Household enterprises: The impact of informality on productivity and profitability

Nesma Ali* Mohamed Ali Marouani†

Sept. 2019

Very preliminary and incomplete. Please do not cite or circulate without author' permission.

Abstract

This paper assesses the impact of commercial registration on the labor productivity and profitability of household enterprises. Based on the 2012 and 2018 rounds of the Egyptian Labor Market Panel Surveys, We use the distance to the local tax office as an instrument for firms' selection into formality. We find a positive effect of formal registration on labor productivity, together with a negative effect on profitability. Our findings also highlight formality heterogeneous effects. The effect of formality on firms' performance differs across firms with different characteristics and depends on the timing of transition to formal status. Formality is more beneficial for micro-sized firms owned by less-educated and female entrepreneurs who do not have a work social insurance.

JEL codes: O17, O12, D22, L26
Keywords: informal sector, entrepreneurship, productivity, Egypt

*Düsseldorf Institute for Competition Economics (DICE), Heinrich-Heine University Düsseldorf, email: ali@dice.hhu.de

†Developpement et societes, IRD and Paris 1 Pantheon Sorbonne University, email: marouani@univ-paris1.fr

1 Introduction

The large size of the informal sector is one of the main features of developing economies. The informal sector comprises small businesses or self-employed entrepreneurs, which produce legal goods and services but evade legal registration, financial reporting and taxation. Informal employment is also part of the informal sector. Understanding informality dynamics and its impact has become a major sub-field of development economics. While Maloney (2004) argues for a new microentrepreneurial view of the informal sector, La Porta and Shleifer (2014) focus on the dualism between the informal and formal sectors where the latter is the main source of growth. Rothenberg et al. (2016) confirm through Indonesian data that informal firms are smaller, relatively unproductive and pay low wages. Other papers investigate the safety net or last resort role of the informal sector and show that both voluntary and last resort segments coexist (Loayza and Rigolini, 2011; Günther and Launov, 2012). Hence, we can argue that the effect of informality is heterogeneous and mostly depends on owners' and firms' characteristics. Ulyssea (2018) presents a unifying framework for the competing explanations of informality encompassing two margins of informality. The decision to register the firm impacts the "extensive margin", while the decision to hire labor formally affects the "intensive margin". His counterfactual analysis on Brazil shows that informality does not necessarily reduce output or productivity.

Amin and Islam (2015) find that small informal firms outperform the larger ones in terms of productivity, based on a survey from seven African capitals. Using the data from the same region, Grimm et al. (2012) empirically identify a new segment of informal entities -"constrained gazelles"- which are characterized by the low levels of capital (similar to the firms in the lower tier of firm size distribution) but high levels of returns. They argue that "constrained gazelles" are potentially the best targets for informal sector interventions aimed at boosting growth.

We use an extensive panel survey of Egyptian household firms to revisit the main results of McKenzie and Sakho (2010) who find that informality is beneficial for medium firms' profitability but has a negative effect for smaller and larger firms in Bolivia. We extend the analysis of McKenzie and Sakho to deal with labor productivity. We also emphasize the heterogeneous effects of informality in terms of firm's size, owner gender and education level. Our data allows us to account for informality intensive margin as done by Ulyssea (2018).

Our paper also contributes to the literature on the impact of formal registration on informal firms' performance. In one of the earlier studies, McKenzie and Woodruff (2006) argue that going formal secures a better access to credit for firms and therefore translates into higher profitability. Fajnzylber et al. (2011) find that formalization increases the revenues and profits for a sample of Brazilian firms. The improvements in firms' performance is attributed to a permanent physical

location which facilitates the accumulation of capital and labor hiring. Similarly, Rand and Torm (2012) find that the increase of formal labor contracts is one of the main channels of better firms' performance in Vietnam. Finally, Benhassine et al. (2018) using an experimental evidence from Benin show that formalization costs exceed their benefits and suggest to target informal firms that have the closest characteristics to formal ones.

We focus on household enterprises operating in Egypt and analyze to what extent firms' formality matters for their productivity and profitability. We also disentangle the main mechanisms driving these effects, accounting not only for the firms' formality and size, but also for the differential impacts of the 2014 policy reforms, e.g. include the devaluation of the Egyptian currency, the removal of subsidies and the introduction of value-added tax. The link between macroeconomic shocks and the informal sector was studied by Fiess et al. (2010) for four Latin American countries. They show that the informal sector behavior can both be pro-cyclical or counter-cyclical depending on the nature of the shock.

Our main identification strategy is similar to that used by McKenzie and Sakho (2010) and relies on the distance to the registration office as an instrument for registration. It assumes that the closer the firm to its corresponding tax office where registration occurs, the easier and cheaper the taxation procedure, the better the access to information on the procedures and the higher the firm probability of being caught. Thus, closer firms are more likely to get a tax ID and operate formally. In addition, we use propensity score nearest neighbor matching and a difference-in-difference estimation to account for possible differential effects.

We find that formality has a positive effect on labor productivity and negative effect on profitability. These effects are heterogeneous and vary across firms with different observable characteristics. They also depend on the timing of the transition to the formal status. The positive effect of formality is stronger for micro-sized firms owned by less educated and female entrepreneurs who do not have a work social insurance. In addition, this positive impact is significantly reduced after the 2014 reforms, except for firms that became formal by 2018.

The remainder of the paper is laid out as follows. Section 2 explores the dataset and the main variables included in the model. Section 3 presents the identification strategy. Section 4 discusses the reported results and section 5 provides the main concluding remarks.

2 Data

This paper uses surveys on non-farm household-enterprises extracted from the Egyptian Labor Market Panel Surveys (ELMPS) for the years 2012 and 2018. These surveys are designed by the

Central Agency for Public Mobilization and Statistics (CAPMAS) and the Economic research forum (ERF).[1] The data covers a representative sample of 3550 unique household firms surveyed in 2012 and 2018 (2236 and 2142 firms respectively), among which 834 household firms were surveyed in both rounds. On average, households report owning one enterprise managed by one member of the household (hereafter "the owner" which is also the survey respondent). 62.5% of the firms included in the sample are one-person firms (self-employed), 31% are micro firms (with 2-4 workers), and 6.5% are small firms (with at least 5 workers).[2] Firms operate in 16 different sectors of activity and are distributed over 6 regions, 22 governorates, 242 Qism (hereafter, cities) and 709 Shyakha (hereafter, town).[3] Therefore, this dataset provides unique and accurate information on firms' location in each administrative division in Egypt which allows us to create an instrument based on the distance to the local tax office where registration occurs. It also provides information on firms' total number of workers, output and costs which we use to compute firms' productivity and profitability.

2.1 Measuring firm-level outcomes

We use the data on the number of workers and costs to calculate the main dependent variables; firms' labor productivity and profitability.

For each firm i in year t, firm-level labor productivity is the ratio of the firm's value added to its total workers (including owner) over the past year, as follows;[4]

$$Labor\ productivity_{it} = \frac{value\ added_{it}}{total\ workers_{it}} \tag{1}$$

Where, value added is the sum of firm-level net earnings defined as the average net earnings of the enterprise per month during the past 12 months, as well as firm-level expenditures in rental of land, machinery and vehicle and hired labor in the past 12 months.

$$\begin{aligned} value\ added_{it} = net\ earnings_{it} + land\ rental_{it} + hired\ labor_{it} \\ + machinery\ rental_{it} + vehicle\ rental_{it} \end{aligned} \tag{2}$$

Firm-level profitability is measured as the ratio between firm's net-earnings and value-added. This measure is a proxy for the percentage of firm's revenues which are turned into profits. In order

[1] The dataset is available and downloadable through the Economic Research Forum data portal: http://erf.org.eg/data-portal/

[2] The number of workers includes the owner. Workers might be members of the household or/& hired from outside

[3] In Egypt, Qism is the second administrative level after Governorate. It refers to a city and can be rural and urban. The third administrative level is Shyakha and refers to a town or village. Shyakha can be either rural or urban. For simplification we will refer to Qism and Shyakha as city and town respectively.

[4] All monetary values are in Egyptian pound and deflated using corresponding CPI-year

to pin-down the treatment effect more specifically we use other outcome variables such as; firm's size (total number of workers), and firm's market shares which are computed as the share of firm i's net-earnings in total net-earnings of all household firms operating in sector j.

2.2 Measuring firm's formality status

Firms' formality status is our independent variable of interest. According to the Egyptian commercial registry law (n°34/1976), investment law (n°72/2017) and tax law (n°91/2005), starting a formal business in Egypt requires the registration of the firm in the commercial registry regardless of the firm type and activity. This registration can be implemented only after the firm's request of a tax-ID (even in case of tax exemption). Firms' commercial registration expires every five years, but its renewal is ensured by the regularity of the firm's annual tax declaration and payment of taxes (if any). Hence, a registered firm (what we observe in the data) is a firm that already acquired a tax ID, respects the annual tax declaration and regularly pays its taxes. Yet, business licensing depends on the firm's type of activity and comes as the third step after firm's obtainment of the tax ID and official registration.

Accordingly, firms' commercial registration is the best proxy measuring firms' formality status. This information is extracted from the survey using the question "does the enterprise have commercial registration?". A "formality" dummy variable is created based on this question and is considered as our main independent variable throughout our estimation process. Furthermore, two other questions can be considered to have more restrictive definitions of firms' formality status; "do you have a business license?" and "do you keep regular accounting books?". Yet, only 18% of the firms in the sample report keeping accounting books in addition to acquiring a commercial registration and a business license.[5] These "more restrictive definitions" will be used as robustness check in the estimation process. The dataset also provides information on the owner formality status by investigating about the incidence of his work social insurance. This information allows us to investigate about the intensive margin of informality (informal employment) as recently done by Ulyssea (2018) who demonstrated the importance of investigating this margin when tackling firm informality effects (the extensive margin).

2.3 Stylized facts on the Egyptian formal/informal sector

Egypt stands as an interesting example when tackling firms' informality and performance. It is the largest country in North Africa and ranked among the highest in the contribution of the informal

[5] Keeping accounting books is a burdensome and biding constraint for small firms in Egypt and especially informal ones

economy in official GDP and of informal employment in total non-agriculture activities (Schneider, 2012; Charmes, 2012). Small-businesses and self-employment play a significant role in the economy and in the labor market (Gollin, 2008). This is due to various push and pull factors as in many other developing countries (Sahnoun et al., 2014; Falco and Haywood, 2016).

Most of the firms in the sample are informal (67.8%) and their rate is increasing throughout the rounds of the survey. This is in line with the reality of the Egyptian private sector and the characteristics of micro and small firms in Egypt (Sahnoun et al., 2014). Nevertheless, table 1 shows that, on average, formal firms are more productive, employ more workers, are larger in size and have a higher rate of survival compared to informal firms. They are also slightly older and managed by relatively older male owners who are more likely to have higher levels of education, and a work social insurance (formal owner). Yet, informal firms are relatively better in terms of profitability. According to figure 1, firms' labor productivity and profitability are increasing throughout the years of the surveys for both types of firms. Formal firms have a relatively higher increasing rate in profitability, while informal ones have a relatively higher increasing rate in labor productivity.

Firms are mostly micro-sized firms (maximum of four workers) with a sole proprietorship. However, most of larger firms and those with shared ownership tend to operate formally. In terms of firms' location, there are more informal firms in upper and lower Egypt compared to other regions in Egypt (Greater Cairo, Alexandria and Suez Canal), and in rural areas compared to urban areas. There are also more informal firms in sectors dealing with household activities, education and real estate, compared to sectors dealing with human health and information and communication. We can deduce from these facts that informal firms are on average operating in high productivity sectors of activity, but prefer to locate in relatively less productive regions and areas in order to hide from the authority.[6]

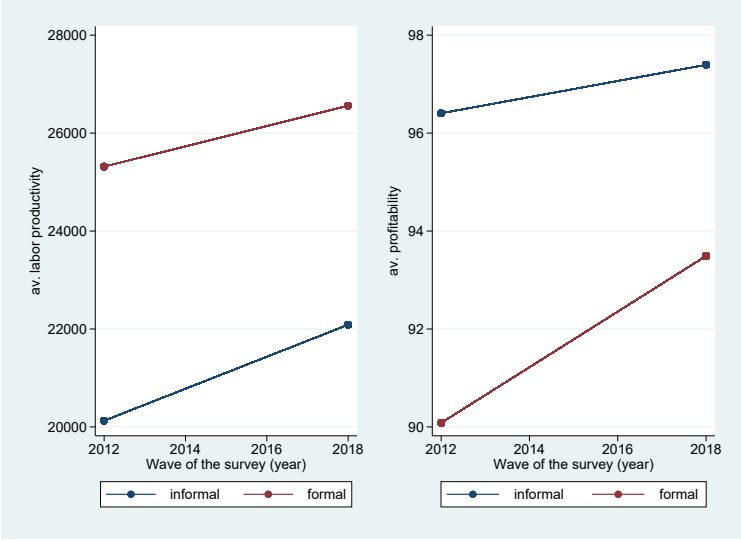
[6] The average labor productivity is higher in the real estate sector compared to human health or information and communication sectors.

Table 1: Firms' characteristics by formality status

	Formal	Informal		
	Mean	Mean	min	max
Total workers	2.73	1.61	1	53
Labor productivity	25812	21115	0	217000
Profitability	91.36	96.89	0	100
Survival (1=firm survives in 2018)	0.46	0.33	0	1
Market share	0.006	0.004	0	1
Firm age	18.44	14.88	1	102
Size (0=self-employed, 1=mid, 2=large)	0.69	0.34	0	2
Ownership (0=sole, 1=partnership)	0.18	0.08	0	1
Informal owners	0.55	0.83	0	1
Owner's age	43.61	40.77	14	86
Owner education (0<intermediate, 1>=intermediate)	0.67	0.46	0	1
Male owner	0.92	0.81	0	1
Total	1397	2669		

Notes. All monetary variables are in Egyptian pounds and deflated using corresponding CPI-year

Figure 1: Evolution of firms' labor productivity and profitability by formality status



3 Empirical strategy

Estimating the effect of firms' formality status on their productivity or profitability should be addressed with caution. Endogeneity bias may arise due to reverse causality existing between formality and firms' performance. Also, firms' decision to operate formally and to better perform depend on the characteristics of the owner such as gender and education; as well as on the characteristics of the enterprise such as its sector of activity and its size. This means that potential omitted variable and self-selection biases could alter our results. In order to overcome these biases, we instrument the formal status of the firms and control for various relevant observable characteristics. Specifically, in spirit of McKenzie and Sakho (2010)- who instrumented formality using the distance of the firm to the tax office where registration occurs - a distance instrument is constructed.

3.1 Instrumental variable approach

Egypt is divided into 27 major administrative divisions (governorates). Each governorate is divided into several cities (Qism) and each city into several towns or villages (Shyakha). In Egypt, tax collection occurs at the public tax office which is distributed in the country across cities.[7] As the dataset provides information about the location of the firms across towns, the distance between the city (or the exact location of the public tax office in each city) and the centroid of its respective town (where firm is located) will be used as an instrument for firms' formality status.[8] This distance variable is in kilometers and based on the longitudes and latitudes coordinates of each location. Using this instrument, we assume the following: 1) firms' distance to their corresponding tax authority in the city explains firms' formality status; 2) there is no direct effect between firms' distance to their corresponding tax authority and firms' performance.

As previously mentioned, according to our first assumption, firms' registration (formality) is applicable after obtaining a tax ID which should be requested by the firm at its local city-level tax authority.[9] Therefore, the closer is the firm to its corresponding tax authority, the higher is its probability to acquire a tax card and register the firm. One can also argue that the closer is the firm to its respective tax office, the lower is the cost of registration, the easier is the access to information on registration and taxation procedures, and the higher is the probability of getting caught due to

[7] The owner of the firm should be at least 21 years old to register a firm. Thus, we restrict our sample to this age limit.

[8] The addresses of city-level public tax offices are extracted from the Egyptian tax authority website (http://www.incometax.gov.eg/map.asp). In 10 cases, corresponding coordinates were not found. For these cases we use the coordinates corresponding to the centroid of the city.

[9] In 2017, a special unit has been created by the government for the development of micro, small and medium enterprises (msme). This unit facilitates firms' creation through a one stop shop that allows the firm to obtain its tax ID and commercial registration from on place. Yet, this unit mostly concerns projects receiving social funds (which corresponds to 4 enterprises in our sample) and is not well spread geographically as tax authorities (almost one unit per governorate).

tax inspection (McKenzie and Sakho, 2010).

Nevertheless, one might think that the second assumption could be violated due to urban agglomeration in Egypt. Most of firms are located in the capital city and big cities surrounding the capital where most of the population density is concentrated, as well as the demand (The World Bank, 2012). This means that more productive firms would be located in the core (residential and industrial zones with sound infrastructures), while less productive firms would be located near the periphery. Yet, this kind of agglomeration resulted in strong diseconomies of scale that pushes firms to diversify their location choice (i.e. city congestion, poor infrastructure, informal settlements).[10] Our data also supports firms' diversification. High or low productive firms are not exclusively concentrated in some specific governorates, cities or towns. They are rather distributed in different locations regardless of their productivity level. Moreover, some of the most productive firms are located exactly in the same governorate-city-town levels as least productive firms.

Yet, In order to ensure that we totally satisfy the exclusion restriction condition, we control in our estimations for the intensity of economic activity in each town using night-lights data extracted from the Earth Observatory group.[11] The latter is normalized with respect to the brightest town per city in order to obtain a variable between 0 (towns with lowest level of economic activity) and 1 (towns with highest level of economic activity). In addition, we control for the population density in each town.[12] Therefore, for two different towns with an equal level of economic activity and population density, firms located in the closest town to the tax office, are more likely to operate formally.

Hence, the effect of firm's formality on firm's labor productivity or profitability is estimated using the following equation (baseline regression);

$$y_i = \beta_0 + \beta_1 formality_i + \beta_2 EcoActivity_s + \beta_3 Pop_s + \beta_4 X_i + d_t + \alpha_j + \theta_g + u_{it} \tag{3}$$

Where, y_i is the dependent variable measured as the logarithm of firm-level labor productivity or profitability as explained in section 2.1 . $formality_i$ is the independent variable of interest and refers to firm i formality status. Formality status is estimated in a first step using $distance_{s,k}$ as an instrument, which is the logarithm of the distance in kilometers between the town s where the firm is located and its corresponding tax office in the city k. $EcoActivity_s$ and Pop_s proxy

[10] See Krugman (1991) and Ellison et al. (2010) for a full review on urban agglomeration and firms' performance.

[11] Several paper used the night-lights data to measure economic activity or GDP such as (Henderson et al., 2011 and Storeygard, 2016). This data is derived from a digital number that falls between 0 and 63. Zero represents the least bright point (i.e. lowest level of economic activity) and 63 represents the brightest point (i.e. highest level of economic activity). The data is downloadable through the following link:https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

[12] data on population density are extracted from the CAPMAS

for town-level economic activity and population density respectively. X_i is a vector of firm-level characteristics including age, size, and ownership status; and owner-level characteristics including age, gender, and education level. Dummies for the round of the survey (d_t), firm's sector of activity (α_j) and governorate (θ_g) are introduced to control for year-governorate-sector specific effects.[13]. All regressions are clustered at the town-level.

4 Results

This section presents the effect of firms' formality on firms' labor productivity and profitability. We start by reviewing the main determinants of firms' formality, highlighting the importance of firms' distance to the tax office (our instrument) using a probit estimation. Next, We compare the estimated coefficients of the main regression provided using OLS, fixed-effects and IV-2sls models. This comparison allows us to tackle formality' heterogeneous effects using a matching approach, and to emphasize timing effects using a difference-in-difference model. Lastly, we present some robustness checks that validate our results using different definitions of formality.

4.1 Firms' formality status and distance to the tax office

The identification strategy used in this paper puts the emphasize on the importance of firms' distance to its corresponding tax office in determining firms' willingness to operate formally. Table 2 reports marginal effects estimated from a probit equation where firms' formality (acquiring a commercial registration) is the dependent variable. The results show that distance has a significant negative impact on firms' formality. According to Column (1), a firm likelihood to operate formally drops by 4 percentage points on average if the firm is located in a town that is 7 km away from the corresponding tax office (the average distance). This negative effect remains valid when controlling for the characteristics of the firm and its owner in Column (2).

Column (2) indicates that, in addition to the distance, firm and owner's characteristics matter significantly. More specifically, the increase in firm's size is associated with 33 percentage points increase in firm's willingness to formalize. Firms' transition into the formal sector is a necessity to ensure its sustainable growth and survival (La Porta and Shleifer, 2014). In addition, firm's probability of operating formally significantly increases when the firm is managed by an older, male and more educated owner who favors shared ownership over sole ownership. Having an intermediate or higher level of education increases the owner's understanding capacity of the formalization's

[13] As only 18% of surveyed firms have been interviewed over the 2 rounds of the survey, firms' transition in formality status is very limited and is subject to measurement error. Therefore, the panel structure of the sample is ignored and the data is treated as a pooled cross section with year fixed effects.

advantages and procedures.

Table 2: Firms' formality status and distance to the tax office

	Pr(Firm is formal)	
	(1)	(2)
Log(distance to tax office)	-0.0408***	-0.0239**
	(0.0111)	(0.00939)
Mid-size firm (ref. self-employed)		0.150***
		(0.0152)
Larger firm (ref. self-employed)		0.337***
		(0.0289)
Firm age		0.00202***
		(0.000751)
Shared ownership (ref. sole)		0.123***
		(0.0223)
Owner age		0.00352***
		(0.000651)
High education level		0.164***
		(0.0149)
Male owner		0.134***
		(0.0241)
Night-light intensity	0.0895	0.0221
	(0.120)	(0.105)
Log (pop.density per town)	0.0110	0.00795
	(0.0108)	(0.00937)
Observations	3,874	3,658
Year FE	Yes	Yes
sector FE	Yes	Yes
Governorates FE	Yes	Yes
Cluster se-level	Town-level	Town-level

Notes. The table reports marginal effect of probit regression where the dependent variable is firm's formality status, which is a dummy variable that equals one if the firm has a commercial registration, and zero otherwise. Robust standard errors are clustered at the town-level and are reported between brackets in all columns.*** p<0.01, ** p<0.05, * p<0.1

4.2 Firms' formality, productivity and profitability

Our baseline regression results for labor productivity are reported in table 3. Column (1) reports estimated coefficients from an OLS estimation with controls for year, sector and governorates specific effects. The results indicate that firm's labor productivity significantly increases by 33% on average when the firm operates formally. It also shows that high educated male entrepreneurs not only have a higher willingness to operate formally (as shown in table 2), but have a significant higher productivity as well. The negative effect associated with firms' size seems counter intuitive as it shows that self-employed are more productive than larger firms. However, it confirms the result

found by Amin and Islam (2015) in seven other African countries.

Column (2) validates the positive effect of formality by using a fixed-effect estimation. However, most of the other covariates lose significance because of limited variation across time. And as explained in section 3.1, OLS estimation suffers from endogeneity and omitted variable bias which leads to the estimation of an instrumental-variable 2SLS model as reported in columns (3-6) of table 3. The effect of formality remains positive and significant when using the 2SLS estimation and across different specifications: baseline regression (column 3), removing year-sector-governorates controls (column 4), and adding firm-owner's education and gender respectively in columns (5 & 6). Also, the estimated effects of the other covariates remain very similar to the OLS estimation's results, except for owner's education level that becomes insignificant.

Table 4 reports the effect of firms' formality status on firms' profitability and other firm-level outcomes. According to column (2), formality marks a reduction of 21% in firms' profitability, which contradicts the findings of McKenzie and Sakho (2010) who found an increase in firms' profit in Bolivia due to tax registration. This result may provide an explanation of Egypt's large informal sector. Formality has a higher social return (increase in labor productivity) but a lower private return (decrease in profitability). Columns (3-4) estimate the effect of formality on the firm's total number of workers and market share. While formality has no significant effect on firms' size, it has a significant positive effect on firm's market share per sector.

Nevertheless, one concern here with the IV results could be the large gap in magnitude between the OLS and 2SLS coefficients. One plausible explanation is that these coefficients show different effects. The OLS reports the average treatment effect (ATE), while the 2SLS reports the local average treatment effect (LATE) which is the causal effect for the group of compliers (i.e. firms that are formal just because they are closer to the tax office). According to Carneiro et al. (2011), this difference mostly arises when the treatment effect is heterogeneous (i.e. the impact of formality differs across firms) or due to measurement errors. One might think that systematic measurement error is the reason as people usually tend to hide the informality status of their firms. However, the data collection process along with the first estimated results allow us to believe that this difference is rather due to heterogeneous effects.

On one hand, columns (5 & 6) show the sensitivity of the estimated effect of formality when introducing observable characteristics as the gender and education of the owner. This argues in favor of heterogeneous effect due to observable characteristics which will be emphasized in the next section. On the other hand, the data used in the paper is primarily a labor market survey administrated for individuals and households. Among the different sections of this survey, there is a form dedicated to households owning a firm and only at that point comes the different questions dealing with "firms".

This means that at this point the respondent will be already aware that investigating on taxation or formality is not the aim of the survey. Second, data collectors are well trained on addressing these types of sensitive questions. Questions on the formality status of the firm are rather indirect questions on the acquisition of commercial registration, licensing or accounting books. Third, in section 4.5, we show that our results remain valid to the use of these different measures, as well as to the use of a combined definition of the intensive and extensive margins of informality.

Table 3: Firms' formality and productivity

	Log(firm labor productivity)					
	OLS	FE	2sls	2sls	2sls	2sls
	(1)	(2)	(3)	(4)	(5)	(6)
Formal firm	0.335***	0.382**	1.865**	2.170***	1.862*	1.470*
	(0.0505)	(0.154)	(0.801)	(0.716)	(0.969)	(0.810)
Mid-size firm (ref. self-employed)	-0.549***	-0.936***	-0.744***	-0.817***	-0.744***	-0.722***
	(0.0488)	(0.152)	(0.146)	(0.127)	(0.163)	(0.138)
Larger firm (ref. self-employed)	-1.120***	-0.958***	-1.570***	-1.660***	-1.570***	-1.509***
	(0.0971)	(0.353)	(0.319)	(0.259)	(0.362)	(0.304)
Firm age	0.0115***	0.0141*	0.0114***	0.00819**	0.0114***	0.00938***
	(0.00226)	(0.00778)	(0.00352)	(0.00382)	(0.00400)	(0.00316)
Shared ownership (ref. sole)	0.0921	0.0594	-0.102	-0.169	-0.102	-0.0803
	(0.0693)	(0.226)	(0.158)	(0.163)	(0.177)	(0.149)
Owner age	-0.00947***	-0.0140	-0.0156***	-0.0165***	-0.0155***	-0.0128***
	(0.00212)	(0.0119)	(0.00269)	(0.00294)	(0.00383)	(0.00275)
High education level	0.185***	-0.406			0.00235	
	(0.0475)	(0.280)			(0.174)	
Male owner	0.768***	0.625**				0.618***
	(0.0651)	(0.278)				(0.143)
Night-light intensity			-0.392	-0.116	-0.392	-0.403*
			(0.259)	(0.313)	(0.256)	(0.241)
Log (pop.density per town)			-0.00380	-0.0193	-0.00375	-0.00578
			(0.0343)	(0.0388)	(0.0360)	(0.0322)
Observations	3,222	814	3,105	3,144	3,105	3,105
Year FE	Yes	Yes	Yes	No	Yes	Yes
Sector FE	Yes	Yes	Yes	No	Yes	Yes
Governorates FE	Yes	No	Yes	No	Yes	Yes
Firm-level FE	No	Yes	No	No	No	No
Cluster se-level	-	-	Town-level	Town-level	Town-level	Town-level
1st stage F-test			15.51	26.48	10.66	12.84

Notes. This table shows the effect of firms' formality on the logarithm of firms' labor productivity. Robust standard errors are reported between brackets in all columns. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Firms' formality, profitability and other outcomes

	(1) Log(labor productivity)	(2) Log(profitability)	(3) Log(total workers)	(4) Log(market share)
Firmal firm	1.865**	-0.210**	0.210	0.0312**
	(0.801)	(0.101)	(0.284)	(0.0143)
Mid-size firm (ref. self-employed)	-0.744***	-0.0209		-0.00403
	(0.146)	(0.0193)		(0.00255)
larger firm (ref. self-employed)	-1.570***	-0.0657		-0.00333
	(0.319)	(0.0472)		(0.00610)
Firm age	0.0114***	0.000811*	0.00214	-8.20e-06
	(0.00352)	(0.000456)	(0.00132)	(5.07e-05)
Shared ownership (ref. sole)	-0.102	0.0373*	0.234***	-0.00520*
	(0.158)	(0.0221)	(0.0709)	(0.00305)
Owner age	-0.0156***	0.00114***	0.00172*	-7.82e-05**
	(0.00269)	(0.000302)	(0.000992)	(3.86e-05)
Night-light intensity	-0.392	-0.0523*	0.0123	-0.00430
	(0.259)	(0.0316)	(0.0928)	(0.00457)
Log(pop. density per town)	-0.00380		0.00639	-0.000718
	(0.0343)		(0.0116)	(0.000510)
Observations	3,105	3,044	3,603	3,232
Year FE	Yes	Yes	Yes	Yes
sector FE	Yes	Yes	Yes	Yes
Governorates FE	Yes	Yes	Yes	Yes
Cluster se-level	Town-level	Town-level	Town-level	Town-level
1st stage F-test	15.51	18.91	20.53	17.17

Notes. This table shows the effect of firms' formality on the logarithm of firms' profitability and other firm-level outcomes using an IV-2sls estimation. Robust standard errors are clustered at the Town-level and reported between brackets in all columns. *** p<0.01, ** p<0.05, * p<0.1.

4.3 Emphasizing heterogeneous effects

This section sheds the light on the heterogeneous effects based on observable characteristics that formality has on productivity. Columns (5 & 6) of table 3 set the gender of the entrepreneur and the education level as potential candidates driving these heterogeneous effects. We will also emphasize those related to the size of the firm as well as to the formality status of the owner of the firm. The size effect is highlighted in most of the papers tackling informality and shows ambiguous results (see for example Maloney, 2004; McKenzie and Sakho, 2010; Amin and Islam, 2015). Similarly, the recent paper of Ulyssea (2018) showed the importance of accounting for the intensive margin of informality (informal workers) when tackling the extensive margin (informal firms). As our data provides incomplete information on workers formality status, the formality status of the firm's owner is used as a proxy.[14]

For that purpose, a propensity score matching is applied using the nearest neighbor matching technique. Results are reported in table 5. We start by matching firms using all of the covariates included in our main regression leaving out the variables firm's size, owner's education, gender and

[14] This is a valid proxy for the intensive margin as 62.5% of the firms in our sample are self-employed (one- person-firms)

formality status. Column (1) reports a higher significant treatment effect compared to the OLS estimation, but still lower effect compared to the 2SLS estimation. We then re-apply this matching by splitting our sample in terms of these last variables (columns 2-10). Columns (2-4) indicate that the effect of formality is positive and significant only for self-employed and micro firms (firms with less than 4 workers). Yet, the treatment effect loses its entire magnitude and significance when considering larger firms (more than 4 workers - column 4). Similarly, the effect of formality remains positive and highly significant for low-educated entrepreneurs comparing to high educated ones. Accordingly, the formality treatment effect remains positive and significant for the other categories, but shows higher magnitudes for firms owned by female (columns 7 & 8), and informal (columns 9 & 10) entrepreneurs, compared to male and formal entrepreneurs. Hence, our estimated results prove that these four variables can explain the potential non-linear relationship existing between formality and productivity, especially in terms of firm size, owner's gender and education level.

Unlike most of the literature which showed a positive correlation between formal firms' productivity and size (see Porta and Shleifer, 2008), our results show that the significant increase in productivity of formal firms decreases with firm's size. Yet, this finding is in line with the results of Amin and Islam (2015) and those of McKenzie and Sakho (2010) who argue that accessing formality benefits in an earlier stage (access to infrastructure, credits and protection of property rights) translates into higher profitability levels for self-employed and micro firms. This might also be explained by the fact that formal smaller firms are on average younger than formal larger ones which is in line with the findings of Fajnzylber et al. (2011). However, larger informal firms have a better knowledge in benefiting from informality by knowing the best efficient way to escape taxation. Hence, these last would have higher productivity if they allocate saved resources towards more investment and economy of scale.[15]

This argument also explains two new contributions presented by this paper. The first concerns owner's formality status indicating that formal firms run by informal owners are significantly more productive compared to those run by formal owners (Columns 9 & 10). The second is about the education level of the owner showing that formality is more beneficial for low educated entrepreneurs comparing to high educated ones (Columns 5 & 6). This might be explained by Egypt's poor education system that makes learning by doing and experience more important than the attained education level, especially that less educated owners in our sample are on average older (as well as their firms).

Another interesting insight of this paper is the heterogeneity of formality effect in terms of the owner's gender. Columns (7 & 8) indicate that formal firms run by female entrepreneurs are almost

[15] Creating economy of scale while staying informal also reflects more corruption practices (bribery) or/& political connections

5 times more productive than those run by male entrepreneurs. These findings add to the existing literature about the gender gap and discrimination in Egypt (El-Hamidi, 2011 and El Hamidi and Baslevent, 2010). Even though females are more likely to be trapped into the informal sector due to barriers set by family and society, they can outperform their male counterparts if they join the formal sector.

Table 5: Formality heterogeneous effects

	(1) Without heterog. controls	(2) Self-employed	(3) Mid-size firm	(4) Small firm	(5) High education	(6) Low education	(7) Male	(8) Female	(9) Formal owner	(10) Informal owner
Formal firm	0.419***	0.512***	0.490***	-0.181	0.0827	0.314***	0.124*	0.604**	0.207*	0.219**
	(0.0722)	(0.0873)	(0.115)	(0.380)	(0.121)	(0.102)	(0.0713)	(0.246)	(0.122)	(0.108)
Observations	3,118	1,914	998	175	1,652	1,445	2,651	414	790	2,303
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Governorates FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster se-level	Town-level	Town-level	Town-level	Town-level	Town-level	Town-level	Town-level	Town-level	Town-level	Town-level

Notes. This table presents average treatment effects of formality on firms' labor productivity from a nearest neighbor propensity score matching. These regressions include other explanatory variables such as firm and owner's age and firm's ownership status. Robust standard errors are bootstrapped (500 replications) and are reported between brackets in all columns.*** p<0.01, ** p<0.05, * p<0.1

4.4 Firms' formality, productivity and reforms

As the aftermath of the Egyptian revolution, starting from 2014, the Egyptian government adopted a dramatic reform program including exchange rate liberalization, fiscal consolidation, energy sector reform and business environment reform. The first wave of these reforms have been translated into increasing inflation through the devaluation of the Egyptian currency (in 2016), the removal of energy subsidies (since 2014), and the introduction of VAT (in 2017). Hence, these reforms could be considered as exogenous shocks that directly affect firms' costs (i.e. transport and energy costs) and firms' capital value (through households' savings that represent firms' main source of capital) from one hand; and consumption and workers' wages from the other hand. As the underlined effects were solely experienced by firms surveyed in 2018, a difference-in-difference estimation (DID) is implemented to estimate the differential in firms' responses to informality after the implementation of these reforms compared to before.

Table 6 reports the estimated results from the difference-in-difference estimation. According to Column (1), the direct effect of formality remains positive and significant and very close to the one found in our main regression. Yet, this positive effect is reduced by almost 20% on average in 2018 compared to before. As explained, this is mainly due to the fact that these reforms induced higher unexpected costs at the firm and owner levels.

Column (2) focuses on the sub-sample of firms that were interviewed in both rounds to account for firms transition in formality status. We account for 4 different categories. The first category includes the "remain formal" group and refers to firms that were formal in 2012 and 2018. The second includes the "remain informal" group and refers to those that were informal in 2012 and 2018. The third includes the "became informal" group and refers to those that were formal in 2012 and became informal in 2018. The fourth and reference category includes the "became formal" group and refers to those that were informal in 2012 and became formal in 2018.

Results from the DID estimation report lower productivity levels for "remain informal" or "became formal" firms compared to "remain formal" firms. However, in 2018 and after waves of reforms, the reported lower productivity is overcompensated only for firms who "became formal". The latter report a net increase in labor productivity of almost 25% from being formal in 2018 compared to 2012. In contrast, For "remain informal" and "became informal" firms, productivity in 2018 has not changed significantly compared to 2012.

Taken all together, operating formally has a net positive impact on firms' labor productivity. This positive impact was significantly reduced after the reforms, except for firms that became formal, who marked a net increase in labor productivity compared to "remain formal" firms. This finding can be explained by the fact that "became formal" firms are the only type of firms who were lucky to

benefit from both informality (via resource saving) and formality (via protection of property rights, access to finance and infrastructure). This was not the case for "remain informal" firms who lost most of the benefits related to informality due to the increase probability of getting caught after the reforms. Therefore, the results of the DID estimation add to our previous results proving that timing has also a role in explaining the non-linearity between formality and firms' productivity.

Table 6: Firms' formality, productivity and reforms

	Difference-in-difference Log(labor productivity)	
	(1)	(2)
2018	-0.0366	-0.256
	(0.0757)	(0.155)
Formal firm	0.425***	
	(0.0601)	
2018*formal	-0.192*	
	(0.101)	
Firm's transition in formality (ref. remain formal)		
Remain informal		-0.419***
		(0.110)
Became informal		0.0285
		(0.121)
Became formal		-0.315**
		(0.130)
2018*Remain informal		0.153
		(0.215)
2018*Became informal		0.0433
		(0.257)
2018*Became formal		0.562**
		(0.232)
Constant	8.965***	10.28***
	(0.237)	(0.439)
Observations	3,274	1,103
R-squared	0.177	0.176
Year FE	Yes	Yes
sector FE	Yes	Yes
Governorates FE	Yes	Yes
Cluster se-level	Town-level	Town-level

Notes. This table shows treatment effect of formality after the 2014 reforms on the logarithm of firm's labor productivity. These regressions include other explanatory variables such as firm's size, age and ownership status, as well as owner's age and gender. Robust standard errors are clustered at the Town-level and reported between brackets in all columns.*** p<0.01, ** p<0.05, * p<0.1

4.5 Robustness checks

So far, the direct positive effect of formality on firms' productivity is validated in all specifications used in this paper; OLS, 2SLS, matching and DID (tables 3, 5 and 6). We also demonstrate that our instrument is strong and relevant even when considering other firm-level outcomes (table 4). We add other robustness checks that mark the validity of our main effect using different definitions for our formality variable as explained in section 2.2. Table 7 demonstrates the significant positive effect that formality has on labor productivity using the following definitions: firm is registered at the commercial registry (column 1), firm is registered at the commercial registry and has a business license (column 2), firm is registered at the commercial registry, has a business license and keeps accounting books (column 3). Column (4) adds the intensive margin of formality to our baseline definition. Here, formality refers to formal firms (acquiring commercial registration and business license), or to formal firms owned by formal entrepreneurs (acquiring a work insurance).[16] Once again the results remain positive and significant.

Table 7: Robustness check: formality definition

	Log(firm labor productivity)			
	(1)	(2)	(3)	(4)
Registration	1.999**			
	(0.906)			
Registration+license		2.023**		
		(0.926)		
Registration+license+accounting			1.746**	
			(0.755)	
Formality firm and owner				1.743**
				(0.826)
Observations	3,154	3,279	3,286	3,621
Year FE	Yes	Yes	Yes	Yes
sector FE	Yes	Yes	Yes	Yes
Governorates FE	Yes	Yes	Yes	Yes
Cluster se-level	Town-level	Town-level	Town-level	Town-level
1st stage F-test	12.84	12.70	23.06	14.93

Notes. This table presents the effect of different definitions of firms' formality on firms' labor productivity using an IV-2SLS estimation. These regressions include other explanatory variables such as firm and owner's age; firm's size and ownership status; and town-level night-light intensity and population density. Robust standard errors are clustered at the town-level and reported between brackets in all columns.*** p<0.01, ** p<0.05, * p<0.1

[16] Combining the intensive and extensive margins of formality create different categories (only firm is formal, only owner is formal, both are formal, and both are informal). According to table 8 in the appendix, the categories that matter the most are "both formal" and "firm is formal". For this reason we decide to define formality in Column (4) of table 7 as mentioned

5 Conclusion

This paper investigates the impact of commercial registration on the labor productivity and profitability of household enterprises in Egypt. Using the Egyptian Labor Market Panel Surveys for the 2012 and 2018 rounds, it also explores the heterogeneous effects of formality on performance and highlights the extent to which household enterprises were deferentially affected by the 2014 economic reforms. In order to overcome the selection into formal registration, which might be driven by unobservable firm characteristics, we instrument formality using the distance to the local tax office where registration occurs.

As expected firms located closer to the tax office are more likely to eventually register. As for the main result, we find a positive effect of formality on labor productivity, which is line with the previous literature covering the effect of formality on different firms' outcomes (Elbadawi and Loayza (2008) for multiple Arab countries, McKenzie and Sakho (2010) for Bolivia, and Fajnzylber et al. (2011) for Brazil). We also show that formality has a positive impact on firms' market share. In contrast, we find that formality has a significant negative impact on firm's profitability and has no effect on firm's employment.

In addition, using a propensity score nearest neighbor matching and a difference-in-difference estimation, this paper proves that the effect of formality differs across firms in terms of observable characteristics, the timing of the effect, and according to firms' transition in formality status. The positive effect of formality is strongr for micro-sized firms (which is in line with the findings of Fajnzylber et al. (2011)) owned by less educated and female entrepreneurs who do not have a work social insurance. The positive effect on labor productivity was significantly reduced after the 2014 economic reforms with an exception of firms that became formal by 2018.

References

Amin, Mohammad and Asif Islam, "Are large informal firms more productive than the small informal firms? Evidence from firm-level surveys in Africa," *World Development*, 2015, *74*, 374–385.

Benhassine, Najy, David McKenzie, Victor Pouliquen, and Massimiliano Santini, "Does inducing informal firms to formalize make sense? Experimental evidence from Benin," *Journal of Public Economics*, 2018, *157*, 1–14.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil, "Estimating marginal returns to education," *American Economic Review*, 2011, *101* (6), 2754–81.

Charmes, Jacques, "The informal economy worldwide: Trends and characteristics," *Margin: The Journal of Applied Economic Research*, 2012, *6* (2), 103–132.

El-Hamidi, Fatma, "How do women entrepreneurs perform? Empirical evidence from Egypt," in "Economic Research Forum (ERF) Working Paper" number 621 2011.

Elbadawi, Ibrahim and Norman Loayza, "Informality, employment and economic development in the Arab world," *Journal of Development and Economic Policies*, 2008, *10* (2), 25–75.

Ellison, Glenn, Edward L Glaeser, and William R Kerr, "What causes industry agglomeration? Evidence from coagglomeration patterns," *American Economic Review*, 2010, *100* (3), 1195–1213.

Fajnzylber, Pablo, William F Maloney, and Gabriel V Montes-Rojas, "Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program," *Journal of Development Economics*, 2011, *94* (2), 262–276.

Falco, Paolo and Luke Haywood, "Entrepreneurship versus joblessness: Explaining the rise in self-employment," *Journal of Development Economics*, 2016, *118*, 245–265.

Fiess, Norbert M., Marco Fugazza, and William F. Maloney, "Informal self-employment and macroeconomic fluctuations," *Journal of Development Economics*, March 2010, *91* (2), 211–226.

Gollin, Douglas, "Nobody's business but my own: Self-employment and small enterprise in economic development," *Journal of Monetary Economics*, 2008, *55* (2), 219–233.

Grimm, Michael, Peter Knorringa, and Jann Lay, "Constrained Gazelles: High Potentials in West Africa's Informal Economy," *World Development*, 2012, *7* (40), 1352–1368.

Günther, Isabel and Andrey Launov, "Informal employment in developing countries: Opportunity or last resort?," *Journal of development economics*, 2012, *97* (1), 88–98.

Hamidi, Fatma El and Cem Baslevent, "The gendered aspects of MSES in MENA: Evidence from Egypt and Turkey," in "Economic Research Forum Working Paper" number 535 2010.

Henderson, Vernon, Adam Storeygard, and David N Weil, "A bright idea for measuring economic growth," *American Economic Review*, 2011, *101* (3), 194–99.

Krugman, Paul, "Increasing returns and economic geography," *Journal of political economy*, 1991, *99* (3), 483–499.

Loayza, Norman V and Jamele Rigolini, "Informal employment: safety net or growth engine?," *World Development*, 2011, *39* (9), 1503–1515.

Maloney, William F, "Informality revisited," *World development*, 2004, *32* (7), 1159–1178.

McKenzie, David and Yaye Seynabou Sakho, "Does it pay firms to register for taxes? The impact of formality on firm profitability," *Journal of Development Economics*, 2010, *1* (91), 15–24.

McKenzie, David J and Christopher Woodruff, "Do entry costs provide an empirical basis for poverty traps? Evidence from Mexican microenterprises," *Economic development and cultural change*, 2006, *55* (1), 3–42.

Porta, Rafael La and Andrei Shleifer, "The unofficial economy and economic development," Technical Report, National Bureau of Economic Research 2008.

Porta, Rafael La and Andrei Shleifer, "Informality and development," *Journal of Economic Perspectives*, 2014, *28* (3), 109–26.

Rand, John and Nina Torm, "The benefits of formalization: Evidence from Vietnamese manufacturing SMEs," *World development*, 2012, *40* (5), 983–998.

Rothenberg, Alexander D, Arya Gaduh, Nicholas E Burger, Charina Chazali, Indrasari Tjandraningsih, Rini Radikun, Cole Sutera, and Sarah Weilant, "Rethinking Indonesia's informal sector," *World Development*, 2016, *80*, 96–113.

Sahnoun, Hania, Philip Keefer, Marc Schiffbauer, Abdoulaye Sy, and Sahar Hussain, *Jobs or Privileges: unleashing the employment potential of the Middle East and North Africa*, The World Bank, 2014.

Schneider, Friedrich, "The Shadow Economy and Work in the Shadow: What Do We (Not) Know?," IZA Discussion Papers 6423, Institute of Labor Economics (IZA) March 2012.

Storeygard, Adam, "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa," *The Review of economic studies*, 2016, *83* (3), 1263–1295.

The World Bank, "Arab Republic of Egypt. Reshaping Egypt's Economic Geography: Domestic Integration as a Development Platform," Technical Report june 2012.

Ulyssea, Gabriel, "Firms, informality, and development: Theory and evidence from Brazil," *American Economic Review*, 2018, *108* (8), 2015–47.

Wahba, Jackline, "Informality in Egypt: a stepping stone or a dead end?," in "Cairo, Egypt: Economic Research Forum (ERF), Working Paper" number 456 2009.

6 Appendix

Table 8: Combining formality intensive-extensive margins

	Log(firm labor productivity)
Firm-owner formality (ref. informal firm & owner)	
Formal firm & owner	0.461***
	(0.0646)
Only firm is formal	0.485***
	(0.0577)
Only owner is formal	0.0705
	(0.0671)
Mid-size firm (ref. self-employed)	-0.483***
	(0.0492)
Larger-size firm (ref. self-employed)	-1.013***
	(0.0882)
Firm age	0.0138***
	(0.00241)
Shared ownership (ref. sole)	0.162**
	(0.0651)
Owner age	-0.0126***
	(0.00223)
Observations	3,404
R-squared	0.083
Year FE	Yes
sector FE	Yes
Governorates FE	Yes

Notes. This table presents the effect of firm-owner's formality categories on the logarithm of firm's labor productivity using an OLS estimation. Robust standard errors are reported between brackets in all columns.*** p<0.01, ** p<0.05, * p<0.1