

# Does Competition from Informal Firms Impact R&D by Formal SMEs?

Evidence Using Firm-Level Survey Data

*Mohammad Amin*



**WORLD BANK GROUP**

Development Economics

Global Indicators Group

November 2021

## 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 it may undermine growth and development of the formal sector. For instance, the growth literature indicates that research and development activity and innovation are a key driver of long-term growth. How does the competition that formal sector firms face from informal sector firms affect research and development activity by the formal firms? The present paper attempts to answer this question using firm-level survey data for small and medium-size enterprises in a large cross-section of mostly developing countries. The results show that higher informal competition leads to greater a likelihood of spending on research and development by formal firms. For the most conservative baseline specification, a one standard deviation increase in informal competition leads to an

increase of 5.2 percentage points in the likelihood of spending on research and development by formal firms. This is a large increase given that less than 18 percent of the firms in the sample engage in research and development activity. Further, consistent with the “parasite” view of informality, the positive impact of informal competition on research and development activity is magnified when the business environment is less conducive to operating in the formal sector compared with informal sector due to factors such as higher corruption, weaker rule of law, more burdensome business regulations, and a higher tax rate on profits. As expected, there is no impact of informal competition on research and development activity among large firms. The main findings are robust to several controls, alternative specifications, and endogeneity checks.

---

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](mailto:mamin@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# **Does Competition from Informal Firms Impact R&D by Formal SMEs? Evidence Using Firm-Level Survey Data**

By: Mohammad Amin\*

-

Keywords: Informality, R&D, Competition, Firm, Developing countries  
JEL: O17, O12, O30

---

\*Senior Economist, Enterprise Analysis Unit, World Bank, Washington DC. Email: [mamin@worldbank.org](mailto:mamin@worldbank.org)

The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

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

I thank the Enterprise Analysis Unit of the Development Economics Global Indicators Department of the World Bank Group for making the data available.

## **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 hurt the growth and development of the formal sector firms. For instance, it is argued in the growth literature that more investment in R&D stimulates technological progress, which provides the foundation for the bulk of economic growth around the world (Grilliches 1998, Acemoglu 2009). A systematic and rigorous analysis of the impact of informal sector on R&D activity by formal sector firms, however, is lacking.

The present paper attempts to fill this gap in the literature by estimating the impact of competition from informal sector firms faced by formal sector firms (henceforth, informal competition) on the formal firms' likelihood of spending on R&D activity. It does so by using firm-level survey data on small and medium enterprises (SMEs) in the formal manufacturing sector of a large cross-section of mostly developing countries. Our results show that greater informal competition leads to a significantly higher likelihood of firms spending on R&D activity. This is consistent with the broader literature that suggests that greater competition forces firms to innovate to differentiate their product and cut costs in order to survive. Further, we find that the positive impact of informal competition on the likelihood of R&D spending by formal firms (henceforth, R&D activity) is much larger (more positive) in countries where the cost of operating in the formal vs. the informal sector is higher and the benefits lower. This finding is along expected lines and as per the predictions of the "parasite" view of informality. As we argue below, it serves as an important check against endogeneity concerns with our estimation.

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.

Another strand of the literature analyzes the impact of greater competition on firms’ R&D and innovation effort. One view going back to Schumpeter (1942) is that greater market power or less competition leads to higher R&D effort by firms. Greater market power encourages R&D by reducing the uncertainty associated with excessive rivalry between competing firms, increasing appropriability of the results of investment in R&D, and increasing the rate of return from R&D activity. The opposite view is also found in the literature, which is that greater market power or less competition discourages R&D effort and innovation due to greater inertia and the creation of

oligopolistic rents that would be destroyed by innovation (Arrow 1962, Scherer 1980, Porter 1990). Empirical evidence on the issue is mixed (discussed in detail in a later section).

One may argue that informal competition is like any other competition. Hence, a separate analysis of the effects of informal competition is not warranted. However, this is not the case. And for several reasons. First, competition between formal firms occurs on a level playing field. This is not the case with informal competition. That is, informal firms tend to enjoy an “unfair” advantage over formal firms since informal firms do not pay taxes and do not comply with costly regulations that formal firms must shoulder. Of course, this “unfair” advantage may be countered at least partly by the benefits of being formal, such as better access to finance, patent protection, etc. The net effect of these contrasting forces which we treat as an empirical issue will determine the nature and level of informal competition faced by formal firms. Second, firms engage in R&D activity to escape competition and maintain their market share. However, it is not clear how much R&D and innovation matters for escaping competition provided by informal firms. It is possible that regular innovations may not be enough for formal firms to overcome the large cost advantage that informal firms typically have over formal firms. Further, competition between formal and informal firms may be more about delivery methods and customized service rather than product or process innovation. Third, there may be limited overlap in the product space between formal and informal firms. This may happen when they target different types of clients and customer needs. Thus, formal firms may compete with informal firms, but the intensity of such competition may be limited and different from the intensity of competition between formal firms that is analyzed in the existing studies.

Theoretically, it is difficult to predict if informal competition has a positive or negative impact (or no impact) on R&D activity of the formal firms. The reason for this is the same as

discussed above while reviewing the literature on competition (between formal firms) and R&D activity. Thus, the issue needs to be resolved empirically.

A rigorous empirical analysis of the impact of informal competition on R&D activity is very limited or non-existent. 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. Mendi and Costamagna (2017) use Enterprise Surveys data to estimate the impact of informal competition on the likelihood of innovation among formal firms. However, the study uses formal firms' perceptions of the informal sector as an obstacle for their operations as a measure of informal competition faced by the formal firms rather than the actual experience with informal competition. Further, the study is restricted to firms in Africa and Latin America.

The impact of informal competition on other aspects of formal firms' performance 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). Gonzales 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. To reiterate, none of these studies assesses the impact of informal competition on formal firms' R&D effort.

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

R&D. 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 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 R&D activity. Moreover, La Porta and Shleifer (2008) use old data collected between 2002 and 2007.

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 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 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 likelihood of spending on R&D by formal manufacturing SMEs. To the best of our knowledge, this is a relatively unexplored area of research.

Second, informal firms are known to be very small and therefore unlikely to pose any problems for 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. We use cross-country comparable firm-level survey data for our analysis. These data allow us to distinguish between SMEs and large firms. Our results show that informal competition affects SMEs but not the large firms.



Third, we pay due attention to endogeneity concerns due to possible reverse causality and omitted variable bias. 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 based on the idea that informality impacts formal firms more when the cost of operating in the formal versus informal sector is higher and the benefits lower; and contrasting the results for SMEs with those for large firms. To provide an example, we briefly discuss here the check based on testing the predictions of the “parasite” view of informality. This is as follows.

One view in the literature is that when formal firms face competition from informal firms, the intensity of competition varies depending in part on how conducive the business environment is to operate in the formal vs. informal sector. For instance, when rule of law is strong and enforcement strict, informal firms fear detection and therefore they are forced to operate on a small scale as fringe players. This blunts the competition that they provide to the formal firms. Similarly, as per the “parasite” view of informality, informal firms enjoy an “unfair” advantage over formal firms since they do not pay taxes and do not comply with costly regulations that formal firms must shoulder. Thus, higher taxes on (formal) businesses, stricter business regulations, and higher corruption exacerbate the “unfair” advantage, allowing informal firms to compete more vigorously with formal firms (Distinguin et al. 2016). Thus, the testable hypothesis is that if the positive impact of informal competition on R&D activity that we find is indeed causal then it should be greater (more positive) when rule of law is weaker, business regulations are stricter, taxes are higher, etc. However, if our empirical results suffer from reverse causality or omitted variable bias

problem, there is no reason why the observed relationship between informal competition and R&D activity should follow the stated predictions.

The empirical exercise reveals a large positive impact of informal competition on the likelihood of R&D by formal firms. Thus, competition from informal firms spurs formal sector firms to spend on R&D to keep abreast of the competitive pressure. Figures 1 and 2 illustrate the point graphically. Figure 3 shows that the positive relationship holds even after accounting for differences in firm-size, industry composition and GDP per capita. According to our most conservative baseline estimate, a one standard deviation increase in the level of informal competition increases the probability of a firm spending on R&D by about 5.2 percentage points. This is a large increase given that the probability of a firm spending on R&D in our sample equals about 18 percent.

Further, as indicated above, the positive relationship between informal competition and R&D activity is much larger (more positive) in countries where the cost of operating in the formal vs. informal sector is higher and benefits lower as captured by higher corruption, more burdensome business regulations, higher profit tax rate, and weaker rule of law. Last and as expected, we find no significant impact of informal competition on R&D by 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 baseline regression results. 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 Indicators (WDI, World Bank), Worldwide Governance Indicators (WGI, World Bank), Doing Business (World Bank), and Economic Freedom of the World (Fraser Institute).

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 (registered with 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.<sup>1</sup> 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.

The sample that we use consists of SMEs in the manufacturing sector of mostly developing countries. Based on the definition used by the ES for stratification purposes, we define SMEs as all firms with fewer than 100 full-time permanent workers employed at the firm three fiscal years ago (lagged employment).<sup>2</sup> From this sample, we exclude firms with 10 or fewer employees. We do so because R&D activity involves substantial fixed costs and is therefore of not much relevance for the very small firms. Large firms (100 or more workers) are excluded because these firms operate in very different product markets than the informal firms. Thus, spillovers from informal sector to the large formal sector firms are unlikely. Results for large firms are shown in separate

---

<sup>1</sup> Details of the sampling methodology and other related information are available at [www.enterprisesurveys.org](http://www.enterprisesurveys.org).

<sup>2</sup> Information is available in the ES on the number of workers at the end of the last fiscal year and 3 fiscal years ago (or 2 fiscal years ago in the case of some countries). We use the latter for defining SMEs in order to avoid simultaneity concerns with the current (last fiscal year end) employment. See for example, Aterido et al. (2011) on the use of lagged employment to avoid the simultaneity problem.

regressions to contrast with the ones for SMEs. Throughout, the sample used is a pure cross-section in that each country and firm is included only once. The most recent round of ES available for the country is used. The firms were surveyed between 2006 and 2018. For our baseline regressions, the sample consists of 12,980 firms spread across 85 countries.

## 2.1 Estimation methodology

The baseline regression exercise involves estimating the following equation:

$$Y_{ij} = \alpha + \beta_1 \text{Informal Competition}_{-i,j} + YFE + IFE + \text{Firm Controls}_{ij} + \text{Country Controls}_j + u_{ij} \quad (1)$$

where subscript  $i$  denotes the firm,  $j$  the country to which the firm belongs.  $Y$  is a dummy variable indicating R&D activity during the last year (defined below); *Informal Competition* is a measure of competition that the firm faces from the informal sector firms. *YFE* is a set of 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 (discussed below).  $u$  is the error term.

Unless stated otherwise, logit estimation method is used for estimating the above equation.<sup>3</sup>

All regressions use Huber-White robust standard errors clustered on the country or on the sub-national level as appropriate.<sup>4</sup> We pay due attention to endogeneity concerns (details below).

---

<sup>3</sup> The only exception is the two-stage instrumental variables (IV) estimation in section 4.4. For this estimation, the IV probit method is used.

<sup>4</sup> In one robustness check (section 4.2), we exploit variation at the sub-national level. In these regressions, the standard errors are clustered at the corresponding sub-national level.

In the online Appendix (included at the end of this paper), Table A1 contains a formal definition of all the variables used in the regressions; summary statistics of the variables used in the baseline regressions are provided in Table A2 while Table A3 provides the correlations between our main explanatory variable, informal competition, and other explanatory variables used in the baseline specification.

## *2.2 Dependent variable*

The dependent variable is an indicator for R&D activity conducted by the firm. It is a dummy variable equal to 1 if the firm spent on R&D activity during the last fiscal year covered by the ES and 0 otherwise (*R & D Activity*). For our baseline sample, the mean value of the variable equals 0.179 and the standard deviation is 0.383. Thus, about 17.9 percent of the firms in our sample spent on R&D activity in the last year.

## *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' R&D activity (reverse causality problem) and firm characteristics (omitted variable bias problem). One solution suggested in the literature is to proxy the firm-level variable, informal competition faced by a firm in our case, by the average value of the variable over all other firms (that is, other than the firm in question) in the same "cell". The cell can be variously defined based on the firms' location, industry, firm-size, and so on. This strategy of using cell averages to overcome endogeneity concerns has been discussed and used in the literature (see for example, Aterido et al. 2011, Amin and Soh 2020, Dollar et al. 2006, Fisman and Svensson 2007, De Rosa et al. 2010).

Applied to this study, the strategy assumes that firms within the same cell are likely to face similar levels of informal competition (confirmed below). Another assumption is that a given firms' R&D activity has no or very little effect on the likelihood of other firms in the cell facing informal competition (reverse causality problem). Similarly, own firms' characteristics that may affect its R&D activity 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 rule 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 the variable of interest (Pounov 2016).

For our main regressions, we use cell averages at the country-level. This follows Distinguin et al. (2016) who also use the proportion of formal firms in the country that compete against informal firms as the instrument for the level of informal competition faced by the firm in question. Thus, 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 where the firm is located 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 0.533 and the standard deviation is 0.209. Thus, on average, about 53.3 percent of the firms in the country-cell face competition from informal firms.

In the robustness section, we show that our main result continues to hold when we use cell averages at more disaggregated (sub-national) levels. These cell averages are at the country times industry times firm-size level and city times industry times firm-size level. Regressions using cell

averages at the country-industry-firm size level control for all country-wide factors (country fixed effects) that may affect R&D activity. Similarly, regressions using cell averages at the city-industry-firm size level account for all city-wide determinants of R&D activity (city fixed effects). Thus, the identification of the relationship between R&D activity and informal competition comes from variations across firms within a country (or city) rather than across countries. For the sub-national cells, industry grouping is at the 2-digit ISIC Rev. 3.1 level (7 industries)<sup>5</sup>; firm-size is defined as small (11-20 employees three fiscal years ago) and medium (21-99 employees); and city is the sub-national sampling region used for stratification purposes by the ES. As above, we exclude all cells with fewer than 5 firms.

#### 2.4 Controls

To further raise our confidence against endogeneity concerns, we control for several determinants of R&D activity that also may be correlated with informal competition. For this, we draw on the literature on the drivers of R&D expenditures - input of R&D activity - and innovation - output of R&D activity. Unless stated otherwise, in this section, the terms R&D activity and innovation are used synonymously.<sup>6</sup>

At a broad level, the determinants of R&D activity can be divided into two groups. In his seminal work, Schumpeter (1942) argued that relatively large firms and firms with greater market power are more likely to do R&D than rest of the firms. This is the first group of factors driving R&D activity. Subsequent analysis identified several other factors affecting R&D activity

---

<sup>5</sup> Using 3-digit instead of 2-digit industry grouping here leads to too few firms in several cells, lowering the reliability of the cell-averages.

<sup>6</sup> One possibility we would like to rule out is that a high level of informality diverts resources away from the formal to the informal sector. This may lower R&D 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 R&D activity between formal firms that compete against informal firms and formal firms that do not do so. Thus, any overall impact of informality on the formal sector does not affect our results.

classified as economic opportunity, technological opportunity and appropriability (Levin et al. 1985, Cohen and Levin 1989). These comprise the second group of factors. Within these broad categories, several firm, industry, and country characteristics have been identified. Details follow.

Schumpeter (1942) argued that greater market power encourages R&D by reducing the uncertainty associated with excessive rivalry between competing firms, increasing appropriability of the results of investment in R&D, and by increasing returns from R&D activity. However, others argue that the opposite is also possible. That is, greater market power may discourage innovation due to greater inertia or the creation of oligopolistic rents that would be destroyed by innovation (Arrow 1962, Scherer 1980, Porter 1990). Empirical evidence on the issue is also mixed. Some studies report a positive impact of greater market power on R&D activity (Crepon et al. 1998, Blundell et al. 1999, Nicholas 2003); other studies report a negative impact (Geroski 1990, Harris et al. 2003, Cohen 2010) or no impact (Broadberry and Crafts 2001). Yet others report a U-shaped relationship (Aghion et al. 2005). Some recent work suggests a positive effect of competition on innovation in firms or industries that are close to the technology frontier and a negative or no effect otherwise (Aghion et al. 2009, Amiti and Khandelwal 2013). As Symeonidis (2019) points out, the mixed results have led several surveys of the empirical literature to conclude that there may be no general relationship between market power and innovation and that industry characteristics such as appropriability conditions, demand and especially technological opportunity explain much more of the cross-industry variation in innovative activity than market power or market structure.

The importance of firm size for innovation activity has been extensively discussed. Following Schumpeter (1942), several arguments have been put forward to support the hypothesis that R&D increases more than proportionately with firm size. This happens because relatively large firms benefit from economies of scale and scope in R&D technology, more efficient



implementation, higher returns from R&D and greater ability to secure funding for risky projects given capital market imperfections. However, the opposite claim is also made. It is argued that small firms are more efficient in performing R&D because they do not suffer from lack of managerial control as large firms do (Holmstrom, 1989); incentives of individual scientists and entrepreneurs are greater in small firms as they have more possibilities to benefit from the results of their work. While there are exceptions, the empirical evidence tends to favor the view that the likelihood of R&D activity rises with the size of the firm (Cohen and Klepper 1996, Nilsen and Schiantarelli 2003, Mairesse and Mohnen 2002, Cohen 2010).

Technological opportunities refer to the possibilities for advancing the knowledge frontier. Such opportunities are typically industry-wide and make innovation easier (less costly) in the relevant industries (Cohen and Levin, 1989). Some caution is necessary here because technological opportunity may raise the average product of R&D without raising its marginal product, and therefore investment in R&D may not rise (Klevorick et al, 1995). Empirical evidence on the impact of technological opportunities on innovation is mixed (see for example, literature review by Barge-Gil and Lopez 2014).

The incentive for R&D activity also depends on economic opportunities and appropriability of investment. Economic opportunities include factors that affect the cost-benefit calculus of investment in R&D while appropriability refers to the ability of innovating firms to appropriate the benefits of innovation. Several firm, industry and country-level factors are included in these two broad groups (discussed below).

Early contributions in the area suggested that economic opportunities depend in part on demand pull factors (Schmookler 1962). These include the scope of future demand, the size and growth of the market, and the consumer's willingness to pay for new or improved products. The

underlying assumption here is that a common pool of knowledge and capabilities is available to all industries. Thus, large and growing markets provide higher incentives to invest in innovation as these markets offer higher returns for the investment (Cohen and Levin 1989). Later studies in the area confirmed this view (Cohen et al, 1987; Brouwer and Kleinknecht, 1989), although there are exceptions. The geographical extent of the firms' main market (national or international as discussed below) and its location (small vs. large cities, etc.), and recent growth performance of the industry and firm are some of the proxy measures used in the literature for demand size. The growth rate of GDP per capita is also used as a proxy for demand pull forces and found to be an important determinant of R&D activity at the aggregate, industry, and firm-level (see Becker 2013 and the literature cited therein).

Overall financial development in the country and firms' financial condition (access to finance) are highlighted in several studies as important factors in influencing R&D activity. Simply put, R&D is a costly and risky activity. Thus, better access to finance due to a more developed financial system or due to firms' own financial condition (availability of internal funds, cash flows, etc.) significantly aids R&D activity (see for example, Cohen 1995, Geroski et al. 2002, Carline and Mayer 2003, Hall 2008, Aghion et al. 2012, and Maskus et al. 2012).<sup>7</sup> The importance of access to finance is also found to vary across firms and economic conditions. It is much more important among the relatively small and young firms and in times of recessions than otherwise (see for example, Bloch 2005, Brown et al. 2009, 2012, and Aghion et al. 2012).<sup>8</sup>

Age of the firm has also been linked to R&D activity, although in contrasting ways. One view is that a firms' age reflects its higher efficiency since only efficient firms can grow and survive in the market (Ericson and Pakes 1995, Jovanovic 1982). Older firms also have more

---

<sup>7</sup> For a survey of the related literature, see for example Hall and Lerner (2010).

<sup>8</sup> Also, see Becker (2015) and the studies cited therein.

experience, accumulated knowledge, and better access to finance due to their past record of successes. Thus, older firms have a higher propensity to engage in R&D activity. However, the opposite claim that propensity to invest in R&D is lower for the older firms is also made. One reason for this is that as per the “product life cycle” model of Klepper (1996), there are fewer technological opportunities available to the relatively older firms. Another reason is that younger firms are likely to have better access to high-skilled workers because of the dynamic and vibrant environment in which they operate (Baron et al. 2002, Hannan et al. 1996, Pallegriano 2018). This increases their propensity to invest in R&D. Further, older firms tend to suffer from inertia and organizational rigidity which may make them less capable of dealing with uncertain nature of demand for innovative products and services (Nelson and Winter 1982). In contrast, younger firms typically operate in an environment characterized by high level of uncertainty and therefore are more inclined to take risks and explore new ideas and technologies (Audretsch 1995, Coad et al. 2016a, b).

Exporting is another factor found to be positively correlated with the propensity to spend on R&D and innovate. One reason for this is that compared to non-exporters, exporters operate in more dynamic and competitive markets. Hence, they are forced to innovate to remain competitive. Another reason is that exporting markets offer better learning and imitation opportunities in terms of new technologies. Additionally, export markets can increase innovation by increasing market size, thus spreading the cost of innovation over a larger market (for example, Grossman and Helpman 1991, Bustos 2011, Aghion et al. 2018, and Lileeva and Trefler 2010). However, it is not clear whether exporting causes firms to innovate more or that it is the innovating firms that self-select into exporting (see for example, literature survey by Harris and Moffat 2011). For instance, in a recent study, using Slovenian microdata over 1996 to 2002, Damijan et al. (2010) test for

causality running from exporting to innovation as well the reverse. They find no evidence of innovation causing firms to export and instead find that exporting causes firms to innovate more. Other studies that suggest that exporting may lead to higher propensity to innovate include for example, Salomon and Shaver (2005), Yang and Chen (2012), Gorodnichenko et al. (2010), Girma et al. (2008), and Criscuolo et al. (2010).

Available evidence suggests that increasing the supply of human capital in the country through expanded university programs and/or relaxed immigration rules is likely to have positive impact on R&D effort and innovation (Bloom et al. 2019, Toivanen and Lotta Väänänen 2016, Valero et al. 2019, and Hausman 2018). Studies at the firm-level also report higher skill-endowment of firms to be associated with greater innovation activity (Leiponen 2005, Piva and Vivarelli 2009). Cohen and Levinthal (1989, 1990) claim that highly qualified employees are a firm's primary vehicle for absorbing external knowledge and, consequently, for enhancing its absorptive capacity. This argument can be extended to the management of the firm (see for example, Koellinger 2008).

The quality of the business environment and institutional development can impact R&D activity in several ways. The business environment affects how firms appropriate the outcomes of R&D efforts, the rate of return from such efforts, and the ability of firms to exploit technological opportunities (OECD/Eurostat 2018). Several elements of the business environment are at play such as the level of corruption, rule of law and the quality of courts, macro-economic stability, patent protection, investor protection, and so on (see for example, Alam et al. 2019 for an overview of the related literature). For instance, Murphy et al. (1993) argue that innovators are particularly vulnerable to expropriation by government officials due to their inelastic need for government services such as permits and licenses, resulting in high risks, uncertainty and vulnerability. Xu and

Yano (2017) confirm that anti-corruption measures in China had a strong positive impact on R&D activity of the firms. Similarly, Srholec (2011) stated that the institutional settings within which firms operate have a significant impact on its likelihood to innovate. Some of the drivers of R&D discussed above such as industry competition, exports, and financial development depend in part depend on regulatory and other business environment elements.

Based on the discussion above we control for firm, industry and country-level variables. Unless specified otherwise, data source for the control variables is Enterprise Surveys. See Table A1 in online Appendix for a formal definition of all the variables used in the regressions.

For our baseline specification, we begin with accounting for all industry wide features such as market structure and demand, technological opportunities, growth potential, and so on. We do so by using dummy variables indicating the industry to which the firm belongs (Industry fixed effects). We use industry grouping at the 3-digit ISIC Rev. 3.1 level. There are 28 industries in our baseline sample.

Market structure and competition may also vary across countries for a given industry, and across firms within a country-industry pair. This variation is not captured up by the industry fixed effects. Thus, our next control is the concentration of output within each country-industry pair as measured by the Herfindahl-Hirschman Index (HHI). This equals the sum of the square of market shares (in terms of output) of all firms in each country-industry pair ( $HHI$ ). Industry here is defined at the 3-digit ISIC Rev. 3.1 level. The variable is based on ES data for formal firms. Hence, it captures competition between formal firms. To account for differences in market power across firms within a country-industry pair, we control for firm-size (defined below).

The ES data we use was collected in different years for the various countries. Thus, we account for the annual global shocks to R&D activity by using dummy variables indicating the year the ES was conducted in the country (year fixed effects).

Next, we control for overall level of development using (log of) GDP per capita (PPP adjusted and at constant 2011 international dollars). We use lagged values (by two years) of the variable and the data source is World Development Indicators, World Bank. Following the study by Lederman and Maloney (2003), GDP per capita is used a broad proxy measure for financial depth, protection of intellectual property rights, government capacity to mobilize resources, and the quality of research institutions that may impact R&D efforts.

Our next set of baseline controls include several firm characteristics. These are firm-size proxied by the (log of) number of permanent full-time workers at the firm (three fiscal years ago, lagged), a dummy variable equal to 1 if the firm is part of a larger organization and 0 otherwise (*Multi-firm*), age of the firm (logs), managerial experience proxied by the (log of) the number of years of experience the top manager has working in the industry, and a dummy variable equal to 1 if the top manager of the firm is a female and 0 otherwise. For access to finance, we use two separate variables. The first is a dummy variable equal to 1 if the firm has overdraft facility and 0 otherwise (*Overdraft*). The second is how severe is (poor) access to finance as an obstacle to firms' current operations (*Finance Obstacle*). The severity level is as reported by the firms in the ES and recorded on a 0-4 scale with higher values implying a more severe obstacle.<sup>9</sup>

We now aim to capture various dimensions of the business environment. Some of the controls here are defined at the firm-level and others at the country level. The firm-level variables

---

<sup>9</sup> Our main results are unaffected if we use other access to finance variables available in the ES. These include: (i) if the firm has an outstanding loan; (ii) a measure of financial constraint based on if the firm applied for a loan during the last year, the reasons for not doing so if it did not apply for a loan, and the outcome of loan application.

include total hours of power outages faced by the firm in a typical month during the last fiscal year (*Power Outages*), country-level cell average of the severity level of (lack of proper) functioning of courts as an obstacle to firms' current operations (*Courts Obstacle*), availability of human capital in the country proxied by the gross enrollment rate in primary education and the same for tertiary education; the annual rate of inflation based on the consumer price index; and the size of the market size proxied by the (log of) total population of the country.

The controls mentioned above are used for our main or baseline estimation (baseline controls). There are additional factors that may impact R&D activity (discussed above) and can also be potentially correlated with informal competition. In the robustness section, we check if our main results survive controls for these factors. Thus, we consider factors associated with differences in economic activity and performance across world regions. 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 follows: Sub-Saharan Africa (SSA), Latin America and Caribbean (LAC), Middle East and North Africa (MENA), Eastern Europe and Central Asia (ECA), South Asia (SAR) and East Asia and Pacific (EAP).

Recent investments in capital stock expand the productive capacity of the firm and could affect the returns to R&D activity and therefore the decision to spend on R&D. Information in the ES is available on capital stock (replacement and book value of machines and equipment). However, due to the non-response, we cannot use this information in the regressions. Hence, we use 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. 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 features could affect the incentives to conduct R&D activity. 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 R&D activity. 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.

We argued above that firms' past performance and scope for future growth (demand pull factors) may impact the likelihood of R&D activity. We proxy for these factors using the annual growth rate of employment (firm size) over the last 3 years (*Employment Growth*) are reported in the ES<sup>10</sup> and the annual growth rate of real GDP per capita (*Growth rate of GDP per capita*, lagged by two years). Data source for the variable is World Development Indicators, World Bank.

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, tax rates, corruption 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

---

<sup>10</sup> Employment growth is computed using the Haltiwanger method to avoid extremely large or small values of the variable (see Table A1 in online Appendix for more details). Thus, the growth rate is bounded between plus/minus 100.



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*). We also experimented with the Distance to Frontier (DTF) measure due to World Bank's Doing Business project. However, this did not any effect on our main results.

For tax rates, we use the tax rate on firms' profits as measured by the World Bank's Doing Business project. 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 rules 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 R&D activity. Last, our baseline specification controls for human capital using gross enrollment rates in primary and tertiary education. Enrollment in secondary education was not included because it is highly correlated with enrollment rate in primary education. However, in the robustness section, we show that including the control for the gross enrollment rate in secondary education does not impact our main results much. The data source for the secondary education enrollment rate is World Development Indicators, World Bank.

### 2.5 Endogeneity checks

To further raise our confidence against endogeneity concerns, we test for several predictions about how informal competition affects R&D activity by formal sector firms. The tests are based on how the effect of informal competition is moderated by the business environment and certain firm characteristics. These are as follows.

### 2.5.1 *Business environment*

The measure of informal competition that we use is the proportion of formal firms in the country that face competition from informal firms. The intensity of such competition may still vary depending on various factors highlighted in the literature. Weaker rule of law that makes it easier for informal firms to operate with less fear of being detected, higher tax rate which reduces profitability of formal firms but not the informal firms, higher corruption that informal firms are able to evade but not the formal firms, and more burdensome business regulations that increase the cost of doing business for formal firms but not the informal firms are some of the factors that allow informal firms, often “unfairly”, to compete more vigorously against formal firms. This point has been made in the literature in discussions of the “parasite” view of informality and the associated “unfair” advantage enjoyed by informal firms because informal firms do not pay taxes and do not comply with costly regulations that formal firms must shoulder.

Our motivation for the endogeneity check closely follows recent work by Distinguin et al. (2016). The authors 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.”

They further note that (page 21):

“High corruption levels not only force governments to impose higher tax rates; they may also lead to poor quality of the judicial system. This may adversely affect formal firms’ investment and innovation appetite, which could result in formal firms losing their competitive edge over informal

firms. Weak control of corruption hence makes it easier for informal firms to inflict an unfair cost advantage over formal firms consistent with the parasite view.”

The discussion above suggests that the impact of informal competition on the formal firms is magnified when the business environment is less conducive (more costly or less beneficial) to operating in the formal compared to the informal sector. As Distinguin et al. (2016) point it, the cost of operating in the formal sector is higher and benefits lower when there is higher corruption in the country, business regulations are more burdensome, tax rates (on formal businesses) are higher and the rule of law (enforcement) is weaker. These predictions are specific to functioning of the informal sector and therefore we can expect them to hold if our result for the relationship between informal competition and R&D activity is truly causal. However, if the result suffers from reverse causality problem or informal competition is a mere proxy for other correlated determinants of R&D activity (omitted variable bias problem), there is no reason for the observed correlation between R&D and informal competition to vary with corruption, business regulations, rule of law, etc. as predicted above. This constitutes our first endogeneity test.

We implement this test by using interaction terms between our informal competition variable and the following business environment 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; Freedom from Business Regulations taken form the Heritage Foundation’s Economic Freedom of the World database; and Distance to Frontier (DTF) measure of regulatory burden on the private firms due to World Bank’s Doing Business project.<sup>11</sup> Note that except for the profit tax

---

<sup>11</sup> The DTF measure is based on 8 sub-components for which data are available for the countries and years in our baseline sample. We focus on 5 of the 8 sub-components that are more directly related to the regulatory environment relevant for the formal-informal interface. See Table A1 in the online Appendix for more details.

rate, higher values of all the other business environment measures mean that the business environment is more conducive to operating in the formal versus the informal sector. So, the test against the endogeneity problem is that the interaction term between informal competition and profit tax rate is positive while all the other interaction terms are negative.

### *2.5.2 Exporting and large firms*

For the second group of endogeneity tests, we argue that for a given increase in informal competition, certain types of firms are likely to suffer more than the others. Thus, we check that a unit increase in informal competition leads to a much smaller increase in R&D activity among large firms compared to SMEs, and among exporting firms vs. non-exporting firms. It is well-known that large firms operate in very different product markets and therefore are 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 the informal firms. We expect the stated predictions to hold if informal competition has a causal impact on R&D activity but there is no strong reason for it to hold otherwise.

## **3. Base Regression Results**

Our base regression results are provided in Table 1. The coefficient values shown in the table are log odds ratios obtained from logit estimation. The associated marginal effects are provided in Table 2 and discussed later in the section.

All the specifications in Table 1 control for industry and year fixed effects. The remaining baseline controls are added sequentially. Table 1 reveals a large and positive relationship between informal competition and R&D activity. The relationship is significant at the 1 percent level in all

the specifications. Without any other controls (except for industry and year fixed effects), the estimated coefficient value of informal competition equals 1.89 (column 1). The coefficient value increases somewhat when we include the various baseline controls in the specification. Adding the control for GDP per capita to the specification causes the estimated coefficient value of informal competition to increase from 1.89 above to 2.02 (column 2) and to 2.08 when we also include the controls for firm-size and HHI (column 3). The coefficient value rises again from 2.08 above to 2.18 when we add the remaining firm-level controls to the previous specification (column 5), and to 2.19 when we also include the remaining country-level controls (column 6).

The marginal effects corresponding to the specifications in Table 1 are provided in Table 2. The marginal effect (of informal competition on R&D activity) is positive, large and statistically significant at the 1 percent level in all the specifications. Without any controls (except for the industry and year fixed effects), the estimated marginal effect equals 0.251 (column 1). This is our most conservative estimate in Table 2. The estimate implies that a one standard deviation increase in informal competition leads to 5.2 percentage points increase in the probability of a firm spending on R&D activity. Alternatively, for each 10 percentage points increase in (the proportion of firms in the country that face) informal competition, the probability that a firm spends on R&D activity increase by about 2.51 percentage points. These are quantitatively large increases in R&D activity given that the probability of a typical firm spending on R&D activity in our sample equals 17.9 percent.

The estimated marginal effect of informal competition becomes larger when we add the baseline controls to the specification. It rises from 0.251 above (column 1) to 0.269 when we add the control for GDP per capita (column 2) and to 0.272 when we also control for firm-size and HHI (column 3). Adding the remaining firm-level controls to the previous specification causes the

estimated marginal effect to increase slightly from 0.272 above to 0.281 (column 5). Thereafter, it remains unchanged when we also control for all the remaining country-level variables (column 6).

Regarding the various controls, larger firm-size, higher exports, and having overdraft facility are associated with a higher probability of R&D activity, significant at the 1 percent level. These findings are along expected lines. For instance, compared to a firm that has no overdraft facility, the probability of spending on R&D activity is higher by about 4.8 percentage points (against a mean value of 17.9 percent) for a firm that has overdraft facility. We find no evidence of any significant impact of market concentration (HHI) on R&D activity. However, in some of the other specifications that we consider below, the variable has a strong positive and significant (at the 5 percent level) relationship with R&D activity.

## **4. Robustness**

### *4.1 Additional controls*

Regression results with the additional controls mentioned in section 2 are provided in Table 3 (marginal effects). Starting with the final baseline specification (column 6, Table 2), the additional controls are added sequentially. In column 1 of Table 3, we add controls for world region dummies (LAC, EAP, etc.); column 2 further includes controls for foreign ownership, purchase of fixed assets dummy, and city-size dummies; controls for time tax and secondary education are added in column 3, controls for the business environment including rule of law, control of corruption, and freedom from business regulations are added in column 4, and controls for growth rate of GDP per capita and firms' employment growth rate are included in column 5.

Our result for the positive relationship between informal competition and R&D activity survives the additional controls. The estimated marginal effect of informal competition remains

positive, large and significant at the 1 percent level in all the specifications in Table 3. The marginal effect is smaller due to the additional controls. That is, starting with the marginal effect of 0.281 in the final baseline specification (column 6, Table 2), it declines to 0.235 with all the additional controls included in the specification (column 5, Table 3). The decline is entirely due to the control for world regions dummies. The remaining controls mentioned in the previous paragraph cause the marginal effect of informal competition to slightly increase compared to the final baseline specification.

#### *4.2 Within-country variation*

The next robustness check involves using cell averages of informal competition at the sub-national level. These cell averages are defined at the country times industry times firm-size level and at city times industry times firm-size level. We make a few changes from the baseline model. First, since firm-size category is used for constructing the cells, we control for dummy variables indicating the firm-size (Firm-size fixed effects). That is, whether the firm is small (11-20 employees) or medium (21-99 employees). Second, when using cell averages at the country-industry-firm size level, we control for dummy variables indicating the country to which the firm belongs (Country fixed effects). Similarly, regressions using cell averages at the city-industry-firm size level control for dummy variables indicating the city (sampling region) where the firm is located (City fixed effects). Third, cell averages of the court obstacle control variable are computed at same sub-national cell level as informal competition.

Table 4 contains the results for the sub-national cell averages. Columns 1 to 4 contain results for country-industry-firm size cell averages while columns 5 to 8 contain the results for city-industry-firm size cell averages. The results in Table 4 show that our main finding of a

positive relationship between R&D activity and informal competition continues to hold. The relationship is significant at close to the 10 percent level in one specification (column 5) and at 5 percent level in the remaining specifications. Compared to the baseline model, the estimated coefficient value of informal competition (log odds ratios) is much smaller. The coefficient value of informal competition ranges between 1.88 and 2.19 in the baseline model (Table 1). The range when using country-industry-firm size cell averages is 0.69 to 0.77. The same for city-industry-firm size cell averages is 0.75 and 1.05. Focusing on the results with cell averages at the country-industry-firm size level and the final specification (column 4, Table 4), the marginal effect of (a unit) increase in informal competition is an increase in the probability of R&D activity by a firm by 8.6 percentage points. This is a large increase given that the probability of R&D activity by a firm in our sample equals 17.8 percent. However, it is much smaller than the corresponding increase of 28.1 percentage points in the final baseline specification (column 6, Table 2).

#### *4.3 Two-stage estimation*

So far, we used cell averages of informal competition directly in the regressions. An alternative is two stage instrumental variables (IV) estimation. Using the IV probit estimation method, 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. In the second stage, we regress dummy for R&D activity on the predicted values of informal competition obtained in the first stage. These two-stage regression results for all the baseline specifications (as in Table 1) are provided in Table A4 in the online Appendix. They 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 positive relationship between (instrumented values of) informal competition dummy and R&D activity. Like the baseline results, the relationship is significant at the 1 percent in all the specifications.

## **5. Endogeneity Checks**

In this section, we discuss checks against endogeneity problem with our main result. For each of the endogeneity checks, we repeat the regression exercise in the baseline model (Table 1) but now add interaction terms as discussed in section 2. For brevity, results are shown for only some of the baseline specifications. Results for the remaining baseline specifications are similar are provided in the online Appendix (indicated below).

### *5.1 Business environment and informality*

The first prediction we test is that the positive relationship between informal competition and R&D activity is stronger (more positive) 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 informal competition and the business environment variables. Regression results for the final specification (with all the baseline controls included) are provided in Table 5. The results are shown with and without controlling for the interaction term between GDP per capita and informal competition to address any concerns that our business environment variables may be spuriously picking up the effects of overall economic development. Results for the other specifications shown in Table 1 are provided in Tables A5-A9 in the online Appendix.

In Table 5, interaction term results are shown for rule of law (column 1 and 6), control of corruption (column 2 and 7), freedom from business regulations (column 3 and 8), DTF (columns

4 and 9), and profit tax rate (column 5 and 10). All the interaction terms carry the expected sign and they are all significant at the 5 percent level or less. That is, the impact of informal competition on R&D activity is significantly larger or more positive in countries where the business environment is less conducive to operating formally due to weaker rule of law, higher corruption, more burdensome business regulations, and higher profit tax rate. As Table 5 reveals, these results hold with or without controlling for the interaction term between informal competition and GDP per capita. As we discussed above, the interaction term results in Table 5 are consistent with the predictions of the “parasite” view of informality. They are unlikely to hold if our informal competition variable were a mere proxy for other correlated drivers of R&D activity. This helps raise our confidence against endogeneity concerns with our main results.

## *5.2 Exporting activity*

Next, we test how the impact of informal competition varies between exporters and non-exporters. To reiterate, if our results above for the impact of informal competition on R&D activity are truly causal, we expect the relationship between informal competition and R&D activity to be larger or more positive for non-exporters than exporters. We test for this prediction 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. The interaction term is added to the baseline specifications (shown in Table 1).

Regression results for the interaction term for exporter dummy are provided in Table A10 in the online Appendix. The results show that as predicted, the interaction term between informal competition and exporter dummy is large, negative, and significant at the 5 percent level or less. Thus, the positive impact of informal competition on R&D activity of the formal firms is significantly larger (more positive) for non-exporters than exporters. We confirm that this result

holds even when we allow for the impact of informal competition to vary with the level of GDP per capita or firm-size (columns 7, 8 and 10, Table A10 in the online Appendix).<sup>12</sup>

### 5.3 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 affected by a given increase in informal competition. Thus, the relationship between R&D activity and informal competition should be weak (insignificant) or much smaller (less positive) than what we found above for SMEs.

Regression results for the baseline specification (as in Tables 1 and 2) for the sample of large firms (100 or more workers) are provided in Table A11 (log odds ratios) and Table A12 (marginal effects) in the online Appendix. As predicted above, there is no significant relationship between informal competition and R&D activity for large firms. The estimated marginal effect of informal competition on R&D activity is negative and small ranging between -0.091 and -0.036. It is statistically insignificant at the 10 percent level or less. This holds for all the baseline specifications. Thus, our prediction that informal firms pose no significant problem for large firms is confirmed.

## 6. Conclusion

The informal sector is large and provides jobs to millions of individuals across the globe. However, the impact it has on the functioning of the formal sector firms is not properly understood. The

---

<sup>12</sup> Since we control for exporter vs. non-exporter dummy, we experimented by dropping the control for exports as a proportion of firms' annual sales (see columns 9 and 10 in Table A10 in the online Appendix). However, this did not change the qualitative nature of the results discussed in this section.

present paper attempted to fill this gap in the literature by analyzing the impact of informal competition on the probability of formal firms spending on R&D activity. One possibility is that competition from informal firms forces formal firms to improve their productivity and differentiate their product from that of the informal firms. The “unfair” advantage enjoyed by informal firms because they do not pay taxes and do not comply with costly regulations that formal firms must shoulder implies that informal competition can be quite intense and formal firms must innovate to escape this competition or perish. Thus, a positive impact of informal competition on likelihood of R&D activity by formal firms is expected. However, another possibility is that the “unfair” advantage enjoyed by the informal sector firms allows them to “steal” market share from the formal firms, lowering production, profitability and therefore the incentive to innovate. Thus, competition from informal firms can hurt R&D activity among formal sector firms. Our empirical results show that the former case holds. That is, higher informal competition leads to higher probability of formal firms doing R&D activity. Further, consistent with some of the predictions of the theoretical models of informality, we find that the positive relationship between higher informal competition and R&D activity by formal firms is larger (more positive) in countries where the business environment is less conducive to operating formally versus informally. That is, in countries where formal firms face higher business tax rates, higher corruption, more burdensome business regulations and weaker rule of law. As discussed above, these findings help raise our confidence against endogeneity concerns with our main result of a positive impact of informal competition on R&D activity by the formal firms.

Several issues remain to be explored. We provide a few examples to illustrate the point. First, our results are restricted to R&D activity by the formal firms. It will be interesting to analyze how informal competition affects other aspects of firm performance such as sales growth rate,

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 covers 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 be a 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 faced by the formal firms. Our motivation for exploring this was to raise our confidence against endogeneity concerns with our main results 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

Acemoglu, D., 2009. In: Introduction to Modern Economic Growth. Princeton University Press, Princeton.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P. (2005). Competition and innovation: An inverted U-relationship. *Quarterly Journal of Economics* 120, 701-728.

Aghion, Philippe, Philippe Askenazy, Nicolas Berman, Gilbert Cette, and Laurent Eymard. 2012. "Credit Constraints and the Cyclicalities of R&D Investment: Evidence from France." *Journal of the European Economic Association* 10, no. 5: 1001–1024.

Aghion, P.; Blundell, R.; Griffith, R.; Howitt, P. and Prantl, S., 2009, 'The Effects of Entry on Incumbent Innovation and Productivity,' *Review of Economics and Statistics*, 91, pp. 20–32.

Alam, Ashraful Moshfique Uddin, and Hassan Yazdifar (2019). "Institutional determinants of R&D investment: Evidence from emerging markets." *Technological Forecasting & Social Change* 138: 34-44.

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>

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.

Amiti, M. and Khandelwal, A., 2013, 'Import Competition and Quality Upgrading,' *Review of Economics and Statistics*, 95, pp. 476–490.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In Nelson, R.R. (Ed.). *The rate and direction of inventive activity*. Princeton University Press. Princeton, NJ.

Audretsch DB (1995) Innovation, Growth and Survival. *Int J Ind Organ* 13(4):441–457. [https://doi.org/10.1016/0167-7187\(95\)00499-8](https://doi.org/10.1016/0167-7187(95)00499-8)

Barge-Gil, Andrés and López, Alberto, (2014), R&D determinants: Accounting for the differences between research and development, *Research Policy*, 43, issue 9, p. 1634-1648.

Baron J.N, Hannan M.T, Burton M.D (2002) Labor Pains: Change in Organizational Models and Employee Turnover in Young, High-Tech Firms. *Am J Sociol* 106(4):960–1012. <https://doi.org/10.1086/320296>

Becker, Bettina (2013). “[The Determinants of R&D Investment: A Survey of the Empirical Research](#),” [Discussion Paper Series 2013\\_09](#), Department of Economics, Loughborough University.

Becker, Bettina. 2015. “Public R&D Policies and Private R&D Investment: A Survey of the Empirical Evidence.” *Journal of Economic Surveys* 29(5): 917–42.

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.

Bloom, Nicholas, Heidi Williams and John Van Reenen. 2019. “A Toolkit of Policies to promote Innovation” *Journal of Economic Perspectives* 33(3): 163–184.

Blundell, R., Griffith, R., Van Reenen, J. (1999). Market share, market value and innovation in a panel of British manufacturing firms. *Review of Economic Studies* 66: 529-554.

Broadberry, S. and Crafts, N., 2001, ‘Competition and Innovation in 1950’s Britain,’ *Business History*, 43, pp. 97–118.

Brown, J.R., Fazzari, S.M and Petersen, B.C. (2009) Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *Journal of Finance* 64: 151-185.

Brown, J.R., Martinsson, G. and Petersen, B.C. (2012) Do Financing Constraints Matter for R&D? *European Economic Review* 56: 1512-1529.

Brouwer, E., Kleinknecht, A. (1999). Keynes-plus? Effective demand and changes in firm-level R&D: an empirical note. *Cambridge Journal of Economics* 23, 385-391.

Bustos, Paula (2011). “Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms,” *American Economic Review* 101 (1): 304–340.

Coad A, Pellegrino G, Savona M (2016a) Barriers to Innovation and Firm Productivity. *Econ Innov New Technol* 25(3):321–334. <https://doi.org/10.1080/10438599.2015.1076193>

Coad A, Segarra A, Teruel M (2016b) Innovation and firm growth: Does firm age play a role? *Research Policy* 45(2):387–400

Cohen, W.M. (2010). Fifty years of empirical studies of innovative activity and performance. In Hall, B., Rosenberg, N. (Eds). *Handbook of the Economics of Innovation* vol 1, 129-213. North-Holland.

- Cohen, W. M., Levin, R.C., Mowery, D.C. (1987) Firm size and R&D intensity: a reexamination, *Journal of Industrial Economics* 35, 543-563.
- Cohen, W. M., Levin, R.C. (1989) Empirical studies of innovation and market structure, In *Handbook of Industrial Organization*, ed. R. Schmalansee, R.D. Willing (eds), 1059-1107. Elsevier Publishers.
- Cohen, W. (1995) Empirical studies of innovative activity. In Stoneman, P. (Ed.) *Handbook of the Economics of Innovation and Technological Change*. Oxford: Blackwell.
- Cohen, W.M., Klepper, S. (1996). A reprise of size and R&D. *Economic Journal* 106: 925-951.
- Cohen WM, Levinthal DA (1989) Innovation and Learning: The Two Faces of R&D. *Econ J* 99(397): 569– 596. <https://doi.org/10.2307/2233763>
- Cohen WM, Levinthal DA (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Adm Sci Q* 35(1): 128–152
- 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.
- Crepon, B., Duguet, E., Mairesse, J. (1998). Research, innovation and productivity: An econometric analysis at the firm level. *Economics of Innovation and New Technology* 7: 115-158.
- Criscuolo, C., Haskel, J.E. and Slaughter, M.J. (2010) Global engagement and the innovation activities of firms. *International Journal of Industrial Organization* 28: 191-202.
- Damijan, J.P., Kostevc, C. and Polanec, S. (2010) From innovation to exporting or vice versa? *World Economy* 33: 374-398.
- 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.
- 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.
- 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.
- Girma, S., Görg, H. and Hanley, A. (2008) R&D and exporting: A comparison of British and Irish firms. *Review of World Economics* 144: 750-773.
- Geroski, P. (1990). Innovation, technological opportunity and market structure”, *Oxford Economic Papers* 42: 586-602.
- Geroski, P., J. Van Reenen and C. Walters. 2002. “Innovations, Patents and Cash Flow,” in A. Kleinknecht and P. Mohnen (eds.): *Innovation and Firm Performance. Econometric Explorations of Survey Data*, Palgrave, London, chapter 2, 31-55.
- Gonzalez, A., & Lamanna, F. (2007). Who Fears Competition from Informal Firms? Policy Research Working Paper Series 4316, World Bank, Washington DC.
- Gorodnichenko, Y., Svejnar, J. and Terrell, K. (2010) Globalization and innovation in emerging markets. *American Economic Journal: Macroeconomics* 2: 194-226.
- Griliches, Z., 1998. R&D and Productivity: The Econometric Evidence. The University of Chicago Press, Chicago.
- Hall BH (2008) The Financing of Innovation. In: Shane S (ed) *Handbook of Technology and Innovation Management*. Blackwell Publishers, Ltd Oxford, pp 409–430.
- Hall, Bronwyn and Josh Lerner. 2010. “The Financing of R&D and Innovation” in *Handbook of the Economics of Innovation*. Vol. 1, edited by Bronwyn Hall and Nathan Rosenberg. Elsevier.
- Haltiwanger, John C., Ron S. Jarmin, and Javier Miranda. (2010). Who Creates Jobs? Small vs. Large vs. Young. NBER Working Paper 16300, NBER.
- Hannan MT, Diane Burton M, Baron JN (1996) Inertia and Change in the Early Years: Employment Relations in Young, High Technology Firms. *Ind Corp Chang* 5(2):503–536. <https://doi.org/10.1093/icc/5.2.503>
- Harris, M.N., Rogers, M., Siouclis, A. (2003). Modelling firm innovation using panel data probit estimators. *Applied Economic Letters* 10(11): 683-686.
- Hausman, Naomi. 2018. “University Innovation and Local Economic Growth.” [https://drive.google.com/file/d/1dfYTyG2zzVvfbRBJQ7arOLt7d\\_7M29Fv/view](https://drive.google.com/file/d/1dfYTyG2zzVvfbRBJQ7arOLt7d_7M29Fv/view).

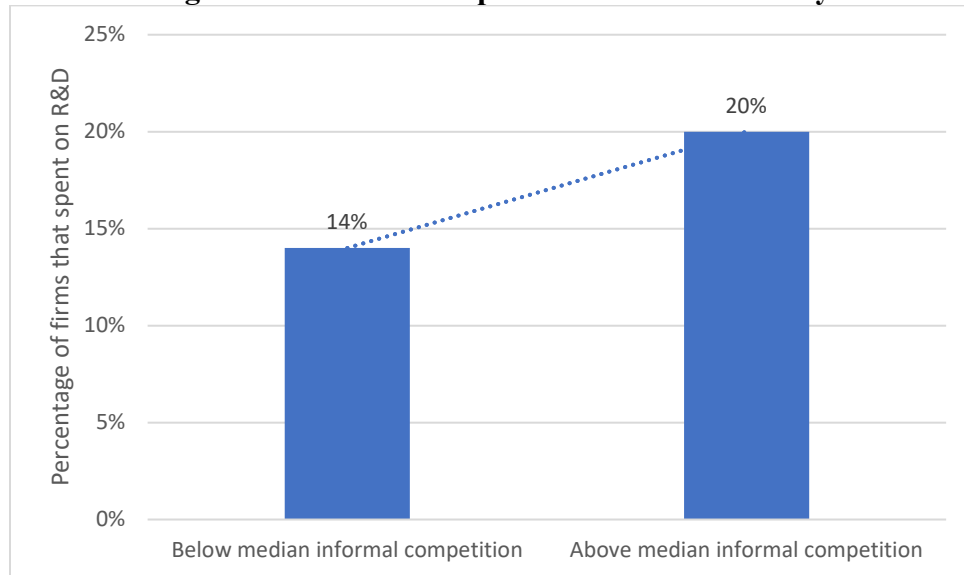
- ILO (2018). *Women and Men in the Informal Economy: A Statistical Picture*. Third edition. Geneva: International Labor Office, International Labor Organization (ILO).
- Islam, A., Palacios Lopez, A., & Amin, M. (2019). Decomposing the Labour Productivity Gap between Migrant-Owned and Native-Owned Firms in Sub-Saharan Africa. *Journal of Development Studies*, 55(9): 2065-2082.
- Klepper, S. 1996. "Entry, Exit and Innovation over the Product Life Cycle," *American Economic Review*, 86(3), 562-583.
- Klevorick, A., Levin, R., Nelson, R., and Winter, S. (1995). On the sources and significance of inter-industry differences in technological opportunities. *Research Policy*, 24: 185–205.
- Koellinger, P., 2008. Why are some entrepreneurs more innovative than others? *Small Business Economics* 31(1): 21–37.
- 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.
- Leiponen A (2005) Skills and Innovation. *Int J Ind Organ* 23(5–6):303–323. <https://doi.org/10.1016/j.ijindorg.2005.03.005>
- Levin, R.C., Cohen, W.M., Mowery, D.C. (1985). R&D, appropriability, opportunity and market structure: New evidence on some Schumpeterian hypotheses. *American Economic Review. Papers and Proceedings* 75(2): 20-24.
- Lileeva, Alla and Daniel Trefler, "Improved access to foreign markets raises plant level productivity... for some plants," *The Quarterly journal of economics*, 2010, 125 (3), 1051–1099.
- 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.
- Mairesse, J. and P. Mohnen (2002), "Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application, *American Economic Review*, **92(2)**, 226-230.
- Maloney, W. F. (2004). Informality Revisited. *World Development*, 32 (7): 1159 – 1178.
- Martin Srholec, 2011. "[A multilevel analysis of innovation in developing countries](#)," *Industrial and Corporate Change*, Oxford University Press, vol. 20(6), pages 1539-1569, December
- Maskus, K.E., Neumann, R. and Seidel, T. (2012) How national and international financial development affect industrial R&D. *European Economic Review* 56: 72-83.

- 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.
- Murphy, K., A. Shleifer and R. Vishny. 1993. Why is rent-seeking so costly to growth? *American Economic Review* 83: 409–417.
- Nelson RR, Winter SG (1982) *An Evolutionary Theory of Economic Change*. Belknap Press of Harvard University Press, Cambridge
- Nilsen, O.A., and F. Schiantarelli (2003), “Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities”, *the Review of Economics and Statistics*, 85, 1021-1037.
- OECD/Eurostat (2018), **Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation**, 4th Edition, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris/Eurostat, Luxembourg.  
<https://doi.org/10.1787/9789264304604-en>
- Pellegrino, Gabriele (2018). Barriers to innovation in young and mature firms. *Journal of Evolutionary Economics* 28: 181–206.
- 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.
- Piva M, Vivarelli M (2009) The Role of Skills as a Major Driver of Corporate R&D. *Int J Manpower* 30(8): 835–852. <https://doi.org/10.1108/01437720911004452>
- Porter, M.E. (1990). *The competitive advantage of nations*. Free Press, New York, NY.
- 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.
- Salomon, R. M. and J. M. Shaver (2005), 'Learning by Exporting: New Insights from Examining Firm Innovation', *Journal of Economics and Management Strategy*, 4, 2, 431–60.
- Scherer, F.M. (1980). *Industrial Market Structure and Economic Performance*. Rand McNally, Chicago.
- Schmookler, J. (1962). Economic sources of inventive activity. *Journal of Economic History* 22: 1-10.
- Schumpeter, J.A. (1942). *Capitalism, Socialism and Democracy*. Harper. New York. NY.
- Srholec, M., 2011. A multilevel analysis of innovation in developing countries. *Ind. Corp. Chang.* 20 (6), 1539–1569.
- Symeonidis, G. (2019), Competition, Innovation and the Use of Innovations. *J Ind Econ*, 67: 565-592. doi:[10.1111/joie.12209](https://doi.org/10.1111/joie.12209)
- Syverson, C. (2011). What determines productivity? *Journal of Economic literature*, 49(2), 326-365.
- Toivanen, Otto, and Lotta Väänänen. 2016. "Education and Invention." *Review of Economics and Statistics* 98(2): 382–96.
- 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.
- Valero, Anna, and John Van Reenen. 2019. "The Economic Impact of Universities: Evidence from Across the Globe." *Economics of Education Review* 68(1): 53–67.
- Xu, Gang and Go Yano. (2017). How does anti-corruption affect corporate innovation? Evidence from recent anti-corruption efforts in China, *Journal of Comparative Economics*, 45(3): 498-519.
- Wagner, J. (2007). Exports and productivity: A survey of the evidence from firm-level data. *World Economy*, 30(1), 60-82.

Yang, C.-H. and Chen, Y.H. (2012) R&D, productivity, and exports: Plant-level evidence from Indonesia. *Economic Modelling* 29: 208-216.

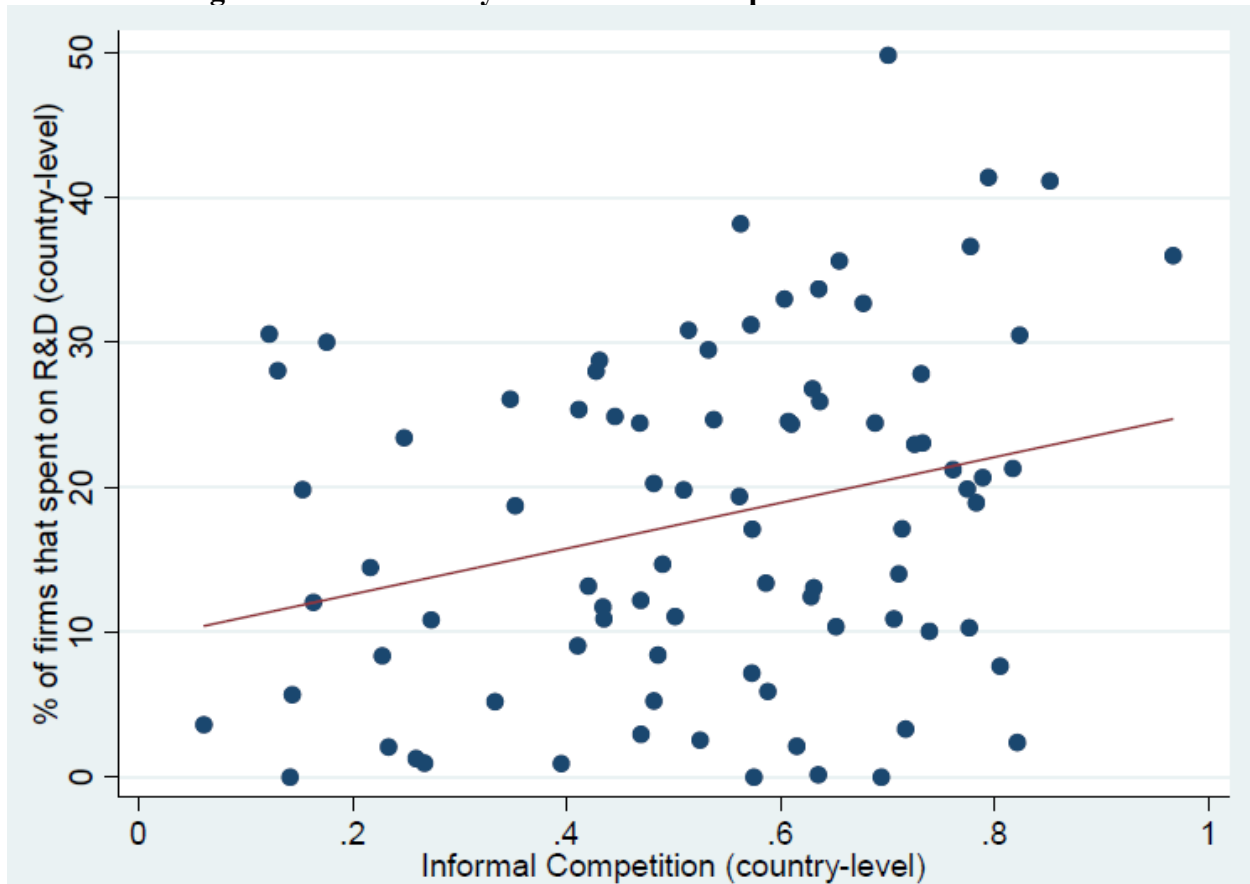
**Figure 1: Informal competition and R&D activity**



Source: Based on author's own calculations using Enterprise Surveys (ES) data.

Note: For the median value of informal competition variable used to define the two groups plotted on the horizontal axis, we first compute the proportion of all firms in the country other than the firm in question who report competing against informal firms (*Informal Competition* variable as defined in section 2). Median value of the variable is used to define the two groups.

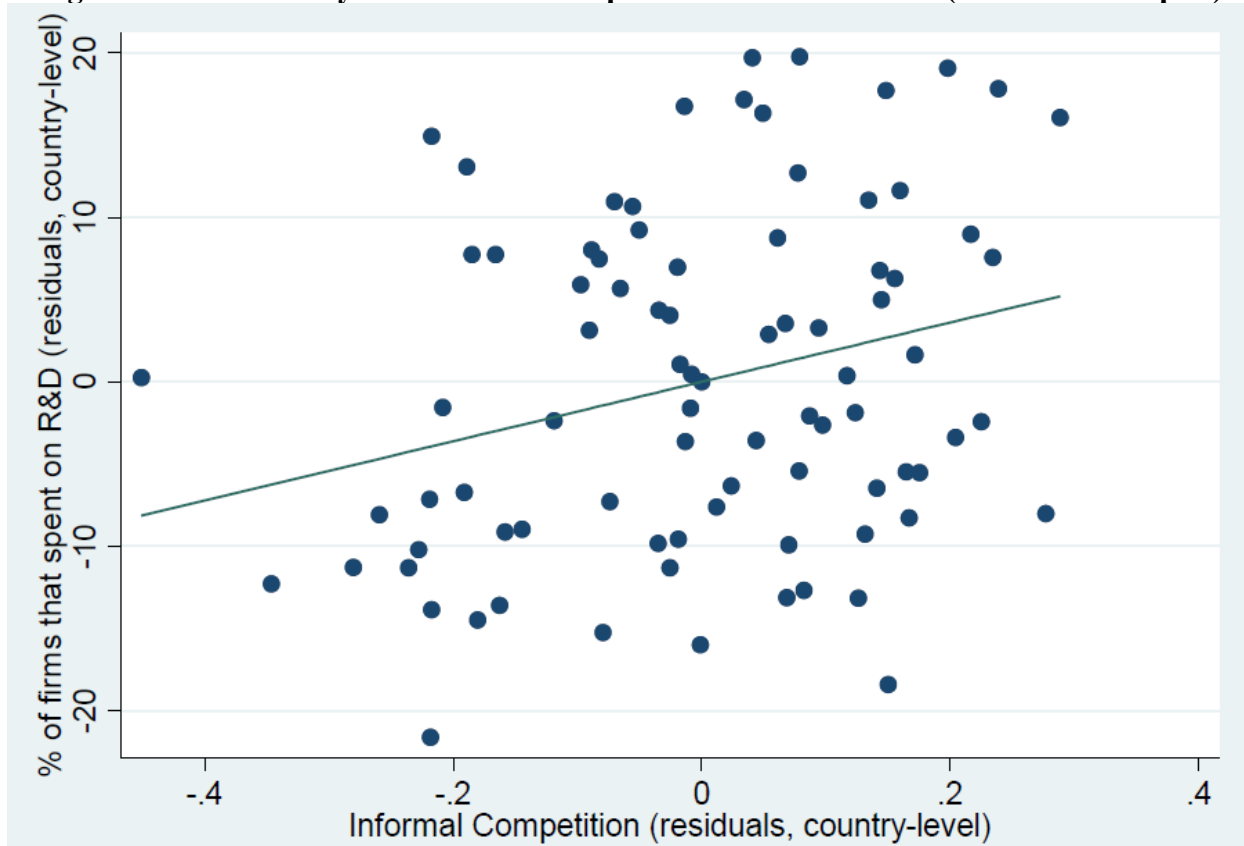
Figure 2: R&D activity and informal competition across countries



Source: Based on author's own calculations using Enterprise Surveys (ES) data.

Note: The horizontal axis plots the *Informal Competition* (country-level cell average) variable as defined in section 2. It equals the proportion of all firms in the country other than the firm in question that report competing against informal sector firms.

**Figure 3: R&D activity and informal competition across countries (Partial scatter plot)**



Source: Based on author's own calculations using Enterprise Surveys (ES) data and World Development Indicators, World Bank.

Note: The figure is a partial scatter plot obtained as follows. First, we compute country-level averages of the following variables: dummy variable equal to 1 if the firm spent on R&D activity during the last year and 0 otherwise (Y-axis variable), Informal Competition (as defined in section 2) which is the proportion of all firms in the country other than the firm in question that report competing against informal sector firms (X-axis variable), log of the number of permanent full time workers employed at the firm 3 fiscal years ago (Firm size), log of GDP per capita (lagged by two years, PPP adjusted and at constant 2011 International Dollars taken from World Development Indicators, World Bank), set of 28 dummy variables indicating the industry (3 digit ISIC Rev. 3.1 level) to which the firm belongs. In the second step, the Y-axis variable was regressed on Firm size, industry dummies and GDP per capita variables using Ordinary Least Squares (OLS) method. The residuals were obtained (Y-axis residuals). Similarly, residuals were obtained for Informal Competition variables (X-axis residuals). The figure is a scatter plot of Y-axis residuals and X-axis residuals. The line of fit shown is significant at the 1 percent level.

**Table 1: Base Regression Results: Log odds ratios from Logit estimation**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)
Informal Competition (Country-level cell average)	1.880*** (0.504)	2.022*** (0.562)	2.076*** (0.564)	2.187*** (0.540)	2.181*** (0.541)	2.192*** (0.570)
GDP per capita (lagged, logs)		0.061 (0.077)	0.080 (0.093)	0.059 (0.094)	0.041 (0.098)	0.030 (0.137)
Firm Size			0.437*** (0.085)	0.391*** (0.094)	0.369*** (0.094)	0.360*** (0.092)
HHI			0.178 (0.555)	0.130 (0.564)	0.032 (0.578)	0.106 (0.571)
Age of Firm (logs)				0.054 (0.093)	0.059 (0.094)	0.051 (0.092)
Manager Experience (logs)				-0.144 (0.103)	-0.144 (0.105)	-0.148 (0.105)
Exports (proportion of sales)				0.764*** (0.260)	0.766*** (0.251)	0.753*** (0.251)
Female Top Manger Y:1 N:0				0.059 (0.160)	0.051 (0.161)	0.074 (0.167)
Multi-firm Y:1 N:0				0.132 (0.161)	0.085 (0.163)	0.090 (0.165)
Power Outages (day hours)					-0.005 (0.014)	-0.008 (0.015)
Finance Obstacle (0-4 scale)					0.040 (0.055)	0.027 (0.055)
Overdraft Y:1 N:0					0.366*** (0.137)	0.378*** (0.141)
Courts Obstacle (cell average)						0.169 (0.143)
Population (lagged, logs)						0.031 (0.060)
Inflation (% , annual)						-0.006 (0.010)
Gross Tertiary Enrollment Rate						0.262 (0.539)
Gross Primary Enrollment Rate						-0.243 (0.439)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.625*** (0.534)	-4.202*** (0.982)	-6.022*** (1.146)	-5.581*** (1.166)	-5.778*** (1.207)	-5.950*** (1.861)
Number of observations	12,980	12,980	12,980	12,980	12,980	12,980
Pseudo R-squared	0.065	0.066	0.076	0.083	0.087	0.089

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



**Table 2: Base Regression Results: Marginal effects from Logit estimation**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)
Informal Competition (Country-level cell average)	0.251*** (0.064)	0.269*** (0.071)	0.272*** (0.070)	0.284*** (0.067)	0.281*** (0.067)	0.281*** (0.071)
GDP per capita (lagged, logs)		0.008 (0.010)	0.010 (0.012)	0.008 (0.012)	0.005 (0.013)	0.004 (0.018)
Firm Size			0.057*** (0.011)	0.051*** (0.012)	0.048*** (0.012)	0.046*** (0.012)
HHI			0.023 (0.073)	0.017 (0.073)	0.004 (0.074)	0.014 (0.073)
Age of Firm (logs)				0.007 (0.012)	0.008 (0.012)	0.007 (0.012)
Manager Experience (logs)				-0.019 (0.014)	-0.019 (0.014)	-0.019 (0.014)
Exports (proportion of sales)				0.099*** (0.034)	0.099*** (0.032)	0.097*** (0.032)
Female Top Manger Y:1 N:0				0.008 (0.021)	0.007 (0.021)	0.010 (0.022)
Multi-firm Y:1 N:0				0.018 (0.022)	0.011 (0.022)	0.012 (0.022)
Power Outages (day hours)					-0.001 (0.002)	-0.001 (0.002)
Finance Obstacle (0-4 scale)					0.005 (0.007)	0.003 (0.007)
Overdraft Y:1 N:0					0.048*** (0.018)	0.049*** (0.019)
Courts Obstacle (cell average)						0.022 (0.018)
Population (lagged, logs)						0.004 (0.008)
Inflation (% , annual)						-0.001 (0.001)
Gross Tertiary Enrollment Rate						0.034 (0.069)
Gross Primary Enrollment Rate						-0.031 (0.056)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	12,980	12,980	12,980	12,980	12,980	12,980

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

**Table 3: Robustness for other controls (Marginal effects)**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)
Informal Competition (Country-level cell average)	0.223*** (0.068)	0.254*** (0.063)	0.269*** (0.060)	0.249*** (0.062)	0.235*** (0.063)
GDP per capita (lagged, logs)	0.018 (0.019)	0.021 (0.018)	0.022 (0.019)	0.027 (0.022)	0.012 (0.025)
Firm Size	0.044*** (0.012)	0.044*** (0.013)	0.034*** (0.013)	0.034*** (0.013)	0.036*** (0.013)
HHI	-0.009 (0.073)	0.021 (0.077)	0.027 (0.076)	0.068 (0.073)	0.059 (0.072)
Age of Firm (logs)	0.006 (0.012)	0.006 (0.013)	0.009 (0.013)	0.015 (0.013)	0.017 (0.013)
Manager Experience (logs)	-0.017 (0.013)	-0.035*** (0.013)	-0.028** (0.013)	-0.037*** (0.013)	-0.035*** (0.013)
Exports (proportion of sales)	0.092*** (0.032)	0.079** (0.031)	0.087*** (0.028)	0.121*** (0.028)	0.114*** (0.029)
Female Top Manger Y:1 N:0	0.005 (0.022)	0.003 (0.024)	-0.011 (0.023)	-0.014 (0.024)	-0.012 (0.022)
Multi-firm Y:1 N:0	0.018 (0.024)	0.002 (0.025)	0.017 (0.025)	0.009 (0.027)	0.006 (0.026)
Power Outages (day hours)	-0.001 (0.002)	-0.002 (0.002)	-0.000 (0.002)	-0.000 (0.002)	-0.000 (0.002)
Finance Obstacle (0-4 scale)	0.002 (0.007)	0.003 (0.008)	0.004 (0.008)	0.003 (0.008)	0.005 (0.008)
Overdraft Y:1 N:0	0.045** (0.018)	0.031* (0.019)	0.028 (0.019)	0.026 (0.019)	0.022 (0.019)
Courts Obstacle (Country-level cell average)	0.018 (0.020)	-0.007 (0.022)	0.010 (0.025)	0.003 (0.025)	0.003 (0.024)
Population (lagged, logs)	0.001 (0.007)	0.004 (0.007)	0.003 (0.007)	-0.008 (0.008)	-0.006 (0.008)
Inflation (% , annual)	-0.001 (0.001)	-0.001 (0.002)	-0.002 (0.002)	-0.003 (0.003)	-0.002 (0.003)
Gross Tertiary Enrollment Rate	0.138* (0.073)	0.175** (0.077)	0.108 (0.082)	0.067 (0.078)	0.092 (0.075)
Gross Primary Enrollment Rate	-0.038 (0.055)	-0.036 (0.058)	-0.061 (0.071)	-0.027 (0.084)	-0.053 (0.086)
Region fixed effects (World regions)	Yes	Yes	Yes	Yes	Yes

Firm Purchased Fixed Assets Y:1 N:0	0.085***	0.080***	0.091***	0.088***
	(0.019)	(0.017)	(0.017)	(0.017)
Foreign Ownership (proportion)	-0.030	-0.021	-0.025	-0.024
	(0.024)	(0.024)	(0.024)	(0.024)
City size fixed effects	Yes	Yes	Yes	Yes
Time Tax (Country-level cell average)		-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)
Gross Secondary Enrollment Rate		0.138*	0.148**	0.169***
		(0.073)	(0.065)	(0.065)
Rule of Law (lagged)			0.141***	0.143***
			(0.045)	(0.044)
Freedom from Regulation (lagged)			-0.034**	-0.040***
			(0.015)	(0.015)
Control of Corruption (lagged)			-0.138***	-0.127***
			(0.042)	(0.039)
Employment Growth (% , annual)				0.001*
				(0.001)
Growth rate of GDP per capita (% , annual)				-0.011*
				(0.006)
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	12,980	11,579	11,117	10,852
			10,852	10,851

Standard errors in brackets. All standard errors are Huber-White robust and clustered on the country. Significance is denoted by \*\*\* (1%), \*\* (5%), \* (10%). Sample size varies due to missing data.

**Table 4: Cell averages at the sub-national level: Log odds ratios**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Informal Competition (Country- industry-firm size level cell average)	0.682** (0.342)	0.766** (0.348)	0.769** (0.342)	0.771** (0.335)				
Informal Competition (City- industry-firm size level cell average)					0.751* (0.391)	0.798** (0.395)	0.791** (0.402)	1.048** (0.420)
HHI		1.051** (0.487)	1.090** (0.493)	1.016** (0.492)		1.080* (0.554)	1.104* (0.570)	0.962 (0.599)
Firm Size		0.456*** (0.160)	0.397** (0.165)	0.402** (0.164)		0.467** (0.213)	0.397* (0.218)	0.371* (0.219)
Age of Firm (logs)			0.092 (0.089)	0.089 (0.088)			-0.026 (0.118)	-0.065 (0.116)
Manager Experience (logs)			-0.126 (0.097)	-0.126 (0.100)			-0.020 (0.119)	-0.015 (0.123)
Exports (proportion of sales)			0.677*** (0.234)	0.656*** (0.240)			1.141*** (0.393)	1.233*** (0.394)
Female Top Manger Y:1 N:0			-0.061 (0.162)	-0.032 (0.161)			-0.004 (0.195)	0.015 (0.194)
Multi-firm Y:1 N:0			0.183 (0.168)	0.137 (0.164)			0.071 (0.219)	-0.027 (0.220)
Courts Obstacle (Country-industry-firm size level cell average)				-0.437** (0.181)				
Power Outages (day hours)				-0.031* (0.018)				-0.019 (0.023)
Finance Obstacle (0-4 scale)				-0.005 (0.048)				-0.012 (0.067)
Overdraft Y:1 N:0				0.291**				0.536***

				(0.146)				(0.192)
Courts Obstacle (City- industry-firm size level cell average)								-0.588***
								(0.184)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes				
City fixed effects					Yes	Yes	Yes	Yes
Firm-size fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-2.586***	-4.191***	-4.084***	-3.545***	-3.112**	-4.881***	-4.541***	-3.834**
	(0.504)	(0.770)	(0.839)	(0.879)	(1.396)	(1.612)	(1.644)	(1.625)
Number of observations	12,522	12,522	12,522	12,522	9,409	9,409	9,409	9,409
Pseudo R-squared	0.144	0.149	0.154	0.160	0.177	0.183	0.191	0.202

Standard errors in brackets. All standard errors are Huber-White robust and clustered on country-industry-firm size (columns 1-4) and on city-industry-firm size (columns 5-8). Significance is denoted by \*\*\* (1%), \*\* (5%), \* (10%)

**Table 5: Interaction term results for business environment variables (Log odds ratios from logit estimation)**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	<b>Without controlling for GDP per capita interacted with Informal Competition</b>					<b>Controlling for GDP per capita interacted with Informal Competition</b>				
Informal Competition (Country-level cell average)	2.020*** (0.538)	1.853*** (0.538)	8.382*** (2.450)	11.114** (4.449)	0.206 (0.954)	-4.758 (4.750)	-3.791 (4.231)	4.003 (3.798)	11.120** (4.838)	3.395 (4.584)
Informal Competition (Country-level cell average)*Rule of Law (lagged)	<b>-0.983**</b> (0.490)					<b>-1.584***</b> (0.606)				
Rule of Law (lagged)	0.814*** (0.262)					1.216*** (0.373)				
Informal Competition (Country-level cell average)*Control of Corruption (lagged)		<b>-1.212**</b> (0.530)					<b>-1.773***</b> (0.613)			
Control of Corruption (lagged)		0.846*** (0.272)					1.187*** (0.344)			
Informal Competition (Country-level cell average)*DTF (lagged)			<b>-0.105***</b> (0.039)					<b>-0.180***</b> (0.051)		
DTF (lagged)			0.038* (0.021)					0.077*** (0.028)		
Informal Competition (Country-level cell average)*Freedom from Regulation (lagged)				<b>-1.276**</b> (0.630)					<b>-1.275*</b> (0.725)	
Freedom from Regulation (lagged)				0.720* (0.415)					0.719 (0.475)	
Informal Competition (Country-level cell average)*Profit Tax Rate (% , lagged)					<b>0.140***</b> (0.050)					<b>0.134**</b> (0.056)
Profit Tax Rate (% , lagged)					-0.084** (0.033)					-0.081** (0.036)

Informal Competition (Country-level cell average)*GDP per capita (lagged, logs)						0.743	0.611	0.986*	-0.001	-0.340
						(0.516)	(0.457)	(0.568)	(0.499)	(0.464)
GDP per capita (lagged, logs)	-0.030	-0.001	0.273	0.014	0.178	-0.487	-0.365	-0.254	0.015	0.365
	(0.142)	(0.146)	(0.192)	(0.143)	(0.138)	(0.370)	(0.325)	(0.387)	(0.354)	(0.293)
Firm Size	0.365***	0.367***	0.381***	0.373***	0.343***	0.363***	0.365***	0.382***	0.373***	0.348***
	(0.092)	(0.092)	(0.091)	(0.094)	(0.099)	(0.093)	(0.093)	(0.092)	(0.094)	(0.099)
HHI	-0.028	0.001	-0.019	0.202	-0.026	0.084	0.101	0.122	0.202	-0.100
	(0.591)	(0.580)	(0.568)	(0.566)	(0.605)	(0.576)	(0.570)	(0.550)	(0.550)	(0.617)
Age of Firm (logs)	0.048	0.050	0.053	0.060	0.062	0.048	0.051	0.051	0.060	0.061
	(0.094)	(0.093)	(0.094)	(0.096)	(0.105)	(0.093)	(0.092)	(0.094)	(0.096)	(0.105)
Manager Experience (logs)	-0.160	-0.157	-0.175*	-0.177	-0.208*	-0.160	-0.155	-0.176*	-0.177	-0.207*
	(0.107)	(0.106)	(0.104)	(0.108)	(0.110)	(0.107)	(0.106)	(0.104)	(0.108)	(0.110)
Exports (proportion of sales)	0.684***	0.677***	0.656***	0.902***	0.783***	0.665***	0.661***	0.627***	0.902***	0.767***
	(0.238)	(0.238)	(0.243)	(0.245)	(0.263)	(0.237)	(0.236)	(0.239)	(0.244)	(0.259)
Female Top Manger Y:1 N:0	0.071	0.073	0.045	0.041	0.009	0.056	0.062	0.026	0.041	0.015
	(0.168)	(0.165)	(0.172)	(0.176)	(0.185)	(0.165)	(0.163)	(0.170)	(0.173)	(0.183)
Multi-firm Y:1 N:0	0.074	0.080	0.074	0.081	0.078	0.088	0.093	0.078	0.081	0.063
	(0.167)	(0.167)	(0.174)	(0.176)	(0.177)	(0.165)	(0.166)	(0.176)	(0.177)	(0.179)
Power Outages (day hours)	-0.002	-0.003	-0.009	-0.004	-0.001	-0.002	-0.003	-0.012	-0.004	-0.002
	(0.016)	(0.016)	(0.015)	(0.015)	(0.019)	(0.016)	(0.016)	(0.015)	(0.015)	(0.019)
Finance Obstacle (0-4 scale)	0.028	0.030	0.036	0.028	0.038	0.034	0.035	0.041	0.028	0.035
	(0.055)	(0.055)	(0.056)	(0.058)	(0.059)	(0.053)	(0.054)	(0.055)	(0.057)	(0.059)
Overdraft Y:1 N:0	0.339**	0.339**	0.382***	0.359**	0.326**	0.342**	0.341**	0.393***	0.359**	0.321**
	(0.138)	(0.140)	(0.142)	(0.146)	(0.148)	(0.139)	(0.140)	(0.144)	(0.148)	(0.149)
Courts Obstacle (Country-level cell average)	0.192	0.160	0.231	0.175	0.264*	0.205	0.163	0.248*	0.174	0.260*
	(0.152)	(0.148)	(0.140)	(0.157)	(0.153)	(0.153)	(0.147)	(0.138)	(0.158)	(0.153)
Population (lagged, logs)	0.041	0.051	0.046	0.029	0.077	0.044	0.056	0.058	0.029	0.070
	(0.058)	(0.061)	(0.058)	(0.067)	(0.091)	(0.057)	(0.060)	(0.059)	(0.066)	(0.091)

Inflation (% , annual)	0.001 (0.008)	-0.003 (0.008)	-0.002 (0.011)	-0.006 (0.021)	-0.001 (0.009)	0.002 (0.008)	-0.004 (0.008)	-0.001 (0.011)	-0.006 (0.021)	-0.001 (0.009)
Gross Tertiary Enrollment Rate	-0.048 (0.483)	0.029 (0.484)	-0.151 (0.549)	0.314 (0.519)	-0.041 (0.591)	0.015 (0.492)	0.092 (0.490)	-0.055 (0.552)	0.314 (0.532)	-0.041 (0.580)
Gross Primary Enrollment Rate	-0.298 (0.450)	-0.231 (0.447)	0.094 (0.485)	0.014 (0.533)	-0.533 (0.450)	-0.359 (0.447)	-0.265 (0.443)	0.144 (0.481)	0.014 (0.548)	-0.493 (0.459)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-5.56*** (1.895)	-5.85*** (1.914)	-10.98*** (2.887)	-11.32*** (3.913)	-6.59*** (2.104)	-1.12 (3.892)	-2.29 (3.341)	-8.51** (3.382)	-11.32*** (4.268)	-8.43** (3.275)
Number of observations	12,980	12,980	12,947	12,683	6,396	12,980	12,980	12,947	12,683	6,396
Pseudo R-squared	0.095	0.094	0.097	0.096	0.102	0.096	0.095	0.099	0.096	0.103

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



## Online Appendix

<b>Table A1: Description of Main Variables</b>	
<b>Variable name</b>	<b>Description of variable</b>
R&D Activity	Dummy variable equal to 1 if the firm spent on R&D activity during the last year and 0 otherwise. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Informal Competition (Country-level cell average)	Proportion of firms in the country where the firm is located excluding the firm in question that report competing against informal firms. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Informal Competition (Country-industry-firm size level cell average)	Proportion of firms in the country-industry-firm size group or cell excluding the firm in question that report competing against informal firms. Country is where the firm is located. Industry is defined at the 2-digit ISIC Rev. 3.1 level (7 industries) and firm-size group is small (11-20 full-time permanent workers) or medium (21-99 full-time permanent workers). The size groups here are based on workers at the firm three fiscal years ago (or two fiscal years ago in some countries) from the date of the Enterprise Survey. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Informal Competition (City-industry-firm size level cell average)	Proportion of firms in the city-industry-firm size group or cell excluding the firm in question that report competing against informal firms. City is the sub-national region or city where the firm is located. Industry is defined at the 2-digit ISIC Rev. 3.1 level (7 industries) and firm-size group is small (11-20 full-time permanent workers) or medium (21-99 full-time permanent workers). The size groups here are based on workers at the firm three fiscal years ago (or two fiscal years ago in some countries) from the date of the Enterprise Survey. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Industry fixed effects	Dummy variables indicating the industry to which the firm belongs. Industry grouping is at the 3-digit ISIC Rev. 3.1 level as provided

	<p>by the ES. In the baseline sample using in Table 1, there are 28 3-digit industries. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Year fixed effects	<p>Dummy variables indicating the year the ES was conducted in the country. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Firm Size (logs)	<p>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. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
GDP per capita (lagged, logs)	<p>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 months covered in the calendar year. Source: World Development Indicators, World Bank.</p>
Age of Firm (logs)	<p>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. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Manager Experience (logs)	<p>Log of the number of years of experience the top manager of the firm has working in the industry. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Exports (proportion of sales)	<p>The proportion of the firms' annual sales made abroad in the last fiscal year. Only direct exports are included. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Female Top Manger Y:1 N:0	<p>Dummy variable equal to 1 if the top manager of the firm is a female and 0 otherwise. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>

Multi-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. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Power Outages (day hours)	Total hours of power outages experienced by the firm in a typical month over the last fiscal year. The figure is divided by 24 to convert to day-hours. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
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. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Overdraft Y:1 N:0	Dummy variable equal to 1 if the firm has overdraft facility and 0 otherwise. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Courts Obstacle (Country-level cell average)	Firms were asked how severe is the (lack of proper) functioning of courts as an obstacle for their current operations. Responses were recorded on 0-4 scale as: no obstacle (0), minor obstacle (1), moderate obstacle (2), major obstacle (3), or a very severe obstacle (4). We take the average value of the reported scores where the average is taken over all other firms (that is, excluding the firm in question) in the same country where the firm is located. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Courts Obstacle (Country-industry-firm size level cell average)	Firms were asked how severe is the (lack of proper) functioning of courts as an obstacle for their current operations. Responses were recorded on 0-4 scale as: no obstacle (0), minor obstacle (1), moderate obstacle (2), major obstacle (3), or a very severe obstacle (4). We take the average value of the reported scores where the average is taken over all other firms (that is, excluding the firm in question) in the same country-industry-firm size group or cell. Country is where the firm is located. Industry is defined at the 2-digit ISIC Rev. 3.1 level (7

	<p>industries) and firm-size group is small (11-20 full-time permanent workers) or medium (21-99 full-time permanent workers).The size groups here are based on workers at the firm three fiscal years ago (or two fiscal years ago in some countries) from the date of the Enterprise Survey.</p> <p>Source: Enterprise Surveys.  <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Courts Obstacle (City-industry-firm size level cell average)	<p>Firms were asked how severe is the (lack of proper) functioning of courts as an obstacle for their current operations. Responses were recorded on 0-4 scale as: no obstacle (0), minor obstacle (1), moderate obstacle (2), major obstacle (3), or a very severe obstacle (4). We take the average value of the reported scores where the average is taken over all other firms (that is, excluding the firm in question) in the same city-industry-firm size group or cell. City is the sub-national region or city where the firm is located. Industry is defined at the 2-digit ISIC Rev. 3.1 level (7 industries) and firm-size group is small (11-20 full-time permanent workers) or medium (21-99 full-time permanent workers).The size groups here are based on workers at the firm three fiscal years ago (or two fiscal years ago in some countries) from the date of the Enterprise Survey.</p> <p>Source: Enterprise Surveys.  <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Population (lagged, logs)	<p>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 months covered in the calendar year.</p> <p>Source: World Development Indicators, World Bank.</p>
Inflation (CPI, annual %)	<p>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</p>

	<p>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 months covered in the calendar year.</p> <p>Source: World Development Indicators, World Bank.</p>
<p>Gross Tertiary Enrollment Rate</p>	<p>Gross enrollment rate in tertiary education. Values of the variable were computed in two steps. In the first step, gross tertiary enrollment rate values for the fiscal year covered by the Enterprise Survey were taken. 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 months 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 tertiary enrollment rate. The final figure is divided by 100 to avoid too small coefficients in the regressions.</p> <p>Source: World Development Indicators, World Bank.</p>
<p>Gross Primary Enrollment Rate</p>	<p>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 were taken. 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 months 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. The final figure is divided by 100 to avoid too small coefficients in the regressions.</p>

	Source: World Development Indicators, World Bank.
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. Source: Enterprise Surveys ( <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a> ) and World Development Indicators, World Bank.
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. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
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. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Foreign Ownership (proportion)	Proportion of the firm that is owned by foreign individuals, companies or organizations. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Time Tax (Country-level cell average)	Average value of the percentage of firms' senior management's time that is spent in dealing with business regulations, where the average is taken over all firms in the country other than the firm in question. Source: Enterprise Surveys. <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a>
Gross Secondary Enrollment Rate	Gross enrollment rate in secondary education. Values of the variable were computed in two steps. In the first step, gross secondary

	<p>enrollment rate values for the fiscal year covered by the Enterprise Survey were taken. 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 months 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 secondary enrollment rate. The final figure is divided by 100 to avoid too small coefficients in the regressions.</p> <p>Source: World Development Indicators, World Bank.</p>
<p>Rule of Law (lagged)</p>	<p>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 months covered in the calendar year.</p> <p>Source: Worldwide Governance Indicators, World Bank.</p>
<p>Freedom From Regulation (lagged)</p>	<p>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 months covered in the calendar year.</p> <p>Source: Economic Freedom of the World, Fraser Institute.</p>

Control of Corruption (lagged)	<p>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 months covered in the calendar year.</p> <p>Source: Worldwide Governance Indicators, World Bank.</p>
Profit Tax Rate	<p>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.</p> <p>The values of profit tax are for the last fiscal year that is covered by the ES. In cases where the fiscal year spanned more than one calendar year, the figure used is the weighted average over the calendar years covered, where the weights used are the proportion of months covered in the calendar year.</p> <p>Source: Doing Business, World Bank.</p>
DTF	<p>A measure of the overall regulatory burden on the private firms. Higher values of the variable imply less burdensome regulations from the firms' point of view. The overall regulatory burden is constructed by taking the simple average over regulatory burden or distance to the frontier in the following five areas: starting a business, obtaining construction permit, paying taxes, trading across borders, and resolving insolvency. The variable is lagged by 2 years from the fiscal year covered by the</p>



	<p>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 months covered in the calendar year.</p> <p>Source: Doing Business, World Bank.</p>
Firm-size fixed effects	<p>Two dummy variables indicating the size group of the firm. Firms are either small or medium. Small firms have between 11 and 20 full-time permanent workers three fiscal years ago (or 2 fiscal years ago in some countries) and the rest are medium firms.</p> <p>Source: Enterprise Surveys.  <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
City fixed effects	<p>A set of dummy variables indicating the city where the firm is located.</p> <p>Source: Enterprise Surveys.  <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>
Growth rate of GDP per capita (lagged)	<p>Annual percentage growth rate of GDP per capita lagged by 2 years from the fiscal year covered by the Enterprise. 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 months covered in the calendar year.</p> <p>Source: World Development Indicators, World Bank.</p>
Employment Growth	<p>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.</p> <p>Source: Enterprise Surveys.  <a href="http://www.enterprisesurveys.org">www.enterprisesurveys.org</a></p>

**Table A2: Summary statistics of the main variables**

<b>Variable</b>	<b>Mean</b>	<b>Standard deviation</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Observations</b>
R&D Activity	0.179	0.383	0	1	12,980
Informal Competition (Country-level cell average)	0.533	0.209	0.002	0.985	12,980
Firm Size	3.255	0.600	2.398	4.595	12,980
HHI	0.112	0.167	0	1	12,980
GDP per capita (lagged, logs)	8.682	1.036	6.493	10.333	12,980
Age of Firm (logs)	2.724	0.659	0	5.242	12,980
Manager Experience (logs)	2.767	0.676	0	4.094	12,980
Exports (proportion of sales)	0.108	0.250	0	1	12,980
Female Top Manger Y:1 N:0	0.137	0.344	0	1	12,980
Multi-firm Y:1 N:0	0.184	0.388	0	1	12,980
Power Outages (day hours)	1.422	4.409	0	30	12,980
Finance Obstacle (0-4 scale)	1.524	1.330	0	4	12,980
Overdraft Y:1 N:0	0.440	0.496	0	1	12,980
Courts Obstacle (Country-level cell average)	0.966	0.639	0	2.987	12,980
Population (lagged, logs)	16.261	1.510	12.526	20.928	12,980
Inflation (% , annual)	6.774	7.546	-1.988	53.229	12,980
Gross Tertiary Enrollment Rate	0.287	0.239	0.005	0.855	12,980
Gross Primary Enrollment Rate	1.004	0.168	0.249	1.462	12,980

**Table A3: Correlation between Informal Competition (Country-level cell average) and other baseline controls**

Variable	Informal Competition (Country-level cell average)
Informal Competition (Country-level cell average)	1
Firm Size	0.011
HHI	0.237
GDP per capita (lagged, logs)	-0.465
Age of Firm (logs)	0.010
Manager Experience (logs)	-0.011
Exports (proportion of sales)	-0.113
Female Top Manger Y:1 N:0	-0.054
Multi-firm Y:1 N:0	0.047
Power Outages (day hours)	0.021
Finance Obstacle (0-4 scale)	0.200
Overdraft Y:1 N:0	0.027
Courts Obstacle (Country-level cell average)	0.479
Population (lagged, logs)	0.117
Inflation (% , annual)	0.067
Gross Tertiary Enrollment Rate	-0.556
Gross Primary Enrollment Rate	0.075

<b>Table A4: Two stage Instrumental Variables estimation (IV probit)</b>						
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>
<b>Panel A: Second stage IV results</b>						
<b>Dependent variable: R&amp;D Activity</b>						
Informal Competition (instrumented values)	1.361*** (0.181)	1.506*** (0.172)	1.526*** (0.167)	1.589*** (0.155)	1.646*** (0.158)	1.800*** (0.156)
GDP per capita (lagged, logs)		0.068** (0.034)	0.075** (0.038)	0.060 (0.038)	0.040 (0.039)	-0.015 (0.051)
Firm Size			0.223*** (0.051)	0.196*** (0.050)	0.185*** (0.050)	0.175*** (0.048)
HHI			0.092 (0.207)	0.041 (0.205)	0.024 (0.208)	0.149 (0.219)
Age of Firm (logs)				-0.034 (0.045)	-0.038 (0.045)	-0.045 (0.043)
Manager Experience (logs)				-0.065 (0.047)	-0.059 (0.046)	-0.059 (0.043)
Exports (proportion of sales)				0.489*** (0.116)	0.473*** (0.115)	0.443*** (0.115)
Female Top Manger Y:1 N:0				0.074 (0.079)	0.069 (0.079)	0.078 (0.075)
Multi-firm Y:1 N:0				0.132 (0.086)	0.103 (0.085)	0.096 (0.083)
Power Outages (day hours)					-0.002 (0.007)	-0.001 (0.007)
Finance Obstacle (0-4 scale)					-0.061** (0.026)	-0.075*** (0.025)
Overdraft Y:1 N:0					0.104 (0.063)	0.084 (0.063)
Courts Obstacle (Country-level cell average)						0.004 (0.075)
Population (lagged, logs)						0.022 (0.021)
Inflation (% , annual)						-0.010** (0.004)
Gross Tertiary Enrollment Rate						0.427* (0.226)
Gross Primary Enrollment Rate						-0.282 (0.186)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.910* (0.548)	-1.366** (0.564)	-2.225*** (0.590)	-1.723*** (0.609)	-1.536** (0.635)	-0.815 (0.854)
<b>Panel B: First stage IV results</b>						
<b>Dependent variable: Firm competes against informal firms Y:1 N:0</b>						
Informal Competition	0.648***	0.599***	0.598***	0.591***	0.545***	0.437***

	(0.065)	(0.073)	(0.073)	(0.073)	(0.073)	(0.080)
Controls above	Yes	Yes	Yes	Yes	Yes	Yes
Number of	12,575	12,575	12,575	12,575	12,575	12,575

---

Standard errors in brackets. Significance is denoted by \*\*\* (1%), \*\* (5%), \* (10%)

**Table A5: Impact of Informal Competition and Rule of Law (Log odds ratios from logit estimation)**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)	(12)
Informal Competition (Country-level cell average)*Rule of Law (lagged)	-1.260*** (0.481)	-1.257*** (0.478)	-1.231*** (0.468)	-1.111** (0.461)	-1.135** (0.477)	-0.983** (0.490)	-1.584*** (0.606)
Informal Competition (Country-level cell average)	2.033*** (0.460)	1.949*** (0.518)	2.001*** (0.521)	2.124*** (0.507)	2.120*** (0.513)	2.020*** (0.538)	-4.758 (4.750)
Rule of Law (lagged)	0.888*** (0.249)	0.923*** (0.247)	0.914*** (0.245)	0.837*** (0.239)	0.803*** (0.254)	0.814*** (0.262)	1.216*** (0.373)
Informal Competition (Country- level cell average)*GDP per capita (lagged, logs)							0.743 (0.516)
GDP per capita (lagged, logs)		-0.051 (0.094)	-0.047 (0.112)	-0.052 (0.112)	-0.046 (0.114)	-0.030 (0.142)	-0.487 (0.370)
Firm Size			0.435*** (0.085)	0.398*** (0.093)	0.379*** (0.093)	0.365*** (0.092)	0.363*** (0.093)
HHI			0.030 (0.576)	-0.006 (0.580)	-0.089 (0.590)	-0.028 (0.591)	0.084 (0.576)
Age of Firm (logs)				0.059 (0.096)	0.062 (0.095)	0.048 (0.094)	0.048 (0.093)
Manager Experience (logs)				-0.161 (0.105)	-0.161 (0.106)	-0.160 (0.107)	-0.160 (0.107)
Exports (proportion of sales)				0.654*** (0.243)	0.669*** (0.236)	0.684*** (0.238)	0.665*** (0.237)
Female Top Manger Y:1 N:0				0.037 (0.160)	0.036 (0.160)	0.071 (0.168)	0.056 (0.165)
Multi-firm Y:1 N:0				0.103 (0.165)	0.060 (0.166)	0.074 (0.167)	0.088 (0.165)
Power Outages (day hours)					0.001 (0.015)	-0.002 (0.016)	-0.002 (0.016)
Finance Obstacle (0-4 scale)					0.042 (0.055)	0.028 (0.055)	0.034 (0.053)
Overdraft Y:1 N:0					0.327** (0.137)	0.339** (0.138)	0.342** (0.139)
Courts Obstacle (Country-level cell average)						0.192 (0.152)	0.205 (0.153)
Population (lagged, logs)						0.041 (0.058)	0.044 (0.057)
Inflation (% , annual)						0.001 (0.008)	0.002 (0.008)

Gross Tertiary Enrollment Rate						-0.048 (0.483)	0.015 (0.492)
Gross Primary Enrollment Rate						-0.298 (0.450)	-0.359 (0.447)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.846*** (0.501)	-3.386*** (1.127)	-5.083*** (1.306)	-4.726*** (1.314)	-5.102*** (1.335)	-5.558*** (1.895)	-1.123 (3.892)
Number of observations	12,980	12,980	12,980	12,980	12,980	12,980	12,980

---

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

**Table A6: Impact of Informal Competition and Control of Corruption (Log odds ratios from logit estimation)**

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
R&D Activity							
Informal Competition (Country-level cell average)*Control of Corruption (lagged)	-1.454*** (0.486)	-1.452*** (0.486)	-1.425*** (0.475)	-1.306*** (0.479)	-1.304*** (0.500)	-1.212** (0.530)	-1.773*** (0.613)
Informal Competition (Country-level cell average)	1.728*** (0.459)	1.699*** (0.519)	1.757*** (0.519)	1.895*** (0.506)	1.905*** (0.508)	1.853*** (0.538)	-3.791 (4.231)
Control of Corruption (lagged)	0.962*** (0.248)	0.971*** (0.243)	0.962*** (0.240)	0.874*** (0.237)	0.826*** (0.254)	0.846*** (0.272)	1.187*** (0.344)
Informal Competition (Country-level cell average)*GDP per capita (lagged, logs)							0.611 (0.457)
GDP per capita (lagged, logs)		-0.016 (0.093)	-0.011 (0.113)	-0.015 (0.112)	-0.010 (0.114)	-0.001 (0.146)	-0.365 (0.325)
Firm Size			0.434*** (0.085)	0.397*** (0.093)	0.377*** (0.093)	0.367*** (0.092)	0.365*** (0.093)
HHI			0.031 (0.571)	0.000 (0.577)	-0.076 (0.585)	0.001 (0.580)	0.101 (0.570)
Age of Firm (logs)				0.059 (0.095)	0.063 (0.095)	0.050 (0.093)	0.051 (0.092)
Manager Experience (logs)				-0.156 (0.105)	-0.155 (0.106)	-0.157 (0.106)	-0.155 (0.106)
Exports (proportion of sales)				0.651*** (0.244)	0.671*** (0.236)	0.677*** (0.238)	0.661*** (0.236)
Female Top Manger Y:1 N:0				0.042 (0.159)	0.042 (0.160)	0.073 (0.165)	0.062 (0.163)
Multi-firm Y:1 N:0				0.106 (0.165)	0.061 (0.165)	0.080 (0.167)	0.093 (0.166)
Power Outages (day hours)					0.000 (0.015)	-0.003 (0.016)	-0.003 (0.016)
Finance Obstacle (0-4 scale)					0.043 (0.055)	0.030 (0.055)	0.035 (0.054)
Overdraft Y:1 N:0					0.333** (0.138)	0.339** (0.140)	0.341** (0.140)
Courts Obstacle (Country-level cell average)						0.160 (0.148)	0.163 (0.147)
Population (lagged, logs)						0.051	0.056



						(0.061)	(0.060)
Inflation (% , annual)						-0.003	-0.004
						(0.008)	(0.008)
Gross Tertiary Enrollment Rate						0.029	0.092
						(0.484)	(0.490)
Gross Primary Enrollment Rate						-0.231	-0.265
						(0.447)	(0.443)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.642***	-3.495***	-5.195***	-4.884***	-5.275***	-5.845***	-2.290
	(0.485)	(1.106)	(1.301)	(1.314)	(1.339)	(1.914)	(3.341)
Number of observations	12,980	12,980	12,980	12,980	12,980	12,980	12,980

---

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

**Table A7: Impact of Informal Competition and DTF (Log odds ratios from logit estimation)**

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
R&D Activity							
Informal Competition (Country-level cell average)*DTF (lagged)	-0.094** (0.038)	-0.108*** (0.039)	-0.110*** (0.038)	-0.099*** (0.036)	-0.095*** (0.036)	-0.105*** (0.039)	-0.180*** (0.051)
DTF (lagged)	0.048** (0.021)	0.044** (0.021)	0.044** (0.021)	0.039** (0.020)	0.036* (0.020)	0.038* (0.021)	0.077*** (0.028)
Informal Competition (Country-level cell average)	7.482*** (2.238)	8.634*** (2.445)	8.817*** (2.379)	8.254*** (2.292)	8.008*** (2.318)	8.382*** (2.450)	4.003 (3.798)
Informal Competition (Country-level cell average)*GDP per capita (lagged, logs)							0.986* (0.568)
GDP per capita (lagged, logs)		0.196* (0.111)	0.217* (0.126)	0.194 (0.127)	0.181 (0.131)	0.273 (0.192)	-0.254 (0.387)
Firm Size			0.457*** (0.087)	0.420*** (0.095)	0.399*** (0.095)	0.381*** (0.091)	0.382*** (0.092)
HHI			0.070 (0.564)	0.041 (0.567)	-0.054 (0.579)	-0.019 (0.568)	0.122 (0.550)
Age of Firm (logs)				0.054 (0.097)	0.060 (0.097)	0.053 (0.094)	0.051 (0.094)
Manager Experience (logs)				-0.170* (0.102)	-0.171* (0.104)	-0.175* (0.104)	-0.176* (0.104)
Exports (proportion of sales)				0.659*** (0.240)	0.668*** (0.235)	0.656*** (0.243)	0.627*** (0.239)
Female Top Manger Y:1 N:0				0.022 (0.165)	0.019 (0.165)	0.045 (0.172)	0.026 (0.170)
Multi-firm Y:1 N:0				0.096 (0.175)	0.052 (0.177)	0.074 (0.174)	0.078 (0.176)
Power Outages (day hours)					-0.005 (0.014)	-0.009 (0.015)	-0.012 (0.015)
Finance Obstacle (0-4 scale)					0.048 (0.054)	0.036 (0.056)	0.041 (0.055)
Overdraft Y:1 N:0					0.372*** (0.136)	0.382*** (0.142)	0.393*** (0.144)
Courts Obstacle (Country-level cell average)						0.231 (0.140)	0.248* (0.138)
Population (lagged, logs)						0.046	0.058

Inflation (% , annual)						(0.058)	(0.059)
						-0.002	-0.001
						(0.011)	(0.011)
Gross Tertiary Enrollment Rate						-0.151	-0.055
						(0.549)	(0.552)
Gross Primary Enrollment Rate						0.094	0.144
						(0.485)	(0.481)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-6.60***	-8.202***	-10.112***	-9.247***	-9.363***	-10.975***	-8.512**
	(1.355)	(1.794)	(1.904)	(1.900)	(1.975)	(2.887)	(3.382)
Number of observations	12,947	12,947	12,947	12,947	12,947	12,947	12,947

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

**Table A8: Impact of Informal Competition and Freedom from Regulation (Log odds ratios from logit estimation)**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Informal Competition (Country-level cell average)*Freedom from Regulation (lagged)	-1.448** (0.641)	-1.424** (0.638)	-1.410** (0.631)	-1.251** (0.585)	-1.204** (0.601)	-1.276** (0.630)	-1.275* (0.725)
Freedom from Regulation (lagged)	0.845** (0.404)	0.823** (0.400)	0.807** (0.402)	0.681* (0.375)	0.641* (0.384)	0.720* (0.415)	0.719 (0.475)
Informal Competition (Country-level cell average)	12.103*** (4.509)	12.053*** (4.528)	12.012*** (4.465)	11.003*** (4.158)	10.674** (4.266)	11.114** (4.449)	11.120** (4.838)
Informal Competition (Country-level cell average)*GDP per capita (lagged, logs)							-0.001 (0.499)
GDP per capita (lagged, logs)		0.049 (0.081)	0.084 (0.098)	0.062 (0.099)	0.053 (0.102)	0.014 (0.143)	0.015 (0.354)
Firm Size			0.456*** (0.086)	0.404*** (0.095)	0.388*** (0.095)	0.373*** (0.094)	0.373*** (0.094)
HHI			0.317 (0.546)	0.276 (0.556)	0.210 (0.563)	0.202 (0.566)	0.202 (0.550)
Age of Firm (logs)				0.060 (0.096)	0.064 (0.096)	0.060 (0.096)	0.060 (0.096)
Manager Experience (logs)				-0.169 (0.106)	-0.172 (0.108)	-0.177 (0.108)	-0.177 (0.108)
Exports (proportion of sales)				0.919*** (0.243)	0.924*** (0.239)	0.902*** (0.245)	0.902*** (0.244)
Female Top Manger Y:1 N:0				0.033 (0.172)	0.028 (0.173)	0.041 (0.176)	0.041 (0.173)
Multi-firm Y:1 N:0				0.123 (0.175)	0.070 (0.177)	0.081 (0.176)	0.081 (0.177)
Power Outages (day hours)					-0.003 (0.015)	-0.004 (0.015)	-0.004 (0.015)
Finance Obstacle (0-4 scale)					0.037 (0.059)	0.028 (0.058)	0.028 (0.057)
Overdraft Y:1 N:0					0.352** (0.145)	0.359** (0.146)	0.359** (0.148)
Courts Obstacle (Country-level cell average)						0.175 (0.157)	0.174 (0.158)

Population (lagged, logs)						0.029	0.029
						(0.067)	(0.066)
Inflation (% , annual)						-0.006	-0.006
						(0.021)	(0.021)
Gross Tertiary Enrollment Rate						0.314	0.314
						(0.519)	(0.532)
Gross Primary Enrollment Rate						0.014	0.014
						(0.533)	(0.548)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-9.817***	-10.131***	-12.000***	-10.598***	-10.557***	-11.319***	-11.322***
	(2.904)	(3.103)	(3.121)	(2.973)	(3.018)	(3.913)	(4.268)
Number of observations	12,683	12,683	12,683	12,683	12,683	12,683	12,683

---

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

**Table A9: Impact of Informal Competition and Profit Tax Rate (Log odds ratios from logit estimation)**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)	(12)
Informal Competition (Country-level cell average)* Profit Tax Rate (%)	0.124** (0.053)	0.131** (0.056)	0.127** (0.055)	0.128** (0.053)	0.126** (0.053)	0.145*** (0.053)	0.143** (0.056)
Profit Tax Rate (%)	-0.079** (0.034)	-0.080** (0.035)	-0.081** (0.034)	-0.079** (0.034)	-0.077** (0.034)	-0.089*** (0.034)	-0.089** (0.036)
Informal Competition (Country-level cell average)	0.155 (0.872)	0.251 (0.903)	0.345 (0.893)	0.450 (0.851)	0.459 (0.868)	0.016 (0.917)	3.172 (3.945)
Informal Competition (Country-level cell average)*GDP per capita (lagged, logs)							-0.346 (0.420)
GDP per capita (lagged, logs)		0.100 (0.086)	0.104 (0.093)	0.091 (0.096)	0.071 (0.099)	0.173 (0.137)	0.367 (0.269)
Firm Size			0.434*** (0.091)	0.383*** (0.099)	0.364*** (0.099)	0.344*** (0.096)	0.347*** (0.096)
HHI			0.099 (0.574)	0.047 (0.582)	-0.024 (0.588)	0.045 (0.583)	-0.024 (0.596)
Age of Firm (logs)				0.066 (0.106)	0.072 (0.105)	0.058 (0.103)	0.058 (0.103)
Manager Experience (logs)				-0.191* (0.107)	-0.193* (0.108)	-0.199* (0.110)	-0.201* (0.110)
Exports (proportion of sales)				0.790*** (0.279)	0.804*** (0.269)	0.824*** (0.261)	0.809*** (0.258)
Female Top Manger Y:1 N:0				0.002 (0.169)	-0.002 (0.170)	0.049 (0.180)	0.055 (0.179)
Multi-firm Y:1 N:0				0.123 (0.172)	0.083 (0.174)	0.090 (0.173)	0.076 (0.175)
Power Outages (day hours)					-0.006 (0.017)	-0.010 (0.019)	-0.009 (0.018)
Finance Obstacle (0-4 scale)					0.056 (0.057)	0.040 (0.057)	0.037 (0.057)
Overdraft Y:1 N:0					0.354** (0.141)	0.365** (0.145)	0.357** (0.147)
Courts Obstacle (Country-level cell average)						0.274* (0.144)	0.264* (0.144)
Population (lagged, logs)						0.075 (0.059)	0.076 (0.060)
Inflation (% , annual)						0.000 (0.009)	0.001 (0.009)

Gross Tertiary Enrollment Rate						-0.128 (0.558)	-0.160 (0.549)
Gross Primary Enrollment Rate						-0.595 (0.437)	-0.538 (0.453)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-2.819*** (0.619)	-3.733*** (0.996)	-5.401*** (1.105)	-4.915*** (1.122)	-5.122*** (1.170)	-6.290*** (1.724)	-8.281*** (3.015)
Number of observations	11,043	11,043	11,043	11,043	11,043	11,043	11,043

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

**Table A10: Exporters vs. non-exporters (Log odd ratios from logit estimation)**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Informal Competition (Country-level cell average)* Firm Exports Y:1 N:0	-1.999*** (0.764)	-2.013*** (0.765)	-2.100*** (0.774)	-2.133*** (0.776)	-2.003*** (0.759)	-2.031*** (0.769)	-1.932** (0.788)	-1.982** (0.798)	-2.029*** (0.770)	-1.981** (0.799)
Firm Exports Y:1 N:0	1.818*** (0.426)	1.833*** (0.431)	1.799*** (0.433)	1.834*** (0.445)	1.747*** (0.435)	1.770*** (0.446)	1.718*** (0.447)	1.739*** (0.449)	1.760*** (0.436)	1.731*** (0.442)
Informal Competition (Country-level cell average)	2.872*** (0.511)	2.828*** (0.559)	2.879*** (0.563)	2.919*** (0.564)	2.886*** (0.568)	2.823*** (0.595)	4.198 (3.454)	2.817 (3.688)	2.824*** (0.596)	2.806 (3.675)
Informal Competition (Country-level cell average)* GDP per capita (lagged, logs)							-0.157 (0.385)	-0.156 (0.383)		-0.155 (0.382)
Informal Competition (Country-level cell average)* Firm Size								0.419 (0.462)		0.421 (0.462)
GDP per capita (lagged, logs)		-0.022 (0.079)	-0.007 (0.100)	0.001 (0.100)	-0.012 (0.104)	0.022 (0.140)	0.113 (0.280)	0.111 (0.278)	0.022 (0.140)	0.110 (0.278)
Firm Size			0.375*** (0.093)	0.382*** (0.097)	0.362*** (0.097)	0.352*** (0.095)	0.352*** (0.095)	0.118 (0.268)	0.351*** (0.095)	0.117 (0.267)
HHI			0.063 (0.603)	0.001 (0.601)	-0.088 (0.612)	-0.019 (0.603)	-0.051 (0.601)	-0.043 (0.600)	-0.017 (0.599)	-0.042 (0.596)
Age of Firm (logs)				0.028 (0.094)	0.032 (0.094)	0.021 (0.091)	0.022 (0.092)	0.020 (0.092)	0.022 (0.090)	0.020 (0.090)
Manager Experience (logs)				-0.154 (0.103)	-0.153 (0.105)	-0.156 (0.105)	-0.157 (0.106)	-0.162 (0.107)	-0.156 (0.105)	-0.162 (0.107)
Exports (proportion of sales)				-0.046 (0.298)	-0.020 (0.295)	-0.026 (0.294)	-0.031 (0.294)	-0.019 (0.295)		
Female Top Manger Y:1 N:0				0.079 (0.160)	0.067 (0.161)	0.093 (0.166)	0.096 (0.164)	0.103 (0.165)	0.093 (0.165)	0.103 (0.165)
Multi-firm Y:1 N:0				0.102 (0.163)	0.056 (0.165)	0.070 (0.166)	0.066 (0.167)	0.053 (0.166)	0.069 (0.166)	0.053 (0.165)
Power Outages (day hours)					-0.002	-0.006	-0.006	-0.006	-0.006	-0.006



						(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Finance Obstacle (0-4 scale)						0.033	0.021	0.021	0.022	0.021	0.022
						(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Overdraft Y:1 N:0						0.333**	0.338**	0.336**	0.330**	0.338**	0.331**
						(0.134)	(0.136)	(0.137)	(0.137)	(0.136)	(0.136)
Courts Obstacle (Country-level cell average)							0.149	0.145	0.136	0.149	0.136
							(0.144)	(0.144)	(0.143)	(0.144)	(0.143)
Population (lagged, logs)							0.045	0.045	0.047	0.045	0.047
							(0.060)	(0.061)	(0.061)	(0.060)	(0.061)
Inflation (% , annual)							-0.006	-0.005	-0.005	-0.006	-0.005
							(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Gross Tertiary Enrollment Rate							-0.031	-0.055	-0.055	-0.031	-0.054
							(0.531)	(0.533)	(0.530)	(0.531)	(0.530)
Gross Primary Enrollment Rate							-0.118	-0.104	-0.109	-0.119	-0.110
							(0.439)	(0.446)	(0.449)	(0.437)	(0.446)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-4.521***	-4.317***	-5.828***	-5.510***	-5.715***	-6.475***	-7.359**	-6.597**	-6.481***	-6.594**	
	(0.520)	(0.963)	(1.191)	(1.198)	(1.249)	(1.895)	(3.053)	(3.130)	(1.898)	(3.129)	
Number of observations	12,980	12,980	12,980	12,980	12,980	12,980	12,980	12,980	12,980	12,980	
Pseudo R-squared	0.085	0.085	0.093	0.094	0.098	0.100	0.100	0.100	0.100	0.100	

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

**Table A11: Base Regression Results for Large Firms: Log odds ratios from logit estimation**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)
Informal Competition (Country-level cell average)	-0.166 (0.542)	-0.197 (0.550)	-0.265 (0.554)	-0.248 (0.545)	-0.274 (0.555)	-0.424 (0.583)
GDP per capita (lagged, logs)		-0.021 (0.122)	0.013 (0.116)	-0.014 (0.115)	-0.013 (0.114)	-0.118 (0.193)
Firm Size			0.353*** (0.105)	0.298*** (0.108)	0.297*** (0.109)	0.298*** (0.109)
HHI			0.315 (0.601)	0.358 (0.647)	0.282 (0.612)	-0.100 (0.595)
Age of Firm (logs)				0.199* (0.118)	0.187 (0.118)	0.177 (0.117)
Manager Experience (logs)				0.104 (0.121)	0.098 (0.125)	0.138 (0.124)
Exports (proportion of sales)				0.083 (0.219)	0.088 (0.217)	-0.016 (0.215)
Female Top Manger Y:1 N:0				-0.083 (0.292)	-0.092 (0.297)	-0.092 (0.273)
Multi-firm Y:1 N:0				0.552*** (0.170)	0.543*** (0.172)	0.574*** (0.170)
Power Outages (day hours)					0.021 (0.020)	0.018 (0.021)
Finance Obstacle (0-4 scale)					-0.037 (0.060)	-0.039 (0.060)
Overdraft Y:1 N:0					0.289* (0.161)	0.282* (0.161)
Courts Obstacle (Country-level cell average)						0.416** (0.167)
Population (lagged, logs)						-0.133** (0.065)
Inflation (% , annual)						0.007 (0.007)
Gross Tertiary Enrollment Rate						0.686 (0.856)
Gross Primary Enrollment Rate						0.519 (0.594)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.226 (0.620)	0.439 (1.428)	-1.646 (1.404)	-2.297 (1.423)	-2.361* (1.407)	-0.679 (2.038)
Number of observations	5,721	5,721	5,721	5,721	5,721	5,721
Pseudo R-squared	0.065	0.065	0.076	0.090	0.094	0.104

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

**Table A12: Base Regression Results for Large Firms: Marginal effects from logit estimation**

Dependent variable: R&D Activity	(1)	(2)	(3)	(4)	(5)	(6)
Informal Competition (Country-level cell average)	-0.036 (0.118)	-0.043 (0.120)	-0.057 (0.121)	-0.054 (0.118)	-0.059 (0.120)	-0.091 (0.125)
GDP per capita (lagged, logs)		-0.004 (0.027)	0.003 (0.025)	-0.003 (0.025)	-0.003 (0.025)	-0.025 (0.041)
Firm Size			0.077*** (0.022)	0.064*** (0.023)	0.064*** (0.023)	0.064*** (0.023)
HHI			0.068 (0.131)	0.077 (0.141)	0.061 (0.132)	-0.021 (0.127)
Age of Firm (logs)				0.043* (0.026)	0.040 (0.026)	0.038 (0.025)
Manager Experience (logs)				0.022 (0.026)	0.021 (0.027)	0.030 (0.026)
Exports (proportion of sales)				0.018 (0.047)	0.019 (0.047)	-0.003 (0.046)
Female Top Manger Y:1 N:0				-0.018 (0.062)	-0.020 (0.062)	-0.019 (0.057)
Multi-firm Y:1 N:0				0.123*** (0.038)	0.121*** (0.039)	0.127*** (0.038)
Power Outages (day hours)					0.005 (0.004)	0.004 (0.005)
Finance Obstacle (0-4 scale)					-0.008 (0.013)	-0.008 (0.013)
Overdraft Y:1 N:0					0.062* (0.034)	0.060* (0.034)
Courts Obstacle (Country-level cell)						0.089** (0.036)
Population (lagged, logs)						-0.028** (0.014)
Inflation (% , annual)						0.002 (0.002)
Gross Tertiary Enrollment Rate						0.147 (0.183)
Gross Primary Enrollment Rate						0.111 (0.128)
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	5,721	5,721	5,721	5,721	5,721	5,721

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