. *American Economic Review*, 92(1):160–180.

Prindle, B., Eldridge, M., Eckhardt, M., and Frederick, A. (2007). The twin pillars of sustainable energy: synergies between energy efficiency and renewable energy technology and policy. *Washington, DC: American Council for an Energy-Efficient Economy.*

Purohit, M. (1975). Sales tax exemptions in india. *Economic and Political Weekly*, 10(10):445–447.

Purohit, M. (1982). Structure of sales tax in india. *Economic and Political Weekly*, 17(34):1365–1375.

Purohit, M. (1993). Adoption of value added tax in india: Problems and prospects. *Economic and Political Weekly*, 28(10):393–397.

Rao, M. G. (2003). Reform in the central sales tax in the context of vat. *Economic and Political Weekly*, 38(7):627–636.

Rasmussen, N. (1957). Studies in inter-sectoral relations. *Revue Economique*, 8(6):1103–1104.

Rauch, J. E. (1991). Modelling the informal sector formally. *Journal of development Economics*, 35(1):33–47.

Ring, R. (1999). Consumers' share and producers' share of the general sales tax. *National Tax Journal*, 52(1):79–90.

Rud, J. P. (2012). Infrastructure regulation and reallocations within industry: Theory and evidence from indian firms. *Journal of Development Economics*, 99(1):116–127.

Sagar, A. D. (2005). Alleviating energy poverty for the world's poor. *Energy Policy*, 33(11):1367–1372.

Shapiro and Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3):443 – 444.

Slemrod, J. (1996). Which is the simplest tax system of them all? *Economic effects of fundamental tax reform*, 355:355–91.

Smart, M. and Bird, R. M. (2009). The impact on investment of replacing a retail sales tax with a value-added tax: Evidence from canadian experience. *National Tax Journal*, pages 591–609.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*.

Tonin, M. (2011). Minimum wage and tax evasion: Theory and evidence. *Journal of Public Economics*, 95(11-12):1635 – 1651.

Ulyssea, G. (2013). Firms, informality and welfare. *Institute for Applied Economic Research (IPEA) Working Paper*.

Urpelainen, J. (2016). Energy poverty and perceptions of solar power in marginalized communities: Survey evidence from uttar pradesh, india. *Renewable Energy*, 85:534–539.

Van Brederode, R. F. (2009). *Systems of general sales taxation: theory, policy and practice*. Number 33. Kluwer Law International.

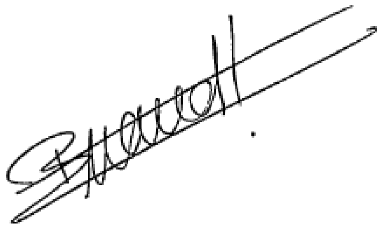
Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbstständig angefertigt und mich keiner anderen als den in ihr angegebenen Quellen und Hilfsmitteln bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Hiermit erkläre ich, dass die Universität meine Dissertation zum Zwecke des Plagiatsablgleichs in elektronischer Form speichert, an Dritte versendet, und Dritte die Dissertation zu diesem Zwecke verarbeiten.

Mannheim, September 2016

Océane Briand



Curriculum Vitae

2010 - present	**University of Mannheim,** **Graduate School of Economics and Social Sciences**, Germany Doctoral Studies in Economics
2009 - 2010	**University of Cergy-Pontoise**, France Master of Research in Economics and Finance
2006 - 2010	**ESSEC Business School**, France *Grandes Écoles Programme*: M.Sc. equivalent Master of Science in Management
2004 - 2006	**Preparatory classes of Lycée Kléber**, France Preparation to the competitive entrance examination to the top French Business Schools
2000 - 2004	**Secondary Education in Lycée Jean Mermoz**, France *Baccalauréat*: French High School Diploma