World Development 173 (2024) 106408

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Finance, informal competition, and expectations: A firm-level analysis ☆

Emanuele Brancati [a], Michele Di Maio [a,*], Aminur Rahman [b]

[a] Department of Economics and Law, Sapienza University of Rome (Italy), and IZA Institute of Labor Economics, Bonn
[b] Asian Development Bank, Manila, Philippines

ARTICLE INFO

JEL classification:
O16
E26
D84
D22

Keywords:
Finance
Informality
Competition
Expectations
MENA countries

ABSTRACT

This paper documents the link between finance and informal competition. Using longitudinal firm-level data, we show that formal firms that are more exposed to the competition of informal firms are less likely to apply for a bank loan. This result is not due to sample selection, omitted variable bias, or reverse causality, and it is robust to different econometric specifications, including the use of an IV strategy. As for the mechanism explaining our result, we show that firms more exposed to informal competition have worse expectations on future sales growth, which in turn are associated with a lower probability of loan application. Finally, we provide suggestive evidence excluding supply-side mechanisms that may explain heterogeneities in firms' access to finance.

1. Introduction

Finance plays a critical role in affecting firms' performance and has a positive impact on investment and employment. However, the connection between firms and banks is weak in most developing countries and it is a possible element contributing to the poor job creation of the private sector in such contexts (Amin, 2021; Bah & Fang, 2015; Betz, Ravasan, & Weiss, 2021). Most of the analyses on the determinants of firms' disconnectedness from the banking system have focused on the obstacles to the supply of credit (Ayyagari, Juarros, Martinez Peria, & Singh, 2021). In this paper, we explore the possibility that there is also a credit demand component explaining the low access to finance characterizing firms in developing countries.

One crucial determinant of credit demand is the firm's past, current, and future economic performance. In most developing countries, an important element affecting formal firms' economic performance is the competition of informal firms. A large informal sector is often a defining characteristic of these economies (Falco, Maloney, Rijkers, & Sarrias, 2015; Ulyssea, 2020), with formal and informal firms coexisting within the same sectors and producing similar products (Ulyssea, 2018). Under these conditions, informal competition can represent an important obstacle to formal firms' operations and to the proper functioning of the overall economy (Distinguin, Rugemintwari, & Tacneng, 2016; Rozo & Winkler, 2021).

This paper documents the link between these two common characteristics of developing countries, namely the existence of informality and the *disconnectedness* of firms from the banking system. Our empirical analysis is guided by a simple conceptual framework that allows us to study how informal competition may explain the disconnectedness between private formal firms and the banking sector. Informal competition may affect credit allocation through several alternative channels. In fact, by distorting the environment of the firm, informal competition potentially affects both the demand and supply of credit. In particular, we focus on the effect of informal competition on expected sales growth, and how the latter affects a firm's loan application, its access to finance, and ultimately credit allocation.

In our analysis, we use longitudinal firm-level data from a confidential version of the World Bank Enterprise Survey (WBES), which also provides information on the geo-localization coordinates for each firm. We restrict our sample to countries in the Middle East and North Africa (MENA) region. There are two main reasons for this choice. First, the characteristics of these countries make them particularly suitable for our analysis. The disconnectedness between the private sector and the banking system is a well-known feature of these economies (ERDB, EIB, & WB, 2016).

☆ We thank the Editor Jampel Dell'Angelo and two anonymous referee for their useful comments. We also thank Roberta Gatti and Asif Mohammed Islam for useful discussions on a preliminary version of this paper, and for comments Margaret McMillan, Valentina Peruzzi, and conference participants at the 2021 EBRD/EIB/World Bank joint Workshop on MENA countries. The findings, interpretations, and conclusions expressed in this paper do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. All errors are our own.

* Corresponding author.
E-mail addresses: emanuele.brancati@uniroma1.it (E. Brancati), michele.dimaio@uniroma1.it (M. Di Maio), amrahman@adb.org (A. Rahman).

https://doi.org/10.1016/j.worlddev.2023.106408
Accepted 21 September 2023
Available online 9 October 2023

At the same time, informality is large, and it is an important obstacle to the operation of formal firms, with the share of firms reporting to be severely affected by the competition of informal firms reaching 40% in some countries. Second, the WBES survey for the MENA region includes questions on the firm's expectations about its future economic performance. This information is the key element to document the mechanism explaining our main result.

Our analysis proceeds in two steps. First, we look at the effect of the (perceived) threat from informal competition on the firm's access to finance. We document that formal firms constrained by informal competition are significantly less likely to apply for a loan. This result is not due to sample selection, omitted variable bias, or reverse causality. It is also robust to the use of alternative estimation strategies, including various matching techniques and an instrumental variable (IV) approach that relies on firms' geolocalization. Next, we explore the possible mechanisms explaining our main result. We show that the negative effect of informal competition on the loan application of formal firms operates through a reduction in their expected future sales. To this end, we first document that firms reporting to be more exposed to informal competition have significantly more pessimistic expectations. Importantly, this effect is not driven by differences in realized past sales. Then, we show that expectations on future sales growth have a positive and significant effect on the firm's probability of applying for a loan. Taken together, these findings support the existence of a credit demand channel that helps explain the disconnectedness of private formal firms from the banking sector. The exposure to the competition of the informal sector worsens the expected growth opportunities of formal firms, which reduces their willingness to apply for credit. To corroborate our argument, we provide evidence ruling out supply-related mechanisms such as differences in the firm's characteristics (including creditworthiness) and in the loan's conditions, or the possibility that these firms prefer alternative sources of funding.

Our paper is related and contributes to three strands of research. The first is the literature on the determinants of firms' access to credit in developing countries. Several studies analyze how access to finance is linked to firms' characteristics (Beck, Demirgüç-Kunt, & Maksimovic, 2005; Betz et al., 2021) and emphasize the existence of obstacles to the supply of credit (Ayyagari et al., 2021; Banerjee & Duflo, 2014; Kersten, Harms, Liket, & Maas, 2017).[1] Our contribution is to provide evidence on the role of credit demand in explaining the low access to finance and the disconnectedness of firms from the banking system. More specifically, we focus on loan application – the very first step in the process of entering a credit relationship – and show how this is affected by the firm's (perceived) level of informal competition.

Second, our paper relates to the vast literature on the effect of the informal sector on the economy (La Porta & Shleifer, 2014; Maloney, 2004). Informality is a distinguishing characteristic of most developing economies, which impacts the behavior and performance of firms operating in the formal sector in various ways. Ulyssea (2018) shows that the coexistence and competition of informal firms with formal ones lead to a misallocation of resources and losses in total factor productivity. Moreover, a number of studies document that informal competition hurts formal firms in terms of output (Rozo & Winkler, 2021), employment (Amin, 2021), and innovation (Avenyo, Konte, & Mohnen, 2021).[2] Distinguin et al. (2016) show that having informal competitors makes formal SMEs more likely to be credit constrained, but this occurs only in countries with weak institutional environments. Our analysis contributes to this literature by showing that the impact of informality on the formal sector depends on the *perceived* threat that formal firms attribute to informal competition. This, in turn, has relevant effects on firms' expectations, borrowing choices, and investment decisions. As such, our paper provides a novel piece to the understanding of the effect of informality on the functioning of the formal economy in developing countries.

Finally, this paper is linked to the small but growing literature on the role of expectations in influencing firms' decisions. Most of this literature looks at expectations on macroeconomic variables (see, for instance, Coibion, Gorodnichenko, & Kumar, 2018) , while only a few studies consider the role played by the firm's expectations on its own future earnings. Among the latter, Gennaioli, Ma, and Shleifer (2016) show that planned and actual investments of US firms are predicted by expected sales, and Boneva, Cloyne, Weale, and Wieladek (2020) look at UK firms to show substantial effects of expectations on employment choices. Finally, Enders, Hünnekes, and Müller (2022) show that changes in expectations of German firms impact their real decisions, even if expectations turn out to be incorrect ex-post. Our paper is the first that, looking at expectations on sales growth for firms in developing countries, shows that they are influenced by the *perceived* level of informal competition and that this effect goes beyond differences in firms' fundamentals or realized performances.

The remainder of the paper is as follows. Section 2 presents the simple conceptual framework that guides our analysis. Section 3 describes the data. Section 4 presents the main results and the possible underlying mechanisms. Section 5 concludes and discusses some policy implications of our findings.

2. Conceptual framework

The allocation of bank credit is determined by the equilibrium between the demand and the supply of loans. On the demand side, a firm's need for external finance has two main motivations. Firms demand credit to finance expansion plans or new investment opportunities. Alternatively, firms may need external finance to cope with the (short-run) effects of negative shocks. On the supply side, banks provide credit to firms based on their evaluation of the riskiness of the borrower (given their information set) and choose the specific loan conditions depending on the firm's characteristics.

One important factor affecting the allocation of credit in developing countries is the level of informal competition faced by formal firms - a common feature of these economies. There is wide evidence that the (unfair) competition of informal firms represents an important obstacle to formal firms' activity by negatively affecting output, employment, and innovation. For instance, informal firms, by avoiding taxation and compliance with regulations (e.g., on labor), can undercut prices and eventually take away market share from formal firms. By distorting the operating environment of the firm, informal competition may potentially influence both the demand and supply of credit.

On the demand side, the effect of informal competition on the firms' need for credit is *a priori* not obvious. Higher competition and unfair pricing from the informal sector worsen firms' expectations about future sales. Yet, this may have two opposite effects on the firm's need for external finance. On the one hand, firms' expected gains are an essential driver of their investment decisions. The unfair pricing competition by informal firms reduces these gains and lowers the return on investments by decreasing, for instance, the value marginal productivity of capital. As a result of the lower desired level of investment, firms have less need for additional funds and apply less for loans. On the other hand, it is also possible that worse expectations about future sales ultimately lead to an increase in firms' credit demand. For instance, firms with poor sales prospects may demand more credit to accumulate a liquidity buffer and weather troubled times. The sign of the relationship between exposure to informal competition, expectations, and firms' demand for credit is, therefore, ultimately an empirical question.

[1] A companion literature uses randomized control trials to explore the effect of interventions alleviating micro-entrepreneurs' financing constraints (Crepon, Devoto, Duflo, & Parienté, 2015; de Mel, McKenzie, & Woodruff, 2008; Quinn & Woodruff, 2019).

[2] Conversely, some papers document a positive contribution of informal firms on overall economic activity in terms of employment and productivity growth (see, for instance, Diao, Kweka, & McMillan, 2018).

On the supply side, banks' lending decisions are based on the evaluation of the creditworthiness of the borrower. The latter may be influenced by the higher competition and unfair pricing of the informal firms. For instance, if the lender internalizes this information, it may revise upward a firm's expected riskiness in light of its worse future performance (which eventually leads to a lower probability of repaying its debt). This may lead to phenomena of credit rationing or less favorable credit conditions, i.e. higher interest rates, smaller loans, or larger collateral requirements.[3]

In the following, we bring these predictions to the data and test how informal competition affects the demand and supply of credit and whether a possible mechanism explaining these changes is the worsening of the firm's expectations on future sales.

3. Data

Our main source of data is the World Bank Enterprise Survey (WBES). The WBES is a firm-level dataset constructed from a standardized and globally comparable survey administrated by the World Bank in 153 countries. The original sample is representative of the population of privately-owned firms with at least five employees operating in the formal (non-agricultural) sector.[4] The survey is conducted face-to-face in different years and at different time intervals. The dataset is a repeated cross-section, but each wave also has a panel component, i.e. some firms are interviewed in more than one wave. One important feature of the version of the WBES dataset we have access to is that – differently from the publicly available one – it also provides information on the firm's geo-localization, which we employ in the construction of our instrument.

In our analysis, we restrict the WBES sample to the Middle East and North Africa (MENA) region. Our main sample thus includes data on formal firms for Egypt, Jordan, Lebanon, Morocco, Tunisia, West Bank and Gaza. This choice is motivated by the relevant characteristics of such countries for our research question. Indeed, while the MENA region has a comparatively deep banking sector, a large number of firms do not have access to bank credit (ERDB et al., 2016). The situation is characterized by limited use of bank finance, which is coupled with weak demand for bank funds (Akbas et al., 2022). About 37% of companies in the region do not have an outstanding loan or access to an overdraft facility, and finance their working capital and investment with internal sources only, i.e. they are financially autarkic. This is especially so in Egypt, Jordan, and West Bank and Gaza Strip (58%, 50%, and 47%, respectively).[5] Most importantly, this gap is even higher when considering voluntary autarkic firms (49%, 47%, and 41% compared to an average of 30% in Low and Medium-income countries, LMI), which implies that the disconnectedness from the banking sector is mainly due to a lack of firms' demand and not driven by credit rationing (as also documented by the low rejection rates, 3%). This is reflected in the substantially lower firms' application rates (16%, and even lower for Egypt, West Bank and Gaza Strip, and Jordan, 4%, 11%, and 13%, respectively) compared to LMI (19%).

Another distinguishing feature of the MENA region is the large size of the informal sector (ILO, 2018). The share of informal economic units (i.e., the sum of informal employers and own-account workers) in Africa and the Arab States (92.4% and 90.8%) is larger than the world average, and than that in emerging and developing countries or developed economies (80.9%, 82.5%, and 55.7%, respectively). Within Africa, the share of informal employers in Northern African countries (95.1%) is significantly above the world and continental averages (50.7% and 77%). This is also paired with a higher percentage of informal own-account workers (96.6% in North Africa vs. 86.1% in the rest of the world). Focusing on our sample of countries, we confirm that the share of informal units is extremely high, with the partial exception of the West Bank and Gaza Strip.[6] Taken together, these figures indicate that the disconnectedness between firms and the banking system and a large presence of informal firms are characterizing features of the MENA region.

Our analysis focuses on 2,227 firms for which we are able to match at least two consecutive waves of the survey, i.e. those firms that belong to the panel component of the sample. In Section 4.1.1, we explain why the estimation strategy requires restricting the analysis to such a sample.[7] We discuss possible sample selection issues below.

Our main measure of interest is *Loan application*, a variable indicating whether or not, at the time of the survey, the firm has applied for a loan or a credit line. The WBES also provides information on the firm's perceived exposure to the competition of informal firms. The survey asks "*to what degree practices of competitors in the informal sector are an obstacle to the current operations of the firm*".[8] The possible answers are: "no obstacle", "minor obstacle", "moderate obstacle", "major obstacle", and "very severe obstacle". We classify a firm to be constrained by informal competition if it declares such practices to be a "major" or "very severe" obstacle (the top two categories).[9] The resulting dummy variable (*Constrained by informal*) takes the value of one for firms perceiving informal competition as an important obstacle to their own activity, and zero otherwise. Crucially, this measure allows us to capture an idiosyncratic element influencing the firm's behavior that goes beyond the existence of an informal sector competing with the formal one.[10] Another important feature of the WBES is that it collects data on the firm's expectations on its future sales. The WBES reports

[3] It follows that informal competition may also affect the funding mix. If firms constrained by informal competition receive less bank loans, private loans to owners and managers may serve as a substitute for bank credit. Similarly, the existence of an overdraft facility may act as a substitution for formal loans from the banking sector, especially for smaller firms.

[4] Firms are selected using random sampling with three stratification levels to ensure representatives across firm size, sector, and subnational region (area).

[5] As a comparison, this share is 44% in Low-Medium Income countries and 31% in upper-middle-income ones. In high-income economies, the share of zero-leverage firms is much lower, e.g. 10% for US listed companies (ERDB, EIB, & WB, 2022).

[6] Note that data are not available for Lebanon. The share of informal employers for each country are: Jordan (94.1%), Tunisia (98.7%), Egypt (92.9%), Morocco (63.2%), and West Bank and Gaza Strip (35.6%). Finally, the share of informal of own account workers are: Jordan (98.5%), Tunisia (99.5%), Egypt (99.0%), Morocco (96.4%), and West Bank and Gaza Strip (48.7%).

[7] The number of survey waves for the MENA countries varies between two and three, from 2007 to 2020. For most countries, there are only two waves which include the panel component: 2013 and 2019 for Jordan, Lebanon, Marocco, West Bank and Gaza, Strip and 2013 and 2020 for Tunisia. The only exception is Egypt, for which there are three waves that include panel firms: 2013, 2016, and 2020.

[8] The rationale behind this question is that competition of informal firms differs from standard competition threats because informal firms operate under different rules and constraints than formal firms and this is likely to influence the characteristics of their competitive behavior (see Section 2).

[9] We combine these two answers because they both indicate that *practices of competitors in the informal sector* are a very relevant obstacle for the firm and there is no clear distinction between the two. The exact wording of the question used to construct this variable is reported in Table C.2. As a robustness check, we show that our findings do not change if we use as dependent the categorical version of this variable (see Section 4.1.1).

[10] This latter situation is captured by another question related to informal competition included in the WBES. The question reads: *"Does this establishment compete against unregistered or informal firms?"*. This measure of informal competition has been used by Distinguin et al. (2016). Our measure differs from that because – in addition to indicating the presence of informal competition – it also indicates the perceived severity of this threat. Our measure is also more suitable to capture the effect of informal competition on the demand for credit. Indeed, while banks may be aware of – and take into account in their decision – the presence of informal firms in the market, the firm can better evaluate

both an ordinal measure of these expectations (*Expected change in sales growth*: *Negative*, *Stable*, or *Positive*) and a continuous measure for firms' expected sales growth in the following year (*Expected value of sales growth*). Expectations on future sales are a critical piece of information for our analysis because it allows us to document a possible mechanism explaining our main findings.

The WBES data also provides a rich set of financial information. For each firm, it reports whether it has an outstanding loan or a credit line, if the bank has rejected its loan application in the past, and the reasons underlying its choice of not applying for a loan. These include (i) the lack of financial needs (an inverse proxy for credit demand); (ii) the excessive complexity of the application procedure; (iii) unfavorable interest rates offered; (iv) collateral requirements that were too high; (v) size or maturity of the loan that was insufficient/inadequate, and (vi) the firm expected that the loan application would be denied. Furthermore, the WBES reports whether the firm is financed through the owner's personal loans, if the firm has an overdraft facility, and the importance of trade credit in financing working capital. Finally, the survey provides information on a large number of firms' structural characteristics, including age (in log), size (log number of employees), exporting and importing status (two dummy variables for whether the firm is exporting or importing), innovativeness (a dummy for firms that introduced new or improved products or services), realized past sales (realized sales growth over the past three years), total number of competitors (the average number of self-declared competitors by *other* firms in the same sector and geographical area), and dummies for the sector and form of proprietorship. For each variable employed in the analysis, Table C.1 reports the corresponding question from the WBES survey and how this is used to define the variable and Table C.2 describes the type of variable obtained.

Descriptives. Table 1 presents descriptive statistics for the main variables employed in our analysis. Around 31% of the firms in our sample have applied for a loan (*Loan application*). Only 19% of firms have an outstanding bank loan or a credit line (*Loan availability*), confirming that most firms in MENA countries are disconnected from the banking sector and have low access to finance. The share of firms that applied and had their application turned down (*Turned Down*) is low and is around 5.6%. Among the reasons for not applying, about 60% of the firms declare that they do not need bank funds (*No need*), 6.6% point to the high-interest rate of the loan (*Interest*), 4.3% to the level of collateral requirements (*Collateral*), 6.6% indicate the complexity of the application procedure (*Complexity*), and less than 1% the insufficient size and the maturity of the loan offered (*Adequacy*). Only 1.3% of firms did not apply for a loan because they expected the request to be rejected (*Expected rejection*). Overall the share of firms rationed by banks (*Rationing*) is 11% (fully) and 16% (partially) (as measured using the methodology in Kuntchev, Ramalho, Rodríguez-Meza, & Yang, 2014). In our sample, 28% of firms regard the competition of the informal sector to be a major or very severe constraint to their activity (*Constrained by informal*). While the WBES only surveys formal firms, almost 89% of them started as unregistered businesses (*Originally informal*), which confirms the relevance of the informal sector in MENA countries.

Sample selection. Since our estimation strategy relies on the panel component of the WBES (see Section 4.1.1 for a discussion of this requirement), before proceeding, we discuss the possible selection issues affecting our estimating sample. The concern is that firms may self-select into the panel by accepting to be re-interviewed in a subsequent wave of the survey. This would imply that firms included in the panel (and in our sample) may be different vis à vis firms that – for various reasons – are not interviewed more than once (and thus are not included in the panel and our sample). Indeed, if such selection is correlated with our main regressors of interest and any of our dependent variables, it may create a bias driving our results. In Table A.1 of the Online Appendix, we tackle this issue by focusing on the sample of firms included in the first wave of each country's survey and test the correlation between the firm's likelihood of being interviewed a second time – i.e., being in the panel and thus in our sample – and the variables employed in the analysis. Our estimates assuage concerns about systematic sample selection bias by showing no correlation between the firm's probability of belonging to the panel, any of the firm's structural characteristics (except for *Age*) (column 1), and any of our main variables of interest (*Constrained by informal, Loan application, Loan availability*) (columns, 2–4). In addition, as a robustness check, we will take care of sample selection issues by employing Heckman-type estimators to deal with endogenous sampling selection (see Section 4.1.1 and Table A.2).

4. Empirical analysis

4.1. Loan application, informal competition, and firm characteristics

Our analysis begins by comparing the characteristics of firms constrained by informal competition (i.e., that perceive the competition of informal firms as a major or very severe constraint to their operations) with unconstrained firms (i.e., firms that do not).

Table 2 presents the conditional means of financial variables, past sales, and expectations on future sales for both groups of firms. Firms constrained by informal competition have a significantly lower probability of applying for a loan (*Loan application*) and of having an outstanding loan or credit line (*Loan availability*). Yet, firms in the two groups are not different in terms of their needs for funds (*No need*) nor their access to other sources of funding, such as overdraft facilities or owner's personal loans (*Overdraft* and *Personal loans*) . Finally, firms in the two groups do not have a different probability of being credit rationed (*Rationing: not rationed, partially rationed*, or *fully rationed*).

As for structural characteristics, while there is no difference in terms of age (*Age*) or international propensity (*Export* and *Import*), constrained firms are somewhat smaller (*Size*) and more concentrated in the manufacturing sector (*Manufacturing*). Importantly, constrained and unconstrained firms are not different in terms of the total number of competitors (*Number of competitors*).[11]

Firms constrained by informal competition do not have worse economic performance (*Past sales growth*) than unconstrained ones. Yet, one dimension along which constrained and unconstrained firms differ is *expectations* on sales growth: firms suffering from informal competition have significantly worse prospects for their future earnings (*Expected value of sales growth*, −3.15% vs +2.13%). This is also confirmed when we look at the categorical variable measuring the expected change in sales growth (*Expected change in sales growth*): constrained firms are significantly less likely to report a positive expected change and more likely to report stable or negative expected changes in future sales growth.

4.1.1. Regression analysis

We study the role played by the exposure to informal competition in the firm's decision to apply for a loan. Our baseline regression model is:

$$\text{Loan application}_{i,t} = \alpha + \beta\ \text{Constrained by informal}_{i,t-1} + \gamma^{\top} X_{i,t-1} + \theta_s + \mu_a + \lambda_t + \varepsilon_{i,t} \quad (1)$$

where $\textit{Loan application}_{i,t}$ is a dummy taking the value of one if firm i at time t has applied for a bank loan or credit line, and zero otherwise.

how much it is potentially affected by the presence of informal competitors. In any case, in Section 4.1.1, we show that when we use this alternative measure to construct our instrument, our results continue to hold .

[11] WBES does not provide the breakdown for the number of total competitors by formal and informal ones.

Table 1
Descriptive statistics.

Variable	Mean	St.Dev	Min	Max
Loan application	0.307	0.462	0	1
Loan availability	0.188	0.391	0	1
Turned down	0.056	0.231	0	1
No need	0.597	0.491	0	1
Reason for not applying				
Interest	0.066	0.248	0	1
Collateral	0.043	0.204	0	1
Complexity	0.066	0.249	0	1
Adequacy	0.007	0.082	0	1
Expected rejection	0.013	0.113	0	1
Rationing: not rationed	0.628	0.483	0	1
Rationing: partially rationed	0.162	0.368	0	1
Rationing: fully rationed	0.109	0.312	0	1
Account	0.828	0.377	0	1
Overdraft	0.314	0.464	0	1
Personal loans	0.075	0.263	0	1
Trade credit	0.074	0.181	0	1
Constrained by informal	0.283	0.450	0	1
Originally informal	0.891	0.312	0	1
Years of formality	3.151	0.573	0.693	5.094
Age	7.597	0.009	7.527	7.609
Size	3.391	1.440	0.693	9.048
Export	0.244	0.430	0	1
Import	0.522	0.500	0	1
Innovation	0.147	0.354	0	1
Manufacturing	0.575	0.494	0	1
Number of competitors	7.828	3.310	0.000	10.980
Manager elected	0.076	0.265	0	1
K investment	0.180	0.385	0	1
R&D investment	0.079	0.270	0	1
Past sales growth	−2.844	21.387	−85.852	99.738
Expected value of sales growth	0.004	0.257	−1.000	1.000
Expected change in sales growth: negative	0.268	0.443	0	1
Expected change in sales growth: stable	0.264	0.441	0	1
Expected change in sales growth: positive	0.468	0.499	0	1

Notes: descriptive statistics for the main variables used in the analysis. Table C.1 provides the description of the question corresponding to each variable as reported in the WBES questionnaire.

Table 2
Firms constrained by informal competition vs unconstrained firms.

Variable	Constrained by informal competition	Unconstrained by informal competition	Diff mean p-value
Loan application	0.259	0.327	0.002
Loan availability	0.155	0.198	0.022
No need	0.603	0.603	0.990
Overdraft	0.315	0.299	0.491
Personal loans	0.078	0.069	0.466
Turned down	0.044	0.051	0.491
Expected rejection	0.016	0.012	0.394
Rationing: not rationed	0.632	0.647	0.519
Rationing: partially rationed	0.159	0.145	0.432
Rationing: fully rationed	0.125	0.101	0.121
Age	2.794	2.772	0.903
Size	3.179	3.385	0.000
Export	0.261	0.257	0.843
Import	0.583	0.567	0.563
Manufacturing	0.619	0.569	0.000
Number of competitors	7.896	7.814	0.585
Past sales growth	−3.665	−3.739	0.875
Expected value of sales growth	−3.158	2.129	0.000
Expected change in sales growth: Negative	0.296	0.240	0.000
Expected change in sales growth: Stable	0.313	0.237	0.000
Expected change in sales growth: Positive	0.390	0.523	0.000

Notes: conditional means. In column 1, we report averages for the sample of constrained firms (i.e., firms that perceive the competition of informal companies as a major or very severe constraint to their operations), while in column 2 we report the averages for the subsample of unconstrained firms (i.e., firms that do not perceive the competition of informal firms as a major or very severe constraint). Column 3 reports the p-value of the t-test on equality of means. Table C.1 reports the question corresponding to each variable as provided by the WBES questionnaire.

Constrained by informal$_{i,t-1}$ is our measure of firm i's perception of informal competition as an important obstacle to its operations. It is a dummy variable that takes the value of one if the firm declares practices of competitors in the informal sector to be a "major" or "very severe" obstacle (the top two categories among the possible answers to this question), and zero otherwise . We use the lagged value of *Constrained by informal* to avoid time inconsistencies due to the fact that *Loan application*$_{i,t}$ is a choice potentially made by the firm before t. To account for this, all controls in the vector $X_{i,t-1}$ are also lagged once.[12] In our main specification, these are the firm's age, size, exporting and importing status, innovativeness, a dummy for the firm not needing a loan, a dummy for the firm having a bank account, the total number of competitors of the firm, and a dummy for managers appointed to political positions. θ_s, μ_a, and λ_t are, respectively, the sector fixed effects, the geographical area fixed effects (41 in total), and the year fixed effects. ε_{it} is the error term. Model (1) is estimated using logit. In all tables, we report White's heteroscedasticity-consistent standard errors, yet, the results are robust to alternative clustering choices.

Table 3 presents the estimates for model (1). Results indicate that firms constrained by informal competition are less likely to apply for a loan. This finding holds across various econometric specifications and samples. Column 1 shows the results when in our regression we only control for the sector, geographical, and time fixed effects: the coefficient of *Constrained by informal* is negative and highly statistically significant. In column 2, we re-estimate the model including an initial set of additional controls for the firm's structural characteristics. Results show that the coefficient of our explanatory variable is only slightly reduced in size and significance. The probability of loan application is significantly and positively associated with the firm's size (*Size*), while the coefficients for the other controls are largely insignificant.

In column 3, we further enrich the specification with two potentially important controls: *No need* (a dummy taking the value of one if the firm has not applied for a loan because it has no financial needs, and zero otherwise) and *Account* (a dummy taking the value of one if the firm has a checking or savings account, and zero otherwise). Results do not change. In column 4, we check that the effect of informal competition on loan applications is not simply capturing a higher degree of competition faced by the firm. When we augment the model with the variable measuring the total number of competitors of the firm (*Number competitors*), the coefficient of *Constrained by informal* is virtually unchanged. This suggests that the effect of the competition of informal firms on the decision of the firm to apply for a loan goes over and beyond the effect of competition *per se*. Finally, in column 5 we also include a dummy variable that takes the value of one if the owner/CEO/top manager/board member has ever been elected or appointed to a political position, and zero otherwise (*Manager Elected*). Results show that politically connected firms are more likely to apply for a loan. Yet, the inclusion of this additional control has practically no effect on the sign, magnitude, or level of significance of our main variable of interest.[13] Based on this specification, firms constrained by informal competition have 5.9 percentage points (pp) lower probability of applying for a loan, which is substantial considering that the average unconditional probability is 31 pp (see Table 1).

[12] The need to use the lagged values of the explanatory and the control variables in our regression is what forces us to restrict our analysis to the panel sample. See for a discussion of sample selection issues.

[13] Egypt is the only country in our sample for which the WBES provides data for three waves. In this case, we can exploit the longitudinal dimension of the data and employ a linear probability estimator with firm-fixed effects. While the sample is small (174 firms), results show that, even accounting for any observable and unobservable firm-level time-invariant characteristic (including any factor that may affect the firm-bank relationship), being constrained by informal competition reduces loan application of formal firms.

Robustness checks. We perform a number of exercises to assess the robustness of our results. A first possible concern with our analysis is the presence of a sample selection bias. Our estimating sample includes firms that have been interviewed at least twice, i.e., all and only those included in the panel sample of the WBES survey. In Section 3, we already showed that the firm's probability of belonging to the panel does not depend on any of our main variables of interest or any of the firm's characteristics, except for *Age* (see Table A.1.). To complement this finding, we explicitly take care of sample selection issues in our regression analysis by employing a Heckman-type estimator. Results reported in Table A.2 show that, even after accounting for the possible selection in our sample of systematically-different firms, our main result still holds: firms constrained by informal competition are significantly less likely to apply for a loan.

Our results are also robust to the use of a categorical version of our main explanatory variable. As discussed in Section 3, *Constrained by informal* is a dummy variable that takes the value of one if the firm declares the practices of firms in the informal sector to be a "major" or a "very severe" obstacle to its operations. These are the top two categories available as a response to this question in the survey. Table A.3 shows that our results do not change if we consider all the response categories separately: the more severe the (perceived) obstacle represented by competition of informal firms, the lower the firm's probability of applying for a loan. This is confirmed across various specifications: the baseline (column 1) and the full model including additional controls (column 2).

Our results do not depend on the clustering choice for the standard errors (see Table A.4). We have also tested alternative proxies for the *Number of competitors*. In Table A.5, we show that our results do not change if we use the self-reported number of competitors by firm i (column 1) or if we compute this variable by taking the raw count of firms interviewed in the WBES survey operating within the same sector and geographical area (minus the firm i) (column 2). Our findings are also robust to the inclusion of a large set of additional controls. Results are reported in Table A.6.

Column 1 shows that the lower probability of applying for a loan for constrained firms is not due to differences in their (past) economic performance, type of destination market, as well as management, ownership structure, and characteristics. Column 2 also indicates that our results are not driven by the geographical and economic characteristics of the firm's location. In column 3, we show that informal competition has an effect that goes over and above other obstacles to firms' operations considered in the WBES survey. Our main result is virtually unaffected by the inclusion of controls for a number of possible constraints to economic activity, including those that could influence the size of the informal sector — such as obstacles related to tax administration, labor regulation, and difficulties in obtaining business licenses and permits. Finally, we consider the possibility that constrained firms have some characteristics which make them less likely to have a connection with a bank. For instance, this would be the case if these are firms that have been operating formally for a shorter period of time. To account for this, we augment our baseline specification with a variable indicating whether the firm was originally operating informally and the number of years since the formal registration. Results reported in column 4 show that our main coefficient of interest is virtually unaffected and that these measures are largely insignificant.

As an additional check to the validity of our findings, we employ matching techniques to test whether our results are driven by systematic differences between firms that are constrained by informal competition and unconstrained ones. To this end, we implement two different estimators for the average treatment effect (ATT), one based on nearest neighbor matching with bias correction and the other based on radius matching. In computing the propensity score, we exploit the full set of firms' characteristics employed so far. Table A.7 reports the balancing properties of the procedure showing no difference in firms'

Table 3
Informal competition and loan application.

Dependent variable	Loan application$_t$				
	(1)	(2)	(3)	(4)	(5)
Constrained by informal$_{t-1}$	−0.0625***	−0.0486**	−0.0517**	−0.0506**	−0.0592**
	[0.0204]	[0.0205]	[0.0208]	[0.0209]	[0.0237]
Age$_{t-1}$		0.00791	0.00850	0.00873	0.00890
		[0.0143]	[0.0145]	[0.0145]	[0.0176]
Size$_{t-1}$		0.0400***	0.0361***	0.0365***	0.0300***
		[0.00672]	[0.00710]	[0.00712]	[0.00854]
Export$_{t-1}$		0.00442	0.00452	0.00402	0.0109
		[0.0225]	[0.0228]	[0.0228]	[0.0259]
Import$_{t-1}$		−0.00622	−0.00569	−0.00783	−0.00117
		[0.0233]	[0.0238]	[0.0239]	[0.0264]
Innovation$_{t-1}$		−0.0253	−0.0243	−0.0247	−0.0508*
		[0.0231]	[0.0233]	[0.0234]	[0.0260]
Account$_{t-1}$			0.0234	0.0243	0.0352
			[0.0252]	[0.0252]	[0.0319]
No need$_{t-1}$			−0.0351*	−0.0363**	−0.0342
			[0.0184]	[0.0184]	[0.0211]
Number competitors$_{t-1}$				−0.00499	−0.00163
				[0.00401]	[0.00438]
Manager elected$_{t-1}$					0.0987***
					[0.0316]
Sector FE	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit	Logit
Pseudo R2	0.181	0.198	0.199	0.200	0.212
Observations	2227	2089	2031	2031	1608

Notes: logit marginal effects. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

characteristics between the treated and the control group after the matching, thus reassuring about the success of the balancing algorithm. Table A.8 shows the estimated ATTs, which confirm the negative effect of informal competition on loan applications.

An instrumental variable strategy. To further take care of endogeneity, we re-estimate our model using an Instrumental Variable (IV) approach. Our IV strategy builds on the widely used "cell-average method", wherein the potentially endogenous variable for firm i is instrumented by its cell average across all other firms (with the exclusion of firm i). This is typically done at the country or sector level (see Amin & Soh, 2021; Distinguin et al., 2016; Fisman & Svensson, 2007). We refine this approach by exploiting the information on the geo-localization of each firm to compute a more precise proxy for the local-level competition threats represented by the informal sector.[14] In particular, we instrument the firm i's perception of informal competition with the share of other firms located in the 10 km radius around firm i and operating in the same sector that declare to be constrained by the competition of informal firms.[15] This measure captures an environmental component in the level of informal competition that is unrelated to the firm's specific characteristics, including its fundamentals, its past availability of bank funds, and its possibly idiosyncratic perception of the level of informal competition. Indeed, it is unlikely that the firm's lack of access to finance drives other firms' perception of the competition threats represented by informal firms, therefore addressing our primary concern about the possibility of reverse causality. Moreover, the granularity of our instrument still allows for the inclusion of sector and geographical area fixed effects, which account for the possible concentration of firms with similar characteristics in certain sectors and locations, as well as for potential heterogeneities in the local features of the banking sector.[16]

Table 4 column 1 presents the estimates of our baseline IV linear probability model. Results confirm our previous findings. Our instrument is positively correlated with the firm's probability of being constrained by informal competition, as shown by the first-stage estimate in the bottom panel. The 2SLS estimates indicate a negative and significant effect of informal competition on formal firms' loan applications (−34 pp). Robustness checks on the instrument are presented in the remaining columns of Table 4. In columns 2, 3, and 4 we show that results also hold if we use alternative buffers around the firm (5 and 25 km, respectively) or we directly employ the WBES sampling geographical areas to construct our instrument. Results hold also if we consider all the firms (interdependent from their sector) in the construction of the averaging cell (see column 5). Finally, in columns 6 and 7, we build the instrument employing the local-level proportion of formal firms reporting to be competing against informal firms.[17] As such, we capture

[14] The information on the firm's geo-localization comes from a (confidential) version of the WBES dataset described and used in Brancati, Di Maio, Gatti, and Islam (2022).

[15] More formally, consider firm i at time t. Define cell $c(i,t)$ as the intersection of the geo-localized operational area (defined accordingly to the chosen distance) and sector in which firm i operates. Call k the general firm in cell $c(i,t)$, with $k = 1, 2, \ldots, N_{c(i,t)}$. Our main instrument is computed as the proportion of firms reporting informal competition as a major constraint to their operations within the same cell: i.e., $Z_{i,t} = \frac{\sum_{k \neq i} \text{Constrained by informal}_{k,t}}{N_{c(i,t)}-1}$. Using a 10 km buffer, the average number of firms in each cell is 50 and the median is 23.

[16] *Geographical area* is the geographical subnational location used by the World Bank as a sampling stratum in the construction of the survey (41 areas in our sample of countries).

[17] See footnote 3 for the survey question related to this variable and for a discussion of how this measure differs from our main explanatory variable.

Table 4
Informal competition and loan application: IV estimates.

Dependent variable:	Loan application$_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constrained by informal$_{t-1}$	−0.344***	−0.264*	−0.360**	−0.710***	−0.218**	−0.330*	−0.372*
	[0.126]	[0.141]	[0.157]	[0.205]	[0.110]	[0.180]	[0.207]
Additional controls	Y	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	N	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y
Model	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Underidentification (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cragg-Donald Wald F	59.55	45.22	38.61	33.75	76.58	29.41	22.99
Stock-Yogo critical value	16.38	16.38	16.38	16.38	16.38	16.38	16.38
Observations	2011	2011	2011	1974	2011	2011	2011
First stage regression							
Instrument	0.918***	0.787***	0.717***	0.385***	0.769***	0.450***	0.408***
	[0.119]	[0.117]	[0.115]	[0.0663]	[0.0879]	[0.0828]	[0.0852]
Averaging variable	Constrained by informal	Constrained by informal	Constrained by informal	Constrained by informal	Constrained by informal	Compete w/ informal firms	Compete w/ informal firms
Averaging buffer (radius)	10 km	5 km	25 km	Area	10 km	10 km	10 km
Averaging sample	By sector	By sector	By sector	By sector	Pooled	By sector	Pooled

Notes: 2SLS estimates. Robust standard errors in brackets. *Instrument* is defined in footnote . *Additional controls* include all the covariates as in Table 3, column 3. Variables are defined in Table C.1. The bottom panel reports the first-stage estimates for our set of instruments. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

the local-level size of the informal sector independently from the firm-specific perception of the competition threats from informal firms. Also in this case, our instrument has a strong power and our main result continues to hold.

Summing up. Firms constrained by informal competition are less likely to apply for a loan. This result is not driven by sample selection bias, differences in firms' fundamentals (including performance and structural characteristics), or reverse causality. The negative impact of informal competition on loan applications also reverberates to the actual availability of bank funds for the firm. In Appendix B, we replicate the analysis presented so far using *Loan availability*, a dummy taking the value of one if the firm has an outstanding bank loan or credit line, as a different dependent variable. Our results consistently show that exposure to higher informal competition also reduces the probability of loan availability for formal firms. Taken together, these results suggest that the size of the informal sector is a contributing factor to the disconnectedness of formal firms from the banking sector.

4.2. Mechanisms

In this section, we explore the possible mechanisms explaining the negative effect of informal competition on the firm's decision to apply for a loan.

4.2.1. Expectations on future sales

Firms apply for a loan for the most diverse reasons, including financing expansion plans, improving production processes, and – more in general – taking advantage of investment opportunities. Because all these choices are driven by firms' expectations (Boneva et al., 2020; Enders et al., 2022; Gennaioli et al., 2016), any element that influences the firm's beliefs about its future economic perspective will also have an impact on its decision to apply for a loan. The degree to which the firm perceives the competition of the informal sector as an obstacle to its operations is one such element. Based on these observations, we argue that a possible mechanism explaining our main result is that more intense informal competition worsens the firm's expectations on future sales growth, which, in turn, have a negative effect on its decision to apply for a loan.[18]

To test for this mechanism, we proceed in two steps. To begin, we provide evidence of a link between informal competition and expected sales growth. To this end, we take advantage of the responses to a question introduced in the most recent wave of the WBES. Specifically, firms are asked whether sales growth for the following year is expected to be negative, stable, or positive, and to provide its expected value.

Table 5 reports the estimates of a multinomial regression in which the categorical variable for the expected change in sales (*Expected change in sales growth*) is regressed on our explanatory variable (*Constrained by informal*), controlling for sector, geographical, and time fixed effects, and for our full set of covariates.[19] Results in columns 1–3 indicate that firms constrained by informal competition are significantly more likely to report negative or stable expected sales growth and substantially less likely to expect an increase in sales for the following year. Column 4 shows that this finding continues to hold if we use the expected value of sales growth (*Expected value of sales growth*) as an alternative outcome variable. To account for any possible concerns about reverse causality, column 5 shows the 2SLS estimates when we employ the same IV strategy based on the "geo-localized" instrument discussed above. Results hold in this case too.[20] Finally, in column 6, we show that results do not change if we only consider firms belonging to the panel sub-sample.

[18] As noted, the sign of the relationship between a firm's financing decisions and expectations is *ex-ante* ambiguous. For instance, an alternative possibility is that the firm could apply for more credit if it believes it has poor sales prospects, as a buffer. Our results below exclude this possibility.

[19] Note that, in this regression, we are not forced to use the lagged variable of *Constrained by informal* as in our main model (1). In that case, we lag the explanatory variable to account for the possibility that the dependent variable (*Loan application*$_{i,t}$) is pre-determined (see the discussion in Section 4.1.1). In this case, there is no such concern. The dependent variable (*Expected change in sales growth*) is forward-looking (it refers to the following fiscal year), while the explanatory variable (*Constrained by informal*) refers to the current situation of the firm. Given the difference in the timing of the two variables, we do not need to lag the explanatory variable and thus to restrict our analysis to the panel subsample. This explains the larger sample in Table 5, columns 1–5. In any case, in column 6, we report the result for the panel subsample and show that our results continue to hold.

[20] We also implement a 2SLS estimation for the categorical variable of expected changes in sales. Results reported in Table A.9 confirm the findings obtained using the continuous variable showing that being exposed to more informal competition significantly reduces the positive expectations on future sales growth.

Table 5
Informal competition and expectations on sales growth.

Dependent variable	Expected change in sales growth (categorical)			Expected value of sales growth (continuous)		
	Negative (1)	Stable (2)	Positive (3)	(4)	(5)	(6)
Constrained by informal	0.0413*** [0.0123]	0.0524*** [0.0146]	−0.0937*** [0.0154]	−0.0417*** [0.00737]	−0.183** [0.0813]	−0.0322** [0.0141]
Past sales growth	−0.00473*** [0.000369]	−0.00138*** [0.000424]	0.00611*** [0.000440]	0.00234*** [0.000194]	0.00228*** [0.000205]	0.00120*** [0.000348]
Additional controls		Y		Y	Y	Y
Sector FE		Y		Y	Y	Y
Geographic area FE		Y		Y	Y	Y
Time FE		Y		Y	Y	Y
Sample		Full		Full	Full	Panel
Model		Multinomial logit		OLS	2SLS	OLS
Pseudo R2 (R2)		0.229		(0.307)	(0.253)	(0.366)
Observations		4313		4191	4190	1265
					First stage	
Instrument					0.389*** [0.0641]	
Underidentification (p-value)					0.000	
Cragg-Donald Wald F					36.882	
Stock-Yogo critical value					16.38	

Notes: multinomial logistic marginal effects (columns 1-3), OLS estimates (columns 4 and 6), and 2SLS (column 5). In columns 1-3, *Expected sales growth* is a categorical variable reporting the expected sales growth in the following year taking the values of −1, 0, and +1 in case of negative, stable, and positive expectations, respectively. In columns 4, 5, and 6 *Expected sales growth* is a continuous measure for firms' expected sales growth in the following year. In columns 1-5, the sample includes all the firms interviewed in the last wave of the WBES. In column 6, the sample is restricted to firms interviewed in the last wave of the WBES which belong to the panel sub-sample. In column 5, we employ the same IV strategy as in column 1 of Table 4. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 6
Expectations on sales growth, loan application, and loan availability.

Dependent variable:	Loan application$_t$ (1)	Loan availability$_t$ (2)
$\mathbb{E}(\widehat{\text{Sales growth}})_{t-1}$	0.810*** [0.309]	0.537** [0.234]
Additional controls	Y	Y
Sector FE	Y	Y
Geographic area FE	Y	Y
Time FE	Y	Y
Model	Logit	Logit
Pseudo R2	0.189	0.181
Observations	1464	1405

Notes: logit marginal effects. *Loan application* is a dummy taking the value of one if firm i at time t applied for a bank loan or credit line, and zero otherwise. *Loan availability* is a dummy taking the value of one if firm i at time t has a bank loan, and zero otherwise. $\mathbb{E}(\widehat{\text{Sales growth}})_{t-1}$ is a continuous variable measuring expected sales growth at time $t-1$. We construct the lagged value of expected sales growth by fitting the values retrieved from the estimated coefficients in Table 5, column 4. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 7
Expectations on sales growth and investment in physical capital or R&D.

Dependent variable:	K investment$_t$ (1)	R&D investment$_t$ (2)
$\mathbb{E}(\widehat{\text{Sales growth}})_{t-1}$	0.431* [0.252]	0.275** [0.120]
Additional controls	Y	Y
Sector FE	Y	Y
Geographic area FE	Y	Y
Time FE	Y	Y
Model	Logit	Logit
Pseudo R2	0.139	0.198
Observations	1428	1355

Notes: logit marginal effects. *K investment* is a dummy taking the value of one if firm i at time t had a positive investment in physical capital, and zero otherwise. *R&D investment* is a dummy taking the value of one if firm i at time t had a positive investment in R&D, and zero otherwise. $\mathbb{E}(\widehat{\text{Sales growth}})_{t-1}$ is a continuous variable measuring expected sales growth at time $t-1$. We construct the lagged value of expected sales growth by fitting the values retrieved from the estimated coefficients in Table 5, column 4. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Across all these various models, estimation procedures, and samples, results consistently indicate that firms perceiving to be more exposed to informal competition report a lower expected value for sales for the following year. Importantly, such effect goes over and beyond that of *realized* sales growth in the previous periods (*Past sales growth*), which we include as a control in all specifications.

As a second step in our argument, we provide evidence of a link between expected sales growth and loan application. To do so, we have to confront some data limitations. Because firms' expectations on sales growth have been collected only in the last wave of the WBES survey, there is only one such observation for each firm. This implies that we do not have the lagged value of the explanatory variable to be used in the regression of *Loan application*$_t$ on expected sales growth. Yet, as discussed in Section 4.1.1, this is needed to avoid a possible time inconsistency between the dependent and the explanatory variables in such a regression. To overcome this limitation, we recover the lagged value of expected sales growth by fitting the values retrieved from the estimated coefficients in Table 5.[21] This procedure provides us with $\mathbb{E}(\widehat{\text{Sales growth}})_{t-1}$, a variable measuring expected sales growth at $t-1$, which can be suitably used as an explanatory variable in a regression for *Loan application*$_t$.

Table 6 presents the estimation results. Column 1 shows that better expectations on sales growth at $t-1$ increase the probability that the firm has applied for a loan at t, controlling for a large set of covariates and for sector, geographic area, and time fixed effects. Column 2 shows

[21] In practice, we construct $\mathbb{E}(\widehat{\text{Sales growth}})_{t-1}$ by fitting the values retrieved from the estimated coefficients from Table 5, column 4.

Table 8
Informal competition, loan application rejections, and credit rationing.

Dependent variable	Turned down$_t$	Rationing$_t$		
		Not rationed	Partially rationed	Fully rationed
	(1)	(2)	(3)	(4)
Constrained by informal	0.00623	0.0198	−0.0112	−0.0194
	[0.0161]	[0.0264]	[0.0237]	[0.0221]
Additional controls	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit
Pseudo R2	0.231	0.103	0.142	0.119
Observations	1177	1436	1426	1455

Notes: logit marginal effects. The dependent variable is reported in the top row. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 9
Informal competition and reasons for not applying for a loan.

Dependent variable	Expected rejection$_t$	Interest$_t$	Collateral$_t$	Adequacy$_t$	Complexity$_t$	Discouraged$_t$
	(1)	(2)	(3)	(4)	(5)	(6)
Constrained by informal	0.00410	0.0111	−0.00311	−0.00255	−0.00234	0.00820
	[0.00662]	[0.0123]	[0.0112]	[0.00969]	[0.0139]	[0.0192]
Additional controls	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit	Logit	Logit
Pseudo R2 (R2)	0.198	0.0982	0.121	0.137	0.0724	0.0704
Observations	1704	1901	1721	863	1797	1976

Notes: logit marginal effects. The dependent variable is listed in the top row. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table 10
Informal competition and alternative sources of funding.

Dependent variable:	Personal loans$_t$	Overdraft$_t$	Trade credit$_t$
	(1)	(2)	(3)
Constrained by informal	−0.00179	−0.0141	−0.0207
	[0.0135]	[0.0208]	[0.0211]
Additional controls	Y	Y	Y
Sector FE	Y	Y	Y
Geographic area FE	Y	Y	Y
Time FE	Y	Y	Y
Model	Logit	Logit	Tobit
Pseudo R2	0.110	0.221	0.111
Observations	1830	1945	1889

Notes: logit and tobit marginal effects. The dependent variable is listed in the top row. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

that this is also the case when we consider *Loan availability*$_t$ as an outcome.

Taken together, results in Table 5 (documenting a link between being constrained by informal competition and expected sales growth) and in Table 6 (providing evidence of a link between expected sales growth and loan application) show that more intense informal competition reduces loan application (and loan availability) for formal firms via its negative effect on their expectations on future sales. As long as the firm's growth depends on the availability and use of external sources of funding, our results provide evidence indicating that the perceived threat from informal competition – by reducing this possibility – is detrimental to formal firms and to overall economic development.

One specific channel through which the worsening of expected sales may lead to a reduction in loan application is the induced decrease in investments.[22] As noted above, the latter may be due to the unfair pricing policies of informal competitors lowering the value marginal productivity of capital or the return on investments of formal firms. Table 7 documents that a deterioration in the expected sales growth reduces both investments in physical capital (column 1) and R&D (column 2). To directly test that this reduction in investments explains the lower use of external financing, we would need to estimate the effect of desired (or planned) level of investment on (future) loan applications. Unfortunately, the WBES does not provide the former variable. Yet, using our data, we can document that there is a strong positive correlation between loan availability or loan application and (realized) investment. Correlations shown in Table A.10 indicate that when investments are larger, it is more likely that the firm has a loan and or has applied for a loan in the past. While this exercise has important limitations due to the timing of the variables, the results provide evidence consistent with the idea that a reduction in investment is expected to lead to a reduction in the use of external funds.

4.2.2. Other possible mechanisms

Firm characteristics, credit rationing, and loan conditions. One alternative explanation for the lower loan application of firms constrained by informal competition is that their fundamentals influence the conditions at which banks offer credit (see Section 2). For instance, if these firms are riskier and less creditworthy, banks may choose to cut their credit, thus leading to rationing or to loans offered at unfavorable conditions. The bank's behavior on the supply side may, in turn, be internalized by the firm on the demand side by choosing not to apply for a loan, expecting that it would be rejected or granted at a high cost.

Table 8 column 1 shows that firms constrained by informal competition do not report higher rejection rates on previous loan applications

[22] We thank one anonymous referee for suggesting to explore this possibility and to conduct this additional analysis.

Table A.1
Sample selection.

Dependent variable:	The firm is included in the WBES panel			
	(1)	(2)	(3)	(4)
Size	0.00284	−0.00103	0.00209	0.00208
	[0.00544]	[0.00609]	[0.00550]	[0.00566]
Age	0.0265***	0.0331***	0.0262***	0.0286***
	[0.00785]	[0.00855]	[0.00785]	[0.00800]
Export	0.0114	0.0125	0.0113	0.0108
	[0.0156]	[0.0170]	[0.0156]	[0.0159]
Import	0.0140	0.00858	0.0130	0.0160
	[0.0180]	[0.0196]	[0.0181]	[0.0185]
Constrained by informal		−0.00762		
		[0.0153]		
Loan application			0.0130	
			[0.0139]	
Loan availability				−0.00285
				[0.0160]
Sector FE	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit
Pseudo R2	0.0606	0.0679	0.0607	0.0619
Observations	5302	4412	5302	5143

Notes: logit marginal effects. The estimating sample is composed of the entire set of firms interviewed in the 2013-wave of the WBES. The dependent variable is a dummy taking the value of one if the firm is included in the panel sample of the WBES (i.e., it is interviewed in the following wave), and zero otherwise. All regressors are timed at the beginning of period. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

(*Turned down*). This suggests that our main explanatory variable does not merely capture less creditworthy firms. This result is confirmed by columns 2–4 showing that such firms do not have a different probability of credit rationing.[23]

To explore in more detail the role of credit conditions in influencing loan application, Table 9 looks at the possible reasons why the firm does not apply for a credit line related to the conditions of the loan or the application process. To begin, column 1 shows that the lower loan application of firms constrained by informal competition is not due to a higher *expected* rejection rate (*Expected rejection*). This result excludes firms' riskiness (as perceived by the bank) as a factor explaining our main finding. Results reported in columns 2–5 show that constrained firms are not more likely to report – as a reason not to apply for a loan – unfavorable conditions on the interest rate (*Interest*), collateral requirements (*Collateral*), the size and maturity of the loan (*Adequacy*), or the complexity of the application procedure (*Complexity*).[24] In sum, as shown in column 6, constrained firms are not more likely to be discouraged borrowers (*Discouraged*).[25]

These results indicate that the loan conditions offered to firms constrained by informal competition are not worse than those offered to unconstrained firms. As long as these conditions reflect the bank's evaluation of the firm's creditworthiness, these results exclude that heterogeneities in firms' fundamentals as perceived by the bank are a likely explanation for our findings. These results are also at odds with the possibility that the lower loan application of firms constrained by informal competition reflects their anticipation of banks' unwillingness to lend to them because they are considered more fragile borrowers.

Substitution with alternative sources of funding. Another possible explanation for our results is that firms more exposed to informal competition substitute bank loans with other sources of funding. This would imply that such firms report a lower probability of loan application because they optimally choose a different composition of their funding sources, with this having no impact on their overall availability of funds. We explore this possibility in Table 10. We begin by testing whether being constrained by informal competition correlates with the use of personal loans (from CEO or managers) to finance the firm's activity . If firms constrained by informal competition are somewhat riskier, private loans to owners and managers may serve as a substitute for bank credit. Results shown in column 1 suggest this is not the case: constrained firms are not more likely to use personal loans. Similarly, the existence of an overdraft facility may act as a substitution for formal loans from the banking sector, especially for smaller firms. Again, no significant difference seems to emerge between constrained and unconstrained firms (column 2). Finally, column 3 shows that being constrained by informal competition does not have any effect on firms' usage of trade credit. Taken together, these results indicate that firms constrained by informal competition are not substituting bank loans with alternative sources of external funds. This implies that the negative effect of informal competition on loan application and loan availability ends up reducing the total amount of resources available to these firms for financing their operation.

5. Concluding remarks

This paper has documented a link between two defining characteristics of several developing economies, namely a large informal sector and the disconnectedness of formal private firms from the banking sector.

Our analysis focuses on how access to finance of formal firms is affected by the existence and practices of informal firms competing

[23] These results also exclude the possibility of adverse bank selection. If firms more exposed to informal competition are somewhat concentrated in banking relationships with worse institutions that are less willing to grant credit, we should observe a higher probability of rejection. As shown by these results, this is not the case.

[24] This last result suggests that constrained and unconstrained firms are unlikely to be different in their managers' degrees of financial literacy, which thus can be excluded as a possible explanation for our main finding.

[25] Following Betz et al. (2021), a firm is *Discouraged* if it does not apply for a loan because it expects that the application would have been rejected or because of the unfavorable loan conditions (high interest rates, high collateral requirements, insufficient size of loan and maturity, or the complexity of application procedures).

Table A.2
Informal competition and loan application: Heckman selection model.

Dependent variable	Loan application				
	(1)	(2)	(3)	(4)	(5)
Constrained by informal$_{t-1}$	−0.0597***	−0.0469**	−0.0494**	−0.0480**	−0.0555**
	[0.0201]	[0.0202]	[0.0206]	[0.0206]	[0.0236]
Age$_{t-1}$		0.0102	0.0106	0.0105	0.00961
		[0.0143]	[0.0144]	[0.0144]	[0.0176]
Size$_{t-1}$		0.0462***	0.0423***	0.0429***	0.0367***
		[0.00763]	[0.00798]	[0.00798]	[0.00959]
Export$_{t-1}$		0.00824	0.00781	0.00689	0.0151
		[0.0239]	[0.0242]	[0.0242]	[0.0271]
Import$_{t-1}$		−0.00884	−0.00813	−0.0114	−0.00568
		[0.0239]	[0.0246]	[0.0246]	[0.0270]
Innovation$_{t-1}$		−0.0253	−0.0235	−0.0243	−0.0538*
		[0.0241]	[0.0244]	[0.0244]	[0.0281]
Account$_{t-1}$			0.0148	0.0153	0.0209
			[0.0242]	[0.0242]	[0.0293]
No need$_{t-1}$			−0.0286	−0.0300	−0.0299
			[0.0194]	[0.0194]	[0.0222]
Number competitors$_{t-1}$				−0.00654*	−0.00321
				[0.00360]	[0.00401]
Manager elected$_{t-1}$					0.111***
					[0.0372]
Additional controls	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Model	Heckman	Heckman	Heckman	Heckman	Heckman
Selected	2132	2047	1992	1992	1578
Not selected	8475	8475	8475	8475	8475
Wald χ^2	620.90	653.68	641.67	646.01	543.15
Inverse Mill's ratio	0.0136	−0.0108	−0.0137	−0.0151	−0.00893
Observations	10607	10522	10467	10467	10053

Notes: Heckman selection model. In this table, we explicitly model the probability of being included in our analysis in a two-step Heckman-type selection model. The selection equation models the firm's probability of belonging to the panel (i.e., being interviewed in two consecutive waves of the WBES survey) depending on firms' age, size, and belonging cell (the intersection of sector and geographical area, excluded in the main specification). The inverse Mill's ratio is included as an additional regressor in the original specification (reported in the bottom panel). All regressors are timed consistently with previous analyses. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

against them. Using firm-level data from various waves of the WBES, we document that formal firms that are more exposed to the competition of informal firms are less likely to apply for a bank loan. We also provide suggestive evidence that a credit demand mechanism is at work: exposure to informal competition worsens the firm's expectations about its future sales growth, which, in turn, reduces its willingness to apply for a loan. Because the same finding also holds for loan availability, we interpret these results as suggesting that a demand-side factor – i.e., the perception of informal competition – contributes to explaining the disconnectedness of private formal firms from the banking sector. In this sense, our results indicate that informal competition has a negative impact on the overall economy by reducing the use of finance and, thus, limiting the investment possibilities of formal firms.

Our finding adds to previous evidence on the negative effect of informal competition on formal firms and the overall economy. Our analysis documents a novel reason why informality may end up hurting formal firms' performance, namely that informal competition increases the disconnectedness of formal firms from the banking sector and thus decreases their possibilities to exploit potential growth opportunities.

The findings of this paper have some direct policy implications. First, our results suggest that the financial disconnectedness of formal firms also has a credit demand component. This is a novel view on an important phenomenon common to several developing countries, especially in the MENA region. Based on our analysis, policies designed to solely increase the supply of credit are unlikely to have a large effect on the use of finance of formal firms, given that an important obstacle to this is on the demand side. Third, our results provide a novel justification for policies aiming at reducing informality. A smaller informal sector would benefit the overall economy by making formal firms more likely to apply for a loan and thus better able to take advantage of investment opportunities and expand production. Fourth, our results show that measures to support firms should be designed taking into account that the *perception* of a constraint is as important as the existence of an actual constraint in driving firms' behavior. As we document in our analysis, firms' decision not to apply for a loan is influenced by the perceived competition threat from informal firms, which is not necessarily correlated with a poorer (actual) performance of the firm or with the existence of an actual threat. Yet, the negative effects of this perception are real.

CRediT authorship contribution statement

Emanuele Brancati: Conceptualization, Methodology, Software, Data curation, Investigation, Writing – review & editing. **Michele Di Maio:** Conceptualization, Methodology, Software, Data curation, Investigation, Writing – review & editing. **Aminur Rahman:** Conceptualization, Reviewing & editing.

Declaration of competing interest

None.

Data availability

Data will be made available on request

Appendix A. Additional tables

See Tables A.1–A.10.

Appendix B. Informal competition and loan availability

In this section, we present the results obtained when we employ loan availability as an alternative dependent variable. Table B.1 presents the baseline specification, while Table B.2 progressively saturates the model with a rich set of additional controls. In Table B.3, we test alternative definitions of the dependent variable by requiring loans to be issued within the last ten, seven, five, two, or one year. In Table B.4, we focus on the subset of firms without a loan in $t-1$, or we explicitly control for past loans in our baseline specification. Finally, Table B.5 presents the Heckman selection model, Table B.6 shows the matching estimator, while Table B.7 reports our IV. All results consistently show that being exposed to higher informal competition reduces loan availability, thus providing additional support to the idea that exposure to the competition of informal firms is an important obstacle to firms' growth.

Appendix C. Data appendix

See Tables C.1 and C.2.

Table A.3
Informal competition and loan application: Categorical measure.

Dependent variable	Loan application$_t$	
	(1)	(2)
Constrained by informal$_{t-1}$: Minor obstacle	−0.0383	−0.0313
	[0.0275]	[0.0415]
Constrained by informal$_{t-1}$: Moderate obstacle	0.0265	0.0802**
	[0.0235]	[0.0374]
Constrained by informal$_{t-1}$: Major obstacle	−0.0490	−0.0663
	[0.0329]	[0.0452]
Constrained by informal$_{t-1}$: Very severe obstacle	−0.0676***	−0.0855**
	[0.0255]	[0.0399]
Additional control	N	Y
Sector FE	Y	Y
Geographic area FE	Y	Y
Time FE	Y	Y
Model	Logit	Logit
Pseudo R2	0.188	0.199
Observations	2237	2011

Notes: logit marginal effects and within estimator. Variables are defined in Table C.1. The excluded category is: *No obstacle*. *Additional controls* includes all the covariates as in Table 3, column 3. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table A.4
Alternative clustering of the standard errors.

Dependent variable	Loan application$_t$				
	(1)	(2)	(3)	(4)	(5)
Panel A:	Clustering: 2-digit sector				
Constrained by informal$_{t-1}$	−0.0642***	−0.0663***	−0.0734***	−0.0731***	−0.0753***
	[0.0160]	[0.0229]	[0.0206]	[0.0207]	[0.0232]
Panel B:	Clustering: Area				
Constrained by informal$_{t-1}$	−0.0625***	−0.0663**	−0.0734***	−0.0731***	−0.0753**
	[0.0136]	[0.0273]	[0.0279]	[0.0280]	[0.0301]
Panel C:	Clustering: 2-digit sector & Area				
Constrained by informal$_{t-1}$	−0.0642***	−0.0663***	−0.0734***	−0.0731***	−0.0753***
	[0.0194]	[0.0231]	[0.0231]	[0.0233]	[0.0249]
Sector FE	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit	Logit
Pseudo R2	0.182	0.189	0.190	0.190	0.212
Observations	2225	1471	1431	1431	1190

Notes: logit marginal effects and within estimator. Variables are defined in Table C.1. Standard errors in brackets are clustered at the 2-digit sector level (in Panel A), along geographical areas (Panel B), and at the intersection of sectors and areas (Panel C). *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table A.5
Alternative definition for the number of competitors.

Dependent variable	Loan application$_t$		Loan availability$_t$	
	(1)	(2)	(3)	(4)
Constrained by informal$_{t-1}$	−0.0813***	−0.0596**	−0.0905***	−0.0715***
	[0.0278]	[0.0236]	[0.0267]	[0.0222]
Number self-declared competitors$_{t-1}$	0.0001000		−0.00904	
	[0.00751]		[0.00665]	
Number counted competitors$_{t-1}$		0.0110		0.0112
		[0.0199]		[0.0187]
Age$_{t-1}$	0.00556	0.00852	0.0127	0.0176
	[0.0212]	[0.0176]	[0.0204]	[0.0164]
Size$_{t-1}$	0.0410***	0.0295***	0.0355***	0.0314***
	[0.0111]	[0.00848]	[0.00998]	[0.00752]
Export$_{t-1}$	−0.00667	0.0115	0.0158	0.0194
	[0.0346]	[0.0258]	[0.0296]	[0.0221]
Import$_{t-1}$	0.0268	0.0000232	0.0107	0.00813
	[0.0352]	[0.0262]	[0.0309]	[0.0231]
Innovation$_{t-1}$	−0.0481	−0.0506*	−0.0518*	−0.0497**
	[0.0317]	[0.0260]	[0.0281]	[0.0228]
Account$_{t-1}$	0.00646	0.0341	0.0373	0.0812**
	[0.0380]	[0.0319]	[0.0380]	[0.0334]
No need$_{t-1}$	−0.0419*	−0.0331	−0.0492**	−0.0425**
	[0.0253]	[0.0211]	[0.0232]	[0.0191]
Manager elected$_{t-1}$	0.177***	0.0991***	0.147***	0.115***
	[0.0428]	[0.0314]	[0.0343]	[0.0256]
Sector FE	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit
Pseudo R2	0.244	0.213	0.216	0.199
Observations	1086	1608	1031	1540

Notes: logit marginal effects. Variables are defined in Table C.1. *Number self-declared competitors* is the log-number of self declared competitors (by firm *i*) reported in the WBES survey. *Number counted competitors* is an alternative proxy for firms' competitors obtained as the log-count of firms in the same sector and region of firm *i* in the WBES survey. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table A.6
Informal competition and loan application: Additional controls.

Dependent variable:	Loan application			
	(1)	(2)	(3)	(4)
Constrained by informal$_{t-1}$	−0.0771**	−0.102***	−0.115***	−0.115***
	[0.0301]	[0.0371]	[0.0374]	[0.0375]
Past sales growth$_{t-1}$	−0.000811	−0.000636	−0.000791	−0.000787
	[0.000499]	[0.000564]	[0.000573]	[0.000585]
Investment$_{t-1}$	−0.0613**	−0.0690*	−0.0674*	−0.0634*
	[0.0304]	[0.0367]	[0.0366]	[0.0371]
Local market$_{t-1}$	0.0261	−0.00341	0.00313	0.00559
	[0.0453]	[0.0539]	[0.0549]	[0.0551]
National market$_{t-1}$	0.00228	0.0107	0.0241	0.0247
	[0.0387]	[0.0483]	[0.0482]	[0.0485]
Years manager experience$_{t-1}$	0.0000556	0.0000941	0.000952	0.000941
	[0.00116]	[0.00130]	[0.00136]	[0.00138]
Government ownership$_{t-1}$	−0.00154	−0.00157	−0.00168	−0.00173
	[0.00246]	[0.00228]	[0.00241]	[0.00242]
Listed company$_{t-1}$	0.00548	0.00530	0.0182	0.0190
	[0.0430]	[0.0503]	[0.0502]	[0.0506]
Sole proprietorship$_{t-1}$	−0.0282	−0.0633	−0.0627	−0.0585
	[0.0370]	[0.0427]	[0.0423]	[0.0430]
Partnership$_{t-1}$	−0.0104	−0.0658	−0.0625	−0.0623
	[0.0423]	[0.0521]	[0.0529]	[0.0534]
Ltd partnership$_{t-1}$	−0.0173	−0.0182	−0.00837	−0.00947
	[0.0439]	[0.0469]	[0.0471]	[0.0481]

(*continued on next page*)

Table A.6 (*continued*).

Dependent variable:	Loan application			
	(1)	(2)	(3)	(4)
$City1_{t-1}$		−0.0437 [0.0906]	−0.0407 [0.0901]	−0.0423 [0.0916]
$City2_{t-1}$		0.0338 [0.0895]	0.0329 [0.0902]	0.0322 [0.0917]
$City3_{t-1}$		−0.0122 [0.0746]	−0.0306 [0.0723]	−0.0308 [0.0730]
$City4_{t-1}$		0.0109 [0.0758]	−0.00747 [0.0759]	−0.00763 [0.0766]
Number of electric $outages_{t-1}$		0.00104 [0.00123]	0.00116 [0.00121]	0.00116 [0.00122]
Bribery $depth_{t-1}$		−0.000850* [0.000489]	−0.000813 [0.000504]	−0.000858* [0.000513]
Loss from $theft_{t-1}$		−0.000809 [0.00479]	−0.00135 [0.00434]	−0.00120 [0.00434]
Constrained by tax $administration_{t-1}$			0.000186 [0.000378]	0.000173 [0.000381]
Constrained by labor $regulation_{t-1}$			−0.000369 [0.000454]	−0.000352 [0.000458]
Constrained by licenses and $permits_{t-1}$			0.000283 [0.000362]	0.000271 [0.000363]
Constrained by $finance_{t-1}$			0.00110*** [0.000338]	0.00110*** [0.000341]
Constrained by $corruption_{t-1}$			−0.0361 [0.0336]	−0.0340 [0.0340]
Constrained by $crime_{t-1}$			−0.000177 [0.000454]	−0.000162 [0.000456]
Constrained by $electricity_{t-1}$			−0.000315 [0.000366]	−0.000321 [0.000373]
Constrained by $transport_{t-1}$			−0.000253 [0.000444]	−0.000247 [0.000446]
Years $informality_{t-1}$				−0.0521 [0.192]
Originally $informal_{t-1}$				0.00803 [0.0513]
Model	Logit	Logit	Logit	Logit
Time FE	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y
Additional controls	Y	Y	Y	Y
Pseudo R2 (R2)	0.200	0.178	0.195	0.187
Observations	1076	799	771	764

Notes: logit marginal effects. All regressors are lagged once. Unreported additional regressors follow the specification in Table 3, column 3, enriched with dummies. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table A.7
Balancing properties of the matching.

Variable	Unmatched	Mean			% Reduct.	t-test	
	Matched	Treated	Control	% Bias	Bias	t	$p>\|t\|$
Egypt	U	0.38732	0.66338	−57.5	–	−8.24	0.000
	M	0.41667	0.41288	0.8	98.6	0.09	0.930
Jordan	U	0.04577	0.03099	7.7	–	1.14	0.254
	M	0.04924	0.04924	0.0	100.0	0.00	1.000
Lebanon	U	0.23239	0.10986	32.9	–	5.02	0.000
	M	0.24242	0.24242	0.0	100.0	0.00	1.000
Marocco	U	0.07394	0.0493	10.2	–	1.52	0.128
	M	0.06439	0.05682	3.2	69.3	0.36	0.716
State of Palestine	U	0.03873	0.06056	−10.1	–	−1.37	0.170
	M	0.04167	0.04545	−1.7	82.6	−0.21	0.832
Tunisia	U	0.22183	0.08592	38.3	–	5.96	0.000
	M	0.18561	0.19318	−2.1	94.4	−0.22	0.825
Age	U	7.5972	7.5971	1.1	–	0.16	0.876
	M	7.5972	7.5961	12.4	−990.8	1.27	0.205

(*continued on next page*)

Table A.7 (*continued*).

Variable	Unmatched	Mean			% Reduct.	t-test	
	Matched	Treated	Control	% Bias	Bias	t	$p>\|t\|$
Size	U	3.1893	3.5172	−24.3	–	−3.27	0.001
	M	3.1889	3.1367	3.9	84.1	0.49	0.621
Export	U	0.31373	0.32042	−1.4	–	−0.18	0.860
	M	0.32474	0.33505	−2.2	−53.9	−0.22	0.830
Import	U	0.59314	0.57923	2.8	–	0.35	0.730
	M	0.60309	0.63402	−6.3	−122.3	−0.63	0.532
Innovation	U	0.25000	0.15669	23.3	–	2.98	0.003
	M	0.24227	0.23711	1.3	94.5	0.12	0.906
Account	U	0.88732	0.80986	21.7	–	2.96	0.003
	M	0.87879	0.84091	10.6	51.1	1.25	0.211
No need	U	0.54577	0.64366	−20.0	–	−2.88	0.004
	M	0.56818	0.60227	−7.0	65.2	−0.79	0.428
Manufacturing	U	0.59155	0.59155	0.0	–	0.000	1.000
	M	0.61364	0.54167	14.6	–	−0.82	0.094
Retail	U	0.09859	0.07042	10.1	–	1.49	0.136
	M	0.0947	0.08712	2.7	73.1	0.30	0.763
Other services	U	0.30986	0.33803	−6.0	–	−0.85	0.116
	M	0.29167	0.37121	−17.0	−182.4	−1.9	0.112
LLC	U	0.25	0.23803	2.8	–	0.40	0.691
	M	0.25	0.25758	−1.8	36.7	−0.20	0.842
Sole proprietorship	U	0.34155	0.38028	−8.1	–	−1.14	0.253
	M	0.33333	0.375	−8.7	−7.6	−1.00	0.318
Partnership	U	0.16901	0.14507	6.6	–	0.95	0.343
	M	0.16667	0.1553	3.1	52.5	0.35	0.723
Ltd Partnership	U	0.16197	0.14085	5.9	–	0.85	0.396
	M	0.17045	0.17045	0.0	100.0	0.00	1.000
Originally informal	U	0.84155	0.89296	−15.2	–	−2.24	0.025
	M	0.83712	0.81818	5.6	63.2	0.58	0.565
Past sales growth	U	−3.4194	−4.5228	5.9	–	0.82	0.414
	M	−3.7487	−3.9356	1.0	83.1	0.13	0.899
Local market	U	0.39789	0.45493	−11.5	–	−1.64	0.102
	M	0.39394	0.37121	4.6	60.2	0.54	0.592
National market	U	0.52817	0.44366	16.9	–	2.42	0.016
	M	0.52652	0.56439	−7.6	55.2	−0.87	0.383
Board of directors	U	0.5493	0.61408	−13.1	–	−1.88	0.060
	M	0.53788	0.56439	−5.4	59.1	−0.61	0.541
Years manager experience	U	25.884	23.604	19.5	–	2.76	0.006
	M	25.867	25.367	4.3	78.1	0.49	0.622
Government ownership	U	0.94014	0.37183	8.7	–	1.43	0.154
	M	0.35606	0.125	3.6	59.3	0.89	0.373
City 1	U	0.14437	0.09296	15.9	–	2.37	0.018
	M	0.14773	0.1553	−2.3	85.3	−0.24	0.809
City 2	U	0.29225	0.19155	23.6	–	3.48	0.001
	M	0.2803	0.32955	−11.6	51.1	−1.23	0.220
City 3	U	0.17958	0.14507	9.4	–	1.36	0.175
	M	0.17424	0.19318	−5.1	45.1	−0.56	0.575
City 4	U	0.30282	0.53239	−47.8	–	−6.69	0.000
	M	0.31818	0.26894	10.3	78.6	1.24	0.215

Notes: Balancing properties from radius matching (0.2 stdev) in Table A.8.

Table A.8
Matching estimator: Average treatment effect.

Outcome variable:	Loan application	
	Abadie and Imbens (2002) estimator (1)	Radius Matching (0.2 stdev) (2)
Constrained by informal$_{t-1}$	−0.107***	−0.0779***
	[0.0328]	[0.0254]

Notes: Average Treatment Effects (ATE) for *Constrained by informal* (i.e., our treatment variable). In the left panel, we perform the Abadie and Imbens (2011) estimator, while in the right panel, we employ radius matching with a 0.2-stdev caliper. Balancing properties are provided in Table A.7 of the Online Appendix. All regressors are timed consistently with previous analyses. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table A.9
Informal competition and categorical expectations on sales growth: IV.

Dependent variable	Expected change in sales growth (categorical)		
	Negative (1)	Stable (2)	Positive (3)
Constrained by informal	0.0284 [0.0552]	0.290*** [0.0708]	−0.319*** [0.0672]
Additional controls		Y	
Sector FE		Y	
Geographic area FE		Y	
Time FE		Y	
Sample		Full	
Model		IV Multinomial logit	
Pseudo R2 (R2)		0.229	
Observations		4313	
		First stage regression	
Instrument		1.389*** [0.0719]	

Notes: IV multinomial logistic marginal effects. *Expected sales growth* is a categorical variable reporting the expected sales growth in the following year taking the values of −1, 0, and +1 in case of negative, stable, and positive expectations, respectively. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table A.10
Investment in physical capital and R&D, loan availability, and loan application.

Dependent variable:	K investment$_t$ (1)	K investment$_t$ (2)	R&D investment$_t$ (3)	R&D investment$_t$ (4)
Loan availability	0.0489** [0.0212]		0.0303** [0.0136]	
Loan application		0.0341* [0.0193]		0.0367*** [0.0127]
Additional controls	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit
Pseudo R2	0.142	0.133	0.207	0.201
Observations	1624	1674	1552	1600

Notes: logit marginal effects. *Capital investment* is a dummy taking the value of one if firm i at time t had a positive investment in physical capital, and zero otherwise. *R&D investment* is a dummy taking the value of one if firm i at time t had a positive investment in R&D, and zero otherwise. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.1
Informal competition and loan availability.

Dependent variable	Loan availability$_t$				
	(1)	(2)	(3)	(4)	(5)
Constrained by informal$_{t-1}$	−0.0756*** [0.0191]	−0.0659*** [0.0191]	−0.0718*** [0.0192]	−0.0719*** [0.0192]	−0.0720*** [0.0222]
Age$_{t-1}$		0.00496 [0.0132]	0.00732 [0.0135]	0.00730 [0.0135]	0.0178 [0.0164]
Size$_{t-1}$		0.0402*** [0.00597]	0.0365*** [0.00633]	0.0364*** [0.00636]	0.0315*** [0.00758]
Export$_{t-1}$		0.00978 [0.0191]	0.0123 [0.0191]	0.0123 [0.0191]	0.0190 [0.0221]
Import$_{t-1}$		0.00359 [0.0206]	0.00717 [0.0208]	0.00741 [0.0209]	0.00844 [0.0233]
Innovation$_{t-1}$		−0.0169 [0.0204]	−0.0204 [0.0204]	−0.0203 [0.0203]	−0.0498** [0.0228]

(*continued on next page*)

Table B.1 (*continued*).

Dependent variable	Loan availability$_t$				
	(1)	(2)	(3)	(4)	(5)
Account$_{t-1}$			0.0397	0.0395	0.0818**
			[0.0250]	[0.0250]	[0.0333]
No need$_{t-1}$			−0.0411**	−0.0410**	−0.0425**
			[0.0165]	[0.0165]	[0.0191]
Number competitors$_{t-1}$				0.000633	0.00143
				[0.00371]	[0.00411]
Manager elected$_{t-1}$					0.116***
					[0.0259]
Sector FE	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Model	Logit	Logit	Logit	Logit	Logit
Pseudo R2	0.143	0.181	0.186	0.186	0.199
Observations	2131	2006	1954	1954	1540

Notes: logit marginal effects and within estimator. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.2
Informal competition and loan availability: Additional controls.

Dependent variable:	Loan availability			
	(1)	(2)	(3)	(4)
Constrained by informal$_{t-1}$	−0.107***	−0.121***	−0.129***	−0.132***
	[0.0261]	[0.0320]	[0.0330]	[0.0334]
Past sales growth$_{t-1}$	−0.0000911	−0.0000681	−0.0000710	−0.0000322
	[0.000448]	[0.000476]	[0.000479]	[0.000491]
Investment$_{t-1}$	−0.0318	−0.0291	−0.0360	−0.0353
	[0.0258]	[0.0294]	[0.0281]	[0.0285]
Local market$_{t-1}$	0.0203	−0.00236	0.000975	0.00136
	[0.0364]	[0.0425]	[0.0410]	[0.0411]
National market$_{t-1}$	−0.00179	−0.00312	0.0108	0.0101
	[0.0312]	[0.0380]	[0.0358]	[0.0360]
Years manager experience$_{t-1}$	0.000711	0.000790	0.00129	0.00132
	[0.000978]	[0.00110]	[0.00116]	[0.00118]
Listed company$_{t-1}$	0.0297	0.0351	0.0460	0.0463
	[0.0360]	[0.0404]	[0.0406]	[0.0413]
Sole proprietorship$_{t-1}$	−0.0187	−0.0640*	−0.0763**	−0.0779**
	[0.0320]	[0.0383]	[0.0369]	[0.0376]
Partnership$_{t-1}$	−0.0353	−0.0574	−0.0703	−0.0677
	[0.0404]	[0.0503]	[0.0501]	[0.0500]
Ltd partnership$_{t-1}$	−0.0193	0.00383	−0.00601	−0.00558
	[0.0364]	[0.0395]	[0.0389]	[0.0398]
City 1$_{t-1}$		0.0224	0.0225	0.0243
		[0.0793]	[0.0795]	[0.0801]
City 2$_{t-1}$		0.0735	0.0748	0.0799
		[0.0842]	[0.0855]	[0.0861]
City 3$_{t-1}$		−0.00462	−0.0241	−0.0208
		[0.0725]	[0.0703]	[0.0717]
City 4$_{t-1}$		−0.00869	−0.0214	−0.0191
		[0.0766]	[0.0776]	[0.0782]
Number of electric outages$_{t-1}$		0.0000313	0.0000237	−0.00000224
		[0.000644]	[0.000619]	[0.000624]
Bribery depth$_{t-1}$		−0.000343	−0.000386	−0.000346
		[0.000401]	[0.000400]	[0.000405]
Loss from theft$_{t-1}$		−0.000625	−0.00122	−0.00144
		[0.00495]	[0.00639]	[0.00628]

(*continued on next page*)

Table B.2 (*continued*).

Dependent variable:	Loan availability			
	(1)	(2)	(3)	(4)
Constrained by tax administration$_{t-1}$			0.0000602	0.0000629
			[0.000336]	[0.000336]
Constrained by labor regulation$_{t-1}$			−0.000386	−0.000402
			[0.000400]	[0.000407]
Constrained by licenses and permits$_{t-1}$			0.000575*	0.000567*
			[0.000308]	[0.000308]
Constrained by finance$_{t-1}$			0.000877***	0.000852***
			[0.000296]	[0.000302]
Constrained by corruption$_{t-1}$			−0.0130	−0.0148
			[0.0295]	[0.0299]
Constrained by crime$_{t-1}$			−0.000527	−0.000528
			[0.000421]	[0.000422]
Constrained by electricity$_{t-1}$			−0.0000522	−0.0000905
			[0.000310]	[0.000315]
Constrained by transport$_{t-1}$			0.000411	0.000399
			[0.000352]	[0.000355]
Years informality$_{t-1}$				−0.134
				[0.146]
Originally informal$_{t-1}$				0.00655
				[0.0455]
Model	Logit	Logit	Logit	Logit
Time FE	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y
Additional controls	Y	Y	Y	Y
Pseudo R2 (R2)	0.233	0.230	0.260	0.258
Observations	1032	756	729	725

Notes: logit marginal effects. All regressors are lagged once. Unreported additional regressors follow the specification in column 3 of Table 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.3
Informal competition and loan availability: Restricting the timing of the issuance.

Dependent variable:	Loan availability				
Issuance:	10 years	7 years	5 years	2 years	1 year
	(1)	(2)	(3)	(4)	(5)
Constrained by informal$_{t-1}$	−0.0886***	−0.0821***	−0.0765***	−0.0527***	−0.0464**
	[0.0230]	[0.0227]	[0.0222]	[0.0202]	[0.0224]
Model	Logit	Logit	Logit	Logit	Logit
Time FE	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y
Additional controls	Y	Y	Y	Y	Y
Pseudo R2	0.193	0.191	0.195	0.211	0.210
Observations	1370	1362	1352	1306	1066

Notes: logit marginal effects. This table replicates the analysis in column 3 of Table B.1, while restricting the availability of loans to an issuance occurring within the last 10, 7, 5, 2, or 1 year (respectively in columns 1, 2, 3, 4, and 5). All regressors are lagged once. Unreported controls follow the specification in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.4
Informal competition and loan availability: Role of past loans.

Dependent variable:	Loan availability	
	(1)	(2)
Constrained by informal$_{t-1}$	−0.0785***	−0.0906***
	[0.0252]	[0.0229]
Loan Availability$_{t-1}$		0.0981***
		[0.0222]
Model	Logit	Logit
Time FE	Y	Y
Geographic area FE	Y	Y
Additional controls	Y	Y
Sample	Loan Availability$_{t-1}$ = 0	All
Pseudo R2 (R2)	0.156	0.212
Observations	1073	1364

Notes: logit marginal effects. In column 1, we restrict the sample to firms with no loans in $t-1$, while in column 2, we exploit the full sample and enrich the baseline specification with past loans. All regressors are lagged once. Unreported controls follow the specification in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.5
Informal competition and loan availability: Heckman selection model.

Dependent variable	Loan availability				
	(1)	(2)	(3)	(4)	(5)
Constrained by informal$_{t-1}$	−0.0719***	−0.0612***	−0.0668***	−0.0662***	−0.0654***
	[0.0179]	[0.0180]	[0.0183]	[0.0183]	[0.0214]
Age$_{t-1}$		0.00273	0.00450	0.00443	0.00987
		[0.0126]	[0.0128]	[0.0128]	[0.0161]
Size$_{t-1}$		0.0435***	0.0397***	0.0399***	0.0346***
		[0.00676]	[0.00708]	[0.00709]	[0.00871]
Export$_{t-1}$		0.0172	0.0198	0.0195	0.0285
		[0.0213]	[0.0216]	[0.0216]	[0.0247]
Import$_{t-1}$		0.000397	0.00546	0.00430	0.000204
		[0.0213]	[0.0219]	[0.0219]	[0.0245]
Innovation$_{t-1}$		−0.0145	−0.0169	−0.0173	−0.0562**
		[0.0215]	[0.0218]	[0.0218]	[0.0257]
Account$_{t-1}$			0.0217	0.0219	0.0471*
			[0.0215]	[0.0215]	[0.0267]
No need$_{t-1}$			−0.0378**	−0.0384**	−0.0417**
			[0.0173]	[0.0173]	[0.0202]
Number competitors$_{t-1}$				−0.00237	−0.00166
				[0.00319]	[0.00363]
Manager elected$_{t-1}$					0.142***
					[0.0338]
Additional controls	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y
Model	Heckman	Heckman	Heckman	Heckman	Heckman
Selected	2055	1985	1934	1934	1532
Not selected	8475	8475	8475	8475	8475
Wald χ^2	388.64	452.46	448.26	448.94	384.66
Inverse Mill's ratio	−0.00263	−0.0275	−0.0276	−0.0281	−0.0331
Observations	10530	10460	10409	10409	9998

Notes: Heckman selection model. In this table, we model the probability of being included in our analysis in a two-step Heckman-type selection model. The selection equation models the firm's probability of belonging to the panel (i.e., being interviewed in two consecutive waves of the WBES survey) depending on firms' age, size, and belonging cell (the intersection of sector and geographical area, excluded in the main specification). The inverse Mill's ratio is included as an additional regressor in the original specification (reported in the bottom panel). All regressors are timed consistently with previous analyses. *Additional controls* includes all the covariates as in Table 3, column 3. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.6
Matching estimator: Loan availability.

Outcome variable:	Loan availability	
	Abadie and Imbens (2002) estimator (1)	Radius Matching (0.2 stdev) (2)
Constrained by informal$_{t-1}$	−0.0850***	−0.0631***
	[0.0278]	[0.0210]

Notes: Average Treatment Effects for Constrained by informal (i.e., our treatment variable). In the left panel, we perform the Abadie and Imbens (2011) estimator, while in the right panel, we employ radius matching with a 0.2-stdev caliper. Balancing properties are provided in Table A.7 of the Online Appendix. All regressors are timed consistently with previous analyses. Variables are defined in Table C.1. Robust standard errors in brackets. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table B.7
Informal competition and loan availability: IV estimates.

Dependent variable:	Loan availability$_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constrained by informal$_{t-1}$	−0.332***	−0.286**	−0.342**	−0.206	−0.133	−0.483***	−0.455**
	[0.113]	[0.125]	[0.141]	[0.158]	[0.0975]	[0.168]	[0.192]
Additional controls	Y	Y	Y	Y	Y	Y	Y
Sector FE	Y	Y	Y	Y	Y	Y	Y
Geographic area FE	Y	Y	Y	N	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y
Model	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Underidentification (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cragg-Donald Wald F	56.21	44.82	36.75	29.17	73.89	30.44	22.81
Stock-Yogo critical value	16.38	16.38	16.38	16.38	16.38	16.38	16.38
Observations	1951	1951	1951	1916	1950	1949	1951
	First stage regression						
Instrument	0.904***	0.792***	0.706***	0.368***	0.765***	0.463***	0.411***
	[0.120]	[0.118]	[0.116]	[0.0682]	[0.0889]	[0.0839]	[0.0861]
Averaging variable	Constrained by informal	Constrained by informal	Constrained by informal	Constrained by informal	Constrained by informal	Compete w/ informal firms	Compete w/ informal firms
Averaging buffer (radius)	10 km	5 km	25 km	Area	10 km	10 km	10 km
Averaging sample	By sector	By sector	By sector	By sector	Pooled	By sector	Pooled

Notes: 2SLS estimates. Robust standard errors in brackets. *Instrument* is defined in footnote . *Additional controls* include all the covariates as in Table 3, column 3. Variables are defined in Table C.1. The bottom panel reports the first-stage estimates for our set of instruments. *, **, *** indicate statistical significance at the 10%, 5%, and 1%, respectively.

Table C.1
Variable definition.
Source: WBES Questionnaire, various years.

Variable name	Survey question and variable definition
Loan application	Question bmk7: "What is the reason for not having a loan or line of credit at the moment?". Answer bmk7a: "Because this establishment did not apply for a loan or line of credit". Loan application = 0 if bmk7a = yes and 1 otherwise (even if k8 = yes).
Constrained by informal	Question e30: "To what degree are practices of competitors in the informal sector an obstacle to the current operations of this establishment?". Available options: no obstacle; minor obstacle; moderate obstacle; major obstacle; very severe obstacle. Constrained by informal = 1 if e30 = "major obstacle" or "very severe obstacle", and 0 otherwise.
Turned down	If k8 = no, question bmk7: "What is the reason for not having a loan or line of credit at the moment?". Answer bmk7b: "Because the last application for a loan or line of credit was turned down". Turned down = 1 if bmk7b = yes and 0 otherwise.
No need	Question k16: "Referring again to the last fiscal year, did this establishment apply for any loans or lines of credit?". If k16 = no, question k17: "What was the main reason why this establishment did not apply for any line of credit or loan?", answer k17a: "No need for a loan – establishment had sufficient capital". No need = 1 if k17a = yes, and 0 otherwise.
Account	Question k6: "Now let's talk about the establishment's current situation. At this time, does this establishment have a checking or savings account?". Account = 1 if k6 = yes and 0 otherwise.
Loan availability	Question k8: "At this time, does this establishment have a line of credit or a loan from a financial institution?". Loan availability = 1 if k8 = yes and 0 otherwise.
Interest	Answer k17c: "Interest rates were not favorable". Interest = 1 if k17c = yes, and 0 otherwise.
Collateral	Answer k17d: "Collateral requirements were too high". Collateral = 1 if k17d = yes, and 0 otherwise.
Adequacy	Answer k17e: "Size of loan and maturity were insufficient". Adequacy = 1 if k17e = yes, and 0 otherwise.

(*continued on next page*)

Table C.1 (*continued*).

Variable name	Survey question and variable definition
Complexity	Answer k17b: "Application procedures were complex". Complexity = 1 if k17b = yes, and 0 otherwise.
Expected rejection	Answer k17f: "Did not think it would be approved". Expected rejection = 1 if k17f = yes, and 0 otherwise.
Discouraged	if k17b = yes, or k17c = yes, or k17d = yes, or k17e = yes, or k17f = yes, and 0 otherwise.
Personal loans	Question k15d: "At this time, does the owner or owners of this establishment have any outstanding personal loans that are used to finance this establishment's business activities?". Personal loans = 1 if k15d = yes, and 0 otherwise.
Overdraft	Question k7: "At this time, does this establishment have an overdraft facility?". Overdraft = 1 if k7 = yes, and 0 otherwise.
Trade credit	Trade credit = "% Purchases on credit from suppliers and advances from customers" (answer k3f).
Rationing	Variable constructed as in Kuntchev et al. (2014). Rationing = 2 (fully constrained) if the firm does not have external sources of finance and applied for a loan and was rejected (question bmk7b) or did not apply because of the terms and conditions (question k17). Rationing = 1 (partially constrained) if the firm has external sources of finance and the loan was approved in part, it was rejected, or because of the terms and conditions. Rationing = 0 (not constrained) otherwise.
Past sales growth	Question d2: "In the last fiscal year, what were this establishment's total annual sales for all products and services?". Question n3: "Three fiscal years ago, what were total annual sales for this establishment?". Past sales growth is measured as a percentage change in sales between the last completed fiscal year and the previous period. All sales values are deflated to 2009 using each country's GDP deflators.
Expected change in sales growth: Positive	Question bmd1a: "Considering the next year, are this establishment's total sales expected to increase, decrease, or stay the same?". Positive expectations = 1 if bmd1a = "increase" and 0 otherwise.
Expected change in sales growth: Stable	Stable expectations = 1 if bmd1a = "stay the same" and 0 otherwise.
Expected change in sales growth: Negative	Negative expectations = 1 if bmd1a = "decrease" and 0 otherwise.
Expected value of sales growth	Question bmd1b: "In percentage terms, what is the expected change in total sales?". Expected value of sales growth = bmd1b if Expected change in sales growth: Positive = 1, Expected value of sales growth = –bmd1b if Expected change in sales growth: Negative = 1, and 0 otherwise.
Age	Question b5: "In what year did this establishment begin operations?". Age = ln(1+T–b5), where T is the year of the survey.
Size	Question l2 "Looking back, at the end of two fiscal years ago, how many permanent, full–time individuals worked in this establishment? Please include all employees and managers". Size = ln(1+l2).
Export	Question d3: "Coming back to the last fiscal year, what percentage of this establishment's sales were: national sales [d3a], indirect exports (sold domestically to third party that exports products) [d3b], direct exports [d3c]?". Export = 1 if d3c $\geq$ 1%.
Import	Question d12b: "In the last fiscal year, what percentage of this establishment's purchases of material inputs or supplies were: of domestic origin [d12a], of foreign origin [d12b] ?". Import = 1 if d12b $\geq$ 1%.
Innovation	Question h1: "During the last three years, has this establishment introduced new or improved products or services?". Innovation = 1 if h1 = 1, and 0 otherwise.
Number of competitors	Question e2: "In the last fiscal year, for the main market in which this establishment sold its main product, how many competitors did this establishment's main product face?". The original answer was a cardinal measure distinguishing the following classes: i. 0, ii. 1, iii. 2–3, iv. 4–5, v. 6–10, vi. 11–180, or vii. too many to count. We generate a continuous measure by imposing the median number of each class and assuming the lower bound of 181 for the last category vii. We then compute the stratum average defined as the mean number of reported competitors faced by other firms in the same area and sector of the firm (in log).
Manager elected	Question bmb5: "Has the owner, CEO, top manager, or any of the board members of this firm ever been elected or appointed to a political position in this country?". Manager elected = 1 if bmb5 = yes, and 0 otherwise.
K investment	Question n5a: "In the last fiscal year, how much did this establishment spend on purchases of new or used machinery, vehicles, and equipment?". K investment = 1 if n5a$\geq$1%, and 0 otherwise.
R&D investment	Question h8: "In last fiscal year, did this establishment spend on formal research and development activities, either in-house or contracted with other companies, excluding market research surveys?". R&D investment = 1 if h8 = yes, and 0 otherwise.
Originally informal	Question b6a: "Was this establishment formally registered when it began operations?". Originally informal = 1 if b6a = yes, and 0 otherwise.
Years of formality	Question b6b: "In what year was this establishment formally registered?". Years of formality = ln(1+T–b6b), where T is the year of the survey.
Local market	Question e1: "In the last fiscal year, which of the following was the main market in which this establishment sold its main product?". Available answers: i. Local (main product sold mostly in same municipality where establishment is located), ii. National (main product sold mostly across the country where establishment is located), and iii. International. Local market = 1 if e1 = i.
National market	National market = 1 if e1 = ii.
Years manager experience	Question b7: "How many years of experience working in this sector does the top manager have?". Years manager experience = log(1+b7).
Government ownership	Question b2: "What percentage of this firm is owned by each of the following". Government ownership = b2c, "% Government or State".
City 1	question a3: "Size of locality". Available answers: i. "City with population above 1 Million", ii. "Over 250.000 to 1 million", iii. "50,000 to 250,000", iv. "Less than 50,000". City 1 = 1 if a3 = iv, and 0 otherwise.

(*continued on next page*)

Table C.1 (*continued*).

Variable name	Survey question and variable definition
City 2	City 2 = 1 if a3 = iii, and 0 otherwise.
City 3	City 3 = 1 if a3 = ii, and 0 otherwise.
City 4	City 4 = 1 if a3 = i, and 0 otherwise.
Electric outages (N)	Question c7: "In a typical month, over the last fiscal year, how many power outages did this establishment experience?". Electric outages (N) = log(1+c7).
Bribery depth	Bribery depth is computed similarly as the Graft Index from Gonzalez, Lopez-Cordova, and Valladares (2007). it is constructed from the following questions. Question c5: "In reference to that application for an electrical connection, was an informal gift or payment expected or requested?". Question c14: "In reference to that application for a water connection, was an informal gift or payment expected or requested?". Question g4: "In reference to that application for a construction-related permit, was an informal gift or payment expected or requested?". Question j5: "In any of these inspections or meetings (with tax officials) was a gift or informal payment expected or requested?". Question j12: "In reference to that application for an import license, was an informal gift or payment expected or requested?". Question j15: "In reference to that application for an operating license, was an informal gift or payment expected or requested?".
Loss from theft	Question i4: "In the last fiscal year, what were the estimated losses as a result of theft, robbery, vandalism or arson that occurred on this establishment's premises either as a percentage of total annual sales?".
Constrained by tax administration	Question j30b: "To what degree is Tax Administration an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. tax admin = 1 if j30b = iv. or v., and 0 otherwise.
Constrained by labor regulation	Question l30: "To what degree are Labor regulations an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. labor reg = 1 if l30 = iv. or v., and 0 otherwise.
Constrained by license	Question j30c: "To what degree is Business Licensing and Permits an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. license = 1 if j30c = iv. or v., and 0 otherwise.
Constrained by finance	Question k30: "To what degree is Access to Finance an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. finance = 1 if k30 = iv. or v., and 0 otherwise.
Constrained by corruption	Question j30f: "To what degree is Corruption an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. corruption = 1 if j30f = iv. or v., and 0 otherwise.
Constrained by crime	Question i30: "To what degree is Crime, Theft and Disorder an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. crime = 1 if i30 = iv. or v., and 0 otherwise.
Constrained by electricity	Question c30: "To what degree is Electricity an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. electricity = 1 if c30 = iv. or v., and 0 otherwise.
Constrained by transport	Question d30b: "To what degree is Transport an obstacle to the current operations of this establishment?". Available options: i. no obstacle, ii. minor obstacle, iii. moderate obstacle, iv. major obstacle, or v. very severe obstacle. Constr. transport = 1 if d30b = iv. or v., and 0 otherwise.

Table C.2
Variable description.

Variable name	Description
Loan application	Dummy for firms that applied for a loan or credit line (independently of the outcome).
Constrained by informal	Dummy for firms identifying practices of competitors in the informal sector as a major or very severe obstacle to their operations
Loan availability	Dummy for firms with an outstanding loan or credit line.
Turned down	Dummy for firms whose application for a loan or credit line was turned down.
No need	Dummy for firms that did not apply for a loan because they did not need funds.
Interest	Dummy for firms that did not apply for a loan because interest rates were not favorable.
Collateral	Dummy for firms that did not apply for a loan because collateral requirements were too high.
Adequacy	Dummy for firms that did not apply for a loan because size of the loan or maturity were insufficient.
Complexity	Dummy for firms that did not apply for a loan because application procedure was too complex.
Expected rejection	Dummy for firms that did not apply for a loan because they thought the loan would be denied.
Discouraged	Dummy for discouraged borrowers that did not apply because of the complexity of the procedure, interest rates, collateral requirements, adequacy of the loan, or expected to be rejected.
Personal loans	Dummy for the existence of owner(s)' personal loans used to finance firms' activity.
Overdraft	Dummy for the availability of an overdraft facility.
Trade credit	Share of working capital financed through trade credit.
Rationing	Categorical measure for firms' rationing. It takes value 0, 1, and 2, depending on whether is not rationed, partially rationed, or fully rationed.
Past sales growth	Realized sales growth over the last three years
Expected value of sales growth	Continuous measure for firms' expected sales growth over the following year.
Expected change in sales growth: Positive	Dummy for firms expecting increasing sales in the following year.
Expected change in sales growth: Stable	dummy for firms expecting stable sales in the following year.
Expected change in sales growth: Negative	dummy for firms expecting decreasing sales in the following year.

(*continued on next page*)

Table C.2 (*continued*).

Variable name	Description
Account	Dummy for firms with a checking or savings account.
Originally informal	Dummy for firms originally starting their activity without being formally registered.
Years of formality	Log–years since the firm was formally registered.
Age	Log–age (1+).
Size	Log–employees (1+).
Export	Dummy for exporting firms.
Import	Dummy for importing firms.
Innovation	Dummy for innovative firms.
Number of competitors	Log-average number of competitors in the sector and geographical area of the firm.
Manager elected	Dummy for CEO, managers, or board members elected/appointed to a political position in the country.
K investment	Dummy for firms investing in physical capital.
R&D investment	Dummy for firms investing in R&D.
Manufacturing	Dummy for firms operating in the manufacturing sector.
Retail	Dummy for firms operating in the retail sector.
Listed company	Dummy for listed companies.
LLC	Dummy for LLC firms.
Sole proprietorship	Dummy for sole proprietorship firms.
Partnership	Dummy for partnership firms.
Ltd Partnership	Dummy for Ltd partnership firms.
Local market	Dummy for firms mainly selling products to local markets.
National market	Dummy for firms mainly selling products to national markets.
Years manager experience	Number of years of experience of the manager (in log).
Government ownership	Share of the firm owned by the government.
City 1	Dummy for firms operating in cities with population below 50,000.
City 2	Dummy for firms operating in cities with population between 50,000 and 250,000.
City 3	Dummy for firms operating in cities with population between 250,000 and 1,000,000.
City 4	Dummy for firms operating in cities with population above 1,000,000.
Electric outages (N)	Number of electric outages experienced in the last year (in log).
Bribery depth	Percentage of instances in which a firm was either expected or requested to provide a gift or informal payment during solicitations for public services, licenses or permits.
Loss from theft	Losses due to theft and vandalism against the firm as a percentage of total sales.
Constrained by tax administration	Dummy for firms identifying tax administration as a major constraint.
Constrained by labor regulation	Dummy for firms identifying labor regulation as a major constraint.
Constrained by license	Dummy for firms identifying business licensing and permits as a major constraint.
Constrained by finance	Dummy for firms identifying access to finance as a major constraint.
Constrained by corruption	Dummy for firms identifying corruption as a major constraint.
Constrained by crime	Dummy for firms identifying crime, theft and disorder as a major constraint.
Constrained by electricity	Dummy for firms identifying electricity as a major constraint.
Constrained by transport	Dummy for firms identifying transportation as a major constraint.

References

Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, *29*(1), 1–11.

Akbas, O., Alper, K., Benincasa, E., Betz, F., Davradakis, E., Gattini, L., & Kadereit, T. (2022). *Access to finance in the Middle East and North Africa: Evidence from the 2019 enterprise survey*: *MENA enterprise survey report working papers: Volume 1*, World Bank, European Bank for Reconstruction and Development, & European Investment Bank.

Amin, M. (2021). *Does competition from informal firms hurt job creation by formal firms? Evidence using firm-level survey data*: *World Bank policy research working paper 9515*, Washington, DC.: World Bank.

Amin, M., & Soh, Y. C. (2021). Does greater regulatory burden lead to more corruption? Evidence using firm-level survey data for developing countries. *World Bank Economic Review*, *35*(3), 812–828.

Avenyo, E. K., Konte, M., & Mohnen, P. (2021). Product innovation and informal market competition in sub-Saharan Africa. *Journal of Evolutionary Economics*, *84*(C), 605–637.

Ayyagari, M., Juarros, P., Martinez Peria, M. S., & Singh, S. (2021). Access to finance and job growth: Firm-level evidence across developing countries. *Review of Finance*, *25*(5), 1473–1496.

Bah, E.-h., & Fang, L. (2015). Impact of the business environment on output and productivity in Africa. *Journal of Development Economics*, *114*, 159–171.

Banerjee, A., & Duflo, E. (2014). Do firms want to borrow more? Testing credit constraints using a directed lending program. *Review of Economic Studies*, *81*, 572–607.

Beck, T., Demirgüç-Kunt, A., & Maksimovic, V. (2005). Financial and legal constraints to growth: Does firm size matter? *The Journal of Finance*, *60*(1), 137–177.

Betz, F., Ravasan, F., & Weiss, C. (2021). *Financial Informality among formal firms: Evidence from small businesses in Egypt*: *Oxford Martin working paper series no. 2021-2*.

Boneva, L., Cloyne, J., Weale, M., & Wieladek, T. (2020). Firms' price, cost and activity expectations: Evidence from micro data. *The Economic Journal*, *130*(627), 555–586.

Brancati, E., Di Maio, M., Gatti, R., & Islam, A. M. (2022). *Conflict and firms' performance: Geo-localised evidence from a global dataset*: *Mimeo*.

Coibion, O., Gorodnichenko, Y., & Kumar, S. (2018). How do firms form their expectations? New survey evidence. *American Economic Review*, *108*(9), 2671–2713.

Crepon, B., Devoto, F., Duflo, E., & Parienté, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics*, *7*(1), 123–150.

de Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in micro enterprises: Evidence from a field experiment. *Quarterly Journal of Economics*, *123*(4), 1329–1372.

Diao, X., Kweka, J., & McMillan, M. (2018). Small firms, structural change and labor productivity growth in Africa: Evidence from Tanzania. *World Development*, *105*, 400–415.

Distinguin, I., Rugemintwari, C., & Tacneng, R. (2016). Can informal firms hurt registered SMEs' access to credit? *World Development*, *84*, 18–40.

Enders, Z., Hünnekes, F., & Müller, G. J. (2022). Firm expectations and economic activity. *Journal of the European Economic Association*, *20*(6), 2396–2439.

ERDB, EIB, & WB (2016). *What's holding back the private sector in MENA? Lessons from the enterprise survey*. European Bank for Reconstruction and Development, European Investment Bank, World Bank Group.

ERDB, EIB, & WB (2022). *Access to finance and investment, in unlocking sustainable private sector growth in the Middle East and North Africa: Evidence from the enterprise survey*. European Bank for Reconstruction and Development, European Investment Bank, World Bank Group.

Falco, P., Maloney, W. F., Rijkers, B., & Sarrias, M. (2015). Heterogeneity in subjective wellbeing: An application to occupational allocation in Africa. *Journal of Economic Behaviour and Organization*, *111*, 137–153.

Fisman, R., & Svensson, J. (2007). Are corruption and taxation really harmful to growth? Firm level evidence. *Journal of Development Economics*, *83*(1), 63–75.

Gennaioli, N., Ma, Y., & Shleifer, A. (2016). Expectations and investment. *NBER Macroeconomics Annual*, *30*(1), 379–431.

Gonzalez, A., Lopez-Cordova, J. E., & Valladares, E. E. (2007). *The incidence of graft on developing-country firms*: *World Bank policy research working paper 4394*, Washington, DC: World Bank.

ILO (2018). *Women and Men in the informal economy: A statistical picture* (3rd ed.). International Labour Organization, Geveve.

Kersten, R., Harms, J., Liket, K., & Maas, K. (2017). Small firms, large impact? A systematic review of the SME finance literature. *World Development*, *97*, 330–348.

Kuntchev, V., Ramalho, R., Rodríguez-Meza, J., & Yang, J. S. (2014). *What have we learned from the enterprise surveys regarding access to credit by SMEs? World Bank policy research working paper 6670*.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, *28*(3), 109–126.
Maloney, W. (2004). Informality revisited. *World Development*, *32*(7), 1159–1178.
Quinn, S., & Woodruff, C. (2019). Experiments and entrepreneurship in developing countries. *Annual Review of Economics*, *11*(1), 225–248.
Rozo, S., & Winkler, H. (2021). Is informality good for business? The impacts of IDP inflows on formal firms. *Journal of Human Resources*, *56*, 1141–1186.
Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, *108*, 2015–2047.
Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, *12*, 525–546.