Received: 10 October 2022 | Revised: 22 January 2024 | Accepted: 27 February 2024

DOI: 10.1111/rode.13101

REGULAR ARTICLE

WILEY

How do small formal and informal firms in Egypt compare? An analysis of firm characteristics and implications for formalization efforts

Caroline Krafft[1] | **Ragui Assaad**[2] | **Khandker Wahedur Rahman**[3,4] | **Maakwe Cumanzala**[1]

[1]Department of Economics and Political Science, St. Catherine University, Saint Paul, Minnesota, USA

[2]Humphrey School of Public Affairs, University of Minnesota, Minneapolis, Minnesota, USA

[3]Oxford Martin School, University of Oxford, Oxford, UK

[4]BRAC Institute of Governance and Development, BRAC University, Dhaka, Bangladesh

Correspondence
Ragui Assaad, Humphrey School of Public Affairs, University of Minnesota, Minneapolis, MN, USA.
Email: assaad@umn.edu

Funding information
World Bank Group

Abstract

Formalizing firms can potentially increase the tax base, expand safety and social protections for workers, create good jobs, and promote firm growth. However, the costs and processes of formality may be too challenging for some firms to bear. Thus, informal firms may not be able to survive the transition to formality, and attempts to expand formality through more intensive enforcement may be harmful and counterproductive to job creation and growth. This paper investigates the potential for currently informal firms to formalize in Egypt. The paper compares the characteristics of formal and informal micro and small nonagricultural firms and identifies the extent of similarities and potential for formalization. The analysis finds that, beyond firm size and whether the firm operates in a fixed establishment, the basic and easily observable characteristics of firms

are not closely linked to formality. Firm age, productivity, and owner characteristics, such as education, are strongly predictive of formality. The predicted probability of being formal is greater than 50% for only about 26% of informal firms, suggesting most firms are not good candidates for formalization. The paper develops profiles (groups and clusters) of similar firms to identify those with a higher potential for formalization.

KEYWORDS

Egypt, firms, formality, microenterprises

JEL CLASSIFICATION

D22, L11, L26

1 | INTRODUCTION

Informal firms play a critical, but controversial, role in the economies of developing countries (La Porta & Shleifer, 2014). These firms may have untapped potential for job creation but may be held back by burdensome government regulations (Besley & Burgess, 2004). Alternatively, informal firms may have an unfair competitive advantage compared with formal firms (due to lower costs) (Ali & Najman, 2016). Formal and informal firms may also be, fundamentally, different, in terms of entrepreneur and firm characteristics, such as productivity (Mishra, 2022). The degree to which this dualism pertains, and which types of informal firms can potentially formalize, has critical implications for policy design and enforcement.

In Egypt, following structural reforms, informal enterprises and informal employment became more common (Assaad et al., 2022; Rizk & Rashed, 2022). There have been ongoing calls to increase firm formality in Egypt, to expand the tax base and generate formal employment (Ali & Najman, 2016; Egyptian Center for Economic Studies, 2005; World Bank, 2014). The assumption underlying these calls is that informal firms, or at least subsets of them, can be formalized through either more intensive enforcement of existing rules or by moderately relaxing these rules. This implies that informal and formal firms are similar to each other in terms of productivity, employment, types of activities, and other characteristics, and that formalization would not result in the firms disappearing altogether.

This paper investigates which firm characteristics are associated with firm formality, and thus which informal firms might be susceptible to formalization. We use household enterprise data from Egypt to compare the characteristics of micro and small nonagricultural establishments (which employ 22% of workers) by formality. The paper models firm formality and quantifies the potential for formalization based on the overlap in predicted probability between formal and informal firms. The paper then develops profiles (groups and clusters) of similar firms to identify types of informal firms with a higher potential for formalization.

Our paper builds on literature on the determinants of firm formality, interventions designed to formalize firms, and the costs and benefits of formalization. The determinants of firm formality vary across contexts, but formal firms are generally larger, more productive, male-owned,

and have more educated owners (Benjamin & Mbaye, 2012; Williams et al., 2016). Information and initial/fixed cost-reduction interventions have limited success in increasing formality (Benhassine et al., 2018; Bruhn & McKenzie, 2014; de Andrade et al., 2016; de Mel et al., 2013; Galiani et al., 2017), while enforcement and other incentives are somewhat more effective (Benhassine et al., 2018; de Mel et al., 2013; Fajnzylber et al., 2011). In some contexts, formalization improves subsequent firm performance, but in other cases, it does not (Ali & Marouani, 2020; Benhassine et al., 2018; de Mel et al., 2013; Demenet et al., 2016; Fajnzylber et al., 2011; McKenzie & Sakho, 2010; Rand & Torm, 2012). Although our focus in this paper is on the determinants of formality rather than the impact of formality on other outcomes, the disparate results in the literature on the impact of formality suggest that only for some firms do the benefits of formalization exceed their costs, making only a select segment of firms potentially formalizable (Benhassine et al., 2018).

Past research comparing formal and informal firms and the potential for informal firms to formalize has often focused on productivity (Benjamin & Mbaye, 2012; de Vries, 2010; Gelb et al., 2009). Only some research has also looked at other owner and firm characteristics as determinants of formality (Bruhn, 2013; Williams et al., 2016). We use data on household nonfarm micro and small enterprises (MSEs), with information not only on the firms but also their owners, to assess an array of potential predictors of firm formality. We also innovate and leverage our multivariate models to predict the probability of firm formality and use the overlap in that probability between formal and informal firms as a new metric of the similarities between informal and formal firms and an upward bound on potential formalizability.

Past research has sometimes used cluster analysis to depict microenterprises in terms of their characteristics (Cunningham & Maloney, 2001) or to analyze heterogeneity among informal firms (Echevin & Murtin, 2009). We build on this work by undertaking cluster and grouping analysis, based on important predictors of firm formality, to identify what firms may be amenable to formalization. Cluster and grouping analyses could potentially inform targeting and policy design to promote formality.

The analysis finds that, beyond firm size and whether the firm operates in a fixed establishment, the basic characteristics of firms that are easily observable for enforcement are not closely linked to formality. Firm age, productivity, and owner characteristics, such as education, are strongly predictive of formality. There is some overlap in the predicted probability of formality between formal and informal firms, suggesting some potential for formalization. However, many informal firms have a low probability of formality and may not survive formalization through more intensive enforcement of existing rules. Only 26% of informal firms have a predicted probability of formality of 50% or greater.

2 | CONTEXT

Since structural adjustment in the 1990s, the Egyptian labor market had a sharp decline in formal employment. The public sector shrank from 39% of employment in 1998 to 26% in 2018 (Assaad et al., 2022). The private sector was unable to replace the shrinking public sector with formal jobs. From 1998 to 2018 the share of formal private wage employment in total employment grew from 8% to just 12%, while the share of informal private employment grew from 53% to 62% (Assaad et al., 2022). Among household nonfarm enterprises, firm formality has also declined over time, from 55% in 1998 to 47% in 2018 (Rizk & Rashed, 2022). There have, therefore, been longstanding calls in policy circles in Egypt to formalize micro and small firms to

expand the tax base and increase access to formal employment (Ali & Najman, 2016; Egyptian Center for Economic Studies, 2005; World Bank, 2014).

The Egyptian government has been attempting to improve the regulatory and business climate facing MSEs for some time. Law 141 of 2004 designated the Social Fund for Development as the lead agency on MSE policy, and in 2017, these responsibilities shifted to the new Micro, Small, and Medium Enterprise Development Agency (MSMEDA) (Tawakol, 2021). MSMEDA, however, had a limited mandate in terms of promoting firm formalization. The new MSME Law (No. 152 of 2020) specifically targets informal firms and formalizing such firms (Shehata & Partners Law Firm, 2020). The law came into effect in late 2022. The law also provides definitions for enterprise sizes based on annual business revenue (microenterprises <LE 1 million, small 1–50 million, medium 50–200 million). Firms of all sizes are required to register and formalize but are allowed to get a temporary 5-year license during which they would not be subject to tax liability as they work on compliance. The law provides lower taxes for MSMEs, reduces costs of commercial registration, establishes an MSMEs service unit, and creates MSME-favorable land allocation.

Even prior to the 2020 MSME law, there had been substantial improvements in the procedures to start a new business (or to formalize an existing business) in Egypt. Although the cost of formally starting a business declined from 69% of income per capita in 2007 to 20% in 2020, it may still be prohibitive for small firms (World Bank, 2006, 2020). Similarly, the time required to start a business fell from 19 days in 2007 to 12 days in 2020. Egypt's rank in starting a business has accordingly improved from 165th out of 175 economies in 2007 to 90th out of 190 economies in 2020 (World Bank, 2006, 2020). However, other aspects of formality, such as enforcing contracts and registering property remained burdensome, with Egypt ranking 166th out of 190 economies in enforcing contracts, and 130th in registering property in 2020 (World Bank, 2020).

3 | DATA

3.1 | Data source

As we are interested in firm informality, we focus on a universe of private nonagricultural firms with fewer than 25 workers, but we situate this universe within employment in Egypt. This focus is because public entities are not at risk of being informal. Formality is not relevant for most agricultural activities, and firm informality among firms of 25 workers or more is rare (Assaad et al., 2022).[1] Given this universe, we use the Egypt Labor Market Panel Survey (ELMPS) conducted by the Economic Research Forum in collaboration with Egypt's Central Agency for Public Mobilization and Statistics. Four waves of this survey were carried out in 1998, 2006, 2012, and 2018.[2] We use the most recent (2018) wave. In Krafft et al. (2020), we compare ELMPS 2018 to ELMPS 2012, as well as to establishment censuses and surveys in Egypt and show that while the exact rates of formality vary, the characteristics of firms by formality are similar across sources. Results comparing ELMPS 2012 and 2018 are particularly similar, which is why in this paper we present results just for the 2018 wave.

The ELMPS is a nationally representative household survey that has tracked a panel of households while continually adding refresher samples (around 2000 additional randomly sampled households each wave) since 1998. Household attrition rates are relatively low across the

long span of the panel (23.5% from 1998 to 2006; 17.3% from 2006 to 2012; 15.0% from 2012 to 2018 (Krafft et al., 2021)). Particularly importantly, those engaged in nonwage work (household enterprises) were not differentially likely to attrite (Krafft et al., 2021). The ELMPS 2018 round, which we use here, has been validated against other representative labor market surveys and censuses in Egypt (Krafft et al., 2021).

The ELMPS includes a nonagricultural household enterprise module. Information on agricultural activities is collected in other modules that are not relevant to our purpose. Any household that reports a member who is self-employed, an employer, or an unpaid family worker in either their primary or secondary jobs must respond to either the nonagricultural or agricultural enterprise module. Households can report on multiple enterprises. There were 2204 enterprises with fewer than 25 workers[3] in 2111 households in the 2018 wave.[4] For each enterprise, there are questions about enterprise characteristics, household members' participation, employment of nonhousehold members, and net revenues.

3.2 | Measuring formality

We follow the 15th International Conference of Labour Statisticians guidance defining the informal sector as household (unincorporated) enterprises without registration (International Conference of Labour Statisiticians, 1993). Therefore, we use a yes/no question asked of all household enterprises in the ELMPS on whether they have a commercial/industrial registration to measure formality.[5]

3.3 | Firm characteristics

The costs and benefits of formality depend on firm characteristics. We categorize firm sizes into one worker (self-employment), two workers, 3–4 workers, or 5–24 workers.[6] The industry is categorized as manufacturing and related trades; construction; wholesale and retail; transportation and storage; accommodation and food service; various professional activities; and other services. The data have the calendar year the firm started. Given the heaping of responses, we organize age by decade for firms older than 7 years. We create a firm age variable that is single years of age from 0 to 7 years old, and then 8 to 17 years old, 18 to 27 years old, 28+, or do not know. We have information on location, which we present in terms of region and urban/rural.

We identify additional details about firm owners (the person who worked the most on the enterprise). We include the owner's sex, age group (<30, 30–49, and 50+), education (illiterate, less than secondary, secondary, higher education), and occupation (professional/technical, other white-collar, blue-collar craft, blue collar noncraft).

In addition to owner characteristics, we use firm characteristics from the ELMPS enterprise module. Specifically, we have current capital (categorically), the type of the workplace (e.g., workshop), and sole proprietorship (ownership by the household) versus collective ownership (partnership). Using the number of workers and the net revenues of the enterprise going to the household, we calculate net revenue/worker as a crude measure of productivity. We also use indicators for whether the firm hires relatives and nonrelatives from outside the household, since the costs to hire nonrelatives are likely to be much higher, particularly if such workers demand social insurance, which is dependent on firm formality.

4 | METHODS

4.1 | Situating our universe

We begin our analyses by situating our focus within employment in Egypt.[7] We then describe employment formality by segment; note that this is distinct from our focus on firm formality, both in terms of the denominator and since a formal firm can have both formal and informal employees.

4.2 | Predictors of formality

We then assess the correlates of formality for private nonagricultural establishments with 1–24 workers, using various models. A key focus of our research is investigating what factors predict formality and whether formal and informal firms are comparable. We run logit models to predict formality. In "Model 1," we include only a few firm characteristics that are readily observable for potential enforcement: firm size, type of workplace, region, and economic activity. With this model, we show that firms outside establishments (including those in private homes) have an extremely low likelihood of being formal and, therefore, focus on firms in fixed establishments. "Model 2" includes additional firm characteristics that are not readily observable but likely to shape the benefits and costs of formality and thus formality decisions: partnership, capital, log net revenue per worker, firm age, and employing (non-)relatives from outside the household. "Model 3" adds the owner's characteristics: sex, education, occupation, and age group.

4.3 | Comparing distributions of predicted probability of formality by formality status

We predict the probability of formality using each model. We compare the distribution of the predicted probability of formality by whether firms are formal. We take the size of the overlap in the distribution functions of predicted probability across formal and informal firms as a test and metric of whether informal firms are similar to formal firms and thus might potentially formalize. This measure of similarity between informal and formal firms based on the overlap of predicted probability is, in part, driven by the number of covariates included in the model. We also look at the proportion of informal firms in different quartiles of the distribution of predicted probability to ascertain the proportion of informal firms that are most eligible for formalization.

4.4 | Grouping analysis

We undertake grouping of firms, by characteristics that were particularly important in our regression models, as one approach to creating profiles of firms that vary in their potential to formalize. We explore the proportion of each group in the different quartiles of the distribution of predicted probability of formality. We undertake this work only for model 3.

We group on firm size, firm age, whether the firm hired a nonrelative outside of the household, quartile of the log of net revenue per worker, and owner's education level. We use a collapsed set of categories for these covariates to have a manageable number of resulting groups and reasonable cell size. We further collapse categories as needed.[8] After this exercise, we have 27 groups, all of which have 10 or more informal firms.

4.5 | Cluster analysis

As another approach to identifying profiles of firms with varying potential for formalization, we undertake cluster analysis. Cluster analysis is a data reduction and classification method, designed to create a small number of clusters (groups) that maximize similarity within the clusters and difference between them (Everitt et al., 2011). We use k-medians clustering and Gower's dissimilarity coefficient (since it is amenable to a mix of binary and continuous data) to create 10 clusters.[9] We use the variables in model 3 for the cluster analysis.

After obtaining the clusters, based on the characteristics of these clusters, we identify the delineating features for each cluster and name clusters accordingly. We present descriptive statistics on categorical variables showing the ratio of each category within the cluster relative to the category overall in the sample. In addition, we standardize the continuous variables in each of the clusters relative to the mean and standard deviation of that variable in the sample.

5 | RESULTS

5.1 | Situating our universe of firms in total employment

We focus on nonagricultural private sector employment, in establishments with 1–24 workers. This is about a fifth of employment in Egypt (22%, Table S4). A quarter of employment is in the public sector (24%), another quarter is outside of establishments (24%) and a further fifth (19%) in agriculture. Finally, 11% of employment is in nonagricultural private firms in establishments that have 25+ employees. Table S4 also measures employment formality in terms of social insurance. Among nonagricultural private firms in establishments with 25+ workers, half (52%) of workers have social insurance coverage. The share is much lower, only 12%, among those working in nonagricultural private sector firms with 1–24 workers, and lower still for those employed outside establishments (8%) and in agriculture (2%). Note that many workers who lack social insurance may still be in firms that are formal, so this is an underestimate of workers who are working in formal firms. Hereafter, we focus on nonagricultural, private-sector employment in enterprises with 1–24 workers (initially both in and out of establishments) and turn to the firm as the unit of analysis.

5.2 | Descriptives of firm characteristics by firm formality

In our initial universe of household enterprises with 1–24 workers, 31% are formal and 69% informal. Table S5 presents the characteristics of firms and owners by formality. Enterprises tend to be small (63% are pure self-employment with only one worker), more so for informal than formal firms. About half are in the wholesale and retail sector (for both formal and

informal firms). Higher capital firms are more likely to be formal, along with older firms and firms that employ outside workers.

The type of workplace is strongly related to formality, with shops overrepresented among formal firms (59% compared with 26% for informal firms). Very few formal firms are outside of establishments, which include the workplace categories of own home, mobile worker, street and kiosk vendors, and vehicles. Only 10% of firms outside of establishments are formal compared with 49% of firms inside establishments (not shown). Women-owned firms are particularly likely to be informal, while more educated or older owners are more likely to have formal firms. The median monthly revenue per worker is 2500 pounds for informal firms and 5000 for formal firms (not shown).

5.3 | Modeling formality

5.3.1 | Predictors of formality

We estimate a series of increasingly rich logit models of predictors of formality. Model 1, shown in Table 1, is a relatively parsimonious model that includes the variables that are most easily observable (potentially observable to government inspectors). Model 2, shown in Table 2, includes additional firm-related characteristics. Model 3, also shown in Table 2, adds firm owner characteristics.

We present, in Table 1, odds ratios from a logit regression for the probability of being formal for firms inside and outside of establishments, as well as for all firms with 1–24 workers (Model 1). As expected, firm size is a strong predictor of formality in all three estimates. According to the pooled model, compared to one-person firms, the odds of being formal are 97% higher for two-worker firms, almost three times as high for 3–4 worker firms, and more than seven times as high for 5–24 worker firms.

Some industry groups have higher odds of formality compared with wholesale and retail, specifically transportation and storage and various professional activities, significantly so in the pooled model. Firms in urban Upper Egypt, in all models, and some other regions for firms outside establishments, are significantly more likely to be formal relative to firms in the Greater Cairo region. The type of workplace is strongly predictive of formality, with nearly every other workplace type, except workshops, significantly less likely to be formal than a shop, which is probably due to the greater visibility of shops to government enforcement agents compared with rooms in buildings or workshops. Categories of workplaces outside of fixed establishments, with their much lower visibility to potential enforcers, are much less likely to be formal (e.g., 97% less likely for a mobile worker), motivating our splitting of results and focusing on firms operating in fixed establishments hereafter.

We now turn to the results of Model 2 and Model 3 (Table 2). The results on the effect of the firm characteristics on the odds of formality do not change appreciably once we add the owner characteristics in Model 3, so for these characteristics we discuss both models together and then move to the results on owner characteristics in Model 3. First, we note that firm size plays the same steadily increasing and highly significant role we observed in Model 1, but adding additional regressors attenuates the effect of firm size.[10] The effects of industry become more mixed and lose significance as more regressors are added to the model. The effect of region on formality also becomes smaller when additional regressors are added, with only urban Upper Egypt being significant when compared with Greater Cairo. Odds ratios for type

TABLE 1 Odds ratios from logit "Model 1" of formality for private household enterprises with 1–24 workers by in and out of establishment status.

	All firms	Out of est.	In est.
Ent. size (one worker omit.)			
2	1.969***	1.171	2.190***
	(0.349)	(0.583)	(0.430)
3–4	2.960***	3.805**	2.854***
	(0.589)	(1.688)	(0.604)
5–24	7.466***	7.013**	8.694***
	(2.087)	(4.934)	(2.841)
Type of workplace (shop omit.)			
Own home	0.195***	6.321***	
	(0.060)	(3.315)	
Office/flat/building/rooms	0.363***		0.340***
	(0.083)		(0.081)
Workshop/factory	0.720		0.757
	(0.182)		(0.195)
Mobile worker	0.026***	0.409*	
	(0.008)	(0.173)	
Street and kiosk vendors	0.312*	1.113	0.846
	(0.182)	(0.684)	(0.536)
Transport based	0.064***		
	(0.023)		
Industry (wholesale/retail omit.)			
Manufacturing and related trades	1.155	1.154	1.145
	(0.279)	(0.549)	(0.308)
Construction	0.847	1.351	0.555
	(0.462)	(0.842)	(0.500)
Transportation and storage	3.596***	6.799***	0.528
	(1.365)	(3.257)	(0.488)
Accommodation and food service	1.003		1.159
	(0.292)		(0.385)
Various professional acts.	1.795*	0.896	2.020*
	(0.467)	(0.694)	(0.590)
Other service	0.739	0.541	0.799
	(0.192)	(0.421)	(0.224)
Region (Greater Cairo omit.)			
Alx. Sz C.	1.443	2.608	1.140
	(0.503)	(2.396)	(0.434)

TABLE 1 (Continued)

	All firms	Out of est.	In est.
Urb. Lwr.	1.292	4.658***	0.884
	(0.346)	(1.863)	(0.265)
Urb. Upp.	2.408***	3.641***	2.026*
	(0.631)	(1.424)	(0.619)
Rur. Lwr.	1.491	3.869**	1.032
	(0.386)	(1.595)	(0.304)
Rur. Upp.	0.838		0.793
	(0.217)		(0.239)
N	2152	944	1155
Pseudo *R*-squared	.227	.152	.0883
Predicted probability of formality for reference category	0.346	0.011	0.389

Note: Empty cells indicate types of workplaces that are not applicable or perfect predictors (Greater Cairo perfectly predicts failure (informality) outside of establishments, $N = 42$, dropped and rural Upper Egypt acts as reference; accommodation and food service perfectly predicts failure (informality) outside of establishments, $N = 11$).
*$p < .05$; **$p < .01$; ***$p < .001$.*Source*: Authors' calculations based on ELMPS 2018.

of workplace do not change much as more regressors are added and are consistent with those obtained in Model 1 (in establishments).

In model 2, we consider firm characteristics that are not necessarily readily observable to government inspectors. Firms with collective ownership are more likely to be formal relative to sole proprietorships. Contrary to expectations, having higher levels of invested capital does not have a significant effect on the odds of formality, except that not knowing the amount of capital increases the odds of being formal (with the result being significant only in model 2). Higher log net revenue per worker is associated with higher odds of being formal, but the direction of causality is unclear. There are only significant differences by firm age for new firms, which are less likely to be formal than 1 year old firms. Although the expected pattern of older firms being more likely to be formal pertains, odds ratios are insignificant. Hiring relatives from outside the household is not associated with higher formality rates, but hiring nonrelatives is.

We now examine results on the characteristics of the firm owner included in Model 3. Having a female as opposed to a male owner is associated with lower odds of being formal but is insignificant. Firms with a more educated owner have substantially higher and often significant odds of being formal, with owners with higher education having 3.5-times higher odds of formality their firms than illiterate owners. Beyond education, occupation does not seem to have an additional effect, except for lower odds of formality for craft blue collar occupations. Finally, older owners are more likely to have formal firms than young owners, but insignificantly so.

5.3.2 | Are informal firms like formal firms?

Using the preceding models, we consider whether the predicted probabilities of firm formality overlap between formal and informal firms, and thus, which informal firms are similar to

TABLE 2 Odds ratios from logit "Model 2" and "Model 3" of formality for private household enterprises with 1–24 workers in fixed establishments.

	Model 2	Model 3
Ent. size (one worker omit.)		
2	1.774*	1.861*
	(0.435)	(0.470)
3–4	1.097	1.171
	(0.358)	(0.396)
5–24	2.551*	2.567*
	(1.106)	(1.162)
Industry (wholesale/retail omit.)		
Manufacturing and related trades	1.324	1.665
	(0.420)	(0.552)
Construction	1.106	1.680
	(0.970)	(1.442)
Transportation and storage	0.444	0.411
	(0.412)	(0.385)
Accommodation and food service	1.575	1.267
	(0.545)	(0.527)
Various professional acts.	1.769	1.181
	(0.562)	(0.427)
Other service	0.816	0.967
	(0.237)	(0.306)
Region (Greater Cairo omit.)		
Alx. Sz C.	1.182	1.472
	(0.461)	(0.597)
Urb. Lwr.	0.911	1.062
	(0.294)	(0.360)
Urb. Upp.	2.040*	2.076*
	(0.682)	(0.744)
Rur. Lwr.	1.353	1.460
	(0.440)	(0.502)
Rur. Upp.	1.033	1.309
	(0.339)	(0.451)
Partnership (sole proprietorship omit.)		
Collective Ownership	1.633*	1.892**
	(0.379)	(0.456)
Capital (none omit.)		
LE 1–499	0.628	0.597
	(0.329)	(0.315)

TABLE 2 (Continued)

	Model 2	Model 3
LE 500–999	0.692	0.661
	(0.368)	(0.386)
LE 1000–4999	0.970	0.724
	(0.434)	(0.337)
LE 5000–9999	0.801	0.660
	(0.354)	(0.306)
LE 10,000–49,999	1.136	0.829
	(0.480)	(0.365)
LE 50,000 or more	1.509	0.999
	(0.667)	(0.461)
Do not know	3.124*	2.611
	(1.705)	(1.486)
Ln net revenue per worker	1.114**	1.125**
	(0.041)	(0.043)
Ent. age (1-year old omit.)		
0	0.245*	0.298*
	(0.137)	(0.168)
2	0.518	0.438
	(0.260)	(0.230)
3	0.521	0.513
	(0.299)	(0.315)
4	0.606	0.593
	(0.323)	(0.336)
5	1.479	1.408
	(0.909)	(0.794)
6	0.703	0.759
	(0.372)	(0.423)
7	1.059	1.290
	(0.578)	(0.758)
8–17	1.467	1.491
	(0.600)	(0.657)
18–27	1.033	1.165
	(0.449)	(0.549)
28+	1.894	2.055
	(0.875)	(1.055)
Do not know	2.228	2.127
	(1.131)	(1.215)
Relatives outside the household employed (no omit.)		
Yes	1.039	0.924
	(0.309)	(0.274)

(Continues)

TABLE 2 (Continued)

	Model 2	Model 3
Nonrelatives outside the household employed (no omit.)		
Yes	2.441**	1.960*
	(0.690)	(0.566)
Type of workplace (shop omit.)		
Office/flat/building/rooms	0.363***	0.354***
	(0.093)	(0.098)
Workshop/factory	0.668	0.757
	(0.186)	(0.221)
Street and kiosk vendors	0.869	1.087
	(0.449)	(0.600)
Sex (male omit.)		
Female		0.622
		(0.180)
Education (illit. omit.)		
Less than sec.		1.577
		(0.469)
Secondary		2.305**
		(0.654)
Higher education		3.472***
		(1.114)
Occupation (prof./tech. omit.)		
Other white collar		0.809
		(0.227)
Blue collar craft		0.417*
		(0.145)
Blue collar noncraft		0.742
		(0.270)
Age group (<30 omit.)		
30–49		1.151
		(0.311)
50+		1.244
		(0.393)
N	1149	1107
Pseudo *R*-squared	0.174	0.209
Predicted probability of formality for reference category	0.148	0.092

Note: For predicted probability of formality for reference category, log net revenue per worker set at zero.
*$p < .05$; **$p < .01$; ***$p < .001$.*Source*: Authors' calculations based on ELMPS 2018.

formal firms. The pseudo-*R*-squareds of the models are one metric of our ability to differentiate formal and informal firms; consistent with our finding that region and industry were poor predictors of formality, Model 1, which includes these, firm size, and workplace type, has a

pseudo-*R*-squared of 0.083 for firms in establishments. Including other characteristics such as the firm age and revenue per worker as well as employment of nonhousehold members in Model 2 increases it to 0.174 and including owner characteristics in Model 3 raises it to 0.209.

Consistent with the low pseudo-*R*-squared from Model 1, Figure 1 shows high overlap between predicted probabilities, ranging from a 0.53 overlap overall to 0.59 for firms outside establishments and 0.73 for those in establishments. Firms in establishments, which are the most likely to be formal, have the most overlap. As these are the easily observable characteristics, this suggests a sizeable "formalizable" group, but bodes poorly for the potential of policymakers to target which firms are likely to formalize based on readily observed characteristics, whether for incentives or enforcement.

Model 2 and Model 3 (shown in Figure 2) for firms in establishments show diminishing overlap as more characteristics are added, with the overlap falling from 0.73 in model 1, to 0.67 in Model 2, to 0.59 in Model 3. In Model 3, the mode of the distribution of predicted probability of formality is slightly below 0.2 for informal firms, but around 0.8 for formal ones. Given that there are a limited number of covariates within the model, and that the overlap includes some formal firms with a low probability of formality, this overlap is likely an overestimate of the share of firms that could easily formalize.

We further summarize these distributions in Table 3, where we present the quartiles of predicted probability of formality by model for informal firms. Probabilities above 50% could be

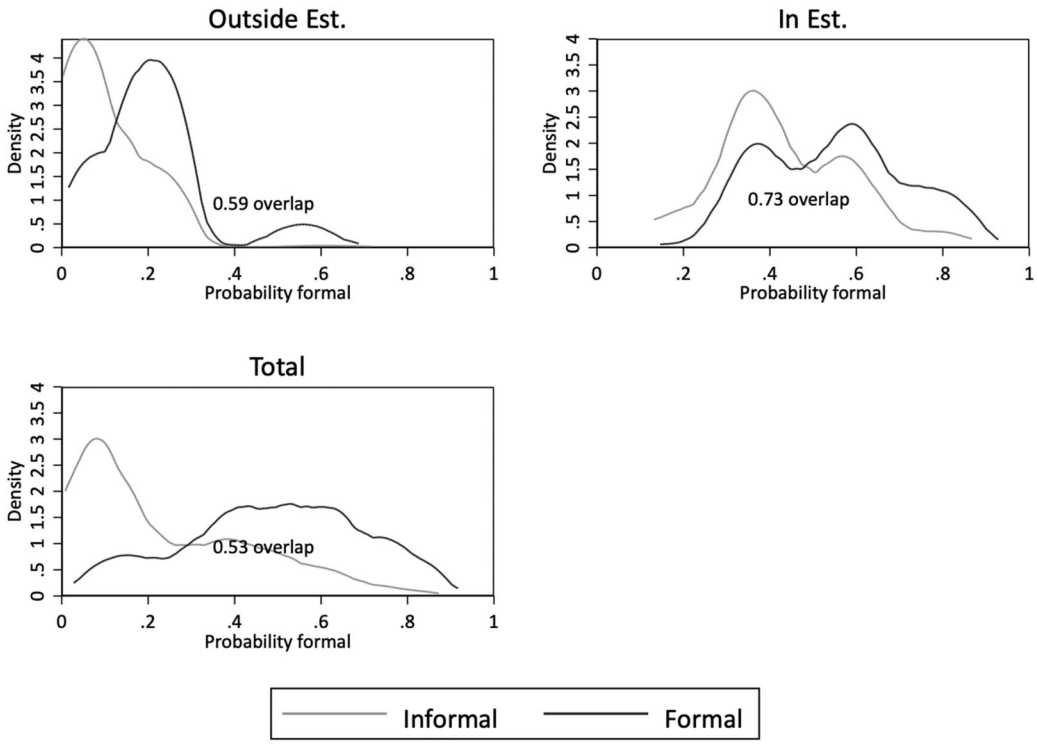

FIGURE 1 Model 1 predicted probability of formality, by actual formality and being in an establishment. *Source*: Authors' calculations based on ELMPS 2018. *Notes*: Kernel density functions with Epanechnikov kernel and bandwidth 0.05, predictions based on models in Table 1. Overlap denotes the proportion of the area beneath the curves that overlaps.

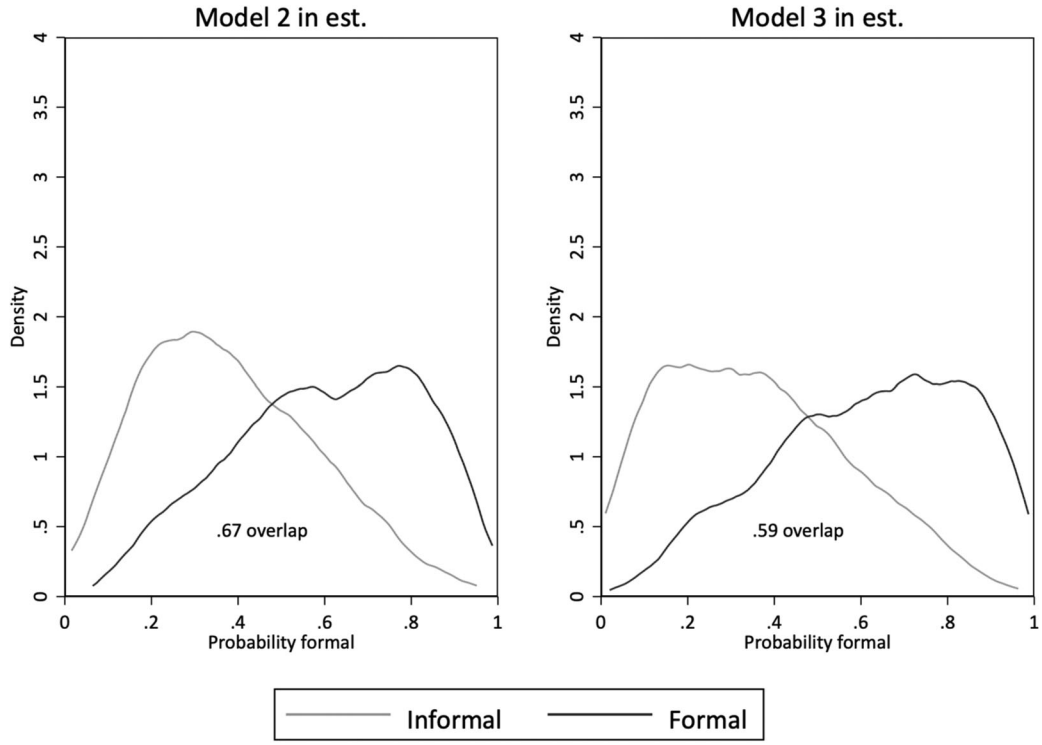

FIGURE 2 Model 2 and 3 predicted probabilities of formality, by actual formality, ELMPS 2018. *Source*: Authors' calculations based on ELMPS 2018. *Notes*: Kernel density functions with Epanechnikov kernel and bandwidth 0.05, predictions based on models in Table 2. Overlap denotes the proportion of the area beneath the curves that overlaps.

TABLE 3 Quartiles of predicted probability of formality by model for informal firms.

Predicted probability of formality	<25%	25% to <50%	50% to <75%	75%+
Model 1 pooled	62.6	24.8	10.9	1.6
Model 1 in est.	9.3	56.0	30.8	3.9
Model 1 out est.	95.1	4.0	0.9	0.0
Model 2 (in est.)	30.1	41.6	23.9	4.5
Model 3 (in est.)	36.6	37.5	21.2	4.7

Note: Predictions based on models in Tables 1 and 2.

construed as representing a high propensity for formalization. In Model 1, this represents 12.5% of all informal firms; with 0.9% of firms outside establishments meeting that condition and 34.7% of firms in establishments. In Model 2, which is restricted to firms in fixed establishments, 28.4% of informal enterprises are above the 50% probability threshold, and in Model 3, this comes down to 25.9%. We can, thus, conclude that while a miniscule proportion of informal firms outside fixed establishments have the potential to formalize, about a quarter of informal firms operating in fixed establishments have such potential.

5.4 | Profiles: Grouping and clustering analysis

In what follows, we group informal firms into different profiles (collections of characteristics that explain formality) based on Model 3. We identify which groups of informal firms are least likely and most likely to be formalizable based on the proportion that falls into each quartile of the distribution of the probability of being formal.

As shown in Table S7, the groups most represented in the bottom quartile of the probability of being formal are one person firms that are relatively young (<5 years of age) and have owners with a below secondary school education (groups 1–3 and 5). These kinds of informal firms have more than a 75% chance of being in the lowest quartile of the probability of formality. Older firms with one worker and a below secondary school owner in the bottom half of revenue per worker (groups 4 and 6) round out the types of firms with 75% or higher chances of being in the lowest quartile of predicted probability. The next group of firms (groups 7–9), with 61%–68% in the lowest quartile of predicted probability, consists of younger lower revenue firms, including two-person firms (with no nonrelatives hired), below secondary school owners, and younger one-worker firms with secondary educated owners. The following set (groups 10–14) is a mix of firms with less educated owners with higher revenues and 1–2 worker firms with secondary educated owners with lower revenues. None of these firms are ever in the top quartile of predicted probability and only some (0%–39%) are in the third quartile. Together, these groups of firms, which constitute 47% of the universe, are unlikely to formalize, and in some cases extremely unlikely.

Turning now to the groups of firms that have a higher chance of formalization, all are either larger (3–24 workers), older, employ nonrelatives, have higher revenue, or have secondary educated owners, and often have a combination of these characteristics. Groups 21–27, which make up 27% of the universe, have less than a 10% chance of being in the lowest quartile of probability. They are made up exclusively of older firms, predominantly larger ones. The small firms among them are either high revenue or hire outside workers and are mostly owned by individuals with secondary education or above. Among them, five groups (groups 21 and 24–27) have more than half their firms in the top two quartiles of predicted probability of formality.

The cluster analysis highlights additional variables that characterize the informal firms that have the lowest predicted probabilities of being formal. As shown in Figure 3, the analysis highlights clusters with a low probability of being formal that can be characterized[11] as "illiterate women-owned retail" (6% of the universe, 78% of which are in the lowest quartile of predicted formality) and "women-owned rural" firms (8% of the universe, 72% of which are in the lowest quartile of predicted formality). These are followed by a cluster we term "blue-collar construction," which also skews lower capital (12% of the universe, 62% of which are in the lowest quartile of predicted formality). All three of these clusters have below-average revenue. The next highest clusters are women-owned firms with high revenues, primarily in urban areas (5% of the universe), and blue-collar workshops, disproportionately in manufacturing (10% of the universe). These two disparate clusters are about half in the lowest quartile of probability and a third in the second quartile, with the remainder (14%) primarily in the third quartile of predicted probability.

The next two clusters of firms have secondary-educated, primarily male owners and are in rural Upper or Lower Egypt (together they constitute 29% of the universe). They also tend to be aged 30–49. Around a quarter of these firms are in the third quartile of predicted probability, but only 2%–4% are in the top quartile. The next two clusters have disparate profiles (blue collar with employees (4% of the universe) and educated high-revenue small (1–2 workers; 14% of universe)), but similar probabilities, with 10%–14% in the lowest quartile, 44%–48% in the second quartile, 35%–38% in the third quartile and 5%–7% in the top quartile of predicted probability.

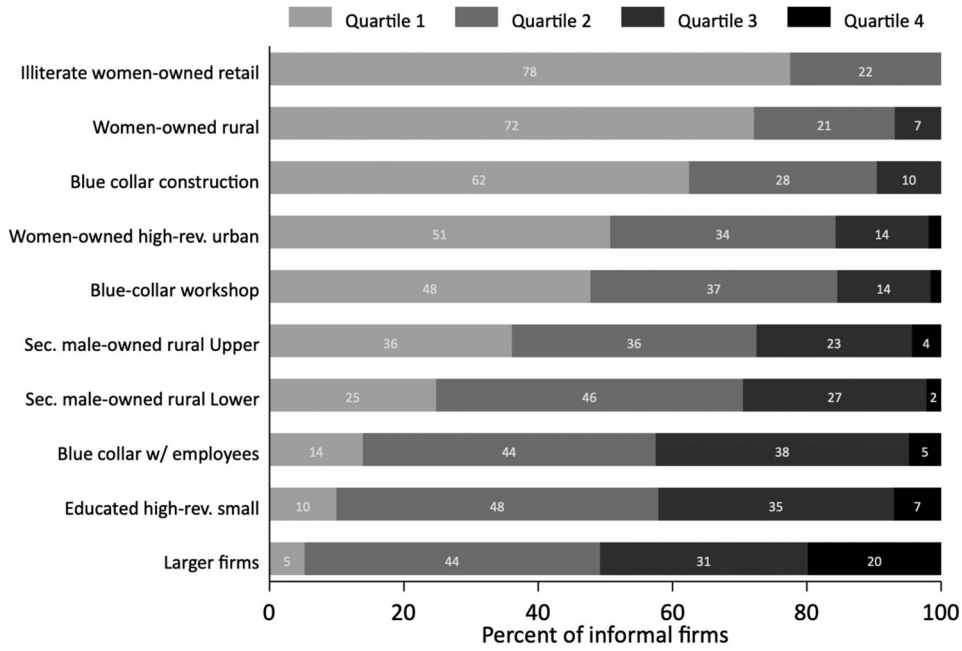

FIGURE 3 Percentage of informal firms in each quartile of predicted probability of formality by cluster, 10 clusters. *Source*: Authors' calculations based on predicted probability from model 3, see Table 2.

These clusters may be more amenable to formalization, although probabilities are still modest. The cluster that is most amenable to formalization is termed "larger firms" (12% of the universe). It includes almost no self-employment and is disproportionately made up of 5–24 worker firms. Only 5% of these firms are in the lowest quartile of predicted probability, 44% in the second quartile, 31% in the third, and 20% in the fourth. Although most amenable to formalization, given the characteristics we observe, we note that only slightly more than half of the firms in this cluster are above the 50% cutoff.

6 | DISCUSSION AND CONCLUSIONS

The informal economy plays a critical but complex role in low- and middle-income countries such as Egypt (La Porta & Shleifer, 2014). Egypt has struggled to create good, formal jobs for its young and growing labor force (Assaad et al., 2022; Gatti et al., 2014). As in other developing countries, there have been recommendations to formalize informal firms in Egypt in order to expand the tax base and increase the availability of good jobs (Egyptian Center for Economic Studies, 2005; World Bank, 2014). Yet such calls assume that firms that are currently informal are formalizable, that they could afford and survive formalization, without negatively impacting job creation.

6.1 | Summary

To answer whether informal firms may be formalizable, in this paper, we compared formal and informal private nonagricultural firms with 1–24 workers. We analyzed which

firm and owner characteristics predict formality. We then used the predicted probability of formality as a metric to quantify which informal firms may be amenable to formalization efforts.

Firms that operate outside fixed establishments were very unlikely to be formal. We thus restricted most of our analyses to firms in fixed establishments. Larger and older firms were more likely to be formal. Industry and region were not reliable predictors of formality. Firms that hired wage workers, especially nonrelatives, were more likely to be formal, as were the ones that were collectively owned or had higher net revenue per worker. Owners with more education were more likely to have formal firms.

When we compared the distributions of formal and informal firms in terms of their predicted probabilities of being formal, we found that there was some overlap. This result suggests that some informal firms are similar to formal firms and thus may be amenable to formalization. The overlap was larger when there was a limited set of observables to predict the probability of formality. As the set of predictors expanded, the overlap between the two distributions declined. The reduced overlap means that, while we are better able to identify informal firms that look like some of their formal counterparts, there are fewer potentially formalizable firms.

We used two different methods to group or cluster informal firms based on their observable characteristics. The grouping analysis suggests that firms that are not amenable to formalization include those that are younger, have only one worker, have less-educated owners, or low revenues. Firms with the highest chances of formalization had one or more of the following characteristics: larger (3–24 workers), employed nonrelatives, had higher revenues, were older, or had owners with at least a secondary education.

Cluster analysis identified three clusters as least amenable to formalization, which we named "illiterate women-owned retail," "women-owned rural," and "blue collar construction." These clusters together make up 25% of informal firms. On the other end of the spectrum, we identified the profiles of firms that have a relatively high probability of formality, larger firms with more educated owners and high revenues.

What do our results imply about the determinants of formality and why they might matter? The benefits and costs of formalization are a key driver of formalization decisions (Ulyssea, 2020). That older and larger firms that hire wage workers are more amenable to formalization is consistent with the fact that formalization imposes certain fixed costs and administrative hurdles (Ulyssea, 2020; World Bank, 2020) that younger and smaller firms are not able to afford. The association of formality with firm age also has two possible explanations: firms formalize later in the life cycle or informal firms are less likely to survive to old age. Moreover, requiring firms to renew their registration every 5 years and the costs that this imposes may contribute to higher rates of informality.

Higher productivity is associated with greater formality, but the direction of causality is unclear. More productive firms are better able to afford the costs of formality, but it is also possible that formality leads to higher productivity. Globally, overlap in productivity between formal and informal firms varies substantially by country context (Benjamin & Mbaye, 2012; Gelb et al., 2009). Having a more educated owner is more conducive to formalization presumably because it makes it easier to negotiate the complex bureaucracy of formality (Almaal, 2020; World Bank, 2020). Our results on the importance of firm and entrepreneur characteristics align with research from Pakistan that finds firm and entrepreneur characteristics matter more than the institutional climate (Williams et al., 2016).

6.2 | Limitations and directions for future research

Given the small number of informal enterprises observed still operating across multiple waves of the ELMPS panel, we do not present analyses of formalization dynamics. However, future research should prioritize creating and analyzing more frequent panel data from sufficiently large samples of MSEs to be able to assess formalization dynamics and their determinants. Ensuring accurate measurement of formality and addressing potential social desirability bias or strategic answers (e.g., responding out of concern for potential legal and tax implications) is an important area for future work, particularly for panel data, where measurement bias will be exacerbated when considering formalization dynamics.

6.3 | Policy implications

Our results have important implications for current debates about extending formality to a broader range of firms in Egypt and elsewhere. Any push toward formalization should carefully target firms that may be able to afford it. In Egypt, this would be somewhat larger and older informal firms that hire wage workers and that operate in fixed establishments. Our estimates suggest that about 15%–30% of informal firms operating in fixed establishments are in groups or clusters where half or more of the firms are in the top two quartiles of probability of formality and, thus, potentially formalizable. Efforts to formalize firms should avoid firms for whom there are few formal counterparts, such as firms operating outside fixed establishments or young/small firms. This and other work suggests that such firms, along with those with younger or female owners, no hired labor, or older low-productivity firms may be more survival self-employment than growth-oriented entrepreneurship and may be hard to formalize (Krafft, 2016; Krafft & Rizk, 2021). It is also important to keep in mind that only some of those potentially formalizable firms are likely to formalize even with the renewed efforts to increase formality in the new MSME law or if any other efforts are rolled out.

Our work adds context to the literature on interventions working to formalize firms. This literature has generally found little success with information interventions or reducing the costs of formalization (Benhassine et al., 2018; de Andrade et al., 2016; de Mel et al., 2013; Galiani et al., 2017). Our results suggest that formal and informal firms are often quite different; differences that information or covering registration fees alone cannot bridge. Even favorable macroeconomic conditions may not be sufficient to induce formality. Panel data on the impact of trade on firm formality, as well as formal and informal employment, demonstrates that export shocks can increase employment in formal establishments, but do not increase firm formalization (Tanaka & Greaney, 2023).

Increased enforcement can lead to increases in formalization (de Andrade et al., 2016). However, such increased enforcement can have negative consequences on firm survival when targeting firms with a low propensity to be formal. In heterogeneity analysis of the impacts of formalization on profits, one study found formalization increased productivity for mid-sized firms but decreased it for smaller firms (McKenzie & Sakho, 2010). Similarly, panel and administrative data on formal firms hiring informal workers have demonstrated that increased inspections raise the probability of informal workers formalizing, but also increased separations, and at the firm level led to persistently lower overall formal employment (Samaniego de la Parra & Fernández Bujanda, 2024).

Whether it makes sense for firms to formalize, or for governments to push formalization, critically depends on the benefits and costs of formalization. In Egypt, a paper using instrumental variables demonstrated that formality had a positive effect on firm performance for a select group of firms (Ali & Marouani, 2020). However, in global experiments, formalization only sometimes led to improved performance, such that in some cases the costs of efforts to formalize were not worthwhile (Benhassine et al., 2018; de Mel et al., 2013; Fajnzylber et al., 2011). Panel data show receiving government support may affect formalization (Nguyen et al., 2014), but also that informal firms that formalize tend to pay higher taxes (Boly, 2020). A key implication of these findings is that targeting informal firms that are more like formal firms could increase the impact and lower the costs of formalizing firms (Benhassine et al., 2018). Our work comparing formal and informal firms shows how more-formalizable clusters or groups of informal firms might be identified.

The new MSME law in Egypt does not target specific firms for formalization. Our results suggest that larger (3–24 workers) firms that employed nonrelatives, had higher revenues, were older, or had owners with at least a secondary education had the highest chances of formalization. While most of these characteristics are hard to observe, we propose that the government use firm size as its primary criterion for targeting enforcement. The law includes provisions that may induce some of the larger, more formalizable firms to self-select into regularizing their status. Including additional firm characteristics, particularly on owners, nonrelative employees, and revenues, in the next Establishment Census might provide the government with better opportunities to target informal firms for formalization efforts.

The international literature has shown that benefits, such as cash incentives, lowering the ongoing costs of formality and the complexity of formalization processes, and linking formalization to business services, training, banking, and tax services can increase formalization (Benhassine et al., 2018; de Mel et al., 2013; Fajnzylber et al., 2011). The 2020 MSME Law helps reduce the ongoing costs of being formal, while enhancing the immediate benefits of formality, such as access to credit and land. While the costs of formal registration declined dramatically in Egypt from 69% to 20% of per capita income from 2007 to 2020 (World Bank, 2006, 2020), and the new law will likely further reduce these costs, they may still be prohibitive for newly formed small firms. While the law simplified the tax payment process, with modest lump sum amounts for microenterprises based on categories of annual turnover, it still depends on estimating revenue, which may be challenging. Other aspects of formality, such as enforcing contracts and registering property, remain prohibitively onerous and expensive.[12]

Reforms designed to lower the ongoing cost and complexity of formality for microenterprises in Brazil did increase formality there (Fajnzylber et al., 2011; Monteiro & Assunção, 2012). Particularly notably, reform-induced increases in formality also led to higher profitability and increased employment. A key channel for this effect was shifting production to a permanent location. Given the rapid growth of employment outside establishments in Egypt—the type of employment with the worst working conditions (Assaad et al., 2022)—reforms that lower costs in ways that shift employment into establishments could be positive. Conversely, increased enforcement of existing regulations, without changing the costs and benefits of formality, might drive employment out of establishments and worsen working conditions as well as other economic outcomes. Efforts to expand formality must, therefore, be designed to maximize the benefits to the economy and to workers and minimize the costs to firms.

ACKNOWLEDGMENTS

This paper was funded by the World Bank and prepared as a background paper for the second Egypt Systematic Country Diagnostic under the task management of Alvaro Gonzalez (Principal Economist, Social Protection and Jobs) and Nistha Sinha (Senior Economist, Poverty and Equity). We appreciate the team's comments on various drafts and the questions and comments of participants in the Strategic Discussion Series seminar. All findings, interpretations, and conclusions in this paper are those of the authors and do not necessarily represent the views of the World Bank. We thank Cassie Miller for her research assistance for the literature review.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The Egypt Labor Market Panel Survey data are publicly available from www.erfdataportal.com. The first author will make replication files in Stata available on her website (www.carolinekrafft.com).

ORCID

Caroline Krafft https://orcid.org/0000-0001-6906-9418
Ragui Assaad https://orcid.org/0000-0002-2608-5335
Khandker Wahedur Rahman https://orcid.org/0000-0003-4285-9190

ENDNOTES

[1] Workers can be informally employed in larger and even formal firms due to the high cost of employing workers formally. We focus in this paper on firm formality rather than employment formality. For a discussion of the effect of labor regulations on the formality of employment in Egypt, see Wahba and Assaad (2017).

[2] For more information on the data, see Krafft et al. (2021). The data are publicly available at www.erfdataportal.com.

[3] Almost all household enterprises detected in the ELMPS 2018 reported fewer than 25 workers: three report 26 and one 53.

[4] Since the ELMPS tracks a panel of households, it is also possible to track a panel of household enterprises in the data. In Krafft et al. (2020), we constructed such a panel. Informal firms often closed over time (62% in 2012–2018). As a result, there was too small a sample ($N = 295$) to reliably assess the predictors of firm dynamics over time for the 27% of informal firms that formalized. Moreover, similar shares of firms formalized as became informal (29% in 2012–2018), which is more difficult, suggesting that measurement error is exacerbated in the panel when trying to assess formality dynamics. We, therefore, do not present formalization dynamics in this paper.

[5] In Krafft et al. (2020), we tested adding payment of taxes to our current definition of formality. Differences in rates of formality were small (5 percentage points higher in establishments) and multivariate results similar.

[6] We generated the firm size categories based on the distribution in the data (see Table S5). Only 6% have 5–24 workers.

[7] Fixed establishments include workplaces such as shops, offices, workshops, or factories. Workplaces that are not fixed establishments include one's own home, a farm or construction site, the street or other open spaces, or a vehicle.

[8] We aggregate groups that have fewer than 10 informal firms (i.e., all variables have the same value except the one that is used for aggregation) using the following algorithm: (i) group quartiles of the log of net revenue per worker on the same side of the median (ii) aggregate all levels of owner's education; (iii) aggregate all quartiles of log of net revenue per worker; and (iv) aggregate across categories of hiring non-relatives from outside the household.

[9] There are a variety of different approaches in the literature to selecting the number of clusters and these different approaches yield different numbers of clusters (Chiang & Mirkin, 2009). The decision to estimate 10 clusters was arbitrary.

[10] In Table S6, we test the sensitivity of our models to our specifications of firm age and firm size, instead estimating linear firm size and quadratic firm age rather than the main model categories. Results are substantively similar.

[11] We assigned names to the clusters after examining their predominant characteristics. These characteristics are summarized in Table S8. Figures S4 and S5 present cluster analysis sensitivity analysis with 5 or 15 clusters instead of 10; results are broadly similar.

[12] Although the Egyptian government has tried to reform the property registration system in recent years, a large fraction (∼95%) of properties remain unregistered (Khalaf, 2021).

REFERENCES

Ali, N., & Marouani, M. A. (2020). Household enterprises: The impact of formality on productivity and profits. *Economic Research Forum Working Paper Series No. 1452*.

Ali, N., & Najman, B. (2016). Informal competition, firms' productivity and policy reforms in Egypt. *Economic Research Forum Working Paper Series No. 1025*. Cairo, Egypt.

Almaal. (2020). *The validity period of the commercial registration in Egypt and the conditions necessary for its extraction (Arabic)*. Almaal. Retrieved April 11, 2020. https://www.almaal.org/commercial-register#i-4

Assaad, R., AlSharawy, A., & Salemi, C. (2022). Is the Egyptian economy creating good jobs? Job creation and economic vulnerability from 1998 to 2018. In C. Krafft & R. Assaad (Eds.), *The Egyptian labor market: A focus on gender and vulnerability*. Oxford University Press.

Benhassine, N., McKenzie, D., Pouliquen, V., & Santini, M. (2018). Does inducing informal firms to formalize make sense? Experimental evidence from Benin. *Journal of Public Economics*, *157*, 1–14.

Benjamin, N. C., & Mbaye, A. A. (2012). The informal sector, productivity, and enforcement in West Africa: A firm-level analysis. *Review of Development Economics*, *16*(4), 664–680.

Besley, T., & Burgess, R. (2004). Can labor regulation hinder economic performance? Evidence from India. *The Quarterly Journal of Economics*, *119*(1), 91–134.

Boly, A. (2020). The effects of formalization on small and medium-sized enterprise tax payments: Panel evidence from Viet Nam. *Asian Development Review*, *37*(1), 140–158.

Bruhn, M. (2013). A tale of two species: Revisiting the effect of registration reform on informal business owners in Mexico. *Journal of Development Economics*, *103*(1), 275–283.

Bruhn, M., & McKenzie, D. (2014). Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, *29*(2), 186–201.

Chiang, M. M.-T., & Mirkin, B. (2009). Intelligent choice of the number of clusters in K-means clustering: An experimental study with different cluster spreads. *Journal of Classification*, *27*, 3–40.

Cunningham, W. V., & Maloney, W. F. (2001). Heterogeneity among Mexico's microenterprises: An application of factor and cluster analysis. *Economic Development and Cultural Change*, *50*(1), 131–156.

de Andrade, G. H., Bruhn, M., & McKenzie, D. (2016). A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *World Bank Economic Review*, *30*(1), 24–54.

de Mel, S., McKenzie, D. D., & Woodruff, C. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, *5*(2), 122–150.

de Vries, G. J. (2010). Small retailers in Brazil: Are formal firms really more productive? *Journal of Development Studies*, *46*(8), 1345–1366.

Demenet, A., Razafindrakoto, M., & Roubaud, F. (2016). Do informal businesses gain from registration and how? Panel data evidence from Vietnam. *World Development*, *84*, 326–341.

Echevin, D., & Murtin, F. (2009). What determines productivity in Senegal? Sectoral disparities and the dual labour market. *Journal of Development Studies*, *45*(10), 1707–1730.

Egyptian Center for Economic Studies. (2005). The case for formalization of business in Egypt. *ECES Policy Viewpoint No. 17*.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. John Wiley & Sons.

Fajnzylber, P., Maloney, W. F., & Montes-Rojas, G. V. (2011). Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program. *Journal of Development Economics*, *94*(2), 262–276.

Galiani, S., Meléndez, M., & Ahumada, C. N. (2017). On the effect of the costs of operating formally: New experimental evidence. *Labour Economics*, *45*, 143–157.

Gatti, R., Angel-Urdinola, D. F., Silva, J., & Bodor, A. (2014). *Striving for better jobs: The challenge of informality in the Middle East and North Africa*. World Bank.

Gelb, A., Mengistae, T., Ramachandran, V., & Shah, M. K. (2009). To formalize or not to formalize? Comparisons of microenterprise data from southern and East Africa. *Center for Global Development Working Paper Series No. 175*.

International Conference of Labour Statisiticians. (1993). Resolution concerning statistics of employment in the informal sector, adopted by the Fifteenth International Conference of Labour Statisticians (January 1993). International Labour Organization.

Khalaf, O. S. (2021). *Egypt: Real estate registration crisis—Policies & scenarios*. Egyptian Institute for Studies.

Krafft, C. (2016). Understanding the dynamics of household enterprises in Egypt: Birth, death, growth, and transformation. *Economic Research Forum Working Paper Series No. 983*. Cairo, Egypt.

Krafft, C., Assaad, R., & Rahman, K. W. (2021). Introducing the Egypt labor market panel survey 2018. *IZA Journal of Development and Migration*, *12*(12), 1–40.

Krafft, C., Assaad, R., Rahman, K. W., & Cumanzala, M. (2020). How do small formal and informal firms in the Arab Republic of Egypt compare? *World Bank Policy Research Working Paper Series No. 9423*. Washington, DC.

Krafft, C., & Rizk, R. (2021). The promise and peril of youth entrepreneurship in MENA. *International Journal of Manpower*, *42*(8), 1500–1526.

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

McKenzie, D., & Sakho, Y. S. (2010). Does it pay firms to register for taxes? The impact of formality on firm profitability. *Journal of Development Economics*, *91*(1), 15–24.

Mishra, A. (2022). What sustains informality? A study of the interactions between formal and informal sector firms. *The Journal of Development Studies*, *58*(7), 1403–1415.

Monteiro, J. C. M., & Assunção, J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. *Journal of Development Economics*, *99*(1), 105–115.

Nguyen, T., Verreynne, M. L., & Steen, J. (2014). Drivers of firm formalization in Vietnam: An attention theory explanation. *Entrepreneurship and Regional Development*, *26*, 574–593.

Rand, J., & Torm, N. (2012). The benefits of formalization: Evidence from Vietnamese manufacturing SMEs. *World Development*, *40*(5), 983–998.

Rizk, R., & Rashed, A. (2022). Trends and determinants of Women's entrepreneurship in Egypt. In C. Krafft & R. Assaad (Eds.), *The Egyptian labor market: A focus on gender and vulnerability* (pp. 174–196). Oxford University Press.

Samaniego de la Parra, B., & Fernández Bujanda, L. (2024). Increasing the cost of informal workers: Evidence from Mexico. *American Economic Journal: Applied Economics*, *16*(1), 377–411.

Shehata & Partners Law Firm. (2020). *S&P Overview of the New SME Laws*. September 11, 2020. Lexology. https://www.lexology.com/library/detail.aspx?g=964f9b2b-94b1-4b98-852e-f2ad1d52d712

Tanaka, K., & Greaney, T. M. (2023). Trade and employment in the formal and informal sectors: A natural experiment from Cambodia. *Journal of Asian Economics*, *90*, 101676.

Tawakol, F. (2021). Enhancing the NILEX growth in Egypt: A study of the relevant laws and regulations. *Journal of Law in the Middle East*, *1*, 100–115.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, *12*, 525–546.

Wahba, J., & Assaad, R. (2017). Flexible labor regulations and informality in Egypt. *Review of Development Economics*, *21*(4), 962–984.

Williams, C. C., Shahid, M. S., & Martínez, A. (2016). Determinants of the level of informality of informal microenterprises: Some evidence from the city of Lahore, Pakistan. *World Development*, *84*, 312–325.

World Bank. (2006). *Doing business 2007: How to reform*. World Bank.

World Bank. (2014). *Arab Republic of Egypt: More jobs, better jobs: A priority for Egypt*. World Bank.
World Bank. (2020). *Doing business 2020: Economy profile: Egypt, Arab*. *Rep. Doing Business*. World Bank.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Krafft, C., Assaad, R., Rahman, K. W., & Cumanzala, M. (2024). How do small formal and informal firms in Egypt compare? An analysis of firm characteristics and implications for formalization efforts. *Review of Development Economics*, *28*(3), 1292–1315. https://doi.org/10.1111/rode.13101