Dynamic labor demand and informality

Armela Mancellari *

7 October, 2023

Abstract

I explore the implications of firms incentives to hire informal labor for resource allocation and anti-informality policies. I build and estimate a dynamic structural model in which firms employ informal labor to evade payroll taxes and to avoid costs of hiring and firing formal workers. First, I show that gains in allocative efficiency accruing to better enforcement are far more modest when firms use informal labor to adjust to shocks. Second, failing to account for informal labor results in an overstatement of formal labor adjustment costs. Third, reducing formal labor market rigidities is as effective as enhanced enforcement in reducing informality.

Keywords: informal employment, firms, labor adjustment costs, misallocation
JEL codes: E20, J46, C61, L11

*Office of the Comptroller of the Currency. Email: armela.mancellari@occ.treas.gov. The opinions expressed here are of the author alone and do not necessarily reflect the views of Office of the Comptroller of the Currency or the U.S. Department of the Treasury. I thank my Ph.D. advisor Stephen Yeaple, my dissertation committee members Jonathan Eaton, James Tybout, and David Abler, colleagues and faculty at the Penn State Department of Economics, participants in the 2021 SED conference, Paul Ko, Nathan Converse, Ahmet Mancellari, Sebastian Infante, Alessandra Peter, and Jay Dennis. I also thank Russell Cooper, Gualiang Hu, and Gee Young Oh for sharing their code. The microdata was made available by the National Statistical Institute of Albania (INSTAT). All errors are my own.

1 Introduction

Informal employment is pervasive in developing countries, as well as in some developed economies. This type of working relationship, where workers have no contracts or social security benefits and pay no taxes, can involve up to three fourths of the labor force.[1] Informal employment is not just a small-firm phenomenon; the practice is rife across the firm size distribution. For example, in Brazil, firms with more than 11 employees hire about half of all informal workers.[2] In Albania, about 10 percent of all manufacturing workers in larger firms are informal.[3] Despite the prevalence of informal hiring in larger firms, nearly all of the existing literature models informal employment in a manner that implies it will be used only by small firms.[4] In this paper, I account for this feature of the data by building a rich model in which formal firms hire informal workers not only to evade payroll taxes, for example, as in Ulyssea (2018), but also to avoid formal labor adjustment costs that stem in part from labor market rigidities. I then employ a novel strategy to infer information about firms' use of informal labor, which allows me to estimate the model and to analyze the allocation of resources within and across firms.

The analytical approach that I adopt overcomes two key challenges that face both policy-makers who must address informal employment and researchers working to understand the phenomenon. The first challenge is assessing the effects of policies toward informal employment when *formal* firms face not only static incentives, as previously studied in the literature, but also dynamic incentives to employ informal workers.[5] When hiring workers informally serves as an adjustment margin, allowing firms to meet demand fluctuations, rigid formal labor markets, coupled with strict enforcement of laws prohibiting informal employment, in effect "punish" firms that receive large shocks. I build a unified structural model that incorporates both types of incentives and thus provides new insights into the trade-offs created by government policies. The second challenge stems from the fact that formal firm-level employment data is systematically distorted in an environment with informal employment. Thus, in a dynamic environment where firms use informal employment to buffer demand

[1] According to Ulyssea (2018) and the references therein, in countries like Brazil, Peru, and Mexico, informal workers are about 35-65 percent of the labor force. In Paraguay more than 70 percent of workers are informal. Hazans (2011) estimates an average of 25 percent of the labor force as being informal in Southern Europe and 10 percent in Northern Europe.

[2] Ulyssea (2018), based on the Monthly Employment Survey in Brazil.

[3] Author's calculations for firms of more than 10 employees, based on data from the Labor Force Survey, 2010-2014.

[4] Ulyssea (2018) is a notable exception.

[5] In this paper, I focus on the effect of policies on firms, but other notable effects include workers' potential welfare costs of not having legal protection when hired informally.

shocks, taking employment data at face value will result in a systematic mismeasurement of the adjustment costs associated with formal labor. To avoid this, I exploit a policy shock to draw inferences about firms' use of informal labor. I then use that information to inform my strategy for estimating the structural model.

I use my framework to answer three questions: First, what is the effect of imperfect enforcement of regulations on the efficiency of labor allocation across heterogeneous firms? I provide new insights on this question with a dynamic model of the intensive margin of informality, in which all firms are formal, i.e. registered with the tax authority, but some decide to hire part of their workforce informally.[6] The dynamic dimension of informal employment is key in generating a policy trade-off that is new in the literature. Recent research shows that imperfect enforcement of anti-informality regulations creates inefficiences in labor allocation by helping less productive firms stay in operation. For example, Ulyssea (2018) shows that when enforcement is more stringent on larger formal firms, the lower productivity ones are affected the most when enforcement is tightened, since they hire more informal workers. However, in this paper I show that when firms face formal labor adjustment costs, in part due to labor market rigidities, imperfect enforcement works in the opposite direction, allowing either contracting or expanding firms to reach their desired sizes faster. This mechanism is another way that firms use informal labor to bypass the inefficiencies created by regulations.[7]

Second, how much are formal labor adjustment costs mismeasured if one ignores informal workers at the firm level? In countries with high informal employment in large firms, even detailed firm-level data generally does not reflect the full resources of the firm. Taking this data at face value would lead one to infer large adjustment costs whenever one observed high inaction rates among firms adjusting their labor. However, with my proposed mechanism in place, observed sluggish changes in formal labor are also a result of firms adjusting their informal labor. Using a sample of firms in Albania, I find that naively assuming firms do not hire informal workers results in estimated adjustment costs that are too high by at least a factor of two.

Third, what is the most effective policy intervention to reduce the aggregate share of informal

[6] I focus on the intensive margin of informal employment for four reasons: i) the mechanism of informal labor serving as a margin of adjustment to avoid formal labor costs only applies to firms that hire both types of workers, ii) the extensive margin of informality, in which firms are either entirely formal or entirely informal, has been well researched in the literature, iii) in Albania, the country I apply the model to, 93 percent of firms appear registered, and hence are formal, in the classic definition, and iv) there are no surveys of informal firms in Albania.

[7] This argument is not new in the literature on informality. Maloney (2004) argues that the informal sector is largely voluntary in developing countries and should be thought of as "the developing country analogue of the voluntary entrepreneurial small firm sector found in advanced countries."

employment: Increasing firms' costs of employing informal labor through greater government enforcement, or reducing the benefits of using informal labor through policies to lessen labor market rigidities? I find that the second type of policy intervention would be just as effective in lowering the informal share of employment.

I address these questions using a structural model that I estimate with detailed firm-level data from Albania, a small developing economy. As noted earlier, pervasive informality creates distorted data since firms do not report informal workers in official data. More specifically, we do not observe firms across the size distribution adjust their formal and informal labor with some regular frequency.[8] To deal with this issue I exploit a 2015 change in policy that forced Albanian firms to report their use of labor truthfully, which allows me to draw inferences about their prior use of informal labor. In September 2015, the government of Albania launched an anti-informality effort that targeted firms of all sectors and sizes.[9] The effort included a public awareness campaign, waiving fines for firms that became compliant by the end of the year, increasing the number of fiscal inspections, and imposing higher fines for non-compliance.[10] Dramatic increases in reported sales and labor suggest that the campaign, which continued through 2017, was effective in the formalization of informal employees.[11]

I use patterns of changes in reported sales and labor around the shock to infer which firms were most likely misreporting their activities. I classify manufacturing firms into four categories: (i) likely liars about both labor and sales, (ii) likely truth tellers about labor, but liars about sales, (iii) likely truth tellers about labor and sales, and (iv) likely liars about labor, but truth tellers about sales. My identification strategy relies on the fact that sales and labor should move in the same direction following a shock. As an example, if I observe a negative change in sales at the end of 2015 but a positive change in formal labor, the firm was most likely underreporting labor in 2014.

Such a classification would be accurate only if there was no other type of measurement error, if all misreporting was underreporting, and all misreporting firms were affected by the 2015

[8] Researchers try get around this issue by either matching aggregate statistics on informal employment using Labor Force Surveys, or by using unique datasets such as data on firm inspections and formalization of informal workers. de la Parra (de la Parra) has access to such inspection data for Mexico, but that is quite unusual. Aggregate informality statistics would not be adequate in our case, since adjustment costs are estimated from firm-level data.

[9] Detailed information on the campaign can be found at Albania's Economic Reform Programme (ERF) 2016-2018, MoF (2018).

[10] European Bank for Reconstruction and Development (EBRD) (2019)., Tabak and Borkovic (2019)

[11] According to ERF 2016-2018, the two tax categories that over performed in 2015 were VAT and payroll taxes.

anti-informality campaign. I therefore relax these assumptions by allowing for misclassification. In particular, I use a multinomial logit framework with misclassification in the spirit of Hausman et al. (1998) to estimate the probability that each firm type is misclassified conditional on observable firm characteristics including foreign ownership, whether it exports, and the capital intensity of the sector of operation.

Equipped with this methodology to exploit an identifying shock, I build and estimate a dynamic structural model of firms that are heterogeneous in their profitability and their type. A fixed mass of firms randomly draw one-time cost shocks associated with being each type, before they enter the market. Having no other information, each firm compares the expected lifetime profits of each type and chooses a type. Firms then draw a persistent revenue profitability shock each period and decide how much to produce using a decreasing returns to scale technology in capital, intermediates, and the number of workers. Formal and informal workers are perfect substitutes in production, but informal workers have no contracts or benefits. Wages are fixed, with formal workers being paid more than informal workers after taxes.[12] Formal labor is subject to quadratic adjustment costs, while informal labor is increasingly costly due to government enforcement of labor laws. A firm that chooses type (iv), hiring some workers informally but accurately reporting sales, faces the trade-off of evading payroll taxes and avoiding adjustment costs when hit by a shock or incurring penalties. In a deterministic steady state, the incentive for hiring informal workers is to avoid payroll taxes, as in Ulyssea (2018), and pay lower wages. In a stochastic steady state with firm heterogeneity, however, firms use informal workers to avoid adjustment costs as well.[13] Firms' informal employment increases in the short term in response to a positive shock and slowly declines to its deterministic steady state value as their formal labor stock is adjusted. When hit by a negative shock, firms face the trade-off between firing informal workers without incurring adjustment costs and paying higher wages to formal workers they hold on to. On the sales side, firms that choose to hide their sales do so for VAT avoidance only.[14]

I estimate the structural model via simulated method of moments (SMM) and match mo-

[12] This choice is motivated by the fact that informal workers in larger formal firms get paid less in Albania, despite having similar observable characteristics to formal workers, as shown in Section 2 of this paper.

[13] Using informal labor to meet increased demand was the most common reason for informal employment revealed to me by managers of a few manufacturing firms in Albania during informal interviews I conducted in October 2017. Other reasons included getting around the mandatory pension age and the mandatory minimum working age, and giving double shifts without overtime pay.

[14] This aspect of informality is not the focus of this paper but is necessary to model due to its pervasiveness in Albania.

ments from manufacturing firms in Albania during the period 2011-2013.[15] I weight each firm's contribution in the calculation of moments for each type by the probability of being classified as that type, which I previously extracted from the data. The key parameters to identify are the shock process, the adjustment cost parameter, the discount factor, the cost of informal labor, the cost of underreporting sales, and the parameters governing the distribution of costs related to being a specific type. Moments of type (iii) firms, the truth tellers, identify the shock process, the adjustment cost parameter, and the discount rate. More specifically, they include the serial correlation in sales and labor, covariance of labor and sales, formal labor turnover rate, and selected percentile cutoffs of the size distribution in terms of employment. Moments calculated for type (iv) firms, liars about labor but not sales, as well as the economy-wide share of informal workers, identify the informal labor cost parameters. They include selected percentile cutoffs of the firm size distribution in terms of employment. Moments for type (ii) firms, liars about sales but not labor, identify the parameter on the cost of hiding sales. Moments include selected percentile cutoffs of the firm size distribution in terms of sales. Lastly, the overall shares of types in the economy identify the parameters associated with the initial draw of costs of each type. I use the estimated model to obtain counterfactual labor allocations under two different policy interventions: i) an increase in the cost of hiring informal workers, and ii) a reduction the cost of adjusting formal workers.

Four key results emerge from the quantitative model. The first two results show that mismeasurement issues can be quite serious in a developing economy setting. The second two show that policies that aim at lessening adjustment costs might be more desirable than those that increase the cost of informality. First, I find that formal labor adjustment costs in Albania are relatively high. I estimate adjustment costs to be about 17 percent of revenues, or 20 percent of the annual compensation of a formal worker.[16] This suggests that the flexibility of informal labor markets plays an important role in firms' decisions to hire formal and informal workers. My estimated costs are different from earlier studies on manufacturing plants in the U.S. and China. Cooper and Willis (2009) find adjustment costs to be about three percent of revenues for the U.S. and Cooper et al. (2015) find adjustment costs to range from 160 percent of average worker compensation in state-owned plants, to 17 percent in private plants in China.[17] Second, I show that taking the data at face value and assuming that

[15] The government of Albania increased the penalties for hiring informal labor about 5-fold at the end of 2013. I exclude 2014 from the estimation because that year is clearly not a steady state of the economy.

[16] To calculate these objects, I simulate 10,000 firms for 103 periods and use the last period to compute total adjustment costs over total revenues, and total adjustment costs over total compensation of formal workers.

[17] While informative, the estimates in the literature are comparable to my estimate only up to a degree.

all firms truthfully report their labor and sales produces parameter estimates that seriously overstate the size of the adjustment costs. In that exercise, I estimate the adjustment costs to be more than double the estimate I obtain when carefully accounting for truth tellers and evaders.

Third, making informal employment prohibitively costly increases the dispersion in sales per worker by about 20 percent, which suggests that allocative efficiency is lower under perfect enforcement.[18] I also find that the mean, median and variance of firm sizes in the baseline economy, where some firms hire informal workers, are higher than in an economy without informal employment. Fourth, the government can achieve a similar reduction in the share of informal employment whether it lowers adjustment costs by 75 percent or increases the cost of hiring informal workers by 50 percent. The former policy, however, comes with the added benefit of lessening the regulation-induced formal labor market distortions in the economy.

This paper is related to several literatures. It contributes to the theoretical and applied literature on informality, firms, and development, with important work by: La Porta and Shleifer (2014), who give a complete characterization of formal and informal sector firms in developing countries; Meghir et al. (2015), who develop an equilibrium wage posting model that generates overlapping formal and informal sector firms in terms of size; Haanwinckel and Soares (2020), who model the interaction of firm and worker heterogeneity with labor regulations in a search model of informal labor markets; Dix-Carneiro et al. (2019), who study the interaction of labor market frictions, trade shocks, and informality; Ulyssea (2018), who show the importance of the intensive margin of informality in assessing policies that target the share of informal workers in the economy; Alvarez and Ruane (2019), who assess Total Factor Productivity (TFP) effects of policies aimed at reducing informal employment when firms face idiosyncratic distortions; and others.[19]In my study, I expand the literature

First, my model is a model of firms, not plants. Second, I do not model non-convex adjustment costs, since non-convex costs are less binding when the data is annual. Third, I do not observe labor hours in the data, so I use the number of employees in the production function instead. If adjusting hours is less costly than adjusting workers, firms would use that margin the most, so the adjustment costs on the number of employees would appear higher. Anecdotal evidence from interviews with firm managers in Albania suggests that most of the adjustment on the hours dimension happens on the informal margin. Firms seem to either hire informal workers on a per-need basis or use current formal workers to work extra shifts against undeclared payments.

[18] In an environment with adjustment costs and idiosyncratic shocks, dispersion in sales per worker arises naturally, and does not necessarily imply misallocation (see. Asker et al. (2014)). However, when adjustment costs reflect not only the technology of the firm, but also government regulations, the dispersion in sales per worker does imply a degree of misallocation. Due to labor market rigidities, productive firms (those that receive positive shocks) cannot reach their optimal size immediately, while unproductive firms (those that receive negative shocks) have to hold on to workers they don't need.

[19] There is also a rich literature that studies the linkages between the formal and informal sectors as well as the response of the informal sector to business cycle fluctuations. An excellent survey of the literature

on the intensive margin of informality, as developed in Ulyssea (2018), by adding dynamic labor demand. The additional mechanism implies that when rigid labor markets make the adjustment of formal labor more costly, informal employment actually improves allocative efficiency, allowing firms to reach their optimal size faster. The mechanism in my model is similar in spirit to the one developed in Goldberg and Pavcnik (2003) in which the "de facto" adjustment costs for formal workers are microfounded in the need to pay them efficiency wages to keep them from shirking. Using reduced-form evidence, they find the mechanism to be significant in Colombia, where labor markets were very rigid before the trade reform, but not in Brazil. My approach is more macroeconomic, and I go a step further than Goldberg and Pavcnik (2003) and estimate a structural model to understand the implications of a regime change.

This paper also contributes to the literature on misallocation of resources with seminal papers by: Restuccia and Rogerson (2007) who show that differences in the allocation of resources across heterogeneous firms are an important factor in accounting for cross-country differences in output per capita; Hsieh and Klenow (2009) who find large effects of misallocation on China's and India's TFP; David and Venkateswaran (2019) who show that idiosyncratic policy distortions explain most of the cross-section variation in marginal revenue products of capital, rather than capital adjustment costs; Asker et al. (2014) who show that more variable transitory firm-level shocks in developing countries, coupled with adjustment costs, explain most of the cross-country differences in TFP; and others.[20] My main departure from the literature is in the definition of labor adjustment costs, which include not only hiring and firing costs related to search and matching frictions but also costs related to the rigidity of the labor market. For example, inflexible contractual arrangements for formal workers are more costly in the face of transitory shocks, and access to a substitute input that alleviates the inflexiblity improves the efficiency of the allocation of total labor in the economy.

In addition, this paper is related to the literature on the dynamics of labor demand in firms and plants, with seminal work by Cooper and Willis (2004), Cooper et al. (2004), and Cooper et al. (2015). My main contribution is to estimate labor adjustment costs in a setting where firms hire employees informally. I show that analyses ignoring informal employment will incorrectly infer large adjustment costs whenever observed changes in reported labor are

can be found in Box 3.1 of the World Bank's "Global Economic Prospects, January 2019".WB (2019). My departure from that literature is that I study formal firms, some of which choose to directly employ informal workers, and how these firms respond to idiosyncratic fluctuations, rather than aggregate business cycle fluctuations.

[20] For more work on misallocation, see Restuccia and Rogerson (2017), who provide an excellent survey of the literature.

sluggish. Lastly, my paper is related to the literature on tax evasion in public finance, with contributions by Kleven et al. (2011), Best et al. (2015), Carrillo et al. (2017), and many more. I contribute to that literature by modeling another incentive for misreporting input costs.

I describe the data, the anti-informality campaign, and my classification of firms into types in Section 2. In Sections 3 and 4, I present the model and the structural estimation procedure. Section 5 depicts the counterfactual analyses and Section 6 concludes the paper.

2 Firms' use of informal labor: Data and inference

This section has three goals. First, I introduce the empirical setting for my analysis by describing firm and worker characteristics, some institutional features of the labor market, and the tax framework in Albania. In particular, I provide evidence that formal firms across the size distribution hire workers informally and that Albania has relatively rigid labor markets. I also show that formal and informal workers have similar observed characteristics but that informal workers earn less than formal workers.

Second, I document the aggregate effects of the 2015 policy shock and describe the strategy I use to infer which firms likely underreport their labor and/or sales. I then classify manufacturing firms into four types and report characteristics for each type: (i) likely liars about both labor and sales, (ii) likely truth tellers about labor, but liars about sales, (iii) likely truth tellers about labor and sales, and (iv) likely liars about labor, but truth tellers about sales. I show that this initial classification is compatible with expected type features. For example, type (iii), truth tellers, have a higher share of foreign firms and exporters. Moreover, I show that truth tellers have a lower average inaction rate when adjusting the number of workers compared to other firm types, as well as higher volatility of reported labor, consistent with truth tellers being less likely to use informal labor.

Third, I relax the assumptions used in the classification process and allow for misclassification of firms into types. I estimate the probability that a firm is correctly classified or misclassified with a multinomial logit framework with misclassification in the spirit of Hausmann et al. (1998). I use the resulting probabilities when I estimate the structural model in Section 4, as they allow me to correctly weigh each firm's contribution to the moments for each type.

2.1 Data and definitions

Following standard practice in the literature, I define informal workers as those workers that are employed by a firm but have no contract, which means their firm pays neither payroll taxes for them nor social security contributions on their behalf.[21] All firms in my theoretical and empirical analyses are formal in the sense that they are registered with the tax authority.[22] However, some firms may hire part of their workforce informally, some may hide part of their sales, and others may do both.

I use three datasets in my analysis: the Structural Business Survey (SBS) collected by the Albanian Statistical Institute (INSTAT), the Labor Force Survey (LFS) also collected by INSTAT, and customs records data made available by INSTAT. The SBS is an annual survey that collects data on the universe of active firms with more than nine employees and a representative sample of firms with less than 10 employees.[23] The LFS covers about one percent of all households in Albania.[24] Most of the analysis covers the period 2011-2015, which is a relatively stable time in terms of policy changes. I exploit for identification the biggest policy shock during the period, the shock to reporting in late 2015.

I also use information from informal interviews with managers of manufacturing firms in Albania. In October 2017, I conducted interviews with general managers of seven firms in the apparel, furniture, and foods sectors in Albania, with the goal of gaining some insights into their operations and the nature of informality in manufacturing. Those interviews revealed that one of the main reasons firms employ workers informally is to meet unexpected increases in demand.[25]

[21] I exclude self-employed persons from the theoretical and empirical analysis because the focus of my paper is on larger firms.

[22] According to the World Bank Enterprise Survey Indicators 2019, 93 percent of firms surveyed were formally registered when they started their operations. Moreover, the same data shows that firms that started their operations while informal registered after 0.3 years on average, which is in line with the European average of 0.2 years. The World Bank Enterprise Survey Indicators 2013 show an even lower share of unregistered firms, at 97 percent of firms.

[23] The survey follows Eurostat standards and is of relatively high quality. A metadata analysis of 2015 shows an overall coverage rate of about 80 percent. http://www.instat.gov.al/media/4890/sbs-esms-2015.pdf

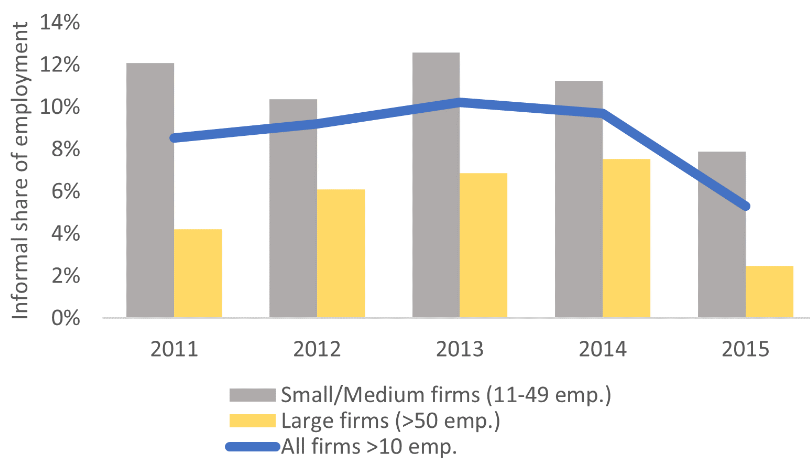
[24] Since 2012, the survey has been conducted quarterly, and a household is kept in the sample for five consecutive quarters. Before 2012, the survey was conducted annually with resampling every year.

[25] Notes with redacted information on firm identifiers are available upon request.

2.2 Firms

In this subsection, I discuss key features of firms, the labor market, and taxes in Albania that motivate my modelling choices in Section 3. I focus my analysis on manufacturing firms in Albania with more than nine employees.[26][27] Figure 1 shows that informal employment is widespread among medium and large firms, adding to the growing evidence that the intensive margin of informality is empirically relevant (see Ulyssea (2018)). Furthermore, as documented by Ulyssea (2018) for Brazil and by Perry et al. (2007) for other Latin American countries, the data in Albania shows that as firms grow larger, they hire a smaller share of their workforce informally.

Figure 1: Informal share of employment by firm size



Source: Author's calculations using Labor Force Survey data. The informal share of employment is calculated by aggregating all formally and informally employed workers in manufacturing firms of more than 10 employees. I exclude workers employed by smaller firms to ensure maximum compatibility with my sample of firm data from the Structural Business Survey, which includes firms with more than 9 employees.

A major reason for larger firms hiring informal workers is to meet unexpected demand shocks. Hiring an employee formally entails costs related to the higher salaries, the taxes paid on their salaries, and the inflexibility of the contract if the demand shock is short-lived. Indeed,

[26] I focus on manufacturing because it has a large number of firms (about 23 percent of all firms in an average year between 2002-2015), hires about 30 percent of all workers, and houses about 44 percent of all foreign firms. In an average year, 11 percent of all workers in manufacturing firms of more than 10 employees are informal.

[27] Based on calculations from the LFS, taking the average across years 2007 to 2014. In the LFS, individuals report their sector of employment as well as the size bracket of their firm. There is a slight inconsistency when comparing LFS data with data from the SBS. In the LFS, the brackets are 1-10 employees, 11-49 employees, and >50 employees. Firms with exactly 10 employees are classified as micro in the LFS, and small in the SBS.

labor market regulation data from the World Bank's Doing Business Surveys show that Albania's labor market is relatively rigid in terms of hiring difficulties and redundancy costs.[28] In particular, fixed-term contracts, which allow firms to better respond to fluctuations in demand, are prohibited for permanent tasks. Furthermore, the notice period for redundancy dismissal is about 10 weeks, higher than about 95 percent of all countries listed in the Doing Business Survey in 2015.[29] Lastly, in an environment where wages are barely above the minimum wage, the downward rigidity of minimum wages increases the opportunity cost of a formal employee even more.

The fiscal costs facing larger firms in Albania include: a 20 percent Value Added Tax (VAT), a 16.7 percent payroll tax, and a 15 percent profit tax. To avoid VAT, firms in Albania hide a significant portion of their sales and/or overreport spending on intermediate inputs.[30] While underreporting sales is not the focus of my paper, I model this aspect of informality in order to be able to take the model to the data.[31]

2.3 Workers

Formal and informal workers in Albania appear very similar along several dimensions, but they earn different wages. Table 1 depicts those characteristics. Mincer regressions suggest that informal workers get paid on average 26 percent less than observationally similar formal workers after taxes, as shown in Table 2. Based on this evidence, in my theoretical analysis I assume that formal and informal workers are perfect substitutes in production but are paid different after-tax wages.[32]

[28] doi (Doing Business 2017). The World Bank.

[29] Only Belgium, Cameroon, Comoros, The Gambia, Kuwait, Slovak Republic, Sweden, and Zimbabwe had higher notice periods.

[30] The misreporting ranges from 10-50 percent of revenues, according to evidence from my interviews. Boka and Torluccio (2014) estimate the overall informal economy to range from 1.3 to 30 percent of GDP in 2011.

[31] I choose to only model the VAT and not the profit tax for two reasons. First, to lower the profit tax burden firms have the incentive to underreport value added and overreport labor costs. The payroll tax counterbalances the incentive to overreport labor, and the VAT avoidance mechanism already captures the underreporting of sales. Second, there is ample evidence across developing countries that firms tend to underreport their labor costs, either by not declaring workers, or by underreporting their salaries. Often times, the reported costs coincide with the minimum wage. See, for example, Horodnic (2016).

[32] A persistent wage gap for observationally equivalent formal and informal workers within the same firm can be explained with the microfounded model of Goldberg and Pavcnik (2003), as an example. In that model, since monitoring costs for formal workers are very high, due in part to the need to keep detailed records on their performance in case of firing them, firms pay them efficiency wages to avoid monitoring them, whereas informal workers get paid their reservation wage.

Table 1: Formal and informal workers characteristics

	Manufacturing workers	
	Formal	Informal
Individual Characteristics		
Hourly wage in log current ALL	4.72 (0.02)	4.48 (0.04)
P90/P10	1.26	1.33
P75/P25	1.11	1.12
Male share	38%	39%
Age, average	40	37
Education (share)		
Elementary (grades 1-4)	1%	1%
Junior high (grades 5-8)	42%	50%
High school (grades 9-12)	48%	44%
University and graduate school	8%	5%
Job Characteristics		
Weekly hours, average	46 (0.12)	48 (0.62)

Source: Pooled Labor Force Survey data of manufacturing workers in firms with more than 10 employees during 2011-2014. The table shows that formal and informal workers in manufacturing have similar observable characteristics. In particular, they have similar education profiles and work similar hours. Standard errors are in parentheses.

2.4 Inferring firms' use of informal labor

In this subsection I describe the anti-informality campaign carried out in late 2015 and how I use this policy shock to infer firms' use of informal labor.

The anti-informality campaign of 2015. In September 2015, the government of Albania launched an anti-informality campaign, targeting larger firms in particular, that was aimed at increasing value-added, profit, and payroll tax compliance.[33] The strategies to increase tax revenue included launching a public awareness campaign, waiving fines for firms that became compliant by the end of the year, increasing the number of fiscal inspections, and imposing higher fines for non-compliance. Figure 2 shows that the campaign had sizeable effects, especially on domestic firms, with sales and employment increasing by almost 50 percent in 2015. The contrast between foreign and domestic firms suggests that the increase in labor and sales was due to greater compliance, rather than larger profitability shocks. The sharp increase among domestic firms occurred even as aggregate manufacturing output grew only about 5 percent in 2015, down from about 10 percent in 2013 and 2014.[34]

[33] The Albania Investment Council provides a detailed timeline of the campaign, its media coverage, and short term results. https://www.investment.com.al/wp-content/uploads/2015/08/Working-Document-on-Informality-A-Common-Government-Business-Challenge-5.pdf

[34] INSTAT data.

Table 2: Mincer regression

	log hourly wage
Formal	0.22***
	(0.045)
Individual characteristics	YES
Time fixed effects	YES
Observations	1,952
R-squared	0.216

Source: Pooled Labor Force Survey data 2011-2014. The Mincer regression shows that informal workers that are observationally similar to formal workers are paid about 22 percent less than formal workers. Individual characteristics I control for include the relationship of the individual to the head of the household, education level, marital status, age, gender, and sector of operation. Time fixed effects are year dummies. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

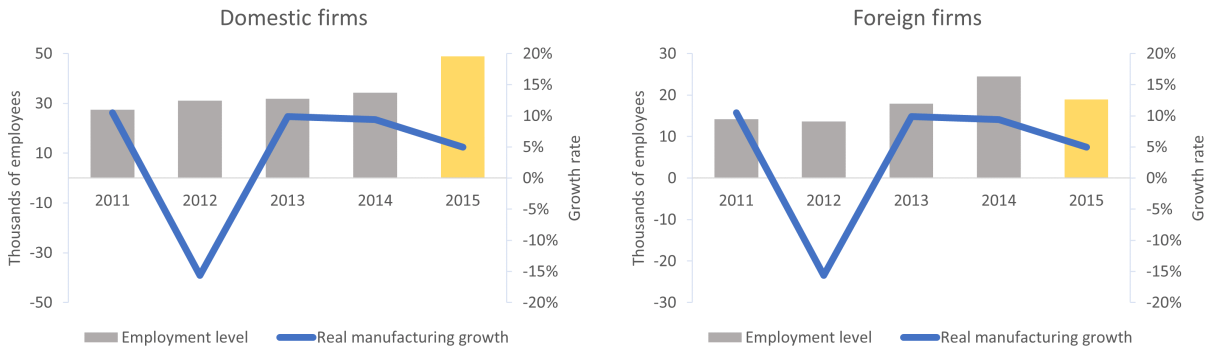
The government reported that as a result of the campaign there were 74,853 more workers registered in November 2015 than in January 2015, about 9 percent of total employment. Furthermore, according to tax administration data, VAT and payroll taxes were the only two tax categories that exceeded projections in 2015, by 2.5 percent and 2.15 percent, respectively.[35] As shown in Figure 3, labor force survey data demonstrates that the probability that informal workers transition into formal jobs reached its highest level in the last quarter of 2015.

Inferring which firms use informal workers. The main idea for the identification of firm type is the following: Whenever firms choose to misreport their sales and/or labor, their incentives are such that they underreport. At the end of 2015, under government pressure, firms revealed their actual revenues and input use. The changes in reporting in 2015 therefore reveal information about firm type. If a firm reported changes in sales and labor in 2015 of opposite signs, they were likely misreporting one or the other in 2014.[36]

[35] Albania's Economic Reform Programme (ERF) 2016-2018, Ministry of Finance and Economy (2018). https://shtetiweb.org/wp-content/uploads/2016/02/Albanias-Economic-Reform-Programme-2016-2018.pdf

[36] I assume that the decisions to misreport labor and misreport sales are independent for two main reasons: First, VAT tax compliance is checked by the General Tax Directorate, whereas compliance with labor laws and social security contributions is checked by the Labor Office. These government bodies were not required to share information with each other at the time period that covers the data. Second, incentives and the ease of undereporting need not be the same for labor and sales. As explained later in the text, certain firm characteristics, such as whether the firm exports, make it easier to misreport labor than sales, and other firm characteristics make it easier to misreport sales than labor.

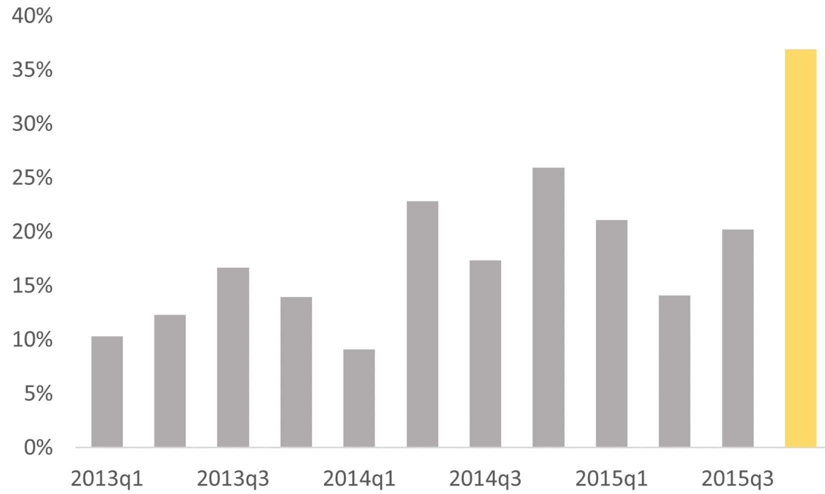
Figure 2: Changes in employment levels and growth in manufacturing



Source: Employment levels are calculated by the author using Structural Business Survey data on manufacturing firms of more than 9 employees. The real manufacturing growth rate is published by the World Bank and can be calculated using the underlying data in ¡https://data.worldbank.org/indicator/NV.IND.MANF.KD?locations=AL¿. The left panel shows domestic firms' employment level (left y-axis) increasing by about 50 percent in 2015 despite aggregate growth (right y-axis) being lower than in the previous year. The right panel shows foreign firms' employment level (left y-axis) declining in 2015, after two years of sustained growth. The contrast between domestic and foreign firms suggests that the positive change in employment in 2015 was most likely a result of the anti-informality campaign.

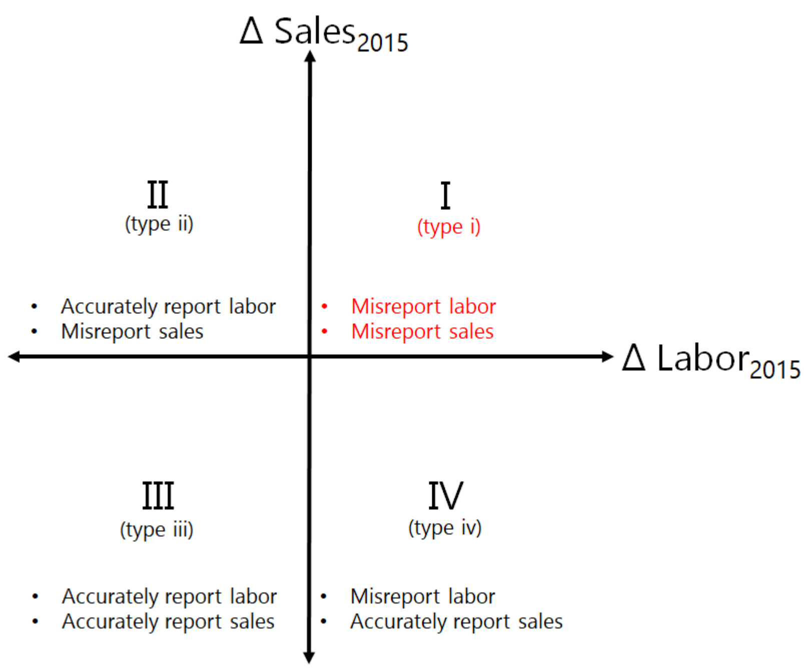
More specifically, a firm is more likely to be type (ii), a liar about sales but truth teller about labor, if it reports a positive change in sales, but a negative change in labor in 2015. The negative change in labor likely reflects an underlying negative profitability shock in 2015, which is expected to cause a negative change in sales. However, sales grew, which indicates underreporting of sales in 2014. Similarly, a firm is more likely to be type (iv), a liar about labor but truth teller about sales if it reported a positive change in labor, but a negative change in sales. Because prior to 2015 incentives were to underreport, firms that reported negative changes in both labor and sales in 2015 most likely received negative shocks. I initially classify these firms as type (iii), truth tellers about both labor and sales. Firms that reported a positive change in both labor and sales may be any of the types, with some likelihood. I initially classify all firms with positive reported changes in labor and sales to be type (i), and I later refine this classification. Plotting the changes in labor and sales in 2015, the quadrants in Figure 4 correspond to this initial classification of firms into types.

Figure 3: Informal-formal transition probability



Source: Author's calculations with data from the Labor Force Survey. The transition probability represents the share of informally employed workers that moved from informal employment in quarter t to formal employment in quarter t+1. The record high informal to formal transition probability in the last quarter of 2015 suggests that the anti-informality campaign of late 2015 resulted in substantial formalization of informally employed workers.

Figure 4: Initial firm type classification



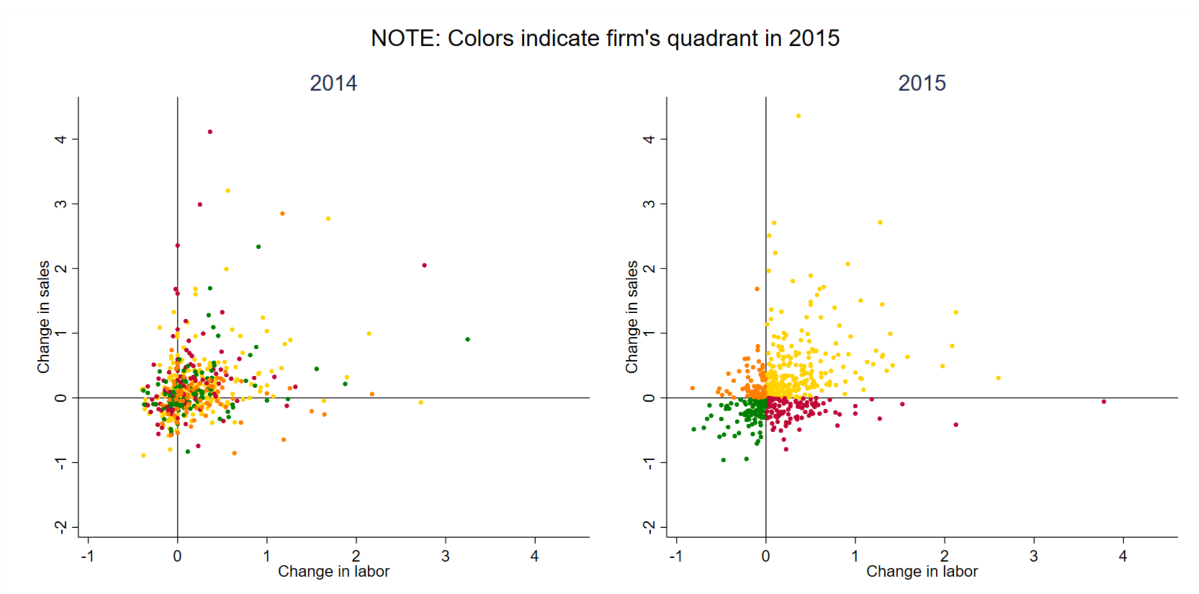
This is a schematic representation of the initial classification of firms into types. I calculate the changes in reported labor and sales in 2015 and observe the sign. Oppositely signed changes in labor and sales indicate misreporting in 2014. For example, a quadrant IV firm, which reported a positive change in labor and a negative change in sales in 2015, likely received a negative profitability shock in 2015. Its underreporting of sales in 2014, however, resulted in a positive change in labor in 2015. I classify this firm as a type 4 firm, likely liar about labor, but truth teller about sales. Please refer to the text for more details about the classification strategy.

Validating the inference strategy. The dramatic changes in aggregate labor and sales in 2015 suggest that the policy shock affected many firms. We expect the aggregate evidence presented above to translate to firm-level evidence in the sense that more firms reported positive changes in sales and employment in 2015 than in other years. Based on this insight, I populate the graph in Figure 4 with data, as shown in Figure 5. The right panel illustrates my classification of firms based on their reporting in 2015. The left panel shows which quadrant each firm occupied in 2014. Importantly, most firms that appear in quadrant IV in 2015 were previously in other quadrants, suggesting that oppositely signed changes in labor and sales are not a feature of these firms' technology but a reaction to the policy shock. As expected, type (iii) firms, truth tellers, seem to be in quadrants I or III in 2014. Lastly, I check whether more firms move toward quadrant I in 2015 compared to other years by running the following regression:

$$\mathbb{I}(Quad1_{it}) = \alpha + \beta_1 \mathbb{I}(Quad1_{it-1}) + \beta_2 \mathbb{I}(Quad1_{it-1}) * \mathbb{I}(t = 2015) + \beta_3 \mathbb{I}(t = 2015) + \epsilon_{it}$$

where $\mathbb{I}(Q1_{it}) = 1$ if firm i was observed in quadrant I in period t. The coefficient on $\mathbb{I}(t = 2015)$ in Table 3 shows that the probability of firms moving into quadrant I is indeed higher in 2015 than in other years. The appendix provides more evidence that the patterns of changes observed in 2015 are compatible with the initial firm type assignment.

Figure 5: Firms' changes in sales and labor in 2015



Source: Author's calculations with Structural Business Survey data. The right panel plots each firm's percent changes in labor and sales in 2015 where the four different colors represent the four different firm types: type 1 firms are in yellow, type 2 firms in orange, type 3 firms in green, and type 4 in red. The left panel plots each firm's changes in labor and sales in 2014, with firms maintaining the 2015 color/type classification. The fact that most type 4 (red) firms appear in quadrants other than IV in 2014, suggests that oppositely signed changes in labor and sales are not a feature of these firms' technology but a reaction to the policy shock.

Table 3: The probability of movement to quadrant I in 2015 versus in other years

Variables	$\mathbb{I}(Quad1_t)$
$\mathbb{I}(Quad1_{it-1})$	.07**
	(.026)
$\mathbb{I}(Quad1_{it-1}) * \mathbb{I}(t = 2015)$	-0.05
	(.049)
$\mathbb{I}(t = 2015)$	.13***
	(.034)
constant	.35***
	(.017)
Obs	1,986
R-squared	0.013

Source: Structural Business Survey data, 2011-2015. Regression results show that the probability that a firm moves from other quadrants to quadrant I is higher in 2015 than in other years, as confirmed by the positive and significant coefficient on $\mathbb{I}(t = 2015)$. This is consistent with the expectation that the policy shock of late 2015 incentivized firms to declare their unreported labor and sales, resulting in observed positive changes in both labor and sales for more firms than in previous years. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Characteristics of firm types. If this initial classification is mostly correct, we should expect firms in each quadrant to exhibit characteristics that are type-specific. Column 2 in Table 4 shows that type (ii) and (iii) firms have a higher share of foreign firms, consistent with the fact that foreign firms are more likely to be audited by parents and third parties, and thus less likely to misreport. Type (iii) and (iv) firms have more exporters among them, as shown in Column 3 of Table 4, consistent with the fact that the Customs Office serves as a cross-check to the reported sales and thus makes it harder to misreport sales. Moreover, Column 1 in Table 4 shows that average firm size in terms of employees is quite similar across the four types, suggesting that firm size (or productivity) and type are uncorrelated. I use this feature of the data when I develop the theoretical model in the next section.

Lastly, Table 5 shows that when I compare truth tellers with other firm types in years prior to 2015 two patterns emerge: first, the standard deviation of the annual change in the number of workers is higher among truth tellers than among other types. This is consistent with our expectation that truth tellers are less likely to use informal workers, and so their reported labor accurately reflects the volatility of the profitability shocks and the adjustment costs they face. Second, the inaction rate, defined as the share of firms that report zero annual net changes to their workforce, is lower among the likely truth tellers than among other types. This is consistent with the expectation that firms that hire informal labor do not adjust their labor as often as truth tellers and therefore appear to face high adjustment costs.

Table 4: Characteristics of firm types in 2015

Variables	(1) log employees	(2) $\mathbb{I}(foreign)$	(3) $\mathbb{I}(exporter)$
	.	-	
$\mathbb{I}(Quad1_{t=2015})$	3.8 (0.05)	0.12 (0.02)	0.68 (0.03)
$\mathbb{I}(Quad2_{t=2015})$	3.7 (0.10)	0.24 (0.04)	0.64 (0.05)
$\mathbb{I}(Quad3_{t=2015})$	3.8 (0.11)	0.24 (0.04)	0.72 (0.04)
$\mathbb{I}(Quad4_{t=2015})$	3.7 (0.08)	0.11 (0.03)	0.72 (0.04)
Obs	717	717	717
R-squared	0.93	0.18	0.69

Source: Structural Business Survey data 2015, manufacturing firms with more than 9 employees. Column 1 shows that average firm size is similar across all four firm types. Column 2 shows that the share of foreign firms is higher among type 2 and type 3 firms. Column 3 shows that the share of firms exporting is higher among type 3 and type 4 firms. Standard errors in parentheses.

Table 5: Characteristics of firm types before 2015

Variables	(1) SD no. of employees	(2) Inaction rate
	.	
$\mathbb{I}(Quad3_{t=2015} = 1)$	58.7	0.08 (0.01)
$\mathbb{I}(Quad3_{t=2015} = 0)$	28.6	0.11 (0.01)
Obs	2,907	2,907
R-squared	-	0.11

Source: Structural Business Survey data 2006-2014, manufacturing firms with more than 9 employees. The table shows characteristics of firms classified as truth tellers, i.e. $\mathbb{I}(Quad3_{t=2015} = 1)$, and non-truth tellers, i.e. $\mathbb{I}(Quad3_{t=2015} = 0)$, in years prior to the policy shock. Column 1 shows that the standard deviation of employment among truth tellers is higher than that among non-truth tellers. This is consistent with the expectation that non-truth tellers use some informal employment to respond to shocks. Column 2 shows that the share of firms that do not adjust their annual labor, the inaction rate, is lower among truth tellers. Standard errors in parentheses.

2.5 Estimating the probability of misclassification

While my initial classification of firms into types is consistent with other features of the data, as shown above, it relies on three strong assumptions: that there is no other type of measurement error, that all misreporting firms were affected by the 2015 anti-informality campaign, and that all firms reporting positive changes are type (i) firms. I relax these assumptions by allowing for misclassification.[37] In particular, I use a multinomial logit

[37] Misclassification may also result from unmodeled firm heterogeneity. For example, if firms face idiosyncratic labor productivity shocks that follow an AR(1) process, a negative change in labor might be the result of more efficient use of labor for that particular firm. As long as that type of shock is uncorrelated with the profitability shock, which is the shock in my structural model, a measurement error of this kind should be random.

framework with misclassification in the spirit of Hausman et al. (1998) to estimate the probability that each firm type is misclassified conditional on observable firm characteristics, including foreign ownership, whether it exports, and the capital intensity of the sector of operation.

Throughout this empirical exercise, and in the structural model that I develop in the following section, I maintain the assumptions that firms choose their type before they learn about their profitability and that firm types do not change. The first assumption is motivated by the fact that there is no correlation between firm size (or productivity/profitability) and type according to my original classification, as shown in Table 4. The second assumption is plausible given the brevity of the period I use for estimation: 2013 for the multinomial logit, and 2011-2013 for the structural parameters that govern firm type and production choices. Furthermore, type (i) and type (iv) firms, which lie about labor, have the option to report their labor truthfully and, thus, to not use any informal labor in a given period. For easier notation moving forward, let the firm types be indexed by j, where a type (i) firm is denoted as $j = 1$, a type (ii) firm as $j = 2$, a type (iii) firm as $j = 3$, and a type (iv) firm as $j = 4$.

Strategy. Firm types are unobserved. As discussed earlier, I exploit the shock to reporting in 2015 to classify firms into one of the four types j. It is important to note that the shock to reporting did not alter the choices of the firm. Firms made choices during the year unaffected by the shock. In the last quarter of the year firms were compelled to reveal their hidden data: informal workers were formalized and hidden sales were declared.

There are three sources of misclassification in type assignment: random measurement error in "true" reporting at the end of the year, random measurement error related to the possibility that some misreporting firms were unaffected by the enforcement shock, and a possible error in classifying all firms that reported a positive change in both labor and sales as type $j = 1$.[38] Following Hausman et al. (1998) I can consistently estimate the probability that a firm is misclassified or correctly classified as type j, conditional on vector X_i of observable characteristics, where $X_i = \{Ownership, Exporter, Sector\ capital\ intensity\}_i$.

Let $d_{ij} = 1$ if I classify firm i as type j, where $j \in \{1, 2, 3, 4\}$. Let $\bar{d}_{ij}$ be the true firm type: $\bar{d}_{ij} = 1$ if type j was chosen by firm i. Let the firm type choice be determined by a latent variable

$$U^*_{ij} = \alpha_j + \boldsymbol{\beta}'_j \boldsymbol{X_i} + z_{ij},$$

[38] While the estimated informal share of employment in firms of more than 10 employees was very low in 2015, at 3 percent, it was not zero. That indicates that some firms continued to keep workers off the books, despite the shock.

with z_{ij} $iid \sim H(x)$ and $H(x) = e^{-e^{-x}}$. As is standard in multinomial logit models, $\bar{d}_{ij} = 1$ iff $U_{ij} \geq U_{i(k \neq j)}$ $\forall$ $j, k \in \{1, 2, 3, 4\}$; otherwise, $\bar{d}_{ij} = 0$.

Let $a_{s,k}$ be the probability that a firm classified as type k is actually type s, $\forall$ $s, k \in \{1, 2, 3, 4\}$. I can write the matrix of the probabilities of misclassification as:

$$\boldsymbol{a} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \tag{1}$$

where by definition, the elements of each row sum to 1.

The expected value of the observed dependent variable d_{ij} for each $j \in \{1, 2, 3, 4\}$ is:

$$\begin{aligned} Pr(d_{ij} = 1|X_i) &= Pr((d_{ij} = 1|\bar{d}_{ij} = 0)|X_i)Pr(\bar{d}_{ij} = 0|X_i) \\ &+ Pr((d_{ij} = 1|\bar{d}_{ij} = 1)|X_i)Pr(\bar{d}_{ij} = 1|X_i) \\ &= Pr((d_{ij} = 1|\bar{d}_{ij} = 0)|X_i)(1 - Pr(\bar{d}_{ij} = 1|X_i)) \\ &+ Pr((d_{ij} = 1|\bar{d}_{ij} = 1)|X_i)Pr(\bar{d}_{ij} = 1|X_i) \\ &= \sum_{k \neq j} a_{k,j}\omega_{k,j} + (1 - \sum_{k \neq j} a_{k,j}\omega_{k,j} - \sum_{k \neq j} a_{j,k}) \frac{exp(\alpha_j + \beta_j' X_i)}{\sum_k exp(\alpha_k + \beta_k' X_i)}, \end{aligned} \tag{2}$$

for $k, s \in \{1, 2, 3, 4\}$, where $\omega_{k,j} = \frac{exp(\alpha_k + \beta_k' X_i)}{\sum_{s \neq j} exp(\alpha_s + \beta_s' X_i)}$.

And the log likelihood is given by:

$$logL = \sum_i \sum_j d_{ij} log Pr(d_{ij} = 1|X_i). \tag{3}$$

Similar to the result shown in Hausman et al. (1998), as long as $\sum_{k \neq j} a_{k,j}\omega_{k,j} + \sum_{k \neq j} a_{j,k} < 1$ the model parameters are identified, and one can consistently estimate them via maximum likelihood maintaining the assumption that misclassification does not depend on the covariates. Intuitively, the identification condition says that firms should on average be correctly classified. If the inequality does not hold, our estimates of α_j and β_j would have incorrect signs, as can be confirmed by inspecting the last line of (2).

When estimating the matrix $\boldsymbol{a}$, I impose the following restrictions:

$$\boldsymbol{a} = \begin{bmatrix} a_{11} & a_{12} = a_{32} & 0 & a_{14} = a_{34} \\ a_{21} & a_{22} & a_{23} = a_{43} & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & 0 & a_{43} & a_{44} \end{bmatrix} \tag{4}$$

The restriction requires that $a_{43} = a_{13} = a_{24} = 0$, so that firms classified as type $j = 2, 3, 4$ can only be misclassified as the adjacent type. For example, a firm classified as type $j = 2$ may be misclassified as a type $j = 1$ or $j = 3$, but not $j = 4$. The three equality restrictions reflect the assumption that misclassification to adjacent cells is symmetric. These kinds of restrictions are common in the literature (see e.g., Hausman et al, (1998)).

Results. The estimated multinomial logit with misclassification shows that firms are on average correctly classified. However, all firm types have some probability of misclassification. We can see in (5) the estimated probablities of classification and the bootsrapped standard errors in parentheses. The results show that quite a few firms classified as type 3 have some likelihood of being either type 2 or type 4 firms. Most firms classified as type 1 are correctly classified. Table 5 shows the multinomial logit coefficients, with truth tellers serving as the base. In particular I note that, as expected, foreign firms are more likely to be truthtellers, and exporters are more likely to be type 4 firms than other types.

$$\hat{\boldsymbol{a}} = \begin{bmatrix} 0.68(0.0000) & 0.14(0.0000) & 0 & 0.18(0.0000) \\ 0.02(0.0000) & 0.64(0.0000) & 0.34(0.0000) & 0 \\ 0.00(0.0005) & 0.14(0.0000) & 0.68(0.0000) & 0.18(0.0000) \\ 0.02(0.0000) & 0 & 0.34(0.0000) & 0.64(0.0000) \end{bmatrix} \tag{5}$$

Table 6: Multinomial logit with misclassication

	Coefficient
Type 1 (liars about both labor and sales)	
Foreign=1	-1.20 (5.44)
Export status=1	-0.01 (1.31)
Two digit sector capital intensity	-0.75 (0.29)
Constant	5.13 (2.23)
Type 2 (liars about sales only)	
Foreign=1	9.18 (33.83)
Export status=1	8.06 (26.41)
Two digit sector capital intensity	-0.49 (3.64)
Constant	-15.45 (27.46)
Type 3 (truth tellers)	-
Type 4 (liars about labor only)	
Foreign=1	-3.31 (32.98)
Export status=1	10.89 (10.29)
Two digit sector capital intensity	-1.55 (2.06)
Constant	-3.70 (10.61)
Log likelihood	-558.2

Source: 2013 Structural Business Survey data, manufacturing firms with more than 9 employees. The table shows the coefficients of the multinomial logit with misclassification, relative to type 3 firms, the truth tellers. I use patternsearch in Matlab to estimate the parameters. I bootstrap the data 100 times with replacement to calculate the standard errors.

3 Structural model

The facts documented in the previous section suggest that to understand how informality and government enforcement of anti-informality regulations affect the allocation of resources within and across firms, one needs to carefully model all four firm types and the decisions facing each type. To that end, I develop a dynamic model where firms differ in their profitability and the cost they must incur to be each of the four types. The dynamics in the model come from the fact that one of the inputs in production, formal labor, is costly to adjust. The adjustment costs are quadratic and are interpreted as reflecting costs related to matching frictions in the market as well as rigidities coming from policies that make it difficult to hire and fire workers. An example of the latter is prohibiting firms from using temporary workers for permanent tasks, a regulation that is in place in Albania. To avoid such costs, firms may choose to hire workers informally. Conditional on type, firms make decisions on actual input use and the amount of revenue and labor to report to the government.

To be able to conduct counterfactual analyses, I model the endogenous selection of firms into types. At the beginning of time, before they learn their profitability, firms draw costs for each type and pick the type with the highest expected value. Firm types remain fixed for the rest of their lives, unless there is a regime change, as in the counterfactual exercises.

3.1 Model environment

I develop a discrete time model of an economy that is populated by a fixed number of heterogeneous firms that produce a homogeneous product in a perfectly competitive output market. The price of the product is normalized to 1. A mass of potential workers provide labor perfectly elastically. In the background there is a government that enforces labor regulations and collects taxes.[39]

Technology. Firms have a decreasing returns to scale technology that uses capital, labor and intermediates. As in Cooper et al. (2015), capital is rented and intermediates are purchased in perfectly competitive markets, and, thus, they are adjusted without friction. Firms can hire workers formally or informally. Informal workers have no contracts or benefits, but are identical to formal workers otherwise. Firms generate net revenues according to:

$$y_{it} = A_{it}(l_{it}^F + l_{it}^I)^\alpha \tag{6}$$

where l_{it}^F is the number of formal workers, l_{it}^I is the number of informal workers, A_{it} is a revenue productivity shock, and α is the curvature of the net revenue function.[40] Firms are subject to revenue profitability shocks A_{it} that follow an AR(1) process, such that

$$log(A_{it}) = \rho log(A_{it-1}) + log(\epsilon_{it}),$$

with $log(\epsilon_{it}) \sim iid\ N(0, \sigma_\epsilon)$.

Costs. Formal labor is subject to quadratic adjustment costs, while informal labor is increasingly costly due to expected government fines that grow with the number of informal workers. The costs of informal workers do not have a dynamic component, reflecting the fact that informal workers can be hired and fired costlessly. Furthermore, since most manufacturing is low skill in Albania and the unemployment rate is high, the adjustment costs associated with formal labor primarily reflect policy-generated rigidites rather than search

[39] I do not explicitly model the government since the focus of my work is not on welfare analysis.
[40] See the appendix for the derivation of the net revenue function.

and matching frictions.[41] Total formal and informal labor costs are:

$$C_{it}^{F} = (1+\tau^{w})w^{F}l_{it}^{F} + \frac{\gamma}{2}(\frac{l_{it}^{F} - l_{it-1}^{F}}{l_{it-1}^{F}})^{2}l_{it-1}^{F} \tag{7}$$

$$C_{it}^{I} = w^{I}l_{it}^{I} + \frac{1}{s}(1+\frac{l_{it}^{I}}{b})l_{it}^{I} \tag{8}$$

where τ^w is the payroll tax, w^F, w^I are the wages of formal and informal workers, and γ is the adjustment cost parameter. Parameters $b > 0$ and $s \geq 1$ govern the expected cost of informal workers. The functional form for the cost of informal workers is similar to Ulyssea (2018), with the modification of multiplying by $\frac{1}{s}$.[42]

At the beginning of their existence, firms draw a vector c of costs associated with four modes of reporting sales and labor data to the government. Upon drawing these costs, a firm compares the expected lifetime profits for each type j weighted by $1/c_{ij}$ and select into one of the four modes. As depicted in Table 7, a firm can underreport both sales and labor, which I term $j = 1$, underreport sales only, $j = 2$, undereport labor only, $j = 4$, or tell the truth about both labor and sales, $j = 3$. I assume that when misreported, labor is underreported by the number of informal workers a firm employs. When misreporting sales, firms choose to underreport to avoid VAT. Once firms have selected into a type, the types remain fixed for the rest of the firms' lives if there are no changes to the structure of the economy. I explore such changes through counterfactual exercises in Section 5.

Firms that choose to underreport sales weigh the benefits of avoiding VAT against the expected costs of underreporting, which increase in the share of sales they decide to hide.

Table 7: Four firm types

		labor	
		lie	truth
sales	lie	$j = 1$	$j = 2$
	truth	$j = 4$	$j = 3$

[41] Firm interviews in Albania corroborate the idea that both formal and informal workers are quite easy to find.

[42] The additional parameter is necessary in this setting because without it expected costs of informality would be so high that most firms would choose to hire no informal workers, contradicting the empirical evidence.

3.2 Firm choices conditional on type

Every period, firms draw profitability A_{it} and make decisions about how much formal labor to hire, how much informal labor to hire, and how much of their sales to hide, depending on their already established type. In this subsection, I describe the optimization problem of each of the four firm types, proceeding in order from most to least truthful. The policy functions that solve the optimization problems are obtained numerically through value function iteration (VFI) as described in Adda and Cooper (2003). I discretize the AR(1) shock process using the method described in Tauchen (1986).[43]

Type 3 (truth teller). Firms that truthfully report both labor and sales maximize their expected lifetime profits, which can be written in terms of the value function:

$$\begin{aligned} V_3(A, l^F_{-1}) = &\, max_{l^F}(1-\tau^q)y(A, l^F) - C^F(l^F, l^F_{-1}) \\ &+ \beta \mathbb{E}_{A_{+1}|A} V_3(A_{+1}, l^F) \end{aligned} \tag{9}$$

where τ^q is the VAT rate and β is the discount rate.

Type 2 (truthfully reports labor, lies about sales). Firms that truthfully report labor but misreport sales solve the following optimization problem:

$$\begin{aligned} V_2(A, l^F_{-1}) = &\, max_{l^F,\eta}(1-(1-\eta)\tau^q)y(A, l^F) - C^F(l^F, l^F_{-1}) \\ &- \underbrace{\frac{\eta}{\delta+\eta}\eta(1+\tau^q)y(A, l^F)}_{expected\ cost\ of\ lying} + \beta \mathbb{E}_{A_{+1}|A} V_2(A_{+1}, l^F) \end{aligned} \tag{10}$$

where η is the share of sales a firm chooses to hide, and δ is the parameter that captures the intensity of government enforcement. When $\delta = 0$, enforcement is perfect, in which case a firm would always choose $\eta = 0$. I set the structure of the fine when caught underreporting to that specified in Albanian law.[44] It can be shown that optimal η is equal to $\delta\left(\sqrt{1+\tau^q}-1\right)$, so all type 2 firms hide the same share of their sales.

[43] For more details, see Appendix A3.

[44] The amount that is owed when caught misreporting is equal to the amount that was misreported plus tax liabilities on it. See "The Law on Fiscal Procedures", item 116.

Type 4 (lies about labor, truthfully reports sales). Firms misreporting labor but truthfully reporting sales solve the following problem:

$$\begin{aligned} V_4(A, l^F_{-1}) = max_{l^F, l^I} & (1-\tau^q) y(A, l^F, l^I) \\ & - C^F(l^F, l^F_{-1}) - C^I(l^I) + \beta \mathbb{E}_{A_{+1}|A} V_4(A_{+1}, l^F) \end{aligned} \tag{11}$$

To understand the firm's decision-making consider the response of a type 4 firm to a positive profitability shock. Since l^F and l^I are perfect substitutes, the firm will the type of labor that is cheaper at the margin. For every current size l_{-1} and desired size $l > l_{-1}$, the firm then weighs the costs of expanding formally, which include formal salaries, taxes, the adjustment costs, and the expected change in next period's value function, against the costs of expanding informally, which include informal salaries and expected fines.

With w^I much smaller than $(1+\tau^w)w^F$, in a deterministic steady state all type 4 firms have some informal labor, $\bar{l}$, as in Ulyssea (2018):

$$(1+\tau^w)w^F = w^I + \frac{1}{s}(1 + \frac{2\bar{l}}{b}) \tag{12}$$

Outside of the deterministic steady state, a positive shock increases the left hand side, adding the marginal adjustment costs and the expected change in next period's value function, which results in a higher $\bar{l}$. Thus firm heterogeneity, coupled with adjustment costs leads to informal employment that varies across firms.

Type 1 (lie about both labor and sales). Finally, firms misreporting both labor and sales have the following value function:

$$\begin{aligned} V_1(A, l^F_{-1}) = max_{l^F, l^I, \eta} & (1 - (1-\eta)\tau^q) y(A, l^F, l^I) \\ & - C^F(l^F, l^F_{-1}) - C^I(l^I) - \\ & - \frac{\eta}{\delta + \eta} \eta (1+\tau^q) y(A, l^F, l^I) + \beta \mathbb{E} V_1(A_{+1}, l^F) \end{aligned} \tag{13}$$

3.3 Sorting into types

In period 0, firms draw costs $\boldsymbol{c_i} = \begin{pmatrix} c_{i1} & c_{i2} & c_{i3} & c_{i4} \end{pmatrix}'$ with $\frac{\mathbf{1}}{\boldsymbol{c_{ij}}}$ drawn independently from the following distribution:

$$G(c) = Pr(c_{ij} \geq c) = exp\{-T_j c^{\theta}\},$$

for $c > 0$.

I model these costs as shifting the expected value for each type.[45] The firm compares all four expected values weighted by the respective cost and chooses a type. Letting $y_{ij} = 1$ if firm i chooses type j, one can calculate the share of type j firms, $\forall\ j, k = \{1, 2, 3, 4\}$:

$$\begin{aligned} Pr(y_{ij} = 1) &= \mathbb{E}\left(\prod_{k \neq j} Pr(\frac{\mathbb{E}V_j}{c_{ij}} \geq \frac{\mathbb{E}V_k}{c_{ik}}) \right) \\ &= \int_0^{\infty} \prod_k G(\frac{\mathbb{E}V_k}{t}) g(\frac{\mathbb{E}V_j}{t}) dt \\ &= \frac{T_j (\mathbb{E}V_j)^{\theta}}{\sum_k T_k (\mathbb{E}V_k)^{\theta}} \end{aligned}$$

These costs arise in part from observed firm characteristics that I do not model and also from unobserved components. For example, a firm in a very capital intensive sector has high costs of hiring informal workers due to a high minimum amount of training needed to operate machines. However, the same firm may easily hide sales. In contrast, a firm that exports most of its production finds it harder to hide sales due to customs records being available. And if a firm is foreign owned, it has a high likelihood of being audited by its headquarters, and, thus, it has a high cost of lying about both labor and sales.

To make the decision about type, firms need to form an expectation about the lifetime profits for each type, $\mathbb{E}_0 V_1$, $\mathbb{E}_0 V_2$, $\mathbb{E}_0 V_3$, and $\mathbb{E}_0 V_4$. For example, the expected value of being a type

[45] This structure can be microfounded with a model in which being a certain type carries reputation effects. For example a firm that is discovered to be a liar might suffer from reputational damage that is difficult to recover from. For publicly traded firms, this would be a loss in stock market value.

3 firm is the following:

$$\mathbb{E}_0 V_3 = \mathbb{E}_0 \Bigg\{ max_{\{l^F\}_0^\infty} \Big\{ (1-\tau^q) A l_0^{F\alpha} - (1+\tau^w) w^F l_0^F \\ + \sum_{t>0} \beta^t \Big((1-\tau^q) A l_t^{F\alpha} - (1+\tau^w) l_t^F - \frac{\gamma}{2} (\frac{l_t^F - l_{t-1}^F}{l_{t-1}^F})^2 l_{t-1}^F \Big) \Big\} \Bigg\}$$

Expressing it in terms of the value function in (1):

$$\mathbb{E}_0 V_3 = \mathbb{E}_0 \Bigg\{ max_{l_0^F} \Big\{ (1-\tau^q) A l_0^{F\alpha} - (1+\tau^w) w^F l_0^F + \beta \mathbb{E}_1 V_3(A', l_0^F) \Big\} \Bigg\}$$

Whenever there is a regime change or any other structural change, the ranking of types in terms of expected value will change causing firms to re-sort into types. For instance, an increase in the expected cost of hiring informal workers will lower $\mathbb{E}_0 V_4$, causing some firms to switch from being type 4 to other types.

4 Estimation

This section has two goals. First, I discuss how I quantify the model. I describe the estimation strategy of parameters that I estimate without using the structure of the model. I then describe the structural estimation strategy, the identification of structural parameters as well as the fit of the model.

Second, I show the consequences of ignoring informal employment and other kinds of evasion in the estimation procedure. I show that naively assuming that all firms truthfully report their activities leads to overstatement of adjustment costs of formal labor by a factor of at least two.

4.1 Quantifying the model

To quantify the model, I divide the parameter vector into three groups: the first group includes parameters that have statutory counterparts in Albania, the second includes parameters that I estimate without using the structure of the model (such as the probability that a firm is correctly classified as a given type), and the third group of parameters is estimated using the structure of the model.

Group 1 parameters. τ^w is equal to 0.167 and includes all taxes on salaries owed by firms, including social security and health insurance contributions. The value added tax, τ^q, is equal to 20 percent.

Group 2 parameters. These parameters include formal and informal wages, w^F and w^I, the probability of correct classification and misclassification for each firm type in the data, and the curvature of the net revenue function for each type.

The formal annual wage w^F is 0.288 million Albanian Lek (ALL), about $2,700, and is equal to the average net annual compensation of formal workers in manufacturing firms in 2011-2013. For comparison, the minimum wage in 2013 was 0.264 million ALL. The informal wage w^I is equal 0.21 million ALL, or 75 percent of the formal wage.

To estimate the curvature of the revenue function, I follow Cooper et al. (2015). The main modification in my setting is that each firm's contribution in the regression is determined by the (mis)classification probabilities of being a truth teller obtained in Section 2. For example, a firm classified as type 2 has a 64 percent probability of being correctly classified and a 14 percent probability that is instead type 2. That firm's contribution in the regression is 14 percent. I estimate the curvature via generalized method of moments (GMM), where I regress gross firm revenues on total labor, using initial firm wages and twice lagged labor to control for endogeneity in inputs.[46] Table 8 shows the parameter estimates obtained using IV estimation.

Table 8: Revenue function estimation

	log(real sales)
log(employees)	0.71***
	(0.1)
Observations	267
R-squared	-

Source: Structural Business Survey data 2011-2013, manufacturing firms with more than 9 employees. The table shows the second stage results of the gmm regression with instrumental variables. The instruments used are firms' wage expenses reported in the first year of available data as well as twice lagged reported employment. There is attrition in the sample due to missing data on firms' initial wages. The first stage R-squared is 0.99. Robust standard errors in parentheses, clustered at the two-digit sector level. *** p<0.01, ** p<0.05, * p<0.1

[46] As in Cooper et al. (2015) I do observe net revenues in the data, because capital is not reported.

Group 3 parameters. Lastly, I estimate the rest of the structural model parameters using an SMM approach. Denoting the vector of structural parameters as

$$\Gamma = (\sigma_\epsilon, \rho, \beta, \gamma, b, s, \delta, \theta, T_1, T_2, T_4).$$

I solve the the following minimization problem:

$$Q(\Gamma) = min_\Gamma (m^d - m^s(\Gamma))W(m^d - m^s(\Gamma))' \tag{14}$$

where m^d is a vector of moments calculated from data, $m^s(\Gamma)$ is a vector of moments calculated from simulated model data, and W is weighting matrix. In the data, each firm's contribution to a moment calculation is weighted by the probability of correct classification and misclassification obtained in Section 2.5. The matrix W is the inverted variance-covariance estimate obtained by bootstrapping the data.[47] Table 9 summarizes all structural parameters.

Moments and identification. Moments of type 3 firms, truth tellers, identify the shock process, the adjustment costs parameter, and the discount rate. I use eight moments from type 3 firms to estimate $\sigma_\epsilon, \rho, \beta, \gamma$. Roughly speaking, the serial correlation of sales identifies the persistence of the shock process, whereas the size distribution identifies the standard deviation of the shock. The turnover rate and variation in the serial correlation of employment identify the adjustment costs. Variation in the discount rate influences all moments.

I use four moments from type 2 firms, those lying about sales only, to identify the expected cost of hiding sales. Any differences in the sales size distribution between type 3 firms and type 2 firms must come from type 2 firms underreporting sales. Thus, the cutoffs of the type 2 sales size distribution identify δ.

To estimate the expected cost of hiring informal workers, I use four moments from type 4 firms, liars about labor only, and the aggregate share of informal employment. Any differences in the employment size distribution between type 3 firms and type 4 firms must come from the fact that type 4 firms hire informal workers, which together with the aggregate share of informal workers identify b and s. Lastly, I match the share of types in the economy with the one generated by the model. Table 10 summarizes the subsamples and moments used for identification.

[47] I bootstrap the data 500 times. As is common in the literature, I use the diagonal elements of the variance-covariance matrix in W. See, for example, Cosar et al. (2016).

Table 9: Model parameters

	Description	Source	Value	S.e.
τ^w	Payroll tax	Statutory	0.167	
τ^q	Value added tax	Statutory	0.2	
w^F	Formal wage	Labor force survey	0.288	
w^I	Informal wage	Labor force survey	0.21	
T_3	Location param. of type 1 costs distr.	Normalized	1	-
α	Curvature of the net revenue function	Estimated	0.71	(0.1)
σ_ϵ	Standard deviation of the shock process	Estimated	0.416	(0.003)
ρ	Persistence of the shock	Estimated	0.902	(0.027)
β	Discount factor	Estimated	0.892	(0.001)
γ	Quadratic adjustment cost parameter	Estimated	3.554	(0.139)
δ	Hiding sales cost parameter	Estimated	6.78	(0.001)
b	Informal hiring cost parameter	Estimated	7.475	(0.320)
s	Informal hiring cost parameter	Estimated	41.998	(0.063)
θ	Dispersion param. of all types' costs distr.	Estimated	1.072	(0.000)
T_1	Location param. of type 1's costs distr.	Estimated	2.569	(0.010)
T_2	Location param. of type 2's costs distr.	Estimated	1.170	(0.011)
T_4	Location param. of type 4's costs distr.	Estimated	1.198	(0.001)

Notes: The table shows all model parameters. The top two rows are tax rates that reflect current regulations in Albania. The next two rows reflect formal and informal manufacturing workers' average wages, calculated from the Labor Force Survey during 2011-2013. The fifth row is a normalization. The sixth row shows the result of the IV regression in Table 8. The rest of the parameters are estimated via simulated method of moments minimizing the objective function as shown in equation (14) above. Standard errors are in parentheses, obtained via bootstrap.

Model fit. Table 11 shows how the model performs in matching targeted and non-targeted moments in the data. The model does fairly well in matching serial correlation in sales and employment. However, it overstates both the covariance of sales and employment and the employee turnover rate. This could partially reflect the fact that I am not modeling non-convex formal labor adjustment costs.[48] I do not expect this limitation to be consequential in the counterfactual analyses because the parameters that govern the costs of informal employment are identified from the cross-sectional differences in formal labor between type 3 and type 4 firms. To the extent that non-convex adjustment costs affect the observed distribution of formal labor among type 3 firms in a similar way to that of type 4 firms, any differences in the size distribution can be attributed to type 4 firms using informal labor.

The size distribution of type 3 firms, truth tellers, is matched relatively well, but the number

[48] In Section 2.4 I show that the inaction rate among truth tellers, while low at 8 percent, is not zero.

Table 10: Identification of structural parameters

Parameters	Subsample	Moments
Adjustment cost (γ)	Truth tellers	Serial corr. in labor
Shock (ρ, σ_ϵ)	Truth tellers	Serial corr. is sales, cov. sales labor, size dist.
Discount factor (β)	Truth tellers	All moments
Informal labor (b, s)	Liars about labor only Economy wide	Size distribution in terms of labor Share of informal employment from LFS
Hiding sales (δ)	Liars about sales only	Size distribution in terms of sales
Frechet (θ, T_j)	All firms	Share of type j firms (corrected for misclass.)

Notes: This table summarizes the identification strategy for the structural parameters. Th estimation time period is 2011-2013. Moments from type 3 firms, truth tellers, identify parameters that are common to all firms, shown in the first three rows. Moments from type 4 firms, liars about labor, as well as the aggregate informal share of employment, identify the costs of hiring workers informally. Moments from type 2 firms, liars about sales only, identify the hiding sales parameter. Lastly, the share of all firm types in the economy identifies the costs distribution for each type.

of larger firms is overstated among type 4 firms. The shares for each firm type are replicated almost exactly, as is the economy wide informal share of employment. Regarding non-targeted moments, I note that moments from type 1 firms are not used in the estimation. The model does well in matching the lower part of the distribution for those firms but overstates the upper part.

Table 11: Model Fit

	Model	Data
Targeted moments		
Type 2 10th percentile sales cutoff	14.77	17.5
Type 2 40th percentile sales cutoff	48.38	51.9
Type 2 60th percentile sales cutoff	102.49	109.7
Type 2 80th percentile sales cutoff	284.76	228.7
Type 3 serial correlation in log(employees)	0.99	0.99
Type 3 serial correlation in log(sales)	0.95	0.93
Type 3 covariance (log(employees), log(sales))	0.94	0.74
Type 3 employee turnover rate	0.14	0.01
Type 3 10th percentile employees cutoff	13	15
Type 3 40th percentile employees cutoff	29	35
Type 3 60th percentile employees cutoff	54	61
Type 3 80th percentile employees cutoff	123	151
Type 4 10th percentile employees cutoff	12	14
Type 4 40th percentile employees cutoff	29	25
Type 4 60th percentile employees cutoff	59	44
Type 4 80th percentile employees cutoff	141	89
Overall share of type 1 firms	52.3%	51.5%
Overall share of type 2 firms	16.5%	17.0%
Overall share of type 3 firms	11.0%	11.5%
Overall share of type 4 firms	20.2%	20.0%
Informal share of employment	10.6%	10.5%
Non-targeted moments		
Type 1 10th percentile employees cutoff	13	13
Type 1 40th percentile employees cutoff	34	23
Type 1 60th percentile employees cutoff	72	40
Type 1 80th percentile employees cutoff	179	74

Notes: This table shows how well the estimated structural model does in matching targeted and non-targeted moments. The model fit is relatively good, with the exception of the upper part of the firm size distribution in terms of labor for type 4 and type 1 firms.

4.2 Consequences of naively estimating the adjustment costs

I now examine the consequences of naively estimating adjustment costs by assuming that all firms in our data truthfully report their labor and sales. I find that it produces estimates that are too high by at least a factor of two. Column 2 in Table 12 presents parameters generated by naively treating all firms as truth tellers. Comparing it with column 1, column 2 shows that both the adjustment cost and the discount factor would be misestimated, suggesting adjustment costs three times higher than my baseline, and a borrowing cost of about 16 percent, or four percentage points more than in Column 1.[49]

To exclude the effect of the lower discount factor on the overstatement of adjustment costs, I run another exercise, fixing the discount factor to the level of truth tellers. Column 3 in Table 12 shows that the adjustment costs remain overstated, now by a factor of two.

Table 12: Truth tellers versus naive estimation of adjustment costs

		Truth tellers	Naive	Naive fixed β
α	Net revenue function curvature	0.71	0.63	0.63
σ_ϵ	Shock process st. dev.	0.416	0.817	0.820
ρ	Shock persistence	0.902	0.822	0.818
β	Discount factor	0.892	0.861	-
γ	Adjustment cost parameter	3.554	10.802	8.721
$Q(\Gamma)$	Minimized distance between data & model	39.87	112.90	113.44

Notes: This table demonstrates the shortcomings of ignoring the fact that firms across the size distribution hire workers informally and evade VAT. Column 1 shows the structural estimation results when I carefully account for truth tellers by applying the empirical strategy described in sections 2.4 and 2.5. Column 2 shows the estimation results when I naively assume that all firms are truthfully reporting their activities. The quadratic adjustment cost parameter is inflated by a factor of three. Column 3 shows the results of repeating the same "naive" estimation, but fixing the discount factor at 0.892, the level of the the truth tellers'. The adjustment cost parameter would be inflated by a factor of two.

5 Counterfactual exercises

Equipped with the estimated parameters, I am able to conduct policy experiments to explore the effects of imperfect enforcement on the firm size distribution and the allocation of labor, as well as the effects of government policies that target the aggregate share of informal employment. I also demonstrate the importance of the dynamic versus static incentives

[49] The borrowing cost is calculated as r=1/β - 1

for hiring informal labor, by shutting down each type of incentive one by one. In all the counterfactual exercises the wages are held fixed at the baseline scenario.

5.1 Perfect enforcement of regulations that prohibit informal employment

I shut down the possibility of informal employment by setting $b = 0$, which makes the expected cost of hiring informal workers infinite. In this counterfactual economy there are now only truth tellers (type 3 and type 4) and firms that lie about sales (type 1 and type 2). Column 2 in Table 11 shows that the mean, median, and variance of the size distribution in terms of labor are all smaller than in the baseline economy (Column 1). Furthermore, dispersion in sales per worker, a measure of misallocation in the economy, is higher than in the baseline economy, suggesting that allowing firms to hire informal workers alleviates some of the distortions created by the interaction of formal labor adjustment costs and profitability shocks.[50] When compared with a benchmark economy with no adjustment costs and no informal labor, imperfect enforcement brings the firm size distribution closer to an "undistorted" distribution, as shown in Figure 6.

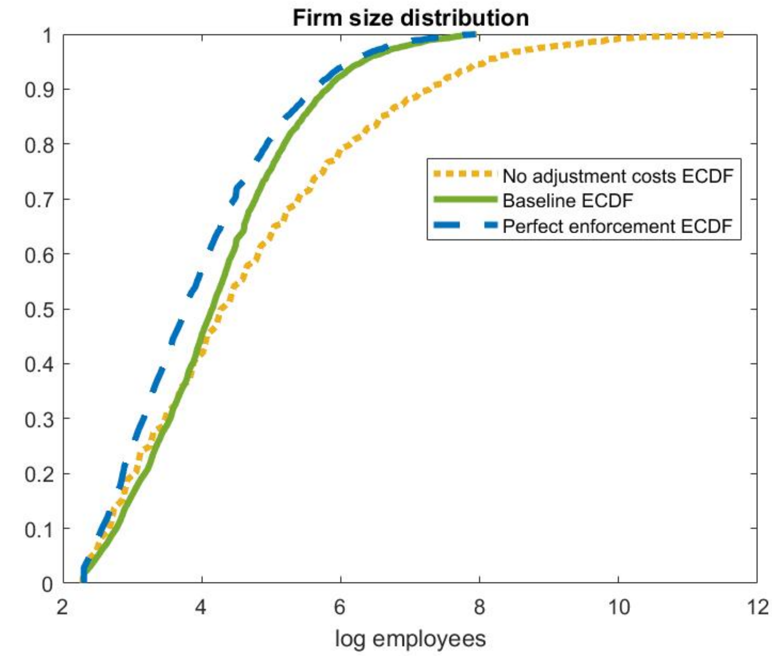
Looking at the change in the distribution of firm types between the baseline scenario and the perfect enforcement scenario, it is worth highlighting that the reduction in the share of type 1 firms, which results from the elimination of the expected benefits of hiring informal workers, is offset by an increase in the share of type 2 firms. This happens because the expected value of lying about sales only, relative to that of other types, is higher when informal employment is prohibited, so more firms sort into type 2 than in the baseline. Therefore, policies that target one type of misreporting affect other misreporting, which is an important consideration when designing such policies.

5.2 Government policies that reduce the share of informal workers

In a second set of experiments, I assess two kinds of government policies that are aimed at reducing the overall share of informal employment in the economy. The first policy doubles the cost of employing workers informally. In the model, the policy is implemented through

[50] Using the dispersion in sales per worker as a measure of misallocation is discussed in Hsieh and Klenow (2009), or Alvarez and Ruane (2019). In those papers misallocation is a result of firm-specific distortions. In my framework misallocation stems from the combination of labor market rigidities that make the adjustment of formal labor more costly, and idiosyncratic shocks,

Figure 6: Distribution of log employees



Notes: This figure shows that shutting down informal employment takes the firm size distribution further away from a benchmark of no adjustment costs. When firms have the option to hire workers informally to respond to demand fluctuations, the allocation of labor improves, as either expanding or contracting firms can reach their desired size faster.

reducing the parameter s from 41.99 to 21. This experiment captures the effects of policies such as increased enforcement through more frequent inspections or higher fines.

The second kind of policy cuts formal labor adjustment costs. These policies are generally discussed in the context of increasing labor market dynamism. However, the intervention has the potential to lower informal labor use due to firms' dynamic incentives to hire informal workers. My model allows me to quantify this effect. In the model, I implement the policy by reducing γ from 3.55 to 1.75, as shown in Column 4 of Table 13, and from 3.55 to 0.8, as shown in Column 5 of Table 13. This experiment captures the effect of policies like allowing the use of temporary contracts for permanent tasks or reducing the redundancy notice period.

Results. The first counterfactual exercise shows that doubling the cost of lying about labor reduces the share of informal employment from 10.5 percent to 5 percent. The decline in informal employment is driven by both the extensive margin, with fewer firms sorting into type 1 and type 4, and the intensive margin, with type 1 and type 4 firms hiring fewer informal workers. Furthermore, the median firm size also decreases while the dispersion in sales per worker increases, as compared to the baseline scenario. This negative effect on allocative efficiency represents a cost of the policy in addition to any explicit fiscal costs associated with, for example, enhanced enforcement.

The second exercise shows that reducing formal labor adjustment costs by half, cuts the informal share of employment from 10.5 percent to 7.8 percent despite unchanged enforcement efforts. The decline in informal employment is mainly driven by the intensive margin, with type 1 and type 4 firms hiring fewer informal workers. In addition, median firm size increases and dispersion in sales per worker decreases. Thus the policy yields an improvement in allocative efficiency relative to the baseline, in addition to its effect on informality.

Lastly, I consider the relative effectiveness of the two policy approaches. I find that a policy that reduces adjustment costs by 75 percent is just as effective in reducing the informal share of workers as one that doubles the expected costs of informal employment. Such a policy has the added benefit of reducing misallocation in the economy even further.

Table 13: Policy experiments and counterfactual exercises

	Baseline	Perfect enforcement	Double inf. costs	50% cut in adj. costs	75% cut in adj. costs
b (Informal labor cost)	7.48	0	7.48	7.48	7.48
s (Informal labor cost)	41.99	41.99	21	41.99	41.99
γ (Formal adj. cost)	3.55	3.55	3.55	1.75	0.8
Mean employment	149	120	135	197	253
Median employment	62	44	52	72	79
Variance of employment	75,821	58,529	68,849	132,326	218,782
Dispersion in MPL	19.21%	23.86%	21.27%	17.26%	14.98%
Informality rate	10.6%	0%	5.1%	7.8%	5.8%
Share of each firm type					
Type 1	52.3%	48.0%	50.2%	51.2%	50.3%
Type 2	16.5%	20.8%	18.6%	17.5%	18.2%
Type 3	11.0%	13.8%	12.3%	11.7%	12.3%
Type 4	20.2%	17.4%	18.9%	19.6%	19.2%

Notes: This table shows various outcomes of interest under different policy regimes. Column 2 shows that under perfect enforcement of anti-informality regulations, the dispersion in the marginal revenue product is higher than in the baseline scenario, indicating more misallocation of labor. Columns three and four compare two types of government policies: doubling the cost of informal employment in Column 3 and reducing the adjustment costs of formal employment in Column 4. Both result in a lower share of informal employment, but the latter alleviates misallocation. Lastly, the last column shows that a 75 percent cut in adjustment costs achieves the same level of reduction in the informal share of employment as stepping up enforcement (Column 3).

5.3 The relative importance of dynamic versus static incentives for informal employment

A key contribution of this paper to the literature on informality and firms is the modeling of dynamic incentives for hiring informal workers. To understand the role of the dynamic and static incentives in generating informal employment, I shut down each type of incentive one by one and calculate the informal share of employment. Table 14 shows the results of four counterfactuals: (1) setting the payroll tax, τ^w, equal to zero, (2) setting the payroll tax equal to zero and equating the formal and informal wages, (3) setting the formal labor adjustment cost γ equal to zero, and (4) setting the formal labor adjustment cost equal to one tenth of its baseline level. The last exercise is a more realistic representation of adjustment costs than setting them equal to zero, since labor market frictions are inevitable.

The results in Table 14 show that the informal share of employment would fall from 10.4% to 6.2% if firms did not have to pay a payroll tax (row (1)), and further to xx% if the salary was the same for formal and informal workers (row (2)). Counterfactual scenarios (1) and (2) demonstrate the importance of the static incentives of informal hiring. In addition, the results of counterfactual scenarios (3) and (4) show that the informal share of employment would fall to xx% and xx%, respectively, if the dynamic incentives for hiring workers informally were reduced or removed. These results suggest that both important in generating informal employment in the economy.

6 Conclusion

This paper develops a model of the intensive margin of informality where firms face a dynamic incentive to hire informal workers to avoid incurring formal labor adjustment costs. These costs reflect not only frictions in the labor market but also the rigidity of labor markets caused in part by government policies, for example, the impossibility of hiring temporary workers for permanent tasks. To overcome the issue of unobserved informal employment at the firm level, I carefully estimate the model with microdata from Albania using a policy shock in late 2015 that induced firms to truthfully report their labor and sales.

Three key findings emerge. First, ignoring informal employment and VAT evasion leads to mismeasurement of adjustment costs by a factor of two in a best case scenario. Second, taking into account firms' dynamic incentive to hire informal workers erodes the improvements

Table 14: Dynamic versus static informal hiring incentives

Experiment	Informality rate
(1) Zero payroll tax	
$\tau^w = 0$	6.2%
(2) Zero payroll tax and equal wages	
$\tau^w = 0$	
$w_F = w_I$	4.4%
(3) Zero adjustment costs	
$\gamma = 0$	2.0%
(4) Drastic reduction in adjustment costs	
$\gamma_{new} = 0.1 * \gamma$	4.4%

Notes: This table demonstrates the relative importance of the static and the dynamic incentives for hiring wokers informally. Counterfactual scenarios (1) and (2) capture the importance of the static incentive to hire workers informally, while scenarios (3) and (4) demonstrate the importance of the dynamic incentive. Scenario (1) shows that the informal share of workers would fall from 10.6% to 6.2% if firms did not have a static incentive to hire informally. Scenario (4) shows that the informal share of workers would fall to 2% if firms did not have a dynamic incentive to hire workers. Results base on the more realistic scenarios (2) and (3) show that dynamic and static incentives are both important in generating informal employment..

in allocative efficiency arising from reducing reliance on informal labor, which previous literature has emphasized. Third, lessening labor market rigidities is effective in reducing the informal share of employment while also improving on the efficiency of labor allocation.

References

(2017). Annex: Labor market regulation. what can we learn from Doing Business data? Technical report, The World Bank.

(2019). *Global Economic Prospects, January 2019: Darkening Skies*. Washington DC: World Bank.

Adda, J. and R. Cooper (2003). *Dynamic Economics, Quantitative Methods and Applications*. The MIT Press.

Alvarez, J. and C. Ruane (2019, November). Informality and Aggregate Productivity: The Case of Mexico. IMF Working Papers 19/257, International Monetary Fund.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014). Dynamic inputs and resource (mis)allocation. *Journal of Political Economy 122*(5), 1013 – 1063.

Best, M. C., A. Brockmeyer, H. J. Kleven, J. Spinnewijn, and M. Waseem (2015). Production versus revenue efficiency with limited tax capacity: Theory and evidence from Pakistan. *Journal of Political Economy 123*(6), 1311–1355.

Carrillo, P., D. Pomeranz, and M. Singhal (2017, April). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics 9*(2), 144–64.

Cooper, R., G. Gong, and P. Yan (2015). Dynamic labor demand in China: public and private objectives. *The RAND Journal of Economics*.

Cooper, R., J. Haltiwanger, and J. Willis (2004). Dynamics of labor demand: Evidence from plant-level observations and aggregate implications. *NBER Working Paper*.

Cooper, R. and J. Willis (2004). A comment on the economics of labor adjustment: Mind the gap. *American Economic Review*.

Cooper, R. and J. Willis (2009). The costs of labor adjustment: Inferences from the gap. *Review of Economic Dynamics*, 632–647.

Cosar, K., N. Guner, and J. Tybout (2016). Firm dynamics, job turnover, and wage distributions in an open economy. *American Economic Review*.

David, J. M. and V. Venkateswaran (2019, July). The sources of capital misallocation. *American Economic Review 109*(7), 2531–67.

de la Parra, B. S. How much is formality worth? the value of social security benefits in Mexico. 2016.

Dix-Carneiro, R., P. Goldberg, C. Meghir, and G. Ulyssea (2019). Trade and informality in the presence of labor market frictions and regulations. Technical report.

Goldberg, P. and N. Pavcnik (2003). The response of the informal sector to trade liberalization. *Journal of Development Economics*.

Haanwinckel, D. and R. Soares (2020). Workforce composition, productivity, and labor regulations in a compensating differentials theory of informality. Working Paper No. 2020-45, Becker Friedman Institute for Economics.

Hausman, J., J. Abrevaya, and F. Scott-Morton (1998). Misclassification of the dependent variable in a discrete response setting. *Journal of Econometrics*.

Hazans, M. (2011). Informal workers acrossEurope: Evidence from 30 countries. *IZA Discussion Paper No. 5871*.

Horodnic, I. (2016). Cash wage payments in transition economies: Consequences of envelope wages. *IZA World of Labor*.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in Denmark. *Econometrica 79*(3), 651–692.

La Porta, R. and A. Shleifer (2014, September). Informality and development. *Journal of Economic Perspectives 28*(3), 109–26.

Maloney, W. (2004). Informality revisited. *World Development 32*(7), 1159–1178.

Meghir, C., R. Narita, and J.-M. Robin (2015). Wages and informality in developing countries. *The American Economic Review 105*(4), 1509–1546.

MoF (2018). Albania economic reform programme (ERF) 2016-2018. Technical report, Ministry of Finance and Economy.

Perry, G., W. Maloney, O. Arias, P. Fajnzylber, A. Mason, and J. Saavedra-Chanduvi (2007). *Informality: Exit and Exclusion*. The World Bank.

Restuccia, D. and R. Rogerson (2007). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*.

Restuccia, D. and R. Rogerson (2017, August). The causes and costs of misallocation. *Journal of Economic Perspectives 31*(3), 151–74.

Tabak, P. and S. Borkovic (2019). Albania diagnostic. Technical report, European Bank for Reconstruction and Development.

Tauchen, G. (1986). Finite state markov-chain approximation to univariate and vector autoregressions. *Economic Letters*.

Ulyssea, G. (2018). Firms, informality and development: Theory and evidence from Brazil. *American Economic Review*.

A1. Empirical Appendix

Data. I identify informal workers based on the following questions in the LFS:

- Q26 in LFS 2011: In your main job, you are an:
    1. Employee
    2. Self-employed with employees
    3. Self-employed without employees
    4. Unpaid family worker
- Q38 in LFS 2011: Are you entitled to the benefits of the social security scheme in this job?
    1. Yes
    2. No
- Q41 in LFS 2011: Does your employer pay social security contributions for you?
    1. Yes
    2. No

All nominal variables are expressed in terms of 2010 prices using appropriate deflators published by INSTAT. Sales are deflated using two-digit sector level Producer Price Indices (PPI). Wages are deflated by the Consumer Price Index (CPI).

Evidence on how the pattern of changes in sales and labor reveals firm types. In this subsection, I check whether my classification of firm types is consistent with expected patterns in the data. First, if more firms were indeed reporting more labor and sales in 2015 than in other years, we should observe more mass in the positive parts of the distribution of changes in sales and labor in 2015. An inspection of the kernel densities of the changes in labor and changes in sales in Figure 7 confirms this hypothesis.

Second, in the model I assume that firms choose their types before they observe their profitability shocks and do not switch types, unless there is a regime change. This assumption is important because it allows me to use the panel dimension of the data to estimate the structural model parameters. Firms classified as liars (of labor, sales, or both) might appear in any quadrants in earlier years, but most truth tellers (up to some measurement error) are

Figure 7: Kernel densities of the changes in labor and changes in sales, 2012-2015

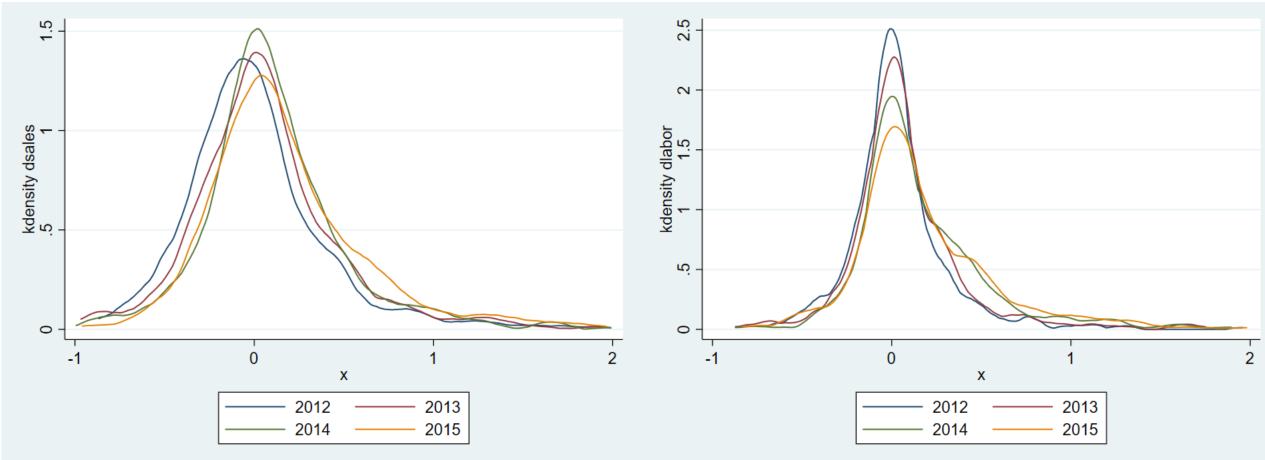


Table 15: Probability that truth tellers appear in quadrants I and III in previous years

VARIABLES	(2) lagged_truth teller
$\mathbb{I}(Q3_{t=2015})$	-.07
	(.05)
constant	.60***
	(0.02)
Obs	717
R-squared	0.003

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

expected to either be in quadrant III or quadrant I in earlier years. To check whether firms classified as truth tellers are likely to be in quadrants I an III, I run the following regression, conditioning on $t = 2015$:

$$\mathbb{I}((Q3_{it-1} \cup Q1_{it-1}) \cap (Q3_{it-2} \cup Q1_{it-2})|Q3_t) = \alpha + \beta\mathbb{I}(Q3_t) + \epsilon_{it}$$

The probability that truth tellers appear in quadrants I and III in earlier years is $\alpha + \beta$. Table 14 shows that about 60% of truth tellers do not appear in quadrants II and IV in the previous years.

A2. Model Appendix

Deriving the net revenue function. Suppose that the output markets are perfectly competitive (as we do throughout this paper). Normalizing the price of output to 1, the revenue function for each firm can be written as:

$$\bar{R}(z,l,k,m) = z(l^{\alpha_l}k^{\alpha_k}m^{1-\alpha_l-\alpha_k})^{\phi} - rk - p_m m$$

The FOCs for k and m:

FOC: $\frac{\partial \bar{R}(z,l,k,m)}{\partial k} = 0: \quad \phi\alpha_k z(l^{\alpha_l}k^{\alpha_k}m^{1-\alpha_l-\alpha_k})^{\phi} = rk$

FOC: $\frac{\partial \bar{R}(z,l,k,m)}{\partial m} = 0: \quad \phi(1-\alpha_k-\alpha_l)z(l^{\alpha_l}k^{\alpha_k}m^{1-\alpha_l-\alpha_k})^{\phi} = p_m m$

One can rewrite the revenue function at the optimal k and m:

$$\bar{R}(z,l,k,m) = (1-\phi(1-\alpha_l))z(l^{\alpha_l}k^{\alpha_k}m^{1-\alpha_l-\alpha_k})^{\phi}$$

After substituting the first order condition for k and m in the equation above, one can write $\bar{R}(z,l,k,m)$ in terms of$z$ and l only:

$$R(A,l) = \underbrace{\psi z^{\frac{1}{\psi}}\phi^{\frac{1-\psi}{\psi}}\left(\frac{\alpha_k}{r}\right)^{\frac{\phi\alpha_k}{\psi}}\left(\frac{1-\alpha_k-\alpha_l}{p_m}\right)^{\frac{1-\phi\alpha_k-\psi}{\psi}}}_{A} l^{\frac{\phi\alpha_l}{\psi}}$$

where $\psi = 1-\phi(1-\alpha_l)$. Denoting $\alpha = \frac{\phi\alpha_l}{1-\phi(1-\alpha_l)}$, the net revenue expression can be expressed in terms of A and l only.

Deriving the share of firms in each type:

$$
\begin{aligned}
Pr(y_{ij}=1) &= \mathbb{E}\left(\prod_{k\neq j} Pr(\frac{\mathbb{E}V_j}{c_{ij}} \geq \frac{\mathbb{E}V_k}{c_{ik}})\right) \\
&= \int_0^\infty \prod_k G(\frac{\mathbb{E}V_k}{t})g(\frac{\mathbb{E}V_j}{t})dt \\
&= \int_0^\infty \theta t^{-\theta-1} T_j(\mathbb{E}V_j)^\theta exp - \{t^{-\theta}\sum_k T_k(\mathbb{E}V_k)^\theta\}dt \\
&= \int_0^\infty \frac{T_j(\mathbb{E}V_j)^\theta}{\sum_k T_k(\mathbb{E}V_k)^\theta} exp - \{u\}du \\
&= \frac{T_j(\mathbb{E}V_j)^\theta}{\sum_k T_k(\mathbb{E}V_k)^\theta}
\end{aligned}
$$

where the second to last equality results from a change of variable: $u = t^{-\theta}\sum_k(\mathbb{E}V_k)^\theta$ so that $du = -\theta t^{-\theta-1}\sum_k(\mathbb{E}V_k)^\theta dt$ for $u \in (-\infty, 0)$. I multiply by -1 to change the limits of integration. The last line follows.

A3. Estimation Appendix

Numerical solution details. To solve for the optimal value function and policy functions for each type j, I use value function iteration (VFI). I discretize the AR(1) profitability shock process using the method in Tauchen (1986), which allows me to compute transition probabilities from/into 100 profitability states. The grid for formal labor, the endogenous state, has 250 points, spaced appropriately, with a non-binding upper bound of 2800 employees.

Estimation details. To implement the simulated method of moments, I construct the loss function in the following way: First, using a fixed panel of firms from the SBS and worker data from the LFS, I calculate moments during the period 2011-2013. For moments calculated from firm level data, I weigh each firm's contribution to the moments of a type by the probability that the firm belongs to that type, as explained in Section 2.5. I then bootstrap the data 500 times and recalculate all moments (keeping the probabilities of correct classification unchanged). I use the diagonal of the inverse of the variance covariance matrix from the bootstrapped data as the weighting matrix in the loss function. Second, I calculate the model equivalent of the data moments. I simulate 10,000 firms, for 103 periods. I discard the first 100 periods and only use the last three periods. The profitability shock draws remain fixed throughout the estimation and counterfactual exercises.

I implement the estimation in Matlab, using the Simulated Annealing algorithm to minimize the loss function. I estimate the shock process, the adjustment costs, and the discount factor using a panel of type 3 firms, truth tellers. I estimate the cost of hiding sales using a panel of type 2 firms, liars about sales. Lastly, I estimate the cost of hiding workers, and the parameters that govern the ex-ante costs of being any one type by using a panel of type 4 firms, liars about labor, as well as the aggregate shares of each type in the economy and the share of informal employment (calculated with LSF data).