Policy Research Working Paper 9515

Does Competition from Informal Firms Hurt Job Creation by Formal Firms?

Evidence Using Firm-Level Survey Data

Mohammad Amin

WORLD BANK GROUP
Development Economics
Global Indicators Group
January 2021

Policy Research Working Paper 9515

Abstract

The informal sector is an important source of livelihoods and jobs for a vast majority of people in developing countries. However, there is concern that the informal sector may undermine job creation in the formal sector. According to the "parasite" view of informality, informal firms can compete against formal firms, and often "unfairly" so as they do not have to comply with costly regulations and pay taxes. This "unfair" advantage makes it difficult for formal firms to compete against informal firms, implying a significant loss of formal sector jobs. Using firm-level survey data for manufacturing small and medium-size enterprises in 109 mostly developing countries, this study estimates the impact of competition from informal firms on the growth rate of employment among formal sector small and medium-size enterprises. The results show that the growth rate of employment declines significantly as competition from informal firms rises. According to the baseline specification, for each one standard deviation increase in informal competition, the employment growth rate declines by 1 percentage point. Consistent with the parasite view of informality, the negative impact on job growth is much larger when the business environment is less conducive to operating formally versus informally due to factors such high corruption, weak rule of law, more burdensome regulations, and high profit tax rate. Several checks are provided against endogeneity concerns.

This paper is a product of the Global Indicators Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The author may be contacted at mamin@worldbank.org.

Does Competition from Informal Firms Hurt Job Creation by Formal Firms? Evidence Using Firm-Level Survey Data

By: Mohammad Amin*

*Senior Economist, Enterprise Analysis Unit, DECEA, World Bank, Kuala Lumpur, Malaysia. Email: mamin@worldbank.org

I would like to thank Jorge Luis Rodriguez Meza for providing useful comments. All remaining errors are my own.

1. Introduction

The informal sector provides livelihoods and jobs to millions of individuals around the globe. According to ILO (2018), about 70 percent of the labor force in the emerging and developing economies is employed in the informal sector. Despite this contribution, there is concern that competition from informal sector firms may impede job creation among formal sector firms. A systematic and rigorous analysis of the issue, however, is lacking. The present paper attempts to fill this gap in the literature by estimating the relationship between competition faced by formal sector firms from informal sector firms (henceforth, informal competition) on the formal firms' growth rate of employment. It does so by using firm-level survey data for SMEs in the formal manufacturing sector of a large cross-section of mostly developing countries. Consistent with the "parasite" view of informality, we find that higher informal competition significantly lowers employment growth among formal firms. The decline is much larger in countries where the cost of operating in the formal vs. the informal sector is higher and the benefits lower. As expected, there is no significant impact of informal competition on the large formal sector firms.

In his celebrated study, Farrell (2004) pointed out that the negative impact of informal competition on the functioning of the formal firms is one of the three key problems with the informal sector. In the related literature, there are two prominent perspectives on how informality impacts the formal firms. These are the "parasite" view and the "dual economy" view (La Porta and Shleifer 2008, 2014). The dual economy view asserts that formal and informal firms are fundamentally different. Informal labor is a by-product of poverty and failure to find good jobs in the formal sector. Informal businesses are small, inefficient, of subsistence nature, and their productivity is too low to allow them to compete and survive in the formal sector. Thus, there is

no spillover from informal to formal sector (La Porta and Shleifer 2008, 2014; Rothenberg et al. 2016).

In contrast, the parasite view considers informal firms as capable of competing against formal firms. Informality arises from a rational decision of firms to operate in the informal sector to avoid costly regulations and taxes—the *rational exit model* (Maloney 2004, De Mel et al. 2011). Such informal firms operate on the fringes of the formal-informal divide, providing stiff competition to the formal sector firms. In general, on a level playing field, competition can lead to higher productivity and growth through industry rationalization and improved efficiency (Djankov and Murrell 2002). However, informal firms do not pay taxes and do not comply with costly regulations whereas formal firms must shoulder the cost of compliance. This "unfair" advantage allows informal firms to charge lower prices and "steal" market share from the formal firms. Thus, formal firms' production declines. Lower prices impinge on their profitability of the formal firms, reducing their incentive and restricting their access to critical funds for investment. Thus, informal competition adversely affects the growth of formal firms.

Another possibility is that high level of informality diverts resources away from the formal to the informal sector. This may lower employment growth among formal firms. Note that this effect is felt by all formal firms irrespective of whether they compete against informal firms or not. In contrast, our empirical analysis is based on the difference in growth rate of employment between formal firms that compete against informal firms and formal firms that do not do so. Thus, any impact of informality that falls equally on all the formal sector firms does not affect our results.

Informal competition and its impact on formal firms' performance other than employment growth has been discussed in the literature. In an early attempt, Tokman (1978) finds that in the city of Santiago, informal foodstuffs commercial establishments can successfully compete with

formal sector counterparts (modern supermarkets). Gonzalez and Lamanna (2007) analyze firm-level survey data on formal manufacturing firms collected by the World Bank's Enterprise Surveys for 14 countries in Latin America. Their findings suggest that about 40 percent of the firms in the region face significant informal competition, with sizeable variation across industries and firm sizes. Using Enterprise Surveys data for Nicaragua, Pisani (2015) explores the firm characteristics that determine the likelihood of formal firms to face informal competition. Mendi and Costamagna (2017) use data from Enterprise Surveys to estimate the impact of informal competition on formal sector firms' innovation. Amin and Okou (2020) also use Enterprise Surveys data and find that informal competition lowers labor productivity of the formal sector firms by 20-24 percent. To reiterate, none of these studies assess the impact of informal competition on formal firms' employment growth.

La Porta and Shleifer (2008) use Enterprise Surveys data to examine the expected effects of informality on the formal sector firms. However, the authors do not investigate the impact on employment growth. Instead, they analyze surveyed formal firms' responses to whether "anticompetitive and informal" practices are an obstacle for their business. The authors acknowledge that the question is noisy because anticompetitive practices can also result from the behavior of the formal sector firms. Firms' view of informality as an obstacle may also be driven by the perceived impact of informal activity on total revenues, profits, or productivity rather than employment growth. Moreover, La Porta and Shleifer (2008) use old data collected between 2002 and 2007.

Perry et al. (2007) argue that informality can have negative effects on formal firms' investment and innovation decisions because it reduces their market share and profitability. However, they do not provide any empirical evidence to support this claim. Rozo and Winkler

(2019) find that in Colombia, an increase in internally displaced persons (IDPs) adversely affected the total output of the formal sector firms. Their suggested interpretation is that an increase in IDPs entails an expansion of the informal sector, which in turn adversely affects the total output of firms in the formal sector. Loayza (1996) suggests that a large informal economy may lower tax revenues and increase public service congestion, which in turn, can negatively affect all firms in the economy. In Loayza's (1996) study however, these negative effects of informality do not explicitly result from the informal competition faced by the formal sector firms.

The present paper contributes to the above literature in several ways. First, it analyzes the impact of informal competition on the growth rate of employment among formal manufacturing SMEs. To the best of our knowledge, this is a first such attempt. The informal sector is often credited with providing employment especially to the poor and in the developing countries. Thus, it is important to properly understand the impact that the informal sector has on employment growth in the formal sector.

Second, informal firms are known to be very small and therefore unlikely to pose any problems for the large firms. Thus, a proper analysis requires that we distinguish between the effects of informality on SMEs and large firms. Studies using macro-level data cannot capture this heterogeneity. Our analysis based on firm-level data confirms that informal competition affects employment growth in SMEs but not the large firms.

Third, we pay due attention to endogeneity concerns. We do so in several ways. These are using the proportion of all other firms (that is, other than the firm in question) that compete against informal firms as a proxy for the level of informal competition experienced by the firm in question (cell average method); using a large number of firm- and country-level controls; showing that our main results follow several predications of the parasite view of informality which is unlikely if our

results suffer from reverse causality or if the informal competition variable we use were a mere proxy for other drivers of employment growth; showing that our results are consistent with some of the other predictions such as informality is more of the parasite type and therefore hurts formal firms more in countries where cheating on taxes is more justified, informality hurts large firms less (or not at all) compared to SMEs, and so on. To provide an example, we briefly discuss here the endogeneity check based on testing the predictions of the parasite view of informality. This is as follows.

Our main result for the relationship between employment growth and informal competition is consistent with the parasite view of informality. According to this view, informality hurts the formal sector firms more when the cost of operating formally vs. informally is higher and the benefit lower. This happens when for example, formal firms face higher corruption, more burdensome regulations, higher taxes, and weaker rule of law (Distinguin et al. 2016). Thus, if our results for the impact of informal competition are truly causal and not spuriously driven, then we should find that informal competition hurts formal firms more when the latter experience a less conducive business environment. However, if our informal competition variable is a mere proxy for other correlated drivers of employment growth, there is no reason for the spurious relationship to follow the stated predictions of the parasite view.

The empirical exercise reveals a large negative impact of informal competition on employment growth rate of formal firms. Figure 1 illustrates the point graphically.[1] For our baseline specification, for each one standard deviation increase in the level of informal competition

[1] Figures 2 and 3 provide variants of Figure 1. Figure 1 is based on mean level of employment growth in countries with high vs. low levels of informal competition. Figure 2 does the same using an alternative measure of employment growth used in our regressions below for robustness purpose. Figure 3 repeats Figure 1 but using median values of employment growth in countries with high vs. low levels of informal competition. For more details, see the note below Figures 1, 2 and 3.

the growth rate of employment declines by 1 percentage point. Further, as predicted above, this negative relationship is much larger (more negative) in countries where the cost of operating formally is higher and benefits lower as captured by higher corruption, more burdensome regulations, higher profit tax rate, and weaker rule of law. We also find no significant impact of informal competition on the growth rate of employment for large firms.

The plan of the remaining sections is as follows. In section 2, we describe the data and the variables used in the regressions. Section 3 contains the regression results for the baseline specification. Section 4 contains some robustness checks while the various endogeneity checks are provided in Section 5. The concluding section summarizes the main findings and suggests scope for future work.

2. Data and Main Variables

The main data source we use is firm-level survey data collected by the World Bank's Enterprise Surveys (ES). We complement the ES with other data sets such as the World Development Indictors (WDI, World Bank), Worldwide Governance Indicators (WGI, World Bank) and so on.

The ES are nationally representative surveys of the non-agricultural and non-financial private economy. The survey includes only the formal or registered firms with 5 or more employees. Thus, informal or unregistered firms and micro firms (fewer than 5 employees) are not covered by the survey. A common sampling methodology – stratified random sampling – is followed in all the surveys along with a common questionnaire.[2] The sample for each country is stratified by industry, firm-size, and location within the country. Weights are provided in the survey and used in all our regressions so that the sample is representative of the target population.

[2] Details of the sampling methodology and other related information are available at www.enterprisesurveys.org.

The sample used in this study is a pure cross-section of SMEs in the manufacturing sector. Based on the size stratification used by ES for sampling purposes, we define SMEs as all firms with fewer than 100 permanent full-time employees at the beginning of the growth accounting period.[3] We focus on the manufacturing sector to keep the sample relatively homogenous Each country and firm is included only once. The most recent round of ES in the country is used. The firms were surveyed between 2006 and 2018. For our baseline results, the sample consists of 20,295 manufacturing SME firms spread out across 109 countries.

2.1 *Estimation Methodology*

The baseline regression exercise involves estimating the following equation:

$$Y_{ij} = \alpha + \beta_1 Informal\ Competition_{-i,j} + \beta_2 Initial\ Employment_{i,j} + YFE + IFE + Firm\ Controls_{ij} + Country\ Controls_j + u_{ij} \qquad (1)$$

where subscript *i* denotes the firm, *j* the country to which the firm belongs. *Y* is the annual growth rate of employment over the last three years; *Informal Competition* is a measure of competition that the firm faces from the informal sector firms. Initial Employment is the (log of) level of employment at the beginning of the growth period. *YFE* denotes dummy variables for the survey year (Year fixed effects) and *IFE* denotes dummy variables for the industry to which the firm belongs (Industry fixed effects). *Firm Controls* and *Country Controls* include various firm- and country-level controls, respectively. *u* is the error term.

[3] The sample of large firms is used in one of the robustness checks (section 5.4) to compare the results for SMEs with those for large firms.

The estimation methodology used in the base regression and most of the other regressions is Ordinary Least Squares (OLS). In some of the robustness checks, we use the two stage least squares estimation, probit estimation, and the Conditional Mixed Process (CMP) estimation due to Roodman (2011). All the regressions use robust standard errors and clustered at the country level.[4] We pay due attention to endogeneity concerns (details below).

A formal definition of all the variables used is provided in Table 1. Summary statistics of the variables used in the baseline regressions and the correlations between them are provided in Tables 2 and 3, respectively.

2.2 *Dependent Variable*

The dependent variable is job creation or the annual growth rate of employment experienced by the firm over the last 3 years.[5] Employment here incudes all full-time permanent workers. The growth rate is computed is two different ways. First, it equals the traditional change in log of employment between the initial and final year (log difference) and divided by the number of years. The ratio is converted to percentage by multiplying by 100 (*Employment Growth*). The mean value of the variable equals 3.5 and the standard deviation is 16.2 percent.

For the second measure of employment growth, we compute the change in employment between the initial and final year and divide it by the average level of employment in the initial and final year. The ratio is divided by the number of years between the initial and final year and multiplied by 100 (to convert to percentage terms) to arrive at *Employment Growth* 1. By construction, the variable is bounded between plus/minus 100 and so is less affected by extreme

[4] In one robustness check (section 4.2), we exploit variation at the sub-national or city times industry level. For these results, we tighten the clustering which is at the city times industry level.
[5] For some countries, information on the growth rate is available over the last 2 years instead of 3 years.

values. In our baseline sample, the mean value of *Employment Growth* 1 equals 3.3 percent and the standard deviation equals 14.7 percent.

2.3 *Main Explanatory Variables*

The ES asked firms if they compete against informal sector firms. Responses to the question cannot be used directly in the regressions as they are likely to be endogenous to the firms' growth experience (reverse causality problem) and firm characteristics (omitted variable bias problem). One solution suggested in the literature is to proxy the firms' response by the average response of all other firms (that is, other than the firm in question) in the same "cell". The cell can be variously defined based on location, industry, firm-size, and so on. This strategy of using cell averages to overcome endogeneity concern has been discussed and used in the literature. See for example, Dollar et al. (2006), Fisman and Svensson (2007), De Rosa et al. (2010), Aterido et al. (2011), and Amin and Soh (2020).

Applied to this study, the strategy assumes that firms within the same cell are likely to face similar levels of informal competition (confirmed below). Further, it is highly unlikely that a given firm's employment growth rate has any effect on the likelihood of other firms in the cell facing informal competition (reverse causality problem). Similarly, own firm's characteristics that may affect its employment growth rate are unlikely to be correlated with the proportion of other firms in the cell that compete against informal firms (omitted variable bias problem), although the possibility cannot be ruled out completely. Thus, using cell averages mitigates the reverse causality and the omitted variable bias problem. The use of cell average also helps to control for potential measurement error if some firms choose not to respond or misreport (Pounov 2016).

Based on the discussion above, we define our main explanatory variable, *Informal Competition* (Country level cell average), as the proportion of all other firms (that is, other than the firm in question) in the country that report competing against informal firms. To ensure adequate thickness within the cells, all cells with fewer than 5 firms are excluded from the sample. In our baseline sample, the mean value of *Informal Competition* equals 0.49 and the standard deviation is 0.18.

In the robustness section, we show that our main result continues to hold when we use cell averages defined at a more disaggregated level. That is, when we proxy informal competition faced by a firm by the proportion of all other firms in the country times industry cell and in the city times industry cell that compete against informal firms (section 4.3). Industry is defined at the 2-digit ISIC Rev. 3.1 level and city is the sub-national region used for stratification purposes by the ES. These regressions exploit variation across firms within a country rather than across countries.

Note that our main regressions use the cell-average for informal competition directly in the regressions. To address any lingering concerns about how well the cell-average proxies for informal competition faced by a firm, we provide results from the two-stage estimation. That is, we first regress informal competition (dummy) at the firm-level on the cell-average of informal competition. Predicted values of informal competition are obtained. In the second stage, employment growth is regressed on the predicted or instrumented values of informal competition obtained in the first stage.

2.4 *Controls*

To further raise the confidence against endogeneity concerns, we show that our results are robust to several firm- and country-level controls. The controls are motivated by existing studies on informality and the broader literature on job creation by firms.

For the baseline regressions, we start with controls for firm-size or the level of employment at the beginning of the growth period, age of the firm, industry fixed effects and dummy variables for the year the ES was conducted (year fixed effects). Studies have shown that the initial firm-size and age of the firm matter in that smaller firms and younger firms tend to grow much faster than the rest (Ayyagari et al. 2011, Haltiwanger et al. 2010). For instance, for the case of the United States, Haltiwanger et al. (2010) find that younger firms grow much faster than the older firms. They also find that smaller firm-size is associated with higher growth, but this result disappears when they control for firm's age. In contrast and for the case of a large cross-section of developing countries, Ayyagari et al. (2011) find that both age and size matter in that younger and smaller firms grow faster than the rest and these effects are independent of each other. Our results for the relationship between employment growth and informal competition faced by the formal firms (henceforth, main result) could be spuriously affected if age and firm size vary systematically with the level of informal competition faced by the formal firms. Thus, we control for firm size and (log of) age of the firm. For firm size, we use the (log of) number of permanent full-time workers at the beginning of the growth period (*Initial Employment*).

Similarly, there may be industry-wide differences in the growth rate of employment as some industries, sunrise industries, offer greater potential to grow. We control for all such industry-wide factors using dummy variables for the industry to which the firm belongs (industry fixed effects). There are 7 industries in our sample defined at the 2-digit ISIC Rev. 3.1 level. The industry grouping is as used for stratification purposes by the ES.

Firm performance has also been linked with exporting activity, although it is not certain if exporting causes improvement in firm performance or the better performing firms self-select themselves into exporting (see for example, Bernard and Jensen 1999, Melitz 2003, Wagner 2007). Benefits may also accrue to firms that are part of a larger parent firm. These benefits may take the form of scale economies, better network, and information sharing. Thus, we control for the porporiton of firms' annual sales made abroad (*Exports*) and a dummy variable equal to 1 if the firm is part of a larger organization and 0 otherwise (*Multi-establishment firm*).

Several studies have analyzed gaps in firm performance stemming from the manager's gender and quality. Firms owned/managed by women tend perform worse than those owned/managed by men (Coleman 2000, Du Rietz and Henrekson 2000, Sabarwal and Terrell 2008). This is likely to be due to differences in the size of the firms managed by women, the difficulty that women face relative to men in obtaining finance, and social attitudes that tend to discourage women's economic participation. Thus, we account for the gender of the top manager with a dummy variable indicating if the top manager of the firm is a female or not (*Female Top Manager*). Differences in the quality of management, for reasons other than the gender of the top manager, have been found to impact firm productivity and performance (see, for example, Bloom and Van Reenen 2007, Syverson 2011, Pfeifer 2015). Education and experience of the top manager could be the possible factors driving management quality. We account for this in our regressions using a proxy measure which equals (log of) the number of years of experience the top manager of the firm has working in the industry (*Top Manager Experience*).

Our next firm-level controls include measures of the business environment. There is a large literature that indicates several aspects of the business environment that impact firm performance. At the broad level, these include physical infrastructure availability, law and order, enforcement

mechanisms for implementing rules and laws, and access to finance. If informal competition experienced by the formal firms varies systematically with the business environment in which the firms operate, our main results could suffer from the omitted variable bias problem. Thus, we control for the following measures of the business environment all defined at the firm-level: total hours of power outages experienced by the firm in a typical month over the last year (*Power Outages*), a dummy variable equal to 1 if the firm suffered losses due to crime during the last year and 0 otherwise (*Crime Losses*), a dummy variable equal to 1 if the firm reported the functioning of courts as a moderate, major or very severe obstacle for its day to day operations and 0 if it reported it is a minor obstacle or no obstacle (*Courts Obstacle*); and two measures of access to finance which include a dummy variable equal to 1 if the firm has overdraft facility and 0 otherwise (*Overdraft*), and how severe is (lack of) access to finance as an obstacle to the firms' day to day operations (*Finance Obstacle*) as reported by the firms. The finance obstacle is defined on a 0-4 scale with higher values implying a more severe obstacle.

Our next set of baseline controls includes macro-level variables. As mentioned above, the ES data we use was collected in different years across countries. Global annual shocks to employment growth could affect our results. Thus, we control for dummy variables indicating the year the ES was conducted in the country (Year fixed effects). Differences in overall economic development and income levels are a cause of concern for our main results. One reason for this is that informality tends to be lower in the richer countries. At the same time, richer countries grow slower than poorer countries (convergence effect). The structure of correlations here implies that failure to control for overall economic development can cause the estimated coefficient value of informal competition to be biased towards zero, weakening our main result (confirmed below). Thus, we control for (log of) GDP per capita, PPP adjusted and at constant 2011 international

dollars (*GDP per capita*). The variable is lagged by two years to allow for lagged effects. The data source is WDI, World Bank.

The availability of human capital is an important driver of growth and overall economic development. One possibility is that greater availability of human capital can push individuals to seek job in the formal sector rather than work in the informal sector. Higher levels of human capital are also likely to contribute positively to firm growth. Thus, failure to control for human capital availability can cause the observed informal competition and employment growth relationship to be more negative than what it is truly. To guard against this problem, we control for a measure of human capital that equals the gross enrollment rate in the country in primary education (*Primary Enrollment*). To avoid too many missing data, average values over three years prior to the year covered by the ES are used (for more details, see Table 1). Data source for the variable is WDI, World Bank.

Our last macro-level control is the level of inflation in the country. While low and moderate level of inflation may reflect fast growing countries, high levels may signal macro-economic instability with adverse effects on employment growth among private firms. Thus, a systematic correlation between employment growth and inflation cannot be ruled out. If informal competition also varies systematically with inflation, our main results could suffer from spurious correlation. Thus, we control for inflation as measured by the annual change in the consumer price index (*Inflation*). Data source for the variable is WDI, World Bank.

The controls mentioned above are used for our main or baseline estimation (baseline controls). There are additional factors that may impact employment growth and can also be potentially correlated with informal competition. In the robustness section, we check if our main result survives controls for these factors. We consider factors associated with differences in

economic activity and performance across world regions such as Africa, Asia, and so on. We control for all such factors using dummy variables for the world region where the country is located (region fixed effects). The regional groups are based on World Bank's classification and are as follows: Sub-Saharan Africa, Latin America and Caribbean, Middle East and North Africa, Eastern Europe and Central Asia, South Asia and East Asia and Pacific.

Recent investments in capital stock and R&D activity expand the productive capacity of the firm are likely to impact growth rate positively. Information in the ES is available on capital stock (replacement and book value of machines and equipment). However, due to the non-response, the information is missing for over half the firms in our sample. Hence, we use an alternative proxy measure of investment which is a dummy variable equal to 1 if the firm purchased fixed assets during the last year and 0 otherwise (*Firm Purchased Fixed Assets*). Amin and Soh (2020) and Islam et al. (2019) also use the same variable as a proxy for capital use. We control for R&D activity using a dummy variable equal to 1 if the firm spent on R&D activity during the last year and 0 otherwise (*Firm Spent On R&D*). Firms with foreign ownership have also been found to enjoy numerous benefits such as better access to modern technology, greater access to international markets, among others. These benefits are likely to impact firms' growth rate. Thus, we control for the proportion of firms' ownership that is with foreign individuals and companies (*Foreign Ownership*). Large cities are typically among countries' largest economic centers and, therefore, can offer agglomeration benefits: larger markets, better infrastructure to access markets and operate, a larger pool of workers, and greater technology spillovers (Rosenthal and Strange 2004). This could impact firm-performance including employment growth. To guard against the implied omitted variable bias problem, we control for a set of dummy variables for city-size

category based on total population: 1 million or more, between 1 million and 250,000, between 250,000 and 50,000, and less than 50,000. Data source for the variable is ES.

The next set of robustness controls include several variables for the quality of the business environment. The rationale for this has already been discussed above. The controls include measures of the regulatory burden on the private firms, corruption in the country and the quality of enforcement of the rules and regulations. For the regulatory burden, we use two proxy measures. These are the country-level average of the firms' senior management's time that is spent in dealing with business regulations (*Time Tax*) as reported by firms in the ES, and the Freedom from Business Regulations measure from Heritage Foundations Economic Freedom of the World (*Freedom From Regulation*). For corruption, we use the Control of Corruption measure taken from the World Bank's Worldwide Governance Indicators (*Control of Corruption*). Similarly, for the quality of enforcement of rule and laws, we use the Rule of Law measure from the World Bank's Worldwide Governance Indicators (*Rule of Law*). Except for *Time Tax* which is available only for the current year, all other business environment indicators are lagged by two years to allow for possible delays in their impact on employment growth.

Last, we control for market size. The assumption is that a larger market size offers greater opportunities for exploiting economies of scale and other scale related benefits. We proxy for market size by the (log of) total population in the country (*Population*) in the year prior to the year covered by the ES.

2.5 *Endogeneity Checks*

To further raise our confidence against endogeneity concerns, we test for several predictions about how informal competition affects the formal sector. The tests can be classified into three groups.

First, our main result above is consistent with the parasite view of informality. This view suggests that that informal firms enjoy an "unfair" advantage over formal firms because unlike formal firms, informal firms do not pay taxes and do not comply with costly regulations. Our first endogeneity check assumes that the "unfair" advantage enjoyed by informal firms is bigger when the business environment is less conducive to operating in the formal vis-à-vis the informal sector. This assumption or prediction is based on the specific mechanisms governing the formal-informal interface. Thus, if it holds in the data, it is most likely because the observed relationship between informal competition and employment growth is causal and not spuriously driven. In contrast, if the informal competition and employment growth relationship suffers from reverse causality problem or the omitted variable bias problem, there is little reason to expect the relationship to be stronger (more negative) when the business environment is less conducive to operating formally.

The idea that informality hurts formal firms more when the business environment is less conducive to operating formally has been discussed in the literature. For instance, Distinguin et al. (2016) estimate the impact of higher competition from informal sector firms on access to finance for formal sector SMEs. They find that the impact is negative and more so when the business environment is less beneficial or more costly to the formal firms due to higher corruption, higher taxes, weaker rule of law, and complicated business rules. They note that (page 19):

"A strong rule of law, however, raises the likelihood of informal activity detection; hence informal firms have to keep their operations small or to cease their activities completely, weakening their capacity to adversely affect formal SMEs' ventures. Moreover, complicated tax rules and a high level of corruption and bureaucracy constrain firms from entering or staying in the formal sector as they both lower the benefits and increase the costs associated with formality. High tax rates may also dissuade formal firms from innovating and investing leading them to lose their competitive edge against informal firms."

Following Distinguin et al. (2016), our first endogeneity check is that if the relationship between informal competition and employment growth that we find is truly causal then a given

increase in informal competition should have a larger negative impact on employment growth in countries with higher corruption, more burdensome (to private formal firms) regulations, higher taxes and weaker rule of law. We test for this by using interaction terms between our informal competition variable and the following variables: Control of Corruption variable from the Worldwide Governance Indicators, World Bank; Rule of Law variable from the Worldwide Governance Indicators, World Bank; Profit tax rate taken from Doing Business project, World Bank; and Freedom from Business Regulations taken form the Heritage Foundation's Economic Freedom of the World database.

Our second check against endogeneity concerns is like in spirit to the previous one. We argue that in countries where there is a relatively high proportion of population that believes that cheating on taxes is justified, informality is more likely to result from the desire to avoid taxes and costly regulations rather than from failure to find good jobs in the formal sector. Thus, the negative spillover effect from informal to formal sector as suggested by the parasite view of informality can be expected to be stronger in countries where cheating on taxes in more strongly justified. Of course, this prediction assumes that the impact of informal competition on employment growth that we observe in the data is truly causal. There is little reason for the prediction to hold if informal competition is a mere proxy for other correlated drivers of employment growth or if our main result suffers from reverse causality. To test for the stated prediction, we use interaction term informal competition variable and a dummy variable equal to 1 if the reported justification level for cheating on taxes is above its median level across countries and 0 otherwise. Some adjustments are made in constructing the variable to avoid too many missing values. For details on how the dummy variable is constructed, see Table 1. Data source for the variable is World Value Surveys.

For the third group of endogeneity tests, we argue that for a given increase in informal competition, certain types of firms are likely to suffer more from informal competition than the others. Thus, we check that a unit increase in informal competition has a much smaller negative impact on large firms vs. SMEs, exporting firms vs. non-exporting firms, and financially unconstrained vs. financially constrained firms. The rationale for these tests is as follows. It is well-known that large firms operate in very different product markets and therefore unlikely to be much affected by informal competition. The logic can be extended to exporting firms that also operate in very different (international) markets than the non-exporting and informal firms. Regarding financial constraints, our endogeneity check is based on the findings of Distinguin et al. (2016). As mentioned above, the study by Distinguin et al. (2016) shows that informal competition worsens access to external finance for SMEs in the formal sector. In other words, one of the ways in which informal competition can hurt SMEs is by reducing their access to finance. Thus, SME firms that do not have proper access to finance are likely to suffer more from a given increase in informal competition than the rest. We expect the stated predictions to hold if informal competition has a causal impact on employment growth but there is no strong reason for it to hold otherwise.

3. Base Regression Results

For all the regressions discussed below, we control for initial employment as this is perhaps one of the most important predictors of employment growth. The remaining baseline controls are added sequentially.

The base regression results are provided in Table 4. These results reveal a large, negative and statistically significant relationship between the growth rate of employment and informal

competition. Without any other controls (except for initial employment), the estimated coefficient value of the informal competition variable equals -5.8. While the coefficient value is large, it is significant only at the 10 percent level. Controlling for GDP per capita, one of our main controls, causes the estimated coefficient value to increase (in absolute value) from -5.8 above to about -6.7 (not shown) and it is now significant at the 5 percent level. The coefficient value increases further to about -7 with controls for the industry and year fixed effects added to the previous specification (column 2). Adding also the remaining baseline controls leads to some decline in the coefficient value of informal competition but it remains large, negative and significant at the 5 percent level. For our final specification that includes all the baseline controls, the coefficient value equals -5.7, significant at the 5 percent level (column 5). The estimate implies that for each 10 percentage points increase in the proportion of firms in the country that compete against informal sector firms, the growth rate of employment declines by about 0.57 percentage point. This is a large increase given that the mean employment growth rate is 3.5 percent. Alternatively, for each one standard deviation increase in informal competition variable (or about 18 percentage points), the associated decline in employment growth rate equals about 1 percentage point.

Results for the various controls are mostly along expected lines. As found elsewhere in the literature, younger and smaller firms (in terms of initial employment) show significantly higher growth rates than the rest. Higher GDP per capita is associated with a significantly lower employment growth rate, confirming the general finding in the literature that poor countries grow faster than the rich countries. Greater exporting activity is associated with higher growth rate while more power outages is associated with lower growth rate. However, the latter result should be treated with due caution as it is significant only at the 10 percent level. For the remaining firm characteristics, we find that firms that are part of a larger parent firm (Multi-establishment firm)

and firms that have overdraft facility grow faster than their respective counterparts, with the difference in the growth rate being significant at the 1 percent level. Last, higher primary education enrollment rates and higher inflation are both associated with higher growth rate (significant at the 10 percent and 1 percent level, respectively).

4. Robustness

4.1 *Alternative growth measure*

For our first robustness check, we repeat the base regressions above using the alternative measure employment growth rate (*Employment Growth* 1). The regression results are provided in Table 5. These results are like the ones discussed above. That is, they show that there is a large and negative relationship between informal competition and growth rate. The relationship is statistically insignificant (but close to the 10 percent level) in column 1 where we do not include any controls (other than initial employment). As above, controlling for GDP per capita causes the estimated coefficient value of informal competition to increase (in absolute value) from -5.4 to -6.4 and becomes significant at the 5 percent level (not shown). It remains significant at the 5 percent when we add the remaining baseline controls to the previous specification (columns 2-5). For our final specification with all the baseline controls included, the estimated coefficient value of informal competition equals -5.1. The estimate implies that for one standard deviation increase in the value of the informal competition variable, the associated decline in the growth rate equals 0.9 percentage points (against the mean growth rate of 3.3 percent).

We can confirm that not just the base regression results, but all the regression results discussed anywhere in this paper hold for both measures of employment growth rate. Robustness

results discussed below using the alternative growth measure (*Employment Growth* 1) are available on request from the author.

4.2 *Additional controls*

Starting with the final baseline specification, we check if the results are robust to additional controls. Thus, in Table 6, we add controls for region dummies (LAC, EAP, etc.) in column 1; in column 2, we include controls for R&D activity, foreign ownership, purchase of fixed assets, and city-size dummies; controls for country-size and several dimensions of the business environment quality (corruption, regulation, etc.) are included in columns 3 and 4. Our result for the negative relationship between informal competition and employment growth survives. That is, the estimated coefficient value of informal competition remains negative, large and it is significant at the 1 percent level in all the specifications. The coefficient value is larger (more negative) due to the additional controls. For instance, starting with the coefficient value of -5.7 in the final baseline specification (column 5, Table 4), it rises to -7.1 with all the additional controls included in the specification (column 4, Table 6).

4.3 *Within-Country Variation*

The next robustness check involves using cell averages of informal competition at a more disaggregated level. We experimented with country times industry and city times industry level, where city is the sub-national sampling region used for stratification purposes by the ES. The industry groupings used is the same as above (2-digit ISIC Rev. 3.1). Regressions using cell averages at the country times industry level control for all country-wide factors via country dummies (Country fixed effects). Similarly, regressions using city-industry level cell averages

control for all city-wide factors through city dummies (City fixed effects). Note that city fixed effects absorb country fixed effects. The identification of the main results here comes from variations across firms within a country rather than across countries. Thus, cross-country differences in factors such as overall economic development, quality of institutions, historical and geographical influences on employment growth, culture, and the quality of the business environment that may lead to omitted variable bias problem are accounted for.

The regression exercise involves repeating the baseline regressions (as in Table 4) with country or city dummies included in all the specifications and using cell averages of informal competition at the country-industry or city-industry level.

Our main result for the negative relationship between informal competition and employment growth continues to hold with country-industry and city-industry cell averages. We present the results for the city-industry level as they are on the conservative side. Table 7 contains the results. The results show that the estimated coefficient value of informal competition is large, negative and significant at the 5 percent level in all the specifications. It is roughly equal in magnitude to the ones in the baseline model. For instance, for the final specification with all the baseline controls included, the coefficient value equals -5.8 (column 4, Table 7) compared to -5.7 in the baseline model (column 5, Table 4).

4.4 *Two-Stage Estimation*

So far, we used the cell averages of informal competition directly in the regressions. An alternative is two-stage estimation where in the first stage we regress the dummy equal to 1 if the firm competes against informal sector firms and 0 otherwise (henceforth, informal competition dummy) on its country level cell average. Predicted values of the informal competition dummy are obtained

and used as instruments for informal competition faced by the firm. In the second stage, we regress employment growth rate on the instrumented values of informal competition obtained in the first stage. This two-stage estimation is slightly complicated because the dependent variable in the first stage is a binary (dummy) variable. The literature offers a few choices here which we use to provide three separate regression estimates. First, we disregard the binary nature of the informal competition dummy. Thus, we perform the usual two stage instrumental variable estimation. Second, we use the probit model in the first stage. The predicted values obtained from the first stage are used in the OLS estimation in the second stage. We do so by independently running the first and second stage regressions. Third, we repeat the previous estimation with the difference that first and second stage estimations are done jointly (instead of independently) using the Conditional Mixed Process (CMP) estimator due to Roodman (2011).

Regression results using the three estimation methods described in the previous paragraph are provided in Tables 8-10. All these results show that in the first stage, informal competition dummy at the firm-level is strongly and positively correlated with its country-level cell average (significant at the 1 percent level). The second stage results are like the baseline results and show a large and negative relationship between (instrumented values of) informal competition dummy and employment growth rate. Like the baseline results, the relationship is significant at the 10 percent level without any controls (except for the initial level of employment). Controlling for GDP per capita causes the relationship to become significant at the 5 percent level and it remains so when the remaining baseline controls are included in the specification. The estimated coefficient value of informal competition is smaller in the CMP estimation and higher in the other two stage estimations. For instance, with all the baseline controls included in the specification, the estimated coefficient value of (instrumented values of) informal competition in the second stage equals -5.7

in the baseline model (column 5, Table 4), -2.7 in the CMP estimation (column 5, Table 10) and about -7.4 in the other two stage estimations (column 5, Table 8 and 9).

5. Endogeneity checks

In this section, we discuss checks against endogeneity problem with our main result. As will be evident below, our results pass the endogeneity checks when we control for the interaction term between our informal competition variable and GDP per capita but not otherwise. The interaction term between informal competition and GDP per capita is found to be negative, implying that a given increase in informal competition hurts formal firms more in the richer countries in our sample. This may seem counter-intuitive since informality is associated with less development and therefor one may expect the adverse effects of informality on the formal sector to be bigger in the poorer countries. However, our result can be explained in several ways. First, the richest countries in our sample are not the high-income OECD countries but the middle and upper middle-income countries. Thus, it is still possible that the adverse effects of informality are lower for the high-income countries compared to the relatively poorer (low-income and middle-income) countries.

The second explanation is based on the availability of resources. That is, the informal sector is known to be under privileged in terms of access to finance, physical infrastructure, etc. This may be particularly so at low levels of income when resources are limited. As development proceeds from very low levels, the informal sector gets better access to resources and more so than the formal sector. This implies that informal sector can compete more vigorously with the formal sector in countries at intermediate income levels than at very low level of income.

The third argument relates to the dynamics of growth at different income levels. It is well known that richer countries have lower growth rates (convergence effect) reflecting limited growth

opportunities compared to the poorer countries. The "unfair" advantage enjoyed by the informal firms over formal firms discussed above hurts more when ceteris paribus, growth opportunities are more limited.

We provide some evidence consistent with the second and third explanations mentioned above. Firms in the ES were asked if informal firms are no obstacle, minor, moderate, major or very severe obstacle to their current operations. If limited resources in the poorer countries or limited growth opportunities in the richer countries magnify the problem posed by informal competition, firms in richer countries should report a larger increase in obstacle severity level than in poorer countries when they compete against informal competition compared to when they do not compete against informal firms. To test this, we define our dependent variable as a dummy equal to 1 if the firm reports informal sector as a major or very severe obstacle and 0 otherwise. The main explanatory variable is the interaction term between GDP per capita and a dummy equal to 1 if the firm competes against informal firms and 0 otherwise. The controls used are as in the baseline model (Table 4) plus the dummy for whether the firm competes against informal firms or not. Logit estimation results provided in Table 11 confirm our prediction. That is, the relationship between informality as a major obstacle and informal competition is positive and it becomes significantly stronger (more positive) as GDP per capita rises.

5.1 *Business Environment and Informality*

The first prediction we test is that the negative relationship between informal competition and employment growth should be stronger (more negative) in countries with weaker rule of law, higher corruption, more burdensome business regulations, and higher taxes. We test for these predictions using interaction terms between the informal competition variable and the business

environment measures. The interaction terms are added to the baseline specifications shown in Table 4. Regression results are provided in Tables 12-15. Note that these results control for the interaction term between GDP per capita and informal competition. As discussed above, our results for the interaction term between the informal competition and business environment measures are weak (insignificant) without the control for the GDP per capita and informal competition interaction term.

Regression results in Tables 12-15 confirm the prediction that informal competition hurts more when the business environment is less conducive to operating in the formal vs. informal sector. The interaction terms between informal competition and the business environment measures carry the expected sign and are significant at the 5 percent level.

5.2 *Exporting and Access to Finance*

The next test is that the negative impact of informal competition on employment growth should be smaller (less negative) for firms that export and for firms that are financially less constrained. As above, we test for these predictions by using interaction terms between informal competition and exporting activity captured by a dummy variable equal to 1 if the firm exports and 0 otherwise; and the interaction term between informal competition and the dummy for firm having overdraft facility. In separate regressions, these interaction terms are added to the baseline specifications (as shown in Table 4).

Regression results for the interaction term with exporters are provided in Table 16. Results for the interaction term with the overdraft facility are provided in Table 17. The results show that as predicted, both the interaction terms are large, positive and significant at the 5 percent level. These results hold even when we allow the impact of informal competition to vary with the level of GDP per capita and/or firm-size proxied by *Initial Employment* (columns 3-6, Tables 16 and

17). This confirms our prediction that a given increase in informal competition hurts non-exporters and the financially constrained firms more.

5.3 *Cheating on Taxes*

Regression results for the interaction terms between informal competition and the proportion of population that believes that cheating on taxes is justified are provided in Table 18. These results use the baseline specifications with the interaction term added. To ensure that our variable for cheating on taxes justification is not picking up the effects of overall economic development, we provide results with and without controlling for the interaction term between informal competition and GDP per capita. As discussed above, the prediction is that informal competition has a larger negative impact on employment growth rate in countries where more people justify cheating on taxes.

The results in Table 18 confirm the stated prediction. That is, the interaction term between informal competition and the dummy for high level of justification for cheating on taxes is negative and large. Without any other controls (except for the initial level of employment), the interaction term is relatively small and insignificant at the 10 percent level. However, controlling for GDP per capita causes the interaction term to increase (in absolute value) from -6 to -19 and become significant at the 5 percent level (column 2). The coefficient remains significant at the 5 percent level and roughly unchanged in value when we add the remaining baseline controls to the specification (column 3). Controlling for the interaction term between informal competition and GDP per capita does not lead to any noticeable change in the results (see columns 4-6).

5.4 *SMEs vs. Large Firms*

Our last endogeneity test involves comparing the impact of informal competition on SMEs vs. large firms. To reiterate, the prediction is that unlike the SMEs, large firms are not affected or at least less adversely affected by a given increase in informal competition than SMEs. Regressions results for the baseline specifications for the sub-sample of large firms (more than 100 workers in the initial period) are provided in Table 19. As expected, the results reveal that there is no significant relationship between informal competition and employment growth. The estimated coefficient value is positive and ranges between 1.3 and 3.5 percent. However, it is statistically insignificant at the 10 percent level. This holds for all the baseline specifications. Thus, our prediction that informal firms pose little problem for large firms is confirmed.

6. Conclusion

The informal sector is large and provides jobs to millions of individuals across the globe. However, there is concern about its impact on job creation in the formal sector. Informal firms do not pay taxes and do not comply with costly regulations that their formal sector counterparts must shoulder. The "unfair" advantage so enjoyed by the informal sector firms allows them to "steal" market share from the formal firms, lowering production and growth of the latter. Thus, competition from informal firms can hurt employment growth in the formal sector. Our empirical results show that this is indeed the case for formal SMEs in the manufacturing sector of a large cross-section of mostly developing countries. We also find that a given increase in informal competition hurts some countries and firms more than others. Specifically, formal firms in countries where the cost of operating in the formal vs. informal sector is higher or benefits lower suffer more from informal competition. Large firms, exporting firms and firms with better access to finance are found to be less adversely affected by informal competition than their respective counterparts. These

heterogeneities in the impact of informal competition on formal firms' employment growth are consistent with the "parasite" view of informality and the general findings in the related literature.

Several issues remain to be explored. We provide a few examples to illustrate the point. First, our results are restricted to employment growth. It would be interesting to analyze how informal competition affects other aspects of firm performance such as sales growth, innovation, investment in physical capital, tendency to export, level and growth of factor (labor, capital and total) productivity, and so on. Second, the firm-level data that we use cover formal firms with 5 or more employees. Thus, micro firms with fewer than 5 employees are excluded. This is an important exclusion given that informal firms are very small and therefore likely to a be bigger challenge for micro firms than the SMEs. Extending the empirical exercise of this paper to micro firms is a fruitful area for future research. Third, while there are several papers that analyze the extent of informal activity in a country, the size of the spillover effect from informal to formal sector firms has not been properly explored. Thus, exploring the factors that cause informal firms to compete more aggressively against formal firms rather than operate independently is another avenue for future research. Last, we found above that the impact of informal competition varies depending on exporting activity and financial constraints faced by the formal firms. Our motivation for exploring these was to raise our confidence against endogeneity concerns and not to explore heterogeneity per se. However, heterogeneity in the impact of informal competition across industry groups, firm-size groups within SMEs, gender of the top manager, and so on can be explored.

References

Amin, M. and Y. C. Soh (2020). Does Greater Regulatory Burden Lead to Higher Corruption: Evidence Using Firm-level Survey Data for Developing Countries. Forthcoming *World Bank Economic Review*. https://doi.org/10.1093/wber/lhaa007

Amin, M. and C. Okou (2020). Casting a Shadow: Productivity of Formal Firms and Informality. *Review of Development Economics*, 24(4): 1610-1630.

Aterido, R., Hallward-Driemeier, M., & Pages, C. (2011). Big Constraints to Small Firms' Growth? Business Environment and Employment Growth Across Firms. *Economic Development and Cultural Change*, 59 (3): 609 – 647.

Ayyagari, Meghana, Asli Demirguc-Kunt, and Vojislav Maksimovic. (2011). Small vs. Young Firms Across the World. Policy Research Working Paper 5631, World Bank, Washington, DC.

Bernard, A. B., & Jensen, J. B. (1999). Exceptional exporter performance: Cause, effect, or both? *Journal of International Economics, 47*(1), 1-25.

Bloom, N., & Van Reenen, J. (2007), "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics* 122 (4): 135 – 1408.

Coleman, S. (2000). Access to capital and terms of credit: A comparison of men-and women-owned small businesses. *Journal of Small Business Management, 38*(3), 37-52.

De Mel, S., McKenzie, D., &Woodruff, C. (2011). What is the Cost of Formality? Experimentally Estimating the Demand for Formalization. Working Paper, University of Warwick, Coventry.

De Rosa, D., Gooroochurn, N., & Görg, H. (2010). Corruption and Productivity: Firm-level Evidence from the BEEPS Survey. World Bank Policy Research Working Paper 5348, The World Bank, Washington DC.

Distinguin, I., Rugemintwari, C., & Tacneng, R. (2016). Can Informal Firms Hurt Registered SMEs' Access to Credit? *World Development*, 84: 18 – 40.

Djankov, S., and Murrell, P. (2002). Enterprise Restructuring in Transition: A Quantitative Survey. *Journal of Economic Literature* 40(3): 739-792.

Dollar, D., Hallward-Driemeier, M., & Mengistae, T. (2006). Investment Climate and International Integration. *World Development*, 34 (9): 1498 – 1516.

Du Rietz, A. & Henrekson, M. (2000). Testing the Female Underperformance Hypothesis. *Small Business Economics*, 14 (1): 1 – 10.

Farrell, D. (2004). The Hidden Dangers of Informal Economy. *McKinsey Quarterly*, 3: 27 – 37.

Fisman, R., & Svensson, J. (2007). Are Corruption and Taxation Really Harmful to Growth? Firm Level Evidence. *Journal of Development Economics*, 83 (1): 63 – 75.

Gonzalez, A., & Lamanna, F. (2007). Who Fears Competition from Informal Firms? Policy Research Working Paper Series 4316, World Bank, Washington DC.

Haltiwanger, John C., Ron S. Jarmin, and Javier Miranda. (2010). Who Creates Jobs? Small vs. Large vs. Young. NBER Working Paper 16300, NBER.

ILO (2018). Women and Men in the Informal Economy: A Statistical Picture. Third edition. Geneva: International Labor Office, International Labor Organization (ILO).

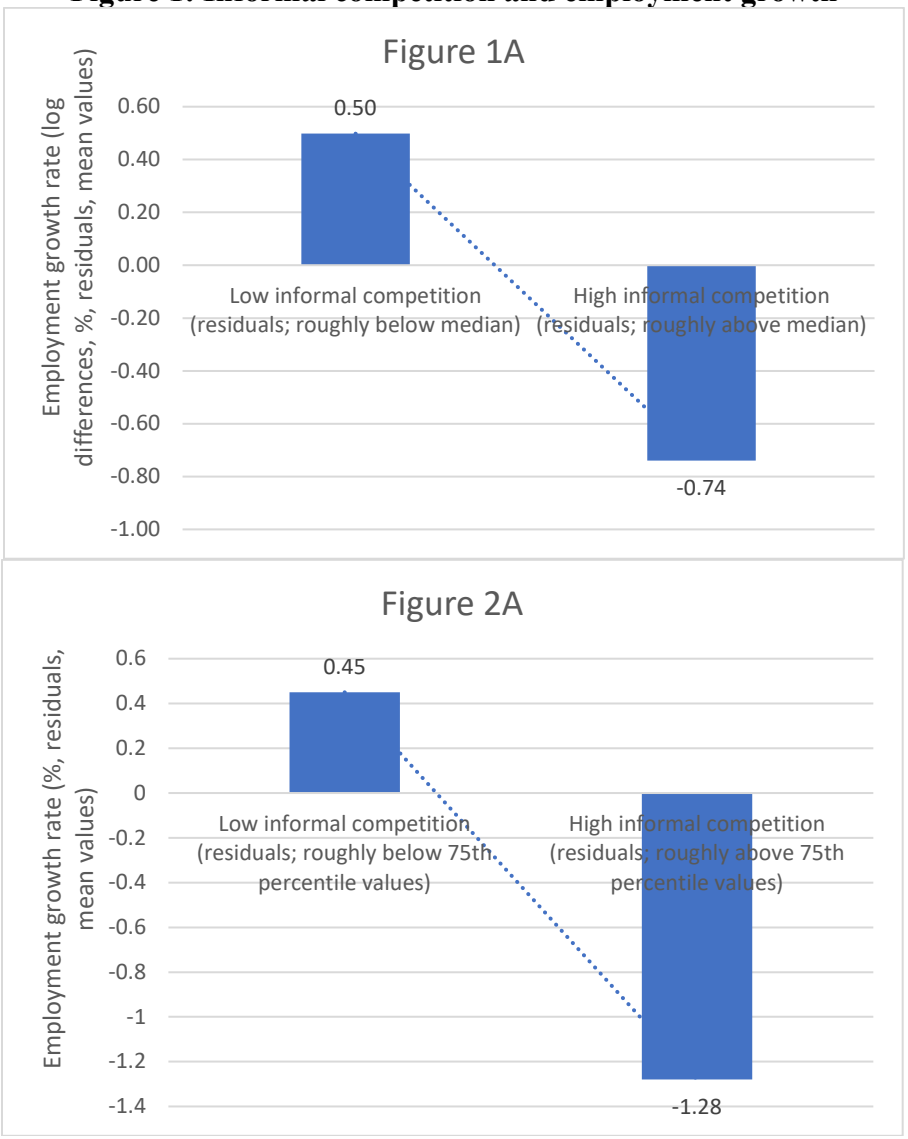
Islam, A., Palacios Lopez, A., & Amin, M. (2019). Decomposing the Labour Productivity Gap between Migrant-Owned and Native-Owned Firms in Sub-Saharan Africa. *Journal of Development Studies*, 55(9): 2065-2082.

La Porta, R., & Shleifer, A. **(**2008). The Unofficial Economy and Economic Development. *Brookings Papers on Economic Activity* Fall: 275 – 352.

La Porta, R., and Shleifer, A. (2014). Informality and Development. *Journal of Economic Perspectives*, 28 (3): 109 – 126.

Loayza, N. (1996). The Economics of the Informal Sector. A Simple Model and Some Empirical Evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45: 129-162.
Maloney, W. F. (2004). Informality Revisited. *World Development*, 32 (7): 1159 – 1178.
Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica, 71*(6), 1695-1725.
Mendi, P., & Costamagna, R. (2017). Managing Innovation Under Competitive Pressure From Informal Producers. *Technological Forecasting and Social Change*, 114: 192 – 202.
Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D., & Saavedra-Chanduvi, J. (2007). Informality: Exit and Exclusion. World Bank, Washington, DC.
Pfeifer, C. (2015), "The Nexus Between Top Managers' Human Capital and Firm Productivity," *Applied Economics Letters* 22 (12): 982 – 986.
Pisani, M. J. (2015). Does Informality Impact Formal Sector Firms? A Case Study From Nicaragua. *Journal of Developing Areas* 49(2): 317-334.
Pounov, C. (2016). Corruption's Asymmetric Impacts on Firm Innovation. *Journal of Development Economics*, 118: 216 – 231.
Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-process Models with CMP. *Stata Journal*, 11(2): 159–206.
Rosenthal, S. S., & Strange, W. C. (2004). Chapter 49—Evidence on the Nature and Source of Agglomeration Economies. *Handbook of Regional and Urban Economics*, 4: 2119 – 2171.
Rothenberg, A.D., Gaduh, A., Burger, N.E., Chazali, C., Tjandraningsih, I., Radikun, R., Sutera, C., Weilant, S. (2016). Rethinking Indonesia's informal sector. *World Development*, 80: 96–113.
Rozo, S., and Winkler, H. J. (2019). Is Informality Good for Business? The Impacts of IDP Inflows on Formal Firms. *Journal of Human Resources*, forthcoming.
Sabarwal, S., & Terrell, K. (2008). *Does gender matter for firm performance? Evidence from Eastern Europe and Central Asia*: The World Bank.
Syverson, C. (2011). What determines productivity? *Journal of Economic literature, 49*(2), 326-365.
Tokman, Victor E., (1978), Competition between the informal and formal sectors in retailing: The case of Santiago, *World Development*, 6, issue 9-10, p. 1187-1198.
Wagner, J. (2007). Exports and productivity: A survey of the evidence from firm-level data. *World Economy, 30*(1), 60-82.

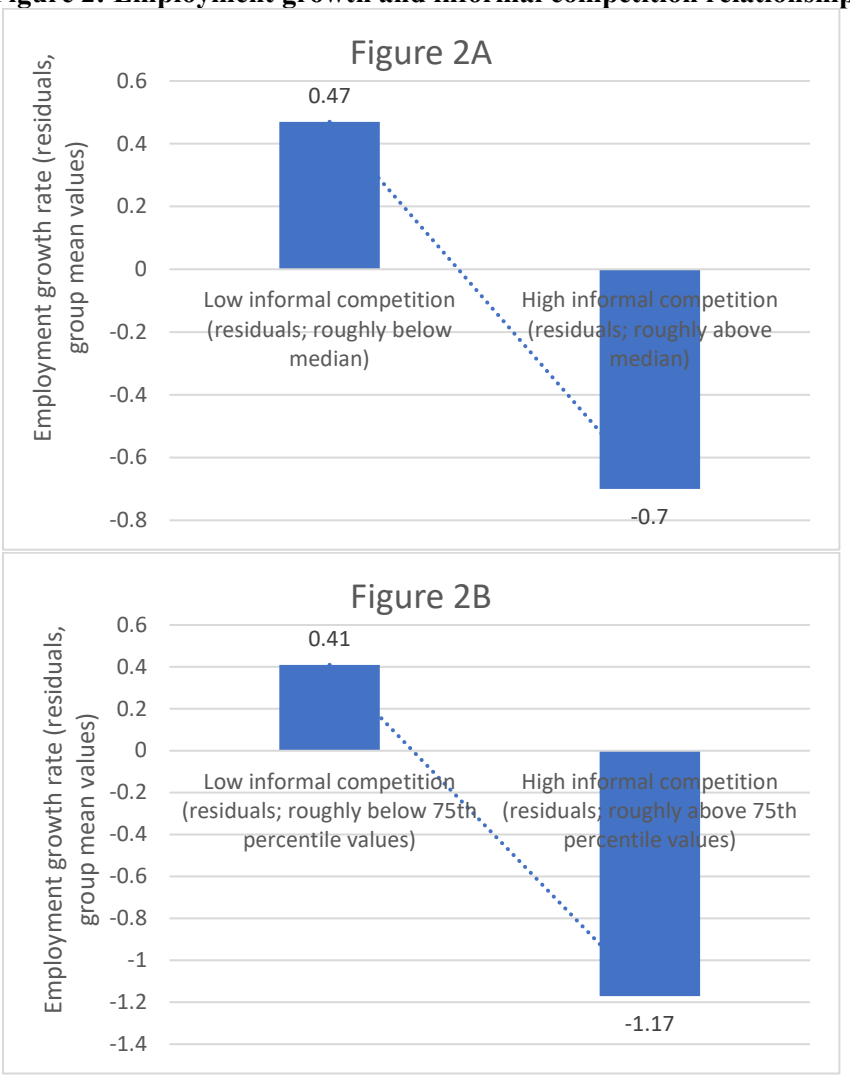
Figure 1: Informal competition and employment growth



Source: Author's own calculations based on Enterprise Surveys data.

Note: 1) Employment growth rate is calculated as the log difference in the level of employment (log of final year employment minus log of initial year employment) divided by the number of years between the initial and last year and multiplied by 100 to convert to percentage terms. The figures are a partial scatter plot of the residuals of employment growth rate (Y axis) plotted against the residuals of informal competition (X-axis). The residuals are obtained by regressing the employment and informal competition (separately) on industry fixed effects, year fixed effects, initial level of employment (logs) and (log of) GDP per capita. Informal competition variable equals the proportion of firms in the country (excluding own firm) that report competing against informal firms. 3) Group averages (mean values) of the residuals are shown. Close to the median value of the informal competition residuals is used to categorize the two groups in Figure 1A. The cut-off value used for the two groups in Figure 1B is close to the 75th percentile value of informal competition residuals.

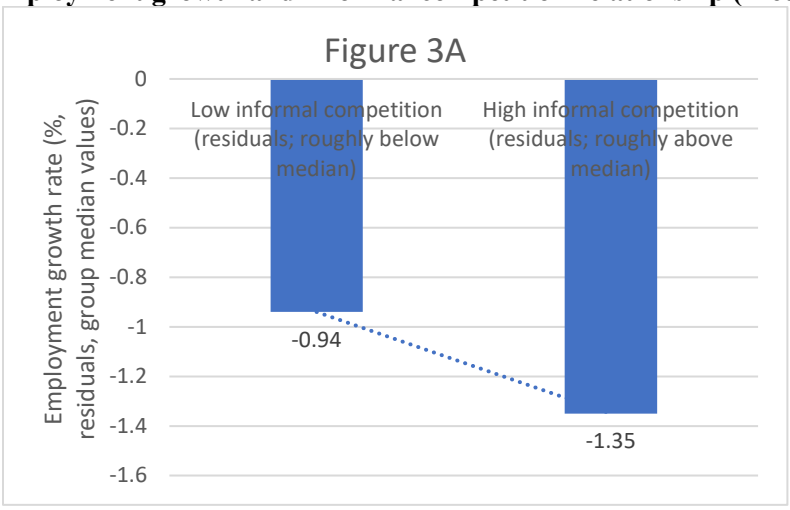
Figure 2: Employment growth and informal competition relationship

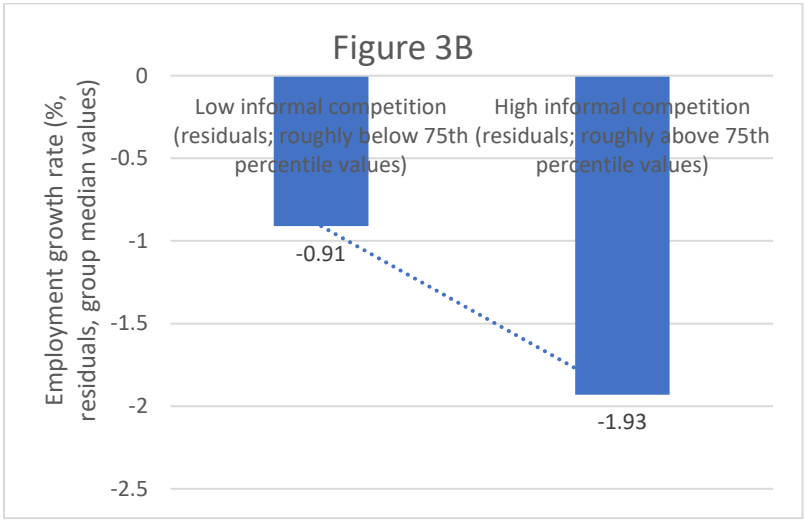


Source: Author's own calculations based on Enterprise Surveys data.

Note: 1) Employment growth rate is calculated as the difference in the level of employment (final year employment minus initial year employment) divided by the simple average (mean) of employment in the final and initial year. The ratio is then divided by the number of years between the final and initial year; the figure is multiplied by 100 to convert to percentage terms (Haltiwanger method). 2) The figures are a partial scatter plot of the residuals of employment growth rate (Y axis) plotted against residuals of informal competition (X-axis). The residuals are obtained by regressing the employment and informal competition (separately) on industry fixed effects, year fixed effects, initial level of employment (logs) and (log of) GDP per capita. Informal competition variable equals the proportion of firms in the country (excluding own firm) that report competing against informal firms. 3) Group averages (mean values) of the residuals are shown. Close to the median value of the informal competition residuals is used to categorize the two groups in Figure 2A. The cut-off value used for the two groups in Figure 2B is close to the 75th percentile value of informal competition residuals.

Figure 3: Employment growth and informal competition relationship (Median values)





Source: Author's own calculations based on Enterprise Surveys data.

Note: 1) Employment growth rate is calculated as the log difference in the level of employment (log of final year employment minus log of initial year employment) divided by the number of years between the initial and last year and multiplied by 100 to convert to percentage terms. The figures are a partial scatter plot of the residuals of employment growth rate (Y axis) plotted against the residuals of informal competition (X-axis). The residuals are obtained by regressing the employment and informal competition (separately) on industry fixed effects, year fixed effects, initial level of employment (logs) and (log of) GDP per capita. Informal competition variable equals the proportion of firms in the country (excluding own firm) that report competing against informal firms. 3) Group averages (median values) of the residuals are shown. Close to the median value of the informal competition residuals is used to categorize the two groups in Figure 3A. The cut-off value used for the two groups in Figure 3B is close to the 75th percentile value of informal competition residuals.

Table 1: Description of Main Variables	
Variable name	**Description of variable**
Employment Growth (log difference, %, annual)	Log of the total number of permanent full-time workers at the firm in the final year (last fiscal year) minus the log of the total number of permanent full-time workers at the firm in the initial year (beginning of the growth period which is 2 fiscal years ago for most countries and 3 fiscal years ago for other countries). The log difference is then divided by the number of years between the initial and final year and the multiplied by 100 to covert to percentage terms. Source: Enterprise Surveys. www.enterprisesurveys.org
Employment Growth 1	Total number of permanent full-time workers in the final year (last fiscal year) minus the same in the initial year (beginning of the growth period which is 2 fiscal years ago for most countries and 3 fiscal years ago for other countries) divided by the average (mean value) number of permanent workers in the final and initial year. The ratio is divided by the number of years between the final and initial year and multiplied by 100 to convert to percentage terms. Source: Enterprise Surveys. www.enterprisesurveys.org
Informal Competition (Country level cell average)	Proportion of firms in the country excluding the firm in question that report competing against informal firms. Source: Enterprise Surveys. www.enterprisesurveys.org
Initial Employment (logs)	Log of total number of permanent full-time workers at the firm in the initial year (beginning of the growth period which is 2 fiscal years ago for most countries and 3 fiscal years ago for other countries). Source: Enterprise Surveys. www.enterprisesurveys.org
GDP per capita (lagged, logs)	Log of GDP per capita lagged by 2 years from the fiscal year covered by the Enterprise. GDP per capita values are in constant 2011 International Dollars and PPP adjusted. In cases where the Enterprise Survey spanned more than one calendar year, the GDP per

	capita figure is the weighted average over the calendar years covered (lagged by 2 years), where the weights used are the proportion of firms covered in the calendar year (lagged by two years). Source: World Development Indicators, World Bank.
Age of Firm (logs)	Log of age of the firm. The age of the firm is defined as the current year minus the year the firm first started operations. Source: Enterprise Surveys. www.enterprisesurveys.org
Manager Experience (logs)	Log of the number of years of experience the top manager of the firm has working in the industry. Source: Enterprise Surveys. www.enterprisesurveys.org
Exports (proportion of sales)	The proportion of the firms' annual sales made abroad in the last fiscal year. Only direct exports are included. Source: Enterprise Surveys. www.enterprisesurveys.org
Female Top Manger Y:1 N:0	Dummy variable equal to 1 if the top manager of the firm is a female and 0 otherwise. Source: Enterprise Surveys. www.enterprisesurveys.org
Multi-establishment firm Y:1 N:0	Dummy variable equal to 1 if the firm is part of a larger parent firm and 0 otherwise. Source: Enterprise Surveys. www.enterprisesurveys.org
Crime Losses Y:1 N:0	Dummy variable equal to 1 if the firm suffered losses due to crime (as reported by the firm) during the last fiscal year and 0 otherwise. Source: Enterprise Surveys. www.enterprisesurveys.org
Courts Obstacle Y:1 N:0	Dummy variable equal to 1 if the firm reports functioning of courts as a moderate, major or very severe obstacle to its current operations and 0 if it reports no obstacle or minor obstacle. Source: Enterprise Surveys. www.enterprisesurveys.org
Power Outages (hours)	Total hours of power outages experienced by the firm in a typical month over the last fiscal year.

	Source: Enterprise Surveys. www.enterprisesurveys.org
Finance Obstacle (0-4 scale)	Severity level of access to finance as an obstacle to firms' day to day operations. The severity level is defined on a 0-4 scale as: no obstacle (0), minor obstacle (1), moderate obstacle (2), major obstacle (3), and a very severe obstacle (4). Source: Enterprise Surveys. www.enterprisesurveys.org
Overdraft Y:1 N:0	Dummy variable equal to 1 if the firm has overdraft facility and 0 otherwise. Source: Enterprise Surveys. www.enterprisesurveys.org
Inflation (CPI, annual %)	Annual percentage rate of inflation based on the consumer price index. Values of the variable used are for the fiscal year covered by the Enterprise Survey. In cases where the fiscal year spanned more than one calendar year, the inflation figure is the weighted average over the calendar years covered, where the weights used are the proportion of firms covered in the calendar year. Source: World Development Indicators, World Bank.
Primary Enrollment (Gross rate)	Gross enrollment rate in primary education. Values of the variable were computed in two steps. In the first step, gross primary enrollment rate values for the fiscal year covered by the Enterprise Survey. In cases where the fiscal year spanned more than one calendar year, the enrollment figure is the weighted average over the calendar years covered, where the weights used are the proportion of firms covered in the calendar year. In the second step, annual values were taken from the first step and averaged over the three years prior to the (final) year covered by the Enterprise Survey. Three years average was taken to avoid too many missing values of primary enrollment rate. Source: World Development Indicators, World Bank.
Industry fixed effects	A set of dummy variables indicating the industry to which the firm belongs. Industry grouping here is at the 2-digit ISIC Rev. 3.1

	level. There are 7 such industries in our baseline sample (used in Table 4). Source: Enterprise Surveys. www.enterprisesurveys.org
Year fixed effects	A set of dummy variables (one for each year) indicating the year the ES was conducted in the country. Source: Enterprise Surveys. www.enterprisesurveys.org
Region fixed effects	A set of dummy variables indicating the region of the country. There are 6 dummy variables one for each world region. The regions are: Sub-Saharan Africa (SSA), Latin America and Caribbean (LAC), Eastern Europe and Central Asia (ECA), Middle East and North Africa (MENA), East Asia and the Pacific (EAP), and South Asia (SAR). These regional groupings are based on World Bank's classification.
City size fixed effects	Dummy variables indicating the size of the city (sampling region) where the firm is located. Based on total population, cities are grouped into the following size categories: more than 1 million population, between 1 million and 250,000 population, between 50,000 and 250,000 population, and less than 50,000 population. Source: Enterprise Surveys. www.enterprisesurveys.org
Firm Spent On R&D Y:0 N:0	Dummy variable equal to 1 if the firm spent on R&D activity during the last fiscal year (prior the year ES was conducted) and 0 otherwise. Source: Enterprise Surveys. www.enterprisesurveys.org
Firm Purchased Fixed Assets Y:1 N:0	Dummy variable equal to 1 if the firm purchased fixed assets during the last fiscal year (prior the year ES was conducted) and 0 otherwise. Source: Enterprise Surveys. www.enterprisesurveys.org
Foreign Ownership (proportion)	Proportion of the firm that is owned by foreign individuals, companies or organizations. Source: Enterprise Surveys. www.enterprisesurveys.org

Time Tax (Country level average)	Country-level average of the percentage of firms' senior management's time that is spent in dealing with business regulations. Source: Enterprise Surveys. www.enterprisesurveys.org
Population (lagged, logs)	Total population in the country. Values of the variable used are one year prior (that is, lagged by 1 year) to the fiscal year covered by the ES. In cases where the fiscal year covered by the ES spans more than one calendar year, the population figure is the weighted average over the calendar years covered (lagged by 1 year), where the weights used are the proportion of firms covered by the ES in the calendar year (lagged by one year). Source: World Development Indicators, World Bank.
Rule of Law (lagged)	Rule of Law indicator from Worldwide Governance Indicators. Higher values of the variable imply better enforcement of rules and laws. The variable is lagged by 2 years from the fiscal year covered by the Enterprise Survey in the country. In cases where the fiscal year spanned more than one calendar year, the Rule of Law figure is the weighted average over the calendar years covered (lagged by 2 years), where the weights used are the proportion of firms covered in the calendar year. Source: Worldwide Governance Indicators, World Bank.
Freedom From Regulation (lagged)	Freedom from Regulation indicator from Fraser Institute's Economic Freedom of the World database. Higher values of the variable imply less regulation (of formal businesses). The variable is lagged by 2 years from the fiscal year covered by the Enterprise Survey in the country. In cases where the fiscal year spanned more than one calendar year, the figure used is the weighted average over the calendar years covered (and lagged by 2 years), where the weights used are the proportion of firms covered in the calendar year. Source: Economic Freedom of the World, Fraser Institute.

Control of Corruption (lagged)	Control of Corruption indicator from Worldwide Governance Indicators. Higher values of the variable imply lower corruption. The variable is lagged by 2 years from the fiscal year covered by the Enterprise Survey in the country. In cases where the fiscal year spanned more than one calendar year, the figure used is the weighted average over the calendar years covered (and lagged by 2 years), where the weights used are the proportion of firms covered in the calendar year. Source: Worldwide Governance Indicators, World Bank.
Profit Tax Rate	The amount of taxes on profits paid by the business expressed as a percentage of total commercial profits of the business as per the Doing Business methodology, World Bank. To make the data comparable across countries, several assumptions are made about businesses. The main assumptions are that they are limited liability companies, they operate in the country's most populous city, they are domestically owned, they perform general industrial or commercial activities, and they have certain levels of start-up capital, employees, and turnover. The variable is lagged by 2 years from the fiscal year covered by the Enterprise Survey in the country. In cases where the fiscal year spanned more than one calendar year, the figure used is the weighted average over the calendar years covered (and lagged by 2 years), where the weights used are the proportion of firms covered in the calendar year. Source: Doing Business, World Bank.
Cheating on Taxes is Justified (dummy)	In the World Value Surveys, respondents were asked the following question: "Do you think cheating on taxes is justified?" Answers were recorded on 1 (Never justified) to 10 (Always justified) scale with higher values implying a higher justification for cheating on taxes. We first compute country-level average of the responses for each year for which data are

	available for the country. Next, for any given year and country, we take the average over the 1ast 10 years for which data are available (10 year moving average). This is done to avoid too many missing values. Last, for the sample of countries and years included in our sample, we define a dummy variable equal to 1 if the 10-year moving average value is above the median level and 0 otherwise. This dummy variable is used in the regressions. Source: World Value Surveys.

Table 2: Summary statistics

Variable	Mean	Std. deviation	Minimum	Maximum	Observations
Employment Growth (%, annual)	3.51	16.07	-170.06	239.37	20,295
Employment Growth 1 (%, annual)	3.34	14.60	-93.55	98.35	20,295
Informal Competition (Country level cell average)	0.49	0.18	0	0.976	20,295
Initial Employment (logs)	2.83	0.88	0	4.61	20,295
GDP per capita (lagged, logs)	9.03	0.72	6.49	10.39	20,295
Age of Firm (logs)	2.74	0.68	0	5.42	20,295
Manager Experience (logs)	2.69	0.71	0	4.09	20,295
Exports (proportion of sales)	0.07	0.21	0	1	20,295
Female Top Manger Y:1 N:0	0.13	0.34	0	1	20,295
Multi-establishment firm Y:1 N:0	0.14	0.35	0	1	20,295
Crime losses Y:1 N:0	0.10	0.30	0	1	20,295
Courts obstacle Y:1 N:0	0.24	0.43	0	1	20,295
Power outages (hours)	21.30	69.92	0	720	20,295
Finance obstacle (0-4 scale)	1.32	1.25	0	4	20,295
Overdraft Y:1 N:0	0.43	0.49	0	1	20,295
Inflation (CPI, annual %)	7.24	5.03	-1.99	53.23	20,295
Primary Enrollment (Gross rate)	102.60	10.93	33.96	145.86	20,295

Source: Author's own calculations based on Enterprise Surveys data and World Development Indicators (WDI, World Bank).

Table 3: Correlations between informal competition (cell averages) and the baseline controls

	Informal Competition (Country level cell average)
Informal Competition (Country level cell average)	1
Initial Employment (logs)	-0.109
GDP per capita (lagged, logs)	-0.137
Age of Firm (logs)	0.050
Manager Experience (logs)	0.074
Exports (proportion of sales)	-0.128
Female Top Manger Y:1 N:0	-0.007
Multi-establishment firm Y:1 N:0	-0.064
Crime losses Y:1 N:0	0.092
Courts obstacle Y:1 N:0	0.143
Power outages (hours)	0.014
Finance obstacle (0-4 scale)	0.178
Overdraft Y:1 N:0	-0.201
Inflation (CPI, annual %)	-0.016
Primary Enrollment (Gross rate)	-0.025

Source: Author's own calculations based on data from Enterprise Surveys and World Development Indicators (WDI, World Bank). Sample size: 20.295 firms used in the baseline regressions.

Table 4: Base Regression Results

Dependent variable: *Employment Growth* (log difference, %, annual)	(1)	(2)	(3)	(4)	(5)
Informal Competition (Country level cell average)	-5.786*	-6.954**	-6.490**	-4.941**	-5.692**
	(3.370)	(2.782)	(2.640)	(2.245)	(2.250)
Initial Employment (logs)	-3.702***	-3.860***	-3.798***	-4.088***	-4.130***
	(0.712)	(0.626)	(0.560)	(0.552)	(0.564)
GDP per capita (lagged, logs)		-1.513**	-1.423**	-1.737***	-1.460***
		(0.625)	(0.563)	(0.486)	(0.474)
Age of Firm (logs)			-3.440***	-3.422***	-3.406***
			(1.047)	(1.035)	(1.026)
Manager Experience (logs)			0.025	0.007	0.082
			(0.359)	(0.349)	(0.359)
Exports (proportion of sales)			2.220**	2.132*	2.213**
			(1.106)	(1.096)	(1.102)
Female Top Manger Y:1 N:0			-1.418	-1.392	-1.388
			(1.602)	(1.542)	(1.556)
Multi-establishment firm Y:1 N:0			3.643***	3.398***	3.424***
			(0.952)	(0.920)	(0.944)
Crime Losses Y:1 N:0				0.120	0.088
				(1.172)	(1.159)
Courts Obstacle Y:1 N:0				-0.454	-0.392
				(0.610)	(0.634)
Power Outages (hours)				-0.007*	-0.007*
				(0.004)	(0.004)
Finance Obstacle (0-4 scale)				-0.353	-0.352
				(0.238)	(0.238)
Overdraft Y:1 N:0				3.300***	3.293***
				(0.680)	(0.703)
Inflation (CPI, annual %)					0.129***
					(0.039)
Primary Enrollment (Gross rate)					0.044*
					(0.024)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	16.843***	21.028***	30.866***	32.888***	22.874***
	(2.044)	(5.648)	(6.893)	(6.523)	(6.828)
Number of observations	20,295	20,295	20,295	20,295	20,295
R-squared	0.042	0.057	0.085	0.096	0.097

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 5: Base regression results using alternative measure of employment growth

Dependent variable: *Employment Growth* 1	(1)	(2)	(3)	(4)	(5)
Informal Competition (Country level cell average)	-5.437	-6.455**	-6.002**	-4.510**	-5.138**
	(3.283)	(2.700)	(2.543)	(2.154)	(2.173)
Initial Employment (logs)	-3.325***	-3.482***	-3.411***	-3.683***	-3.719***
	(0.634)	(0.551)	(0.485)	(0.476)	(0.485)
GDP per capita (lagged, logs)		-1.515**	-1.421**	-1.723***	-1.488***
		(0.608)	(0.547)	(0.461)	(0.452)
Age of Firm (logs)			-3.218***	-3.196***	-3.179***
			(0.992)	(0.975)	(0.968)
Manager Experience (logs)			0.001	-0.014	0.053
			(0.344)	(0.332)	(0.340)
Exports (proportion of sales)			1.956*	1.873*	1.927*
			(1.043)	(1.034)	(1.037)
Female Top Manger Y:1 N:0			-1.266	-1.230	-1.249
			(1.494)	(1.433)	(1.444)
Multi-establishment firm Y:1 N:0			3.257***	3.018***	3.035***
			(0.789)	(0.752)	(0.774)
Crime losses Y:1 N:0				-0.033	-0.069
				(1.100)	(1.090)
Courts obstacle Y:1 N:0				-0.521	-0.459
				(0.575)	(0.596)
Power outages (hours)				-0.007*	-0.007*
				(0.003)	(0.004)
Finance obstacle (0-4 scale)				-0.363	-0.357
				(0.239)	(0.240)
Overdraft Y:1 N:0				3.136***	3.111***
				(0.643)	(0.661)
Inflation (CPI, annual %)					0.098***
					(0.033)
Primary Enrollment (Gross rate)					0.042*
					(0.023)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	15.434***	19.746***	28.849***	30.908***	22.113***
	(1.876)	(5.444)	(6.549)	(6.111)	(6.475)
Number of observations	20,295	20,295	20,295	20,295	20,295
R-squared	0.041	0.058	0.087	0.099	0.101

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 6: Robustness for other controls				
Dependent variable: *Employment Growth* (log difference, %, annual)	**(1)**	**(2)**	**(3)**	**(4)**
Informal Competition (Country level cell average)	-6.584***	-6.448***	-6.718***	-7.072***
	(2.403)	(2.034)	(2.132)	(2.535)
Initial Employment (logs)	-4.123***	-4.335***	-4.356***	-4.244***
	(0.570)	(0.707)	(0.722)	(0.715)
GDP per capita (lagged, logs)	-1.465	-0.674	-0.718	0.309
	(0.929)	(0.861)	(0.860)	(1.102)
Age of Firm (logs)	-3.328***	-2.800***	-2.807***	-2.761***
	(1.027)	(0.724)	(0.729)	(0.729)
Manager Experience (logs)	0.059	0.250	0.278	0.292
	(0.352)	(0.383)	(0.388)	(0.395)
Exports (proportion of sales)	2.153*	1.156	1.239	1.170
	(1.122)	(1.149)	(1.141)	(1.178)
Female Top Manger Y:1 N:0	-1.495	-0.235	-0.237	-0.366
	(1.565)	(0.877)	(0.880)	(0.889)
Multi-establishment firm Y:1 N:0	3.643***	3.365***	3.340***	3.318***
	(0.978)	(0.717)	(0.707)	(0.712)
Crime Losses Y:1 N:0	-0.097	0.207	0.237	-0.238
	(1.151)	(1.032)	(1.024)	(1.031)
Courts Obstacle Y:1 N:0	-0.345	-0.498	-0.487	-0.737
	(0.653)	(0.700)	(0.705)	(0.656)
Power Outages (hours)	-0.007*	-0.007*	-0.007*	-0.007**
	(0.004)	(0.004)	(0.004)	(0.004)
Finance Obstacle (0-4 scale)	-0.354	-0.234	-0.228	-0.261
	(0.232)	(0.192)	(0.185)	(0.195)
Overdraft Y:1 N:0	3.337***	2.335***	2.363***	2.366***
	(0.721)	(0.545)	(0.537)	(0.525)
Inflation (CPI, annual %)	0.135***	0.253***	0.251***	0.173
	(0.036)	(0.040)	(0.041)	(0.158)
Primary Enrollment (Gross rate)	0.059	0.060*	0.059*	0.073*
	(0.037)	(0.035)	(0.034)	(0.037)
Firm Spent On R&D Y:0 N:0		0.016	0.016	0.016
		(0.011)	(0.011)	(0.011)
Firm Purchased Fixed Assets Y:1 N:0		3.316***	3.354***	3.327***
		(0.730)	(0.733)	(0.745)
Foreign Ownership (proportion)		2.257	2.311	2.450
		(2.035)	(2.036)	(2.105)
Time Tax (Country level average)			-0.024	-0.026
			(0.061)	(0.059)
Population (lagged, logs)			0.236	0.386

			(0.310)	(0.366)
Rule of Law (lagged)				-1.596
				(2.948)
Freedom From Regulation (lagged)				-0.020
				(0.782)
Control of Corruption (lagged)				0.706
				(2.949)
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
City size fixed effects	No	Yes	Yes	Yes
Constant	14.466	2.892	0.318	-1.824
	(10.153)	(9.055)	(9.726)	(13.017)
Number of observations	20,295	18,058	18,058	17,482
R-squared	0.098	0.102	0.102	0.100

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 7: Results using cell averages at the City times Industry level

Dependent variable: *Employment Growth* (log difference, %, annual)	**(1)**	**(2)**	**(3)**	**(4)**
Informal Competition (City-Industry level cell average)	-6.997**	-6.747**	-5.803**	-5.796**
	(3.386)	(3.419)	(2.533)	(2.455)
Initial Employment (logs)	-4.100***	-4.100***	-4.174***	-4.361***
	(0.375)	(0.374)	(0.367)	(0.369)
Age of Firm (logs)			-2.973***	-2.963***
			(0.637)	(0.618)
Manager Experience (logs)			0.123	0.039
			(0.403)	(0.402)
Exports (proportion of sales)			2.669*	2.486*
			(1.424)	(1.416)
Female Top Manger Y:1 N:0			-1.275	-1.262
			(1.390)	(1.330)
Multi-establishment firm Y:1 N:0			4.380***	4.209***
			(0.885)	(0.864)
Crime losses Y:1 N:0				0.344
				(1.258)
Courts obstacle Y:1 N:0				-0.391
				(0.673)
Power outages (hours)				-0.004
				(0.005)
Finance obstacle (0-4 scale)				-0.219
				(0.232)
Overdraft Y:1 N:0				2.942***
				(0.589)
Industry fixed effects	No	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes
Constant	18.576***	11.524***	18.266***	18.227***
	(1.890)	(2.830)	(3.459)	(3.565)
Number of observations	19,513	19,513	19,513	19,513
R-squared	0.157	0.158	0.180	0.186

Huber-White robust standard errors and clustered at the city times industry level in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 8: Two stage estimation using linear probability model (OLS) in the first stage

Panel A: Second stage regression results

Dependent variable: *Employment Growth* (log difference, %, annual)					
Informal Competition (dummy, predicted values)	-6.493*	-8.752**	-8.171**	-6.361**	-7.398**
	(3.846)	(3.777)	(3.630)	(3.181)	(3.275)
Initial Employment (logs)	-3.870***	-4.086***	-3.956***	-4.216***	-4.283***
	(0.700)	(0.649)	(0.579)	(0.569)	(0.585)
GDP per capita (lagged, logs)		-1.615**	-1.434**	-1.715***	-1.410***
		(0.722)	(0.625)	(0.531)	(0.521)
Age of Firm (logs)			-3.543***	-3.505***	-3.498***
			(1.177)	(1.139)	(1.147)
Manager Experience (logs)			-0.052	-0.055	0.021
			(0.387)	(0.379)	(0.406)
Exports (proportion of sales)			1.284	1.387	1.338
			(1.097)	(1.035)	(1.035)
Female Top Manger Y:1 N:0			-1.663	-1.596	-1.650
			(1.912)	(1.781)	(1.834)
Multi-establishment firm Y:1 N:0			3.595***	3.351***	3.367***
			(1.140)	(1.074)	(1.128)
Crime losses Y:1 N:0				0.420	0.425
				(1.355)	(1.384)
Courts obstacle Y:1 N:0				-0.438	-0.358
				(0.539)	(0.563)
Power outages (hours)				-0.007*	-0.007*
				(0.004)	(0.004)
Finance obstacle (0-4 scale)				-0.189	-0.157
				(0.216)	(0.214)
Overdraft Y:1 N:0				3.377***	3.362***
				(0.705)	(0.735)
Inflation (CPI, annual %)					0.127***
					(0.039)
Primary Enrollment (Gross rate)					0.053**
					(0.026)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	17.648***	15.883**	25.295***	27.104***	16.242**
	(2.352)	(7.232)	(8.240)	(7.618)	(7.411)
Number of observations	20,295	20,295	20,295	20,295	20,295

Panel B: First stage estimation results (OLS)

Dependent variable: Informal Competition (firm-level dummy)					
Informal Competition (Country level cell average)	0.891***	0.795***	0.794***	0.777***	0.769***
	(0.028)	(0.059)	(0.063)	(0.068)	(0.067)

Other controls (as above)	Yes	Yes	Yes	Yes	Yes
F-statistic	988***	183***	159***	129***	131***
Number of observations	20,295	20,295	20,295	20,295	20,295

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 9: Two stage estimation using probit model in the first stage

Panel A: Second stage regression results					
Dependent variable: *Employment Growth* (log difference, %, annual)					
Informal Competition (dummy, predicted values)	-6.604*	-8.550**	-8.338**	-6.465**	-7.517***
	(3.726)	(3.356)	(3.266)	(2.793)	(2.812)
Initial Employment (logs)	-3.872***	-4.078***	-3.958***	-4.216***	-4.283***
	(0.695)	(0.626)	(0.562)	(0.552)	(0.563)
GDP per capita (lagged, logs)		-1.607**	-1.451**	-1.727***	-1.424***
		(0.643)	(0.568)	(0.478)	(0.460)
Age of Firm (logs)			-3.554***	-3.516***	-3.511***
			(1.060)	(1.044)	(1.036)
Manager Experience (logs)			-0.048	-0.052	0.024
			(0.354)	(0.344)	(0.354)
Exports (proportion of sales)			1.245	1.360	1.307
			(0.992)	(0.999)	(0.996)
Female Top Manger Y:1 N:0			-1.679	-1.614	-1.670
			(1.650)	(1.584)	(1.601)
Multi-establishment firm Y:1 N:0			3.596***	3.352***	3.369***
			(0.949)	(0.918)	(0.942)
Crime losses Y:1 N:0				0.418	0.423
				(1.164)	(1.160)
Courts obstacle Y:1 N:0				-0.443	-0.364
				(0.611)	(0.635)
Power outages (hours)				-0.007*	-0.007*
				(0.004)	(0.004)
Finance obstacle (0-4 scale)				-0.185	-0.152
				(0.249)	(0.251)
Overdraft Y:1 N:0				3.373***	3.357***
				(0.678)	(0.702)
Inflation (CPI, annual %)					0.128***
					(0.038)
Primary Enrollment (Gross rate)					0.053**
					(0.025)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	17.708***	35.536***	45.104***	43.630***	32.187***
	(2.284)	(8.308)	(9.381)	(8.161)	(7.950)
Number of observations	20,295	20,293	20,293	20,293	20,293
Panel B: First stage estimation results (Probit model)					
Dependent variable: Informal Competition (firm-level dummy)					
Informal Competition (Country level cell average)	2.434***	2.177***	2.181***	2.151***	2.130***
	(0.103)	(0.179)	(0.190)	(0.213)	(0.209)

Other controls (as above)	Yes	Yes	Yes	Yes	Yes
Number of observations	20,295	20,293	20,293	20,293	20,293

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 10: Conditional Mixed Process (CMP) estimation results

Panel A: Second stage IV results

Dependent variable: Employment growth (log difference, %, annual)	(1)	(3)	(5)	(7)	(9)
Informal Competition (Country level cell average)	-2.378*	-3.194**	-2.976**	-2.297*	-2.672**
	(1.437)	(1.404)	(1.345)	(1.173)	(1.211)
Initial Employment (logs)	-3.864***	-4.083***	-3.952***	-4.212***	-4.278***
	(0.706)	(0.655)	(0.581)	(0.571)	(0.587)
GDP per capita (lagged, logs)		-1.633**	-1.452**	-1.733***	-1.427***
		(0.730)	(0.633)	(0.538)	(0.527)
Age of Firm (logs)			-3.553***	-3.513***	-3.506***
			(1.194)	(1.154)	(1.162)
Manager Experience (logs)			-0.045	-0.050	0.025
			(0.388)	(0.379)	(0.407)
Exports (proportion of sales)			1.247	1.361	1.307
			(1.111)	(1.040)	(1.042)
Female Top Manger Y:1 N:0			-1.664	-1.600	-1.655
			(1.939)	(1.808)	(1.865)
Multi-establishment firm Y:1 N:0			3.602***	3.354***	3.371***
			(1.151)	(1.081)	(1.136)
Crime losses Y:1 N:0				0.433	0.439
				(1.370)	(1.402)
Courts obstacle Y:1 N:0				-0.438	-0.358
				(0.541)	(0.566)
Power outages (hours)				-0.007*	-0.007*
				(0.004)	(0.004)
Finance obstacle (0-4 scale)				-0.191	-0.159
				(0.217)	(0.216)
Overdraft Y:1 N:0				3.378***	3.364***
				(0.708)	(0.737)
Inflation (CPI, annual %)					0.128***
					(0.040)
Primary Enrollment (Gross rate)					0.052**
					(0.026)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	14.396***	11.456*	21.164***	23.897***	12.484*
	(1.739)	(6.102)	(7.050)	(6.766)	(6.786)
Number of observations	20,295	20,295	20,295	20,295	20,295

Panel B: First stage IV results (Probit; log odds ratios)

Dependent variable: Firm faces informal competition Y:1 N:0 (dummy)	(1)	(2)	(3)	(4)	(5)

Informal Competition (country level average)	2.434***	2.177***	2.181***	2.150***	2.130***
	(0.103)	(0.178)	(0.189)	(0.213)	(0.209)
Other controls (As above)	Yes	Yes	Yes	Yes	Yes
Number of observations	20,295	20,295	20,295	20,295	20,295

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 11: Informality is a major obstacle (logit estimation; log odds ratios shown)					
Dependent variable: Informality is a major obstacle Y:1 N:0 (dummy)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Firm faces informal competition Y:1 N:0 (dummy, infor1)	-0.935	-1.339	-1.273	-0.991	-1.004
	(1.365)	(1.229)	(1.211)	(1.222)	(1.213)
GDP per capita (lagged, logs)	-0.330**	-0.325**	-0.370***	-0.260**	-0.275**
	(0.129)	(0.129)	(0.129)	(0.130)	(0.130)
Firm faces informal competition (dummy)* GDP per capita (lagged, logs)	**0.341****	**0.383*****	**0.376*****	**0.344****	**0.345****
	(0.154)	(0.139)	(0.137)	(0.140)	(0.139)
Initial Employment (logs)	-0.044	-0.014	-0.021	-0.033	-0.031
	(0.053)	(0.046)	(0.050)	(0.050)	(0.050)
Age of Firm (logs)			0.070**	0.058	0.058
			(0.034)	(0.040)	(0.040)
Manager Experience (logs)			0.167**	0.142*	0.140*
			(0.078)	(0.075)	(0.075)
Exports (proportion of sales)			0.022	-0.036	-0.042
			(0.257)	(0.233)	(0.236)
Female Top Manger Y:1 N:0			0.089	0.054	0.054
			(0.184)	(0.142)	(0.143)
Multi-establishment firm Y:1 N:0			-0.036	0.011	0.010
			(0.229)	(0.266)	(0.265)
Crime losses Y:1 N:0				0.694***	0.694***
				(0.195)	(0.196)
Courts obstacle Y:1 N:0				0.585***	0.584***
				(0.091)	(0.089)
Power outages (hours)				0.001	0.001
				(0.001)	(0.001)
Finance obstacle (0-4 scale)				0.310***	0.310***
				(0.056)	(0.056)
Overdraft Y:1 N:0				-0.094	-0.096
				(0.110)	(0.106)
Inflation (CPI, annual %)					-0.007
					(0.009)
Primary Enrollment (Gross rate)					-0.002
					(0.005)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	0.445	0.607	0.299	-1.870	-1.386
	(1.139)	(1.015)	(1.033)	(1.140)	(1.401)
Number of observations	19,909	19,907	19,907	19,907	19,907

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 12: Interaction terms results for Rule of Law						
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Employment Growth (log difference, %, annual)	Without GDP per capita interaction term control			With GDP per capita interaction term control		
Informal Competition (Country level cell average)*Rule of Law (lagged)	**6.064***	**5.456**	**4.008**	**9.998****	**14.669*****	**12.316*****
	(3.654)	(4.315)	(3.622)	(4.370)	(4.472)	(3.782)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)				-8.436***	-13.097***	-11.638***
				(2.935)	(3.097)	(2.965)
Informal Competition (Country level cell average)	-4.252	-6.029**	-5.387**	76.313***	117.407***	104.273***
	(3.458)	(2.966)	(2.412)	(27.464)	(28.591)	(27.685)
Rule of Law (lagged)	-2.449	-2.099	-1.984	-3.294	-6.610***	-6.167***
	(1.838)	(2.201)	(1.761)	(2.176)	(2.361)	(2.033)
GDP per capita (lagged, logs)		-1.426**	-1.276**	2.348	4.677***	4.186***
		(0.709)	(0.558)	(1.446)	(1.498)	(1.514)
Initial Employment (logs)	-3.782***	-3.908***	-4.165***	-3.871***	-3.891***	-4.160***
	(0.693)	(0.639)	(0.579)	(0.647)	(0.648)	(0.589)
Age of Firm (logs)			-3.377***			-3.248***
			(1.027)			(0.999)
Manager Experience (logs)			0.058			-0.017
			(0.349)			(0.338)
Exports (proportion of sales)			2.226**			2.150*
			(1.114)			(1.116)
Female Top Manger Y:1 N:0			-1.425			-1.511
			(1.544)			(1.528)
Multi-establishment firm Y:1 N:0			3.442***			3.489***
			(0.949)			(0.954)
Crime losses Y:1 N:0			0.103			-0.073
			(1.155)			(1.135)
Courts obstacle Y:1 N:0			-0.356			-0.408
			(0.652)			(0.642)
Power outages (hours)			-0.007*			-0.007*
			(0.004)			(0.004)
Finance obstacle (0-4 scale)			-0.347			-0.370

			(0.235)			(0.234)
Overdraft Y:1 N:0			3.282***			3.322***
			(0.662)			(0.660)
Inflation (CPI, annual %)			0.112**			0.096**
			(0.046)			(0.041)
Primary Enrollment (Gross rate)			0.047*			0.046*
			(0.025)			(0.026)
Industry fixed effects	No	Yes	Yes	No	Yes	Yes
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Constant	16.686***	15.912**	17.938**	-6.201	-54.677***	-44.533**
	(2.184)	(7.064)	(7.809)	(12.658)	(16.491)	(17.470)
Number of observations	20,295	20,295	20,295	20,295	20,295	20,295
R squared	0.044	0.058	0.098	0.050	0.062	0.101

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 13: Interaction terms results for Control of Corruption						
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Employment Growth (log difference, %, annual)	Without GDP per capita interaction term control			With GDP per capita interaction term control		
Informal Competition (Country level cell average)*Control of Corruption (lagged)	**5.354**	**5.143**	**3.970**	**12.011****	**15.115****	**13.260****
	(3.657)	(4.545)	(3.738)	(3.732)	(4.604)	(3.761)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)				-9.207***	-12.296***	-11.414***
				(3.174)	(2.918)	(2.958)
Informal Competition (Country level cell average)	-4.848	-5.041	-4.511*	83.655***	112.204***	104.357***
	(3.710)	(3.207)	(2.531)	(29.255)	(27.172)	(27.912)
Control of Corruption (lagged)	-2.746	-1.418	-1.469	-3.987**	-6.506***	-6.259***
	(1.797)	(2.377)	(1.787)	(1.948)	(2.457)	(1.982)
GDP per capita (lagged, logs)		-1.662**	-1.433**	2.351	4.176***	4.004**
		(0.751)	(0.628)	(1.428)	(1.389)	(1.546)
Initial Employment (logs)	-3.765***	-3.891***	-4.153***	-3.836***	-3.884***	-4.160***
	(0.687)	(0.637)	(0.577)	(0.647)	(0.647)	(0.587)
Age of Firm (logs)			-3.398***			-3.283***
			(1.025)			(1.003)
Manager Experience (logs)			0.062			-0.002
			(0.348)			(0.343)
Exports (proportion of sales)			2.207**			2.170*
			(1.113)			(1.121)
Female Top Manger Y:1 N:0			-1.394			-1.483
			(1.550)			(1.536)
Multi-establishment firm Y:1 N:0			3.447***			3.483***
			(0.948)			(0.954)
Crime losses Y:1 N:0			0.091			-0.086
			(1.148)			(1.132)
Courts obstacle Y:1 N:0			-0.348			-0.410
			(0.654)			(0.643)
Power outages (hours)			-0.007*			-0.007*
			(0.004)			(0.004)

Finance obstacle (0-4 scale)			-0.352			-0.381
			(0.237)			(0.237)
Overdraft Y:1 N:0			3.237***			3.307***
			(0.650)			(0.649)
Inflation (CPI, annual %)			0.128***			0.121***
			(0.039)			(0.035)
Primary Enrollment (Gross rate)			0.044*			0.042*
			(0.024)			(0.025)
Industry fixed effects	No	Yes	Yes	No	Yes	Yes
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Constant	16.656***	17.371**	18.760**	-6.276	-49.978***	-43.393**
	(2.118)	(7.357)	(7.990)	(12.534)	(15.499)	(17.853)
Number of observations	20,295	20,295	20,295	20,295	20,295	20,295
R squared	0.043	0.058	0.098	0.051	0.061	0.100

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 14: Interaction terms results for Freedom from Regulation						
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Employment Growth (log difference, %, annual)	Without GDP per capita interaction term control			With GDP per capita interaction term control		
Informal Competition (Country level cell average)*Freedom from Regulation (lagged)	**8.018****	**4.548**	**4.321**	**8.175****	**7.241***	**7.353****
	(3.192)	(3.476)	(2.977)	(3.251)	(3.747)	(3.043)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)				-5.359**	-6.895**	-7.472***
				(2.585)	(2.779)	(2.455)
Informal Competition (Country level cell average)	-60.383**	-38.655	-36.331*	-11.599	6.712	11.965
	(23.367)	(24.890)	(21.391)	(26.147)	(28.147)	(25.668)
Freedom from Regulation (lagged)	-3.037*	-1.668	-2.263	-2.920	-3.189	-4.027**
	(1.650)	(1.776)	(1.467)	(1.823)	(1.981)	(1.620)
GDP per capita (lagged, logs)		-1.469**	-1.267**	1.034	1.634	2.149*
		(0.734)	(0.571)	(1.268)	(1.348)	(1.232)
Initial Employment (logs)	-3.706***	-3.798***	-4.042***	-3.770***	-3.751***	-3.998***
	(0.696)	(0.634)	(0.562)	(0.647)	(0.631)	(0.561)
Age of Firm (logs)			-3.296***			-3.244***
			(1.014)			(1.005)
Manager Experience (logs)			-0.039			-0.116
			(0.332)			(0.326)
Exports (proportion of sales)			1.919*			1.854*
			(1.110)			(1.101)
Female Top Manger Y:1 N:0			-1.590			-1.600
			(1.583)			(1.567)
Multi-establishment firm Y:1 N:0			3.375***			3.356***
			(0.928)			(0.923)
Crime losses Y:1 N:0			-0.349			-0.408
			(1.122)			(1.114)
Courts obstacle Y:1 N:0			-0.544			-0.576
			(0.620)			(0.615)
Power outages (hours)			-0.007			-0.006
			(0.004)			(0.004)

Finance obstacle (0-4 scale)			-0.356			-0.358
			(0.239)			(0.239)
Overdraft Y:1 N:0			3.312***			3.361***
			(0.698)			(0.688)
Inflation (CPI, annual %)			0.045			0.040
			(0.142)			(0.134)
Primary Enrollment (Gross rate)			0.052*			0.058*
			(0.030)			(0.032)
Industry fixed effects	No	Yes	Yes	No	Yes	Yes
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Constant	37.920***	27.710**	35.468**	27.219**	2.189	7.852
	(11.723)	(13.854)	(14.642)	(10.757)	(14.688)	(16.152)
Number of observations	19,641	19,641	19,641	19,641	19,641	19,641
R squared	0.046	0.057	0.097	0.051	0.058	0.098

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 15: Interaction terms results for Profit Tax Rate						
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: Employment Growth (log difference, %, annual)	Without GDP per capita interaction term control			With GDP per capita interaction term control		
Informal Competition (Country level cell average)*Profit Tax Rate (%, lagged)	**-0.149**	**-0.420**	**-0.381**	**-0.256**	**-0.493***	**-0.437****
	(0.282)	(0.272)	(0.240)	(0.249)	(0.251)	(0.212)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)				-4.144	-5.505**	-4.417**
				(2.743)	(2.356)	(2.084)
Informal Competition (Country level cell average)	4.279	5.098	4.700	45.978*	58.654***	47.521**
	(5.805)	(6.715)	(6.055)	(24.942)	(20.679)	(19.077)
Pofit Tax Rate (%, lagged)	0.030	0.191	0.156	0.095	0.233	0.187
	(0.147)	(0.182)	(0.143)	(0.133)	(0.176)	(0.134)
GDP per capita (lagged, logs)		-0.901	-1.068	2.495*	1.771	1.066
		(0.880)	(0.788)	(1.467)	(1.155)	(1.212)
Initial Employment (logs)	-4.732***	-4.745***	-4.813***	-4.721***	-4.715***	-4.780***
	(0.365)	(0.387)	(0.402)	(0.369)	(0.386)	(0.406)
Age of Firm (logs)			-3.646***			-3.654***
			(0.793)			(0.793)
Manager Experience (logs)			-0.015			0.022
			(0.534)			(0.535)
Exports (proportion of sales)			1.703			1.734
			(1.426)			(1.420)
Female Top Manger Y:1 N:0			-0.418			-0.405
			(1.182)			(1.175)
Multi-establishment firm Y:1 N:0			2.531**			2.446*
			(1.257)			(1.282)
Crime losses Y:1 N:0			1.272			1.236
			(1.185)			(1.189)
Courts obstacle Y:1 N:0			0.039			-0.031
			(0.672)			(0.678)
Power outages (hours)			-0.007			-0.007
			(0.011)			(0.011)

Finance obstacle (0-4 scale)			-0.417*			-0.432*
			(0.237)			(0.236)
Overdraft Y:1 N:0			3.076***			3.030***
			(1.043)			(1.060)
Inflation (CPI, annual %)			0.119***			0.125***
			(0.038)			(0.037)
Primary Enrollment (Gross rate)			0.053*			0.049*
			(0.029)			(0.029)
Industry fixed effects	No	Yes	Yes	No	Yes	Yes
Year fixed effects	No	Yes	Yes	No	Yes	Yes
Constant	15.334***	5.751	8.518	-9.857	-26.480**	-16.759
	(2.675)	(9.769)	(9.873)	(14.408)	(12.410)	(13.320)
Number of observations	11,406	11,406	11,406	11,406	11,406	11,406
R squared	0.058	0.069	0.102	0.059	0.071	0.103

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 16: Interaction with exporter dummy

Dependent variable: Employment Growth (log difference, %, annual)	(1)	(2)	(3)	(4)	(5)	(6)
Informal Competition (Country level cell average)*Firm Exports Y:1 N:0	**11.315****	**10.650****	**12.770*****	**12.049*****	**12.810****	**12.287****
	(4.587)	**(4.414)**	**(4.600)**	**(4.500)**	**(4.998)**	**(4.990)**
Informal Competition (Country level cell average)	-8.817***	-7.515***	43.834*	44.099**	43.994*	45.062*
	(2.829)	(2.447)	(23.780)	(20.614)	(25.035)	(22.877)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)			-5.642**	-5.532**	-5.642**	-5.533**
			(2.684)	(2.284)	(2.683)	(2.284)
Informal Competition (Country level cell average)*Initial Employment (logs)					-0.061	-0.365
					(3.222)	(3.098)
Initial Employment (logs)	-4.150***	-4.274***	-4.114***	-4.242***	-4.084**	-4.060**
	(0.664)	(0.594)	(0.664)	(0.594)	(1.686)	(1.576)
GDP per capita (lagged, logs)	-1.820***	-1.592***	0.649	0.833	0.650	0.837
	(0.600)	(0.453)	(1.143)	(0.999)	(1.138)	(1.001)
Firm Exports Y:1 N:0	-2.262	-1.456	-2.963	-2.140	-2.982	-2.247
	(2.111)	(2.140)	(2.115)	(2.198)	(2.284)	(2.378)
Age of Firm (logs)		-3.424***		-3.391***		-3.390***
		(1.017)		(1.013)		(1.012)
Manager Experience (logs)		0.020		0.010		0.012
		(0.349)		(0.353)		(0.353)
Exports (proportion of sales)		-1.751		-1.707		-1.719
		(1.767)		(1.776)		(1.773)
Female Top Manger Y:1 N:0		-1.381		-1.386		-1.395
		(1.547)		(1.538)		(1.568)
Multi-establishment firm Y:1 N:0		3.360***		3.364***		3.374***
		(0.956)		(0.956)		(0.959)
Crime losses Y:1 N:0		-0.052		-0.156		-0.154
		(1.170)		(1.171)		(1.163)
Courts obstacle Y:1 N:0		-0.478		-0.531		-0.529

		(0.618)		(0.614)		(0.605)
Power outages (hours)		-0.007*		-0.007*		-0.007*
		(0.004)		(0.004)		(0.004)
Finance obstacle (0-4 scale)		-0.336		-0.350		-0.351
		(0.240)		(0.241)		(0.243)
Overdraft Y:1 N:0		3.143***		3.167***		3.168***
		(0.677)		(0.680)		(0.681)
Inflation (CPI, annual %)		0.129***		0.139***		0.138***
		(0.037)		(0.035)		(0.035)
Primary Enrollment (Gross rate)		0.048**		0.045*		0.045*
		(0.024)		(0.025)		(0.025)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	25.354***	25.261***	-9.099	-7.745	-9.192	-8.262
	(5.775)	(6.959)	(13.091)	(12.783)	(13.572)	(13.549)
Number of observations	20,295	20,295	20,295	20,295	20,295	20,295
R squared	0.062	0.101	0.063	0.102	0.063	0.102

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 17: Interaction with Overdraft dummy						
Dependent variable: Employment Growth (log difference, %, annual)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Informal Competition (Country level cell average)*Overdraft Y:1 N:0	**9.231****	**8.589****	**11.402****	**10.561****	**11.388****	**10.531****
	(4.459)	**(4.168)**	**(4.390)**	**(4.097)**	**(4.578)**	**(4.364)**
Informal Competition (Country level cell average)	-8.829***	-8.589***	56.125**	51.787**	56.004**	51.519**
	(3.130)	(2.978)	(24.465)	(21.324)	(25.868)	(23.561)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)			-7.005**	-6.512***	-7.005**	-6.513***
			(2.788)	(2.388)	(2.788)	(2.386)
Informal Competition (Country level cell average)*Initial Employment (logs)					0.048	0.104
					(3.160)	(3.000)
Initial Employment (logs)	-4.219***	-4.186***	-4.189***	-4.156***	-4.213**	-4.208***
	(0.633)	(0.574)	(0.633)	(0.575)	(1.609)	(1.542)
GDP per capita (lagged, logs)	-1.514***	-1.322***	1.552	1.543	1.551	1.542
	(0.558)	(0.465)	(1.231)	(1.067)	(1.230)	(1.069)
Overdraft Y:1 N:0	-0.862	-0.704	-1.855	-1.593	-1.848	-1.579
	(2.005)	(1.928)	(1.986)	(1.913)	(2.063)	(2.025)
Age of Firm (logs)		-3.377***		-3.333***		-3.333***
		(1.006)		(0.998)		(0.997)
Manager Experience (logs)		0.055		0.036		0.035
		(0.342)		(0.342)		(0.342)
Exports (proportion of sales)		2.167*		2.097*		2.102*
		(1.098)		(1.081)		(1.076)
Female Top Manger Y:1 N:0		-1.420		-1.431		-1.428
		(1.547)		(1.534)		(1.561)
Multi-establishment firm Y:1 N:0		3.391***		3.386***		3.383***
		(0.932)		(0.929)		(0.927)
Crime losses Y:1 N:0		0.106		-0.006		-0.007
		(1.164)		(1.163)		(1.153)
Courts obstacle Y:1 N:0		-0.389		-0.458		-0.459
		(0.637)		(0.631)		(0.623)
Power outages (hours)		-0.007*		-0.007*		-0.007*
		(0.004)		(0.004)		(0.004)

Finance obstacle (0-4 scale)		-0.320		-0.331		-0.331
		(0.235)		(0.236)		(0.237)
Inflation (CPI, annual %)		0.127***		0.138***		0.138***
		(0.036)		(0.033)		(0.033)
Primary Enrollment (Gross rate)		0.049**		0.047*		0.047*
		(0.025)		(0.025)		(0.025)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	20.836***	22.206***	-19.952	-15.711	-19.881	-15.563
	(5.128)	(6.798)	(13.754)	(13.317)	(14.423)	(14.216)
Number of observations	20,295	20,295	20,295	20,295	20,295	20,295
R-squared	0.068	0.099	0.070	0.100	0.070	0.100

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 18: Cheating on Taxes and Informality						
Dependent variable: Employment Growth (log difference, %, annual)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Informal Competition (Country level cell average)*Cheating on Taxes is Justified (dummy)	**-5.984**	**-18.979****	**-15.545****	**-6.861**	**-18.884****	**-15.630****
	(8.295)	**(8.091)**	**(6.109)**	**(6.207)**	**(8.278)**	**(6.256)**
Informal Competition (Country level cell average)	-4.490	-2.633	-4.252	-40.759	27.078	11.011
	(5.547)	(4.630)	(3.660)	(45.691)	(55.300)	(38.435)
Initial Employment (logs)	-3.476***	-3.701***	-3.995***	-3.654***	-3.703***	-3.994***
	(0.747)	(0.737)	(0.685)	(0.738)	(0.738)	(0.685)
Cheating on Taxes is Justified (dummy)	2.926	5.530	3.550	2.576	5.483	3.610
	(4.340)	(3.837)	(3.045)	(3.589)	(3.955)	(3.144)
Informal Competition (Country level cell average)*GDP per capita (lagged, logs)				3.779	-3.169	-1.617
				(4.647)	(5.916)	(4.051)
GDP per capita (lagged, logs)		-2.957***	-2.615***	-3.414*	-1.682	-1.979
		(0.937)	(0.547)	(1.888)	(2.363)	(1.747)
Age of Firm (logs)			-3.350***			-3.350***
			(1.228)			(1.226)
Manager Experience (logs)			0.355			0.365
			(0.425)			(0.427)
Exports (proportion of sales)			1.586			1.581
			(1.323)			(1.320)
Female Top Manger Y:1 N:0			-1.481			-1.478
			(1.839)			(1.837)
Multi-establishment firm Y:1 N:0			4.250***			4.251***
			(1.339)			(1.339)
Crime losses Y:1 N:0			0.215			0.219
			(1.448)			(1.446)
Courts obstacle Y:1 N:0			-0.059			-0.067
			(0.842)			(0.841)
Power outages (hours)			-0.009***			-0.009***
			(0.003)			(0.003)
Finance obstacle (0-4 scale)			-0.348			-0.353

			(0.277)			(0.277)
Overdraft Y:1 N:0			3.160***			3.162***
			(0.697)			(0.701)
Inflation (CPI, annual %)			0.164***			0.167***
			(0.033)			(0.033)
Primary Enrollment (Gross rate)			0.082**			0.079**
			(0.031)			(0.031)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	15.488***	18.009	18.663	48.327**	6.144	12.927
	(3.488)	(26.266)	(25.317)	(18.870)	(32.847)	(29.340)
Number of observations	14,398	14,398	14,398	14,398	14,398	14,398
R squared	0.040	0.061	0.104	0.045	0.061	0.104

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).

Table 19: Large firm sample					
Dependent variable: Employment Growth (log difference, %, annual)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Informal Competition (Country level cell average)	1.667	1.312	3.490	3.079	2.500
	(4.327)	(4.313)	(4.286)	(4.384)	(4.521)
Initial Employment (logs)	-1.110*	-0.964*	-1.336***	-1.497***	-1.518***
	(0.579)	(0.503)	(0.495)	(0.456)	(0.493)
GDP per capita (lagged, logs)		-1.654**	-1.722**	-1.942**	-1.798**
		(0.675)	(0.725)	(0.772)	(0.764)
Age of Firm (logs)			-0.886	-0.853	-0.866
			(0.630)	(0.632)	(0.639)
Manager Experience (logs)			-0.158	-0.270	-0.174
			(0.617)	(0.574)	(0.570)
Exports (proportion of sales)			4.307***	4.182***	4.320***
			(1.457)	(1.344)	(1.407)
Female Top Manger Y:1 N:0			-0.544	-0.645	-0.619
			(0.949)	(1.050)	(1.026)
Multi-establishment firm Y:1 N:0			2.093***	2.040***	2.106***
			(0.780)	(0.708)	(0.733)
Crime losses Y:1 N:0				2.592**	2.647**
				(1.104)	(1.106)
Courts obstacle Y:1 N:0				-2.090***	-1.995***
				(0.761)	(0.753)
Power outages (hours)				0.000	0.001
				(0.004)	(0.005)
Finance obstacle (0-4 scale)				-0.962***	-0.969***
				(0.284)	(0.275)
Overdraft Y:1 N:0				1.832***	1.878***
				(0.581)	(0.530)
Inflation (CPI, annual %)					0.130***
					(0.048)
Primary Enrollment (Gross rate)					0.038
					(0.050)
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Constant	6.647*	13.719	16.710**	21.724**	15.691
	(3.997)	(8.505)	(8.326)	(8.276)	(10.593)
Number of observations	5,326	5,326	5,326	5,326	5,326
R-squared	0.004	0.034	0.053	0.072	0.075

Huber-White robust standard errors and clustered on the country in brackets. Significance is denoted by *** (1%), ** (5%), * (10%).