WPS7900

Policy Research Working Paper 7900

Can Enhancing the Benefits of Formalization Induce Informal Firms to Become Formal?

Experimental Evidence from Benin

Najy Benhassine
David McKenzie
Victor Pouliquen
Massimiliano Santini

WORLD BANK GROUP
Development Research Group
Finance and Private Sector Development Team
&
Trade and Competitiveness Global Practice Group
November 2016

Policy Research Working Paper 7900

Abstract

Governments around the world have introduced reforms to attempt to make it easier for informal firms to formalize. However, most informal firms have not gone on to become formal, especially when tax registration is involved. A randomized experiment based around the introduction of the *entreprenant* legal status in Benin is used to provide evidence from an African context on the willingness of informal firms to register after introducing a simple, free registration process, and to test the effectiveness of supplementary efforts to enhance the presumed benefits of formalization by facilitating its links to government training programs, support to open bank accounts, and tax mediation services. Few firms register when just given information about the new regime, but 9.6 percentage points more register when they were visited in person and the benefits were explained. The full package of supplementary efforts boosts the impact on the formalization rate to 16.3 percentage points, demonstrating that enhancing the benefits of formalization does induce more firms to formalize. Firms that are larger, and that look more like formal firms to begin with, are more likely to formalize, providing guidance for better targeting of such policies. However, formalization appears to offer limited benefits to the firms, and the costs of personalized assistance are high, suggesting that such enhanced formalization efforts are unlikely to pass cost-benefit tests.

This paper is a product of the Finance and Private Sector Development Team, Development Research Group and the Trade and Competitiveness Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at dmckenzie@worldbank.org.

Can Enhancing the Benefits of Formalization Induce Informal Firms to Become Formal? Experimental Evidence from Benin[#]

Najy Benhassine, *World Bank Group*
David McKenzie, *World Bank Group*
Victor Pouliquen, *World Bank Group* and *Paris School of Economics*
Massimiliano Santini, *World Bank Group*

Keywords: Informality, Small Enterprises, Regulatory Simplification
JEL codes: O17, O12, D21, L26, H25.

[#] The authors owe particular appreciation to all the experts and colleagues who contributed to the design and implementation of the *entreprenant* program and its impact evaluation: Louis Akakpo, Theodore Anthonioz, Zoubir Benhamouche, Julien Bornon, Laurent Corthay, Tonagnon Dadjo, Matina Deen, Magueye Dia, Benedicta Houetchenou, Ferdinand Ngobounan, Adrien Pawlik, Dolele Sylla, Kjartan Sorensen, Hamidou Sorgo, Adama Tiendrebeogo, and Alain Traore. The authors would also like to thank the Government of Benin for the cooperation and support during the impact evaluation, CGA (*Centres de Gestion Agrées),* GUFE (*Guichet Unique de Formalisation des Entreprises)*, Bank of Africa, Orabank and IERPE (Institute for Empirical Research in Political Economy). Research for this paper has been supported by (1) the International Finance Corporation, the Facility for Investment Climate Advisory Services (FIAS), the U.K. Department for International Development (DFID), the Foreign Affairs, Trade and Development Canada (DFATD), and the U.S. Agency for International Development through the Impact Program managed by the World Bank Group (WBG)'s Trade and Competitiveness Global Practice; (2) the Research Support Budget (RSB), the Strategic Research Program (SRP), and the Impact Evaluation to Development Impact (i2i) through the WBG's Development Economics Department; and (3) the Private Enterprise Development for Low-Income Countries (PEDL) initiative managed by the Centre for Economic Policy Research (CEPR) and DFID. For questions or comments, please contact Massimiliano Santini at <msantini@ifc.org>.

1. Introduction

A large majority of micro, small, and medium-sized firms throughout the developing world operate in the informal sector (La Porta and Shleifer, 2014a). This is certainly the case in Benin, where the national statistics agency has estimated that the informal sector represents up to 70 percent of GDP and 95 percent of employment (INSAE, 2009). A high level of informality is often seen to be costly for governments (who lose out on tax revenues and information on the firm sector), formal firms (who may suffer from unfair competition), and for the informal firms themselves (who may not be able to access bank financing, public contracts, or government programs, may face corruption or intimidation from tax inspectors, and as a result have low productivity) (e.g. Levy 2008; Farrell, 2004; Perry et al, 2007; La Porta and Shleifer, 2014b). In response, many countries have implemented business entry regulation reforms in order to reduce informality, spurred by the work of De Soto (1989) and the *Doing Business* project of the World Bank (World Bank, 2016).

However, a review of the existing evidence suggests that easing entry regulations and providing information on the formalization process has only a very limited impact on the formalization of existing informal firms (Bruhn and McKenzie, 2014). The largest impacts have come from reforms which make it easier and cheaper to register for a status not directly linked to tax registration. In Peru, Alcãzar et al. (2010) find that offering a subsidy for the cost of obtaining a municipal license led to 10 to 12 percent of informal businesses obtaining it, while in Malawi, Campos et al. (2015) find that offering assistance with registering in a company registry had a large impact on business registration, with 75 percent of those offered assistance obtaining a business registration certificate. In contrast, impacts have been much smaller when registration for taxes is involved. Providing information and removing the upfront cost of registration had no effect on tax registration in randomized experiments in Sri Lanka (de Mel et al. 2013), Bangladesh (de Giorgi and Rahman, 2013), Brazil (Andrade et al, 2013), Malawi (Campos et al, 2015), or Colombia (Galiani et al, 2015). One interpretation of this evidence is that burdensome regulations are not the main reason these firms are informal, but instead they are rationally choosing to be informal because the benefits of formalizing are low for them compared to the tax and other costs (Maloney, 2004).

The limited success of these studies in getting firms to formalize has meant there have been few opportunities to actually measure what the benefits to informal firms of tax registration actually are. Some evidence is available from Sri Lanka, where de Mel et al. (2013) paid firms to

formalize, and from Brazil, where Andrade et al. (2013) used tax inspectors to force formalization. In neither case were firms able to benefit from many of the purported advantages of formal status, including access to business banking, participation in government training programs, receiving government contracts, or increased certainty over taxes. De Mel et al. (2013) find some impact of formalization on firm profitability, but this impact appears to be driven by a handful of firms for which profit increased substantially, with most firms experiencing no change. This evidence shows the link from formalization to the benefits of formalizing is not automatic, and suggests the need for supplementary services to enhance formalization assistance. There is also little evidence from Africa, where development levels are lower, and the informal sector even larger, than in the Latin American and Asian contexts where most of the existing studies have been done.

In this paper, we test the effectiveness of offering supplementary services to enhance the take-up and returns to formalization in the context of a randomized experiment in Benin. We do this in the context of the launch of the *entreprenant* legal status, a simplified regime being offered to small informal businesses in 17 African states with the goal of making it easier for them to enter the formal economy. This status includes tax registration, and in principle offers the key benefits of formalization in terms of access to bank accounts, government programs and contracts, and tax certainty. In the pilot phase of launching this regime, we worked with the Government of Benin to test experimentally three different programs with a sample of 3,600 informal businesses.

Firms were randomized into a control group and three treatment groups. The first treatment group received in-person visits in which the new status was explained, the potential benefits of formalization verbally described, and advisors helped firms with the paperwork as needed. The second treatment aimed to enhance the benefits of formalization by also offering business training and support opening a business bank account if they formalized. The third treatment built on the second by offering tax mediation services, with the goal of providing protection and assurance against fear of the tax administration. A supplementary treatment provided information in the form of leaflets and a verbal explanation to see whether information alone had an impact.

We use administrative data on formalization coupled with two rounds of follow-up surveys to measure the impact of these treatments. Only 2 percent of the control group formalized over a

two year period, showing that in the absence of any intervention most informal firms stay informal. All three treatments had significant impacts on formalization, with the impacts larger as more supplementary services were offered: there was a 9.6 percentage points increase in registration in the first treatment group, 13 percentage points in the second, and 16.3 in the third, with these differences between groups all statistically significant. In contrast, information leaflets alone had no impact on formalization. We investigate whether governments can target these programs in a way to achieve even greater take-up levels by examining heterogeneity in impact according to key characteristics specified in a pre-analysis plan.[1] We find impacts are higher for male business owners, those with more education, those operating outside the biggest market in Cotonou (Dantokpa), and those that we classified ex ante as being more similar to businesses already formal using species classification (De Mel et al., 2010). Targeting on these characteristics could increase formalization rates to up to 27 percent, meaning that the majority of those offered the program would remain informal, even with targeting.

We then measure the consequences of formalizing for these firms. Formalizing leads to increased participation in business training, more formal accounting, lower tax harassment, and less taxes paid (due to a tax exemption in the year after formalizing). However, formal firms are not significantly more likely to obtain business bank accounts or loan financing, do not gain more customers, and have no significant gains in sales, profits, or standard of living. While the benefits of formalizing are thus modest, the cost of the intervention is not. We calculate an average cost of US$1,200-2,200 per firm formalized without targeting, and at least $600 per firm formalized even if targeting were used. This is large relative to the average monthly profits of these firms of only $79 and to the tax collection the government can expect to receive from such firms. As such, our analysis suggests that while introducing a simplified registration system offers at least time-saving benefits for firms that want to formalize, adding additional services or in-person visits to explain this new status is unlikely to pass a cost-benefit test.

The remainder of the paper is organized as follows: Section 2 describes what is meant by formalization in Benin, and the potential benefits and costs associated with becoming formal; Section 3 details the intervention, the sampling, the study design and the data; Section 4 describes program implementation and take-up on program components; Section 5 presents the

[1] This study was registered in the AEA RCT Registry on October 7, 2014, prior to any follow-up survey data being collected https://www.socialscienceregistry.org/trials/515

theory behind the program and our empirical strategy; Section 6 details the program impact on formalization; Section 7 shows the impact on business performances; and Section 8 concludes.

2. Formalization in Benin

The seventeen OHADA (*Organisation pour l'Harmonisation en Afrique du Droit des Affaires*) member countries adopted a revised General Commercial Law in December 2010, which came into effect in May 2011. The new law, immediately applicable to all OHADA members, introduced the *entreprenant* status, a simplified legal regime specifically designed for small entrepreneurs, whose intended objective is to facilitate the migration of businesses operating in the informal sector into the formal sector. However, the law did not make explicit how the *entreprenant* status practically functioned, nor the specific combination of incentives that it would include, instead allowing each country to fill in the vacuum through *ad-hoc* secondary legislation and institutional changes. Benin, as a member of OHADA, was the first OHADA country to implement the *entreprenant* legal status.

The *entreprenant* status can apply to a physical person running a micro or small business involved in any type of activity. Formalization with this new status is easy, free of charge and takes only one business day. The introduction of the *entreprenant* status is part of a broader effort from the Government of Benin to simplify and reduce the costs of formalization. Reforms of other existing legal status were implemented a few months before the creation of *entreprenant* status, and included the creation of a one-stop shop for business registration, and a significant reduction of the registration costs associated with the main existing legal status. The registration cost for individual enterprises dropped from CFAF 65,000 (USD109[2]) to CFAF 10,000 (USD17) and from CFAF 225,000 (USD378) to CFAF 17,000 (USD29) for limited liability companies (only the *entreprenant* status is totally free of charge). For all statuses the time to register was reduced to one business day. As these reforms (including the creation of the *entreprenant* status) were implemented recently, information on the new conditions to formalize was not likely to be known by the majority of informal businesses operating in Cotonou at the time of the start of the program.

Formalizing in Benin means to choose a legal status and register at the chamber of commerce (GUFE, *Guichet Unique de Formalisation des Entreprises*). It offers some potential benefits

[2] Exchange rate on June 1, 2016 on *oanda.com*: 1 USD= CFAF 596.

(presented in Table A1) depending on the type of status chosen. Most of these potential benefits are related to the possibility to apply for bank services, or to access new markets like government and large companies' contracts. The *entreprenant* status gives access to all advantages except the rights to export and to access large public contracts. It explicitly targeted micro and small businesses managing one type of activity with a limited turnover.[3] Businesses with multiple activities or with turnover greater than a threshold in two consecutive years will lose *entreprenant* status and have to adopt the individual enterprise status.

When they formalize, businesses get a unique fiscal identifier and are registered with the tax administration. Accordingly, the main potential cost of formalization is related to taxes. In Benin, the link between formalization and taxes is complex and varies according to the business. In theory, all businesses with a fixed location would pay taxes even if they were informal. Before the reform of the tax system affecting microentrepreneurs was introduced in 2015,[4] which will be used to calculate the tax owed in 2016, there were four different tax regimes that could apply to informal businesses in Cotonou, depending on their location and economic activity.[5] The regime most commonly applicable to micro, informal businesses was the TPU (*"Taxe Professionnelle Unique")* and was calculated based on the rental value of the business premises. However, in the majority of cases taxpayers did not have a lease contract, the only official and opposable proof of rental value. As a result, the law assigned the tax administration the responsibility for assessing the rental value. This assessment often left a door open for discretion. In practice, tax inspectors estimated businesses' ability to pay based on their appearance and on discussion with business owners.

Data from our baseline survey (see Table 1) show that more than 70 percent of firms think it is difficult to know in advance how much taxes they will have to pay, with this being the case even for formal firms. Slightly more than half of the informal firms in our study (55 percent)

[3] The OHADA General Commercial Law defines the *entreprenant* as having an annual turnover below CFAF 30 million (USD 50,400) for trading activities, CFAF 20 million (USD 33,600) for crafting activities (artisans), and CFAF 10 million (USD 16,800) for services. Once the small business adopts the *entreprenant* status, the turnover threshold should not be exceeded for more than two consecutive years.

[4] In December 2014, the Beninese Parliament adopted a new MSE tax regime. This regime introduced the Synthetic Professional Tax (TPS: *Taxe Professionnelle Synthètique*) which replaces the four taxes that micro and small businesses were subject to before the reform. This new tax introduces a major shift by changing the basis of tax calculation from the rental value to the use of turnover. This reform creates more predictability and transparency in the calculation of the amount of tax due and prevents small businesses from abuses of tax officers. MSEs will start paying the TPS in 2017 based on their 2016 turnover. All *entreprenants* will pay the TPS.

[5] The four tax regimes were the following: "*Taxe Professionnelle Unique*" (TPU), "*Taxe Unique sur les Transports Routiers*" (TUTR), "*Régime du forfait des revendeurs de tissus et divers*", and "*Régime du bénéfice réel simplifié*".

paid some taxes in 2013, with the average amount paid equivalent to 9 percent of average annual profits. Formal firms were more likely to be paying taxes at all (84 percent paid), and paid a higher amount of taxes conditional on paying (an average of 17 percent of profits).

However, in the short-term, the main objective of the Government of Benin with the *entreprenant* program is not to increase the tax collected, but rather to (i) introduce a channel to formalization for micro and small businesses, which may at a later stage grow enough to be able to substantially contribute to the tax revenues, and (ii) create a culture of legality, whereby businesses are encouraged to abide the law, in the belief that it will be ultimately beneficial for the society at large. When they formalize, businesses can benefit from tax exemptions under certain conditions. Businesses which also register to the CGA (an association providing business counseling and account certification) can benefit from a full tax exemption for the first year after formalizing, in addition to a reduction of 40% in the amount of taxes due for the following 3 years. As a result, the amount of taxes paid by firms which formalize may actually decrease in the short-term.

3. Evaluation design

3.1. The Intervention

Given the flexibility provided by the OHADA framework as to how the *entreprenant* status should be implemented, the Government of Benin was interested in knowing the most impactful and efficient way to operationalize the legal status. We worked with the government to design and test the following three packages of incentives to formalization, with the goal of understanding what would be the best combination of incentives:

3.1.1. Package A – Information on the *entreprenant* status and assistance in registering

The *Centres de Gestion Agréés* (CGA) a semi-public organization that focused on providing small and medium enterprises with business management, accounting, and tax consulting services provided advisors who would visit selected firms in person. They explained the benefits of becoming an *entreprenant*, and provided (i) a leaflet describing the *entreprenant* status, its advantages and requirements, (ii) one leaflet explaining the registration process at GUFE, and (iii) one leaflet explaining the different tax regimes applicable to *entreprenants* and how to calculate taxes due within each regime (see section 2). The informal businesses that decided to formalize needed to submit an application at GUFE to obtain the *entreprenant* card.

When necessary, CGA advisors helped *entreprenants* with the formalization process at GUFE, including filling in the declarations and preparing all the required accompanying documents.

3.1.2 Package B – Provision of business services and trainings, and assistance in opening a bank account

The second package aimed to supplement the basic help in package A by facilitating access to the training services and to commercial banks, which are potential benefits of formalizing, but which many firm owners may not otherwise benefit from in practice. Following the first visit to each business, CGA advisors organized a second visit to deliver a 1-2 hour personalized training session. They then noted a variety of additional training sessions that business owners could access conditional on receiving the *entreprenant* card. They could sign up for training at CGA which included four workshops: three mandatory and one optional. The mandatory workshops were: (a) basic accounting, (b) initiation to tax obligations, and (c) financial education. For the optional workshop, businesses were invited to choose between (i) basics of microenterprise management, (ii) initiation to sales development and access to markets, and (iii) basic of business plan development. Each workshop lasted three consecutive half-days. Once the business owner completed the fours workshops with the CGA, he/she received an official diploma, and a sticker acknowledging that he/she received the training.

Firms receiving this package were also offered support from CGA to open a business bank account. The bank partners of the impact evaluation (Orabank and Bank of Africa) designed a specific banking product for the *entreprenant*, with dedicated services and simplified banking access conditions, including a debit card, bank account consultation with mobile phone, cash transfers, SMS-banking, internet banking and mobile money. The *entreprenant* bank accounts in both banks are cheaper than what businesses can usually get (around CFAF 1,000 per month, or USD 1.7, against CFAF 2,000, or USD 3.4) and do not require any initial deposit, whereas business bank accounts usually do in Benin. CGA advisors assisted the *entreprenant* to open a bank account and provided instructions on how to use it.

3.1.3 Package C – Provision of tax preparation support and tax mediation services

The third package aimed to address the uncertainty and concerns that entrepreneurs had about taxes. Firms which formalized under the third group were offered help in preparing tax forms (including tax returns and supporting documentation). However, given that most businesses were subject to the TPU, and that the amount of TPU to be paid by a given business is

determined by the tax administration without any form being filled by the business, this "offer" was not technically implemented. The advisors also left their contact information in case the *entreprenant* had any complaints about future tax payments and inspections, and offered mediation services in case of a dispute between the firm and the tax administration.

Appendix 1 provides more detail on how these three packages were implemented.

3.2 Sample selection and study population's characteristics

A listing survey was conducted in Benin's largest city of Cotonou in March and April 2014. This survey was designed in order to obtain a representative sample of all businesses operating in Cotonou, including Dantokpa market.[6] All businesses with fixed location, except international and nationwide companies and liberal professions, were targeted. Overall, 19,246 businesses were listed, of which a sample of 7,945 were surveyed. We then dropped businesses which were already formal, and which had very high or very low profits and sales to arrive at a sample of 3,596 for the study. Appendix 2 provides details on the sampling protocols and this selection process.

Table 1 provides descriptive statistics for businesses selected in the sample, and compares them to the overall set of informal businesses and to formal businesses. Businesses selected for the study have very similar characteristics to the whole population of informal businesses surveyed, and the overall study shows good external validity for the whole city of Cotonou. Formal businesses had on average 3.2 employees and monthly profits of around CFAF 210,000 (USD 352), while informal businesses had 1.1 employees and a monthly profit of CFAF 46,000 (USD 77). About 60% of businesses were connected to the electricity network, 55% of businesses were involved in trade activities, 26% worked in services, and 16% were craftsmen. 63% of businesses sampled for the study were owned by women. This reflects the high share of female owners in Dantokpa market. Approximately 30% of business owners never went to school, and less than 20% of the businesses were keeping some type of accounting. Formal businesses had higher access to the banking system: 80% of them owned a bank account, whereas only 20% of informal businesses did.

[6] The largest market in Cotonou and one of the largest in West Africa.

In comparison to similar studies in other contexts, the businesses in this study are smaller in size, reflecting the less developed nature of the country and small size of most informal businesses. In the study in Malawi (Campos et al, 2015), businesses had on average two employees and monthly profit of USD 214, while in the study in Sri Lanka (de Mel et al. (2013)), businesses had on average three employees and monthly profit of USD 300.

3.3 Experimental Design

The 3,596 informal businesses[7] were randomly allocated into three treatment groups and one control group. The first group of informal businesses received package A of incentives, the second group packages A and B of incentives, and the third group packages A, B and C.

The randomization was done in the office using STATA and the following methodology was used for stratification:

(1) 16 strata were created using the following variables: business owner gender, business operating in Dantokpa market, trader, and business owns a bank account.
(2) Inside each stratum a Z-score was created as the average of standardized profits, turnover and number of employees. Based on this Z-score, triplets of businesses were created and inside each triplet, businesses were randomly allocated to 3 groups, each of 1,200 firms.
(3) The 1,200 businesses in one group were then randomly allocated further into a first treatment group with 301 businesses, and second treatment group with 899 businesses.

As a result, 301 businesses were allocated to receive package A (treatment group 1), 899 to receive packages A and B (treatment group 2), 1,199 to receive packages A, B, and C (treatment group 3), and 1,197 to the control group. Figure A2 describes the organizational chart of the interventions. We decided to allocate fewer firms to the first treatment based on the existing literature which had shown limited impact of simplification of business registration procedures and cost reduction alone. Our goal was to retain more firms in the treatment groups that we thought would have higher impacts on formalization, in order to provide sufficient power to estimate the impact of formalization on firm performance.

[7] The sample was initially composed of 3,600 businesses, but 4 businesses were in fact duplicates of other businesses in the sample and were dropped from the sample.

Table A2 presents the results of balance checks of baseline characteristics across the different treatment groups and control group. Overall, it shows that all groups are relatively well balanced with respect to observable characteristics: the number of tests that are statistically significant is close to what should be expected due to chance (2 out of 15 tests for the joint tests of all coefficients are equal to zero are significant at the 10% level).

3.4 Data

Three main sources of data are used for this study: administrative data on formalization and program implementation, in-person quantitative surveys with business owners, and qualitative data with study participants and implementing agencies.

Our main measure of formalization is based on monthly administrative data on business registration provided by the GUFE. This database includes the complete list of all newly registered businesses for all legal statuses. Since most businesses in the control group would not have been aware of the new *entreprenant* status, this measure will capture any alternative legal status they registered under. Appendix 3 describes the matching process used to identify whether firms in the GUFE database came from our sample.

Other main outcomes on business performances (profits and turnover) and intermediate outcomes like business knowledge and practices, taxes and banking were measured through in-person interviews with business owners. The baseline survey of the selected sample of businesses was conducted in March-April 2014 prior to program implementation. Two follow-up surveys were conducted in April-June 2015, and in May-June 2016. Attrition rates at first and second follow-up surveys were 11.8 percent and 15.9 respectively and were not correlated with treatment status. Two years after the baseline survey, 8.6 percent of the businesses had closed their operations, and business closure was also not correlated with treatment status. Table A3 presents survey rates, closure rates and attrition rates by groups.

Balance checks of baseline characteristics across the different groups on the sample of businesses successfully surveyed at the two years follow-up survey show that, overall, the post-attrition sample is relatively well balanced with respect to observable characteristics. Results are presented in table A4 and are very close to those presented in table A2 on the whole sample. The number of statistically significant tests is also close to what should be expected due to

chance (2 out of 15 tests for the joint tests of all coefficients are equal to zero are significant at the 10% level).

Program implementation data were also collected to better understand the quality of services delivered. These included detailed monitoring data from CGA and qualitative surveys with implementing agencies and program participants. 37 semi-structured qualitative interviews were conducted with program participants at different stage of the program, a qualitative surveyor was also regularly sent with the CGA advisors to assess if the study design was respected (29 surveys). In addition, 61 qualitative interviews were conducted with business owners not selected for the program to monitor potential externalities of the program. Finally, focus groups were conducted with the main implementing agencies (CGA, GUFE and both commercial banks).

4 Program implementation and take-up

Data from the implementing partner, in addition to quantitative and qualitative data, suggest that the program was implemented consistently with the study design. Treatment allocation was respected for all businesses, and all components of the program were effectively offered to almost all program beneficiaries. The program was implemented on a rolling basis: CGA advisors started to reach out to informal businesses in April 2014, and completed both visits in February 2015.

Between April 2014 and January 2015, 2,399 "first visits" (100% of total) were attempted by CGA. The take up rate for the first visit was remarkably high and 2,344 visits were completed with success (98% of total). First visits were considered as not completed successfully when CGA advisers were not able to locate the business. Between April 2014 and February 2015, all businesses who received a first visit in treatment groups 2 and 3 were offered a second visit by CGA. Only 932 of these second visits were completed with success (44% of total). According to qualitative surveys and focus groups with the CGA, the main reasons for this relatively low take-up rate were that many businesses were not interested by the second visit, or did not have time to receive it. This finding is consistent with McKenzie and Woodruff (2014) who find an average attendance rate of only 65 percent for business training programs in developing countries. Most of the group training sessions applicable to businesses in treatment groups 2

and 3 were conducted after September 2014. The time lapse between the first and the second visits was much greater than originally planned (3 months in average instead of 2 weeks) because of logistical constraints, and because it often required several trips to the business to complete the second visit successfully.

During the two years following program launch, 302 businesses registered with CGA (13 percent of the total in treatment group 2 and 15 percent of the total in treatment group 3), and 272 businesses participated in a group training session at CGA (12 percent of the total in treatment group 2 and 14 percent of the total in treatment group 3). Since businesses had first to register for the *entreprenant* with GUFE in order to be eligible to register at the CGA, and thus receive the trainings, the percentage of eligible businesses that did register with CGA is sizeable. In fact, 83 percent of the businesses in groups 2 and 3 that formalized (362 businesses in total) decided to register with the CGA, and 75 percent decided to obtain trainings. Business owners in groups 2 and 3 who decided to register as *entreprenant* had also the possibility to open a bank account at BoA or Orabank. After two years, 131 businesses opened an *entreprenant* bank account (6.2 percent of total).[8] Panel A of Table 2 summarizes achievement for each program implementation step.

Qualitative information collected with beneficiaries during program implementation suggests that the program was implemented following the study protocol and in particular that the formalization process with the *entreprenant* status was considered as simple and fast (in addition to being free of charge). Panel B of Table 2 shows quantitative data from the follow up survey and confirms that the formalization process was fast and cheap for businesses in treatment groups. 82 percent of businesses that benefited from the program and formalized declared that they did not pay anything in the process (those who paid something in the treatment groups formalized with a different status than the *entreprenant* status).

Qualitative work conducted few days or weeks after the businesses received a visit from the CGA suggests that the program understanding was relatively good, given the complexity of the intervention. However, data from our endline survey suggest that one and a half to two years

[8] Bank data did not include sufficient information besides names that could be used for the matching. As a result, matching between study data and bank data was not perfect and only 70 percent of the *entreprenant* accounts were found in the study data. Therefore, 6.2 percent represents a lower bound of the number of *entreprenant* bank accounts opened by study participants.

later, most businesses had forgotten about the program. Only 36 percent of businesses in treatment groups 2 and 3, and 32 percent of those in group 1, remembered the *entreprenant* program. Moreover, only 23 percent in groups 2 and 3, and 22 percent in group 1, were able to describe correctly what it is.

In the control group, only 13 percent of the businesses declared that they had heard about the *entreprenant* program, and 5 percent were able to describe it correctly. It suggests that only marginal externalities were generated by the program on those not directly targeted. This is consistent with qualitative interviews conducted with informal businesses not targeted by the program.[9]

In practice, tax mediation services were implemented by CGA for all businesses registered with the CGA (even for those in treatment group 2). Some *entreprenants* reported to the CGA that the tax administration requested tax payments that were higher than expected, or that the tax exemption offered during the first year after registration to the CGA was not implemented. The CGA advisors helped them to solve these issues as they arose. The CGA reported that 29 mediation cases (2.4 percent) happened during the two years of program implementation and that all these cases were solved in favor of the *entreprenant* (i.e. the tax exemption was respected by the tax administration).

5 Theory and Empirical Strategy

We begin by sketching a simple organizing framework for how we should think of firms deciding on whether or not to formalize, and how the different interventions may change this decision. This is followed by a description of our empirical strategy.

5.1 Theory: How might the *entreprenant* program impact formalization and business performance?

A firm owner will formalize if the expected discounted value of the net benefits from doing so exceeds the upfront costs. That is, if:

[9] None of the 61 business owners not in the study population that were interviewed some weeks and months after the program started had ever heard of the *entreprenant* status or of any program related formalization.

$$\sum_{t=1}^{T} \beta \delta^t EU\left(\pi_{F,t} - \pi_{I,t}\right) > C_{Money} + C_{Time} + C_{Information} + \lambda_{liquidity} \quad (1)$$

Where $\pi_{F,t}$ denotes the firm's profits if it is formally registered at time t, and $\pi_{I,t}$ denotes the firm's profits if it is not formally registered at time t. C_{Money}, C_{Time}, and $C_{Information}$ denote the monetary, time, and information costs from registering. The shadow value of capital for liquidity-constrained firms is given by $\lambda_{liquidity}$.

In this framework, firms decide whether or not to become formal after weighing these costs and benefits. The basic introduction of the *entreprenant* status then influences this decision by lowering the monetary costs of registering since the registration itself becomes free (which results in both a direct reduction in C_{Money}, as well as in lowering the liquidity costs $\lambda_{liquidity}$) and by lowering the tax obligations associated with formality, especially in the first three years, therefore boosting $\pi_{F,t}$. This should induce formalization by informal firms who were at the margin of formalizing. Our three interventions can then be viewed as changing additional aspects of this decision. Package A further lowers the time and information costs of registering, package B aims to further increase the profitability benefit ($\pi_{F,t} - \pi_{I,t}$) from formalizing by linking it to training and banking services, and package C aims to increase the expected returns from formalizing by reducing uncertainty about tax payments and also lowering the chance of being overcharged taxes relative to informal status.

This framework also offers three predictions which we can test within our experiment. The first is that not all informal firms will formalize following the reform, only those which were close to the margin and for which these changes tip the balance. In particular, while the registration cost is zero, firms which lack personal identification such as a birth certificate or legal title may still face high monetary and time costs of obtaining the documentation necessary for registering, and so not register.

Second, the framework suggests that those who formalize will have been much closer to the margin of formalizing beforehand than those who do not. We test this through examining heterogeneity of response with respect to several pre-specified characteristics of the owners and businesses which are likely to proxy for closeness to the formalization margin. The first is gender. If women are more likely to be running small businesses as a way of working while also taking care of family responsibilities, they may have fewer plans to grow their business to the size where many of the benefits of being formal attain. This would suggest they are further

from the margin of formalizing and will have lower treatment effects. Second, some businesses already have access to other forms of registration that offer partial benefits and for which the added benefits of the *entreprenant* status will be lower. This includes two groups – those in the Dankopta market who are registered with the public company in charge of all markets ("*Société de Gestion des Marchés Autonomes,*" or SOGEMA), and traders who have a access to a "trader card". Third, we use our baseline data on formal and informal firms together with the species classification technique of de Mel et al. (2010) to identify which informal firms look similar to the formal "species", and predict that they will be closer to this formalization margin. Fourth, we consider directly size and owner education, believing smaller, less productive firms are likely to be further from the margin where formalization can benefit them, so will respond less. Finally, if avoiding problems with tax inspections is a benefit of formalizing, we predict that firms that are less frequently inspected will see less benefit from formalizing.

Finally, the framework predicts that the informal firms that formalize as a result of our added interventions will be further from the margin than those who are already formal and those who would formalize without the added help. That is, the interventions should be bringing in smaller, and less like the formal type to begin with, firms.

5.2 Estimation

To analyze the impact of the program on formalization rates, our estimation is at the firm level and involves the following specification for firm *i*:

$$Y_{i,t=1} = \beta_0 + \beta_1 T1_i + \beta_2 T2_i + \beta_3 T3_i + X'_{k,i} + \varepsilon_{i,t=1} \quad (2)$$

Where $Y_{i,t=1}$ is the outcome variable (formalization), $T1_i$ is an indicator for being assigned to treatment group 1, $T2_i$ an indicator for being assigned to treatment group 2 and $T3_i$ an indicator for being assigned to treatment group 3. X_k is a vector of strata dummy variables (one dummy variable for each triplet of businesses) (Bruhn and McKenzie, 2009) and $\varepsilon_{i,t=1}$ is the error term. β_1, β_2 and β_3 provide the intent-to-treat effect of being assigned to treatment groups 1, 2 and 3, respectively. This is the effect of being a business assigned to treatment 1, 2 or 3 relative to being a business in the control group.

To estimate the intent-to-treat impacts of the interventions on business performances and practices, we pool data from the two follow-up surveys to run panel regressions with the following specifications:

$$Y_{i,t} = a + b_1(T1_i * F1) + b_2(T2_i * F1) + b_3(T3_i * F1) + c_1(T1_i * F2) + c_2(T2_i * F2) + c_3(T3_i * F2) + \pi Y_{i,t=0} + \gamma M_{i,t=0} + X'_{k,i} + \varepsilon_{i,t} \quad (3)$$

Where $Y_{i,t}$ is the outcome variable measured post-treatment for business *i* in year *t* (*t*=1,2), $Y_{i,t=0}$ is its baseline value and $M_{i,t=0}$ a dummy variable indicating whether or not this baseline value is missing, $(Tj_i * Fk)$ is the interaction of being assigned to treatment group *j* (*j*=1, 2, 3) with a dummy for the follow-up survey *k* (*k*=1, 2). X_k is a vector of strata dummy variables and $\varepsilon_{i,t}$ is the error term clustered at the business level. b_1, b_2 and b_3 give the intent-to-treat effect at the first follow-up survey of being assigned to treatment groups 1, 2 and 3 respectively. Similarly, c_1, c_2 and c_3 provide the intent-to-treat effect at the second follow-up survey of being assigned to treatment groups 1, 2 and 3 respectively. We then test whether impacts are constant over time (e.g. $b_1 = c_1$), whether they are constant across treatments ($b_1 = b_2 = b_3$), and whether all program impacts are jointly zero ($b_1 = b_2 = b_3 = c_1 = c_2 = c_3 = 0$).

In order to estimate the effect of formalization on business performances and behaviors, we use panel regressions with the following specification:

$$Y_{i,t} = \alpha_0 + \alpha_1 \widehat{FORMAL}_i + \pi Y_{i,t=0} + \gamma M_{i,t=0} + X'_{k,i} + \varepsilon_{i,t} \quad (4)$$

Where $\widehat{FORMAL}$ is an indicator for being formal, which is instrumented respectively by $(T1_i * F1)$, $(T2_i * F1)$, $(T3_i * F1)$, $(T1_i * F2)$, $(T2_i * F2)$ and $(T3_i * F2)$.

Heterogeneous treatment effects are estimated by interacting treatment status and the lagged dependent variable in (2), (3) and (4) with the variable of interest Z.

In cases where an outcome variable was not collected at baseline, these same specifications are estimated without the control for baseline outcome.

6 Impact on formalization

6.1 Overall impact on formalization

As discussed in section 3.4, our main measure of formalization is registration of the business with the chamber of commerce at GUFE (i.e. the registration was found in GUFE data). We think that this definition of formalization is preferable over others that use follow-up survey data because administrative data included information on the whole study population, whereas survey data only have information on those who were surveyed. Moreover, survey data are subject to declaration bias. However, the correlation between survey data and administrative data was high (0.7), and we show similar results using the survey data as well.

Table 3 presents the results on formalization two years after the program started. The impact of the program on the formalization rate was 9.6 percentage points in group 1, 13 percentage points in group 2, and 16.3 percentage points in group 3. All these effects are statistically significant at one percent level. The effects in treatment groups 2 and 3 are higher than in treatment group 1 (although the test is only statistically significant for group 3), and the effect in treatment group 3 is significantly higher than in group 2: both sets of additional incentives included in package B (counseling, trainings and bank services) and in package C (tax mediation) seemed to be valued by informal businesses as incentives to register.

The formalization rate in the control group was only 2.3 percent. Therefore, in the absence of the program, only a few businesses would have formalized. Alternative measures of formalization that combine survey and administrative data show consistent results on businesses surveyed during the follow-up survey. Impact rates in groups 2 and 3 are always significantly higher than in group 1 (for group 2 the test is only significant for declared formalization), and impact rates in group 3 higher than in group 2.

Figure 1 presents trajectories of impacts in time with formalization rates by group in the months following the first visit received by the CGA.[10] It shows that most businesses that choose to formalize because of the program did it relatively quickly after the first visit. For all treatment groups, most of the impact arises during the first month following the first visit. Then for groups 2 and 3, some businesses took more time to formalize but we don't see any significant additional impact five months after the first visit.

[10] For the control group, the date of the first visit was set at the mode of the first visit date in the other groups (i.e. three months).

6.2 A supplementary information experiment

These impacts are higher than has been observed in similar studies in other contexts when formalization also has some tax implications. Importantly, this is the case not only for the groups providing additional incentives to formalization but also for the group only providing (in-person) information. One key question is then whether the relatively high impact measured in group 1 is due mainly to information (i.e. firms decided to formalize when they learned that registration is free of charge and easy to do) or to the fact that the information was delivered in-person by highly trained and qualified CGA advisors who tried to convince business owners of the benefits of formalizing, and provided assistance with forms and the process as needed.

To answer this question, we designed an additional experiment that was implemented during the two year follow-up survey. Fifty percent of the control group (600 firms) was randomly selected[11] to receive two program leaflets just after the completion of the survey (so we are sure that survey answers were not affected by the "leaflets intervention"). The two program leaflets were identical to the leaflets given to group 1 firms when the program started and were introduced by the surveyor with a short script mentioning that the *entreprenant* status is now available for free and in one day to all businesses, and explaining the location of the one-stop shop for business registration. This small intervention tests whether surveyors only providing information on the new status but not in charge of convincing the business of the benefits of formalizing or assisting them with forms can have similar impact on formalization rate. Table A5 presents the results of this "leaflets intervention".

It shows that the leaflets intervention had no significant impact on formalization decision. It means that simply providing (in-person) information on the new status was not sufficient to increase formalization and that the impact measured for group 1 is also due to the fact that the information was provided by trained and qualified staff who took time to convince business owners to formalize.

6.3 Heterogeneity of impact and usefulness of targeting

Table 4 examines heterogeneity in the impact of our interventions by pre-specified business characteristics. We find that male business owners were significantly more likely to formalize

[11] With stratification on the following variables: gender, operates in Dantokpa market and trader.

than female business owners: 9, 12 and 15 percent of businesses owned by women formalized in groups 1, 2 and 3 respectively (2.1 percent in the control group), compared to 18 for those owned by men in group 1 and almost 25 percent for those in group 2 and 3 (4.7 percent in the control group). This result could be correlated with the fact that a large majority of businesses operating in Dantokpa market are owned by women. However, column 8 of the table shows that it is also true outside Dantokpa market for women not operating in trade.

In all groups, formalization rates were 5-10 percentage points higher outside Dantokpa market than inside the market. One potential explanation is that formalization could be less attractive in the market as businesses are already registered with the public company in charge of all markets in Cotonou (SOGEMA). They also usually have representatives in the market they can address in case of problems with the administration. Businesses operating in the trade sector had lower formalization rates than in other sectors. This result is correlated with the fact that almost all businesses in Dantokpa market are traders, but it is also true outside the market.[12] One possible explanation which was mentioned during qualitative interviews is that before the program implementation, traders already had access to a "trader card" that provides a formal status with specific benefits (see Table 1), whereas no such specific card existed for other sectors.

The program was more effective on businesses with an owner who went to at least secondary school, but is not significantly different with firm size per se. Using species classification techniques (de Mel et al, 2010) we classified 18 percent of the businesses in the sample as "looking more like formal businesses before the program".[13] Formalization rates were 3-10 percentage points higher among informal businesses that were similar to formal businesses before program implementation. Finally, businesses that received more than one visit from a

[12] Results not shown but available upon request.

[13] Looking like a formal business owner is based on the predicted probability of being formal from a logit of formality status on baseline characteristics. This logit uses the data collected during the listing/baseline survey on 7,829 businesses who accepted the survey. Among them, 608 (7.8%) were formal at the time of the survey. We used the following baseline characteristics in the logit: operating in Dantokpa market, gender, age, only primary education, only JHS or SHS level, higher level of education, operating in services, craftsman, business created less than 1 year ago, firm connected to electricity network, total number of employees, firm is doing some accounting, have done any advertising in the last 6 months, log of total amount of sales in an average week, log amount of last month profit, firm owner owns a bank account, the firm pays taxes, have done any advertising in the last 6 months (and controls for missing levels of these variables). Using the "predict" command in STATA, we end up classifying as "looking more like formal" 654 (18.2%) businesses out of the 3,596 in the study sample. This classification was done before we got access to any follow-up data and was mentioned in the pre-analysis plan on the AEA social science registry website.

tax inspector in the year prior to program implementation were more likely to formalize. This result, which is only significant for group 3, may suggest that the program was perceived as a way to limit tax harassment.

These results show that the program was more effective on some sub-populations like male business owners, those operating outside Dantokpa market, with at least secondary education and those which look more like formal businesses before the intervention. Targeting these sub-populations could therefore improve program effectiveness.

6.4 How do the formalized firms compare to the already formal and to those who would formalize anyway?

Table A7 compares the baseline characteristics of the firms formalizing through our various interventions to those who were already formal at baseline, and to the few control group firms that formalized. As expected from the theoretical discussion in section 5, the program brought in smaller firms, and firms that looked less like firm already formal at baseline. Firms in the control group that formalized have characteristics that are closer to firms that were already formal at baseline. Differences between newly formalized firms in the control group and in other treatment groups are all going in the expected direction. For example, firms that formalized in the control group had significantly higher level of baseline sales than firms that formalized in the treatment groups (CFAF 90,000, or USD150 against CFAF 54,000 to 62,000, or USD 90 to 104). Most other statistical tests comparing formal firms in control and treatment groups are not statistically significant but this is not surprising given the small number of firms that formalized in the control group (27).

6.5 Cost effectiveness for the impact on formalization

Data on program costs during the two years of program implementation are presented in Table 5. Total program costs were high and the program as it was implemented for the 2,399 firms in a treatment group costed around CFAF 370 million (USD 620,000). Out this total, CFAF 50 million (USD 84,000) were used to made the *entreprenant* status available at the one-stop shop for business registration for any firm who wants to come along and do it, and CFAF 320 million (USD 537,000) to pay for the additional interventions to encourage take-up (in-person visits, business trainings, etc.). This corresponds to a total cost per business included in the program

that ranged from CFAF 71,000 (USD 119) for group 1 to CFAF 171,000 (USD 288) for group 2, which was slightly more expensive than group 3.[14]

Using the program impact on formalization rates, we can then calculate the costs per formalization in each group. The costs per additional formalization were CFAF 737,000 (USD 1,237) in group 1, CFAF 1.3 million (USD 2,217) in group 2 and CFAF 1 million (USD 1,678) in group 3. Even when only considering variable costs of the program, that is the costs that a government would face once all the initial investment will be amortized, the costs per formalization were also very high. For the first group, which shows the best ratio, the variable cost per formalization was CFAF 540,000 (USD 904), which represents more than eleven times the average of baseline monthly profits (CFAF 47,000 or USD 79).

If we assume that the program could be targeted to sub-populations more likely to respond to it and to formalize (as seen in Section 6.3), this cost would be lower, but would still be very high in comparison to business profits. For example, if the program was targeting only business owners with secondary education, the cost per formalization would be CFAF 345,000 (USD 579) witch still represents about 7 times baseline monthly profits.

These costs do, however, incorporate the fact that the experimental design involved some non-negligible tracking costs due to the fact that the CGA had to find and visit a sample of businesses selected by the research team and spread all over the city of Cotonou. Additional economies of scales could be attained without the tracking costs and if the CGA could target businesses located close to one another.

Finally, we can also benchmark these results with results from a program in Sri Lanka offering cash as an incentive to formalization. Del Mel et al. (2012) found that directly paying firms the equivalent of one month of the median firm's profits leaded to registration of one-fifth of firms. This proportion increased to one-half when payments were increased to two months of the median firm's profits. The firms in their study were larger, and so may have been closer to the margin of formalizing to begin with. Nevertheless, this comparison suggests that directly paying

[14] Costs per firm included in the treatment were slightly higher for group 2 than for group 3 because the CGA allocated proportionally more staffs to group 2 than to group 3. This is due to the limited number of firms that a given CGA advisor was able to handle and to organizational constraints (CGA advisors had to be grouped in pairs responsible for firms that belongs to the same group).

firms to formalize may be more cost-effective than the interventions here which instead provided services and support to firms.

6.6 Why don't more firms formalize?

In the third treatment group, which combined all packages of incentives and in which the impact was the greatest, the formalization rate was 18.6 percent (16.3 percentage points more than in the control group). This impact is greater than for similar programs in other contexts (Bruhn and McKenzie, 2014) in which formalization is also linked to taxes. But it means that even though this type of program had a significant impact, the majority of the informal firms still remain informal. This is true even if we consider specific sub-population like businesses operating outside the market or with an owner who went to secondary school, for which impact on formalization remained below 23 percentage points.

Why do most firms remain informal? A first potential explanation is the presence of other legal barriers to formalizing. Data from our midline survey reveals that only 54 percent of informal business owners have legal identification needed to formalization (either a passport or a Beninese ID card). In contrast, 85 percent have a birth certificate and 75 percent an electoral card, so amending the process to allow these alternative forms of identification to be used would alleviate this constraint for many firms. However, lack of identification does not seem to be the binding constraint to formalizing for most informal firms: only 0.6 percent of the control group said this was one of the two main reasons for not formalizing (table A8).

Our endline survey asked informal business owners the two main reasons why they were still informal (table A8). The most common responses in the control group were that firms did not see any benefits from doing so (32 percent), or that they do not want to have to pay more taxes (26 percent). The other main reason was that they viewed the process as too costly, complicated, or time-consuming (31 percent). These responses are similar among those who remain informal in the treatment groups, despite the visits by CGA advisors to explain the new simpler process of registration and the potential benefits of registering. It is consistent with the idea that many of these informal firms are so far from the formalization margin that they consider this information irrelevant – and indeed, as noted before, two years after program launch, only 20-25 percent of businesses in the treatment groups could even remember what the *entreprenant* program is.

7 Impact on firm performances

We first examine the extent to which formalization resulted in any of the key purported benefits of formalization, and then turn to examining impacts on the main outcomes of profitability, sales growth, and employment.

7.1 Impact on intermediate outcomes

Table 6 examines whether formalizing is leading firms to be more likely to access banks, improve accounting and other business practices, be less harassed for taxes, or access new customers. It does this through estimation of equations (3) and (4) using our two rounds of follow-up surveys. The top of the table presents the yearly intent-to-treat impacts of the different interventions, while the bottom of the table presents the impact of formalization for those who respond to treatment.

Despite the facilitation of access to bank accounts in treatments 2 and 3, and the creation by banks of a special account for *entreprenants*, column 1 shows no significant impact of formalizing on whether the business has a bank account. 25 percent of the control group had accounts, suggesting that in practice the requirement to be formal was not always binding, and that those who signed up for the accounts through our intervention were substituting from accounts they would have opened anyway. Treatment group 2 is 5 percentage points more likely to have received a loan in the second year, but there are no other significant impacts on loan usage. As a result, the overall instrumented impact of formalizing on loan receipt is positive, but not statistically significant.

Columns 3 and 4 do show significant impacts of formalizing on the likelihood of attending business training in the past year (67 percentage points), with this impact coming from treatments 2 and 3 who were offered this service. Formalized firms are more likely to be doing any form of accounting (15 percentage points), but this did not translate into improved overall business practices.[15] One possible explanation is that there was some crowding out effects, and better accounting practices were offset by worst marketing and stock control practices.

Formalization also reduced significantly perceived tax harassment. This result is interesting as it is also valid for businesses in group 1 and 2. It means that it was not due mainly to the tax mediation performed by the CGA but instead that all newly formalized businesses faced less

[15] Measured using the same 26 questions on business practices as in McKenzie and Woodruff (2015).

tax harassment. In contrast, we see no significant impacts the likelihood of selling to public institutions or to clients requesting receipts. We examine further the impact on other potential channels such as advertising, business presentation, investment, the number of customers, innovation, trust in institutions, and subjective standards of living, in the appendix table A9. Formalization does not seem to be changing significantly these other intermediate outcomes. There are few coefficients that are significant, in particular on the total value of inventories and row materials, but it does not survive correction for multiple hypothesis testing (Anderson, 2008 and Benjamini et al., 2006).

7.2 Impact on main outcomes

Taken together, the evidence in the previous section shows only limited impacts of formalizing on intermediate channels that might affect firm growth and profitability. We turn to examining these outcomes directly in table 7. One important caveat to note here is that the limited impact the program had on formalization (even though this is large relative to the literature) lowers our power to find impacts of formalizing.

No significant impacts were measured on our main measures of business performances: the amount of sales, level of profits, number of employees and a summary index of sales and profits. Standard errors are however quite large. This is particularly the case when we examine levels of profits or sales as an outcome, given the long tails in these variables. For example, a 95 percent confidence interval for the impact on profits is (CFAF -36,000, CFAF +16,000), relative to a control mean of CFAF 54,000, so includes halving profits or up to a thirty percent gain in profits.

We therefore include other transforms of the data which are less sensitive to outliers, considering the inverse hyperbolic sine transformations of sales and profits (table A10), and plotting the cumulative distribution functions of profits and sales in Figure 2, and quantile regressions of the business profit effect in Figure 3. These confirm a lack of impact on profits and sales across the distribution. Likewise we see no significant impact on a summary standardized index of sales and profits, nor on employment.

However, formalization had a strong and significant negative impact on the likelihood of paying taxes, and on the amount of taxes paid. Newly formalized firms paid almost CFAF 19,000 (USD 32) less in taxes due to formalization. This result can be related to the previous result on tax

harassment, and similarly we note that it holds for all groups. In practice, all newly formalized firms appear to have benefited from the tax exemption, not only those who registered with the CGA as written in the law.

Should we expect this tax exemption to show up as higher profits? There are two ways it could have an effect. The first is a direct effect, as one less business expense. The total reduction in taxes paid is equivalent to 2.9 percent of average monthly profits. Second, if we consider the tax reduction as a windfall cash grant for the business which they re-invest, then even at a monthly return to capital of 5 percent (c.f. de Mel et al, 2008), this would have a FCFA 950 (USD 2) impact on monthly profit, which is equivalent to only 1.7 percent of the control group profits. So the potential impact on profitability through the tax channel is of the order of 4.6 percent, which lies well within our confidence interval for the treatment effect and is too small to detect.

Table A11 presented in the appendix shows the heterogeneous impact of formalization on business outcomes. All the heterogeneous variables used in this table were pre-specified in a pre-analysis plan registered before any follow-up data were collected. For each of this variable, we looked at heterogeneous impact on profits, on an index of profit and sales and on the number of employees. The point estimates and individually significant interactions suggest that formalization had more positive effects for businesses that are run by more educated owners and those which look more like formal firms to begin with. However, none of these interactions survive corrections for multiple testing.

8 Conclusions

Informality is the most common form of business operation in Benin. The new *entreprenant* status was introduced with the goal of offering a faster, cheaper, and easier way for small firms to become formal for tax purposes, and to enable them to access many of the potential benefits of being formal. When this status was introduced, there was a question as to whether the legal change was enough, by itself, to get informal firms to formalize, or whether additional efforts and services were needed.

Our randomized experiment tested three such approaches to encourage informal firms to take up the new *entreprenant* status. While few informal firms registered for this new status after

the legal status was launched, our interventions were successful in getting more informal firms to become formal. Personalized visits to firms coupled with an explanation of benefits and assistance filling out forms induced 9.6 percent of informal firms to formalize, and adding supplementary services in the form of access to business training, bank accounts, and tax mediation services increased this to 16.3 percent. Overall, the majority of businesses that formalized as a result of the program did it relatively soon after the first visit.

However, such efforts are costly, and we find that firms which formalize do not appear to benefit much from this status in the first two years afterwards. They access more business training and pay lower taxes due to a tax exemption, but are not more likely to have business bank accounts, gain new customers, have higher profits or sales, or hire additional workers. As such it appears that the costs of the program are large relative to the benefits for firms.

Our analysis also highlights the potential importance of targeting. The rate of formalization can be doubled by focusing interventions on firms with characteristics which place them closer to the margin of formalizing on their own. In Benin, we find these to be male-operated firms, run by more educated owners, operating outside of the main market and not in retail, as well as firms which we would ex ante classify as looking more similar to formal businesses. From a public policy perspective, we notice as positive outcome that firms are also more likely to formalize, since larger and more productive informal firms may be more likely to be competing with formal firms for customers, and would be liable for more tax payments. However, even with our suggested targeting, we estimate that the cost per firm formalized would still be several multiples of monthly profits for these firms. It may therefore be more cost effective to set in place the new, easy-to-register system, but then to directly pay firms to formalize, as suggested by de Mel et al. (2012), or rely on enforcement efforts to get targeted firms to become formal (Andrade et al, 2013).

References

Alcãzar, L., Andrade, R. and Jaramillo. M. (2010): "Panel/tracer study on the impact of business facilitation processes on enterprises and identification of priorities for future business enabling environment projects in Lima, Peru – Report 5: impact evaluation after the third round." Report to the International Finance Corporation. World Bank Group, Washington, DC.

Anderson, M. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects", *Journal of the American Statistical Association*, 103(484), 1481-1495

Andrade, G. H., Bruhn, M., and McKenzie, D. (2013): "A helping hand or the long arm of the law" Experimental evidence on what governments can do to formalize firms", *World Bank Economic Review*, forthcoming.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006): "Adaptive Linear Step-up Procedures that Control the False Discovery Rate", *Biometrika*, 93(3), 491-507

Bruhn, M. (2013): "A Tale of Two Species: Revisiting the Effect of Registration Reform on Informal Business Owners in Mexico" *Journal of Development Economics* 103: 275–83

Bruhn, M., McKenzie, D. (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments", *American Economic Journal: Applied Economics*, 1(4): 200-32.

Bruhn, M., McKenzie, D. (2014): "Entry regulation and formalization of microenterprises in developing countries", *World Bank Research Observer*, 29(2): 186-201.

Campos, F., Goldstein, M., and McKenzie, D., (2015): "Short-term impacts of formalization assistance and a bank information session on business registration and access to finance in Malawi." *Policy Research Working Paper Series* 7183, The World Bank.

De Giorgi, G., and R. Rahman. (2013): "SME's Registration: Evidence from an RCT in Bangladesh." Stanford University, Stanford, CA, and World Bank, Washington, DC.

De Giorgi, G., Ploenzke, M., and R. Rahman., (2015): "Small Firms' Formalization: The Stick Treatment." *Policy Research Working Paper Series* 7318, The World Bank.

De Mel, S., D. McKenzie, and C. Woodruff. (2013): "The demand for, and consequences of, formalization among informal firms in Sri Lanka." *American Economic Journal: Applied Economics* 5 (2): 122–50.

De Mel, S., D. McKenzie, and C. Woodruff. (2010): "Who are the Microenterprise Owners?: Evidence from Sri Lanka on Tokman v. de Soto" in *International Differences in Entrepreneurship*, J. Lerner and A. Schoar (eds.), pp.63-87, 2010.

De Mel, S., D. McKenzie, and C. Woodruff. (2008) "Returns to Capital: Results from a Randomized Experiment", *Quarterly Journal of Economics*, 123(4): 1329-72

De Soto, H. (1989): *The Other Path*, New York: Harper and Row Publishers.

Farrell, D. (2004). "The Hidden Dangers of the Informal Economy." *McKinsey Quarterly* 2004, no. 3:26–37.

Galiani, S., M. Meléndez and C. Navajas (2015) "On the Effect of the Costs of Operating Formally: New Experimental Evidence", NBER Working Paper no. 21292.

INSAE, (2009): "Comptes nationaux," INSAE, Cotonou.

La Porta, R., Shleifer, A. (2014): "Informality and Development", *Journal of Economic Perspectives* 28(3): 109-26.

La Porta, R. Shleifer, A. (2014): "The Unofficial Economy in Africa," NBER Chapters, in: African Successes: Government and Institutions National Bureau of Economic Research, Inc.

Levy, S. (2008): "Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico", *Washington: Brookings Institution Press*

Maloney, W. (2004): "Informality Revisited", *World Development* 32(7): 1159-78

McKenzie, D. (2012): "Beyond Baseline and Follow-up: The Case for more T in Experiments", *Journal of Development Economics*, 99(2): 210-21.

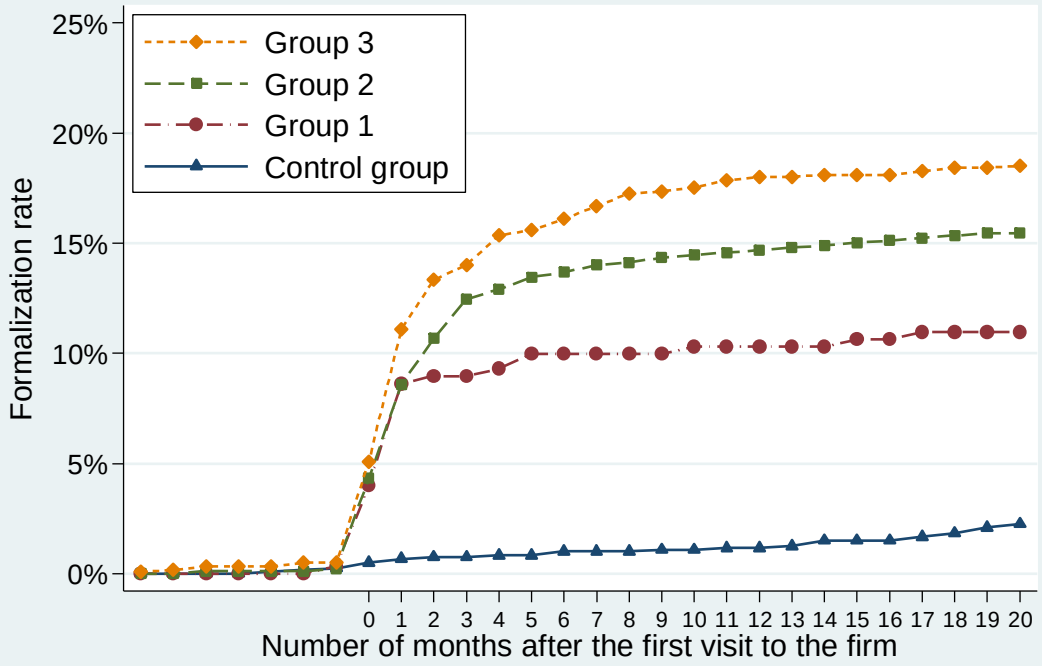
McKenzie, D., Woodruff, C. (forthcoming): "Business Practices in Small Firms in Developing Countries," *Management Science*, forthcoming.

McKenzie, D., Woodruff, C. (2014): "What are we learning from business training evaluations around the developing world?", *World Bank Research Observer*, 29(1): 48-82..

Perry, G., Maloney, G. Arias, O. Fajnzylber, P. Mason, and A., Saavedra, J. (2007): "Informality: Exit and Exclusion". World Bank Latin America and Caribbean Studies: World Bank, Washington D.C.

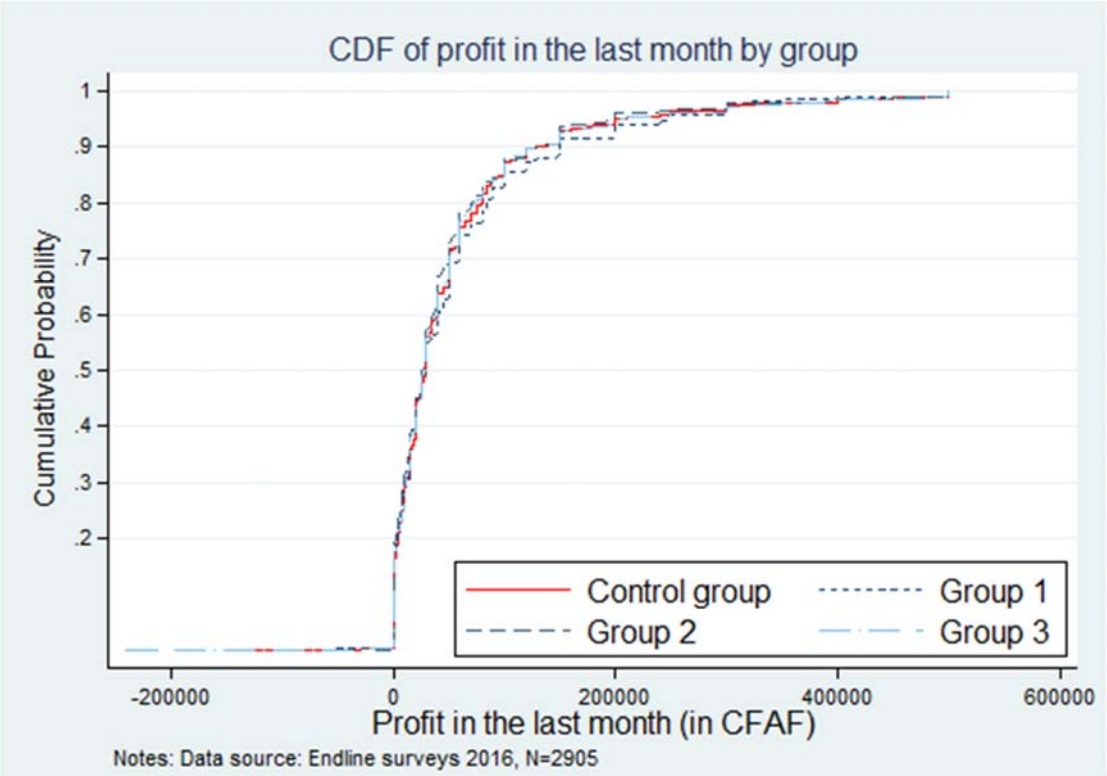
World Bank (2016): "Doing Business 2016: Measuring Regulatory Quality and Efficiency."

Figure 1: formalization rates over time



Notes: N=3,596
For the control group, date of visit 1 is set at the mode of the visit 1 date for other firms (3 months after program start)

Figure 2: Cumulative distributions of profits and sales at endline



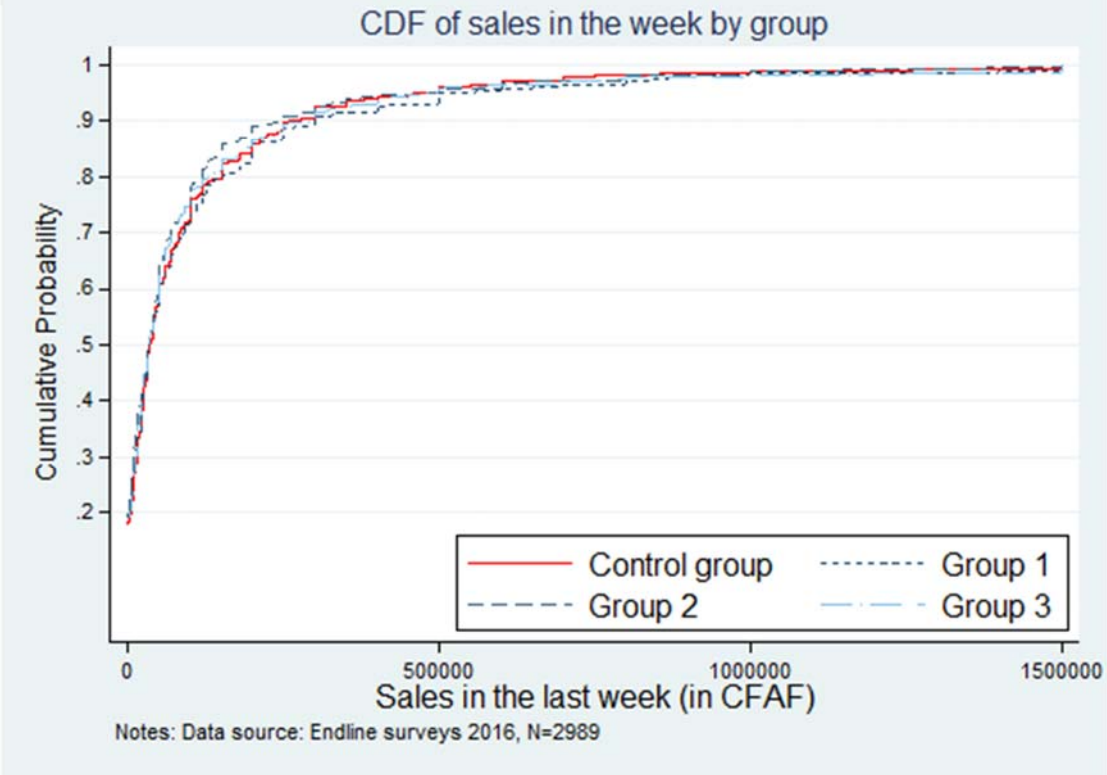
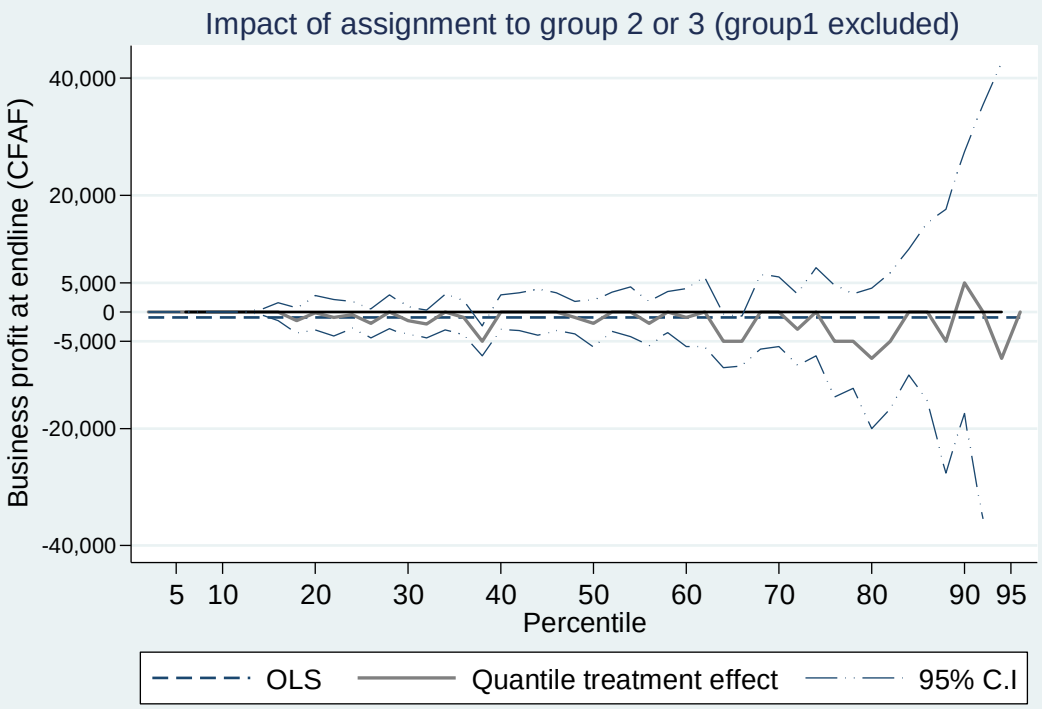


Figure 3: Quantile Regression on Business profit



Notes: Data source: Endline surveys 2016, N=2905

Table 1: Descriptive statistics on study population

	(1) SELECTED Sample		(2) All informal businesses		(3) Formal businesses	
	Mean [SD]	N	Mean [SD]	N	Mean [SD]	N
Firm owner characteristics						
Female owner	0.629 [0.483]	3,596	0.632 [0.482]	7,089	0.419 [0.494]	608
Age of the owner	39.5 [10.4]	3,557	39.4 [11.2]	6,955	43.6 [10.5]	589
Business owner has some formal education	0.712 [0.453]	3,591	0.708 [0.455]	7,081	0.884 [0.32]	606
Business owner has some secondary education	0.409 [0.492]	3,596	0.38 [0.486]	7,090	0.74 [0.439]	608
Firm characteristics						
Trade	0.55 [0.498]	3,596	0.518 [0.5]	7,090	0.584 [0.493]	608
Services	0.262 [0.44]	3,596	0.277 [0.447]	7,090	0.26 [0.439]	608
Craft	0.16 [0.366]	3,596	0.17 [0.375]	7,090	0.09 [0.287]	608
Firm area in m²	18.7 [43.5]	3,590	18.3 [50.8]	7,078	52.5 [106.5]	606
Business connected to electricity network	0.619 [0.486]	3,594	0.605 [0.489]	7,085	0.898 [0.303]	608
Number of employee	1.175 [1.687]	3,596	1.03 [1.603]	7,090	2.961 [4.59]	608
The firm does any form of accounting	0.179 [0.383]	3,594	0.156 [0.363]	7,089	0.642 [0.48]	604
Amount of sales in an average week	60,561 [56,508]	3,596	82,630 [298,695]	6,639	542,167 [4,434,990]	528
Amount of profit in the last month	46,698 [46,578]	3,596	46,434 [141,423]	6,358	223,041 [726,068]	490
Firm owner owns a bank account	0.222 [0.416]	3,514	0.194 [0.395]	6,928	0.789 [0.409]	582
Firm pays taxes	0.547 [0.498]	3,560	0.466 [0.499]	7,005	0.836 [0.371]	597
Amount of taxes paid in the previous year	18,732 [27,265]	3,482	16,649 [30,727]	6,827	316,636 [2,591,065]	533
Thinks that it's difficult to know in advance how much taxes she will have to pay	0.744 [0.437]	2,665	0.764 [0.424]	4,921	0.725 [0.447]	520
Ratio tax/ annual profit for all businesses	0.051 [0.089]	3,482	0.072 [0.174]	6,174	0.128 [0.221]	445
Ratio tax/ annual profit for businesses paying taxes	0.094 [0.104]	1,870	0.165 [0.286]	2,859	0.169 [0.313]	372

Notes: sources: listing-baseline survey March 2014

Table 2: Program Implementation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean [SD] in Control Group	Difference between […] and Control group				P-value for difference…			P-values joint tests G1=G2=G3=0
		Group 1	Group 2	Group 3	N	Group 1 and 2	Group 1 and 3	Group 2 and 3	
PANEL A: Administrative data from CGAs									
Step 1: First visit done successfully	0	0.991***	0.973***	0.976***	3,596	0.064*	0.085*	0.594	0.000***
	[0]	(0.009)	(0.006)	(0.005)					
Step 2: Second visit done successfully	0	-0.008	0.415***	0.466***	3,596	0.000***	0.000***	0.003***	0.000***
	[0]	(0.027)	(0.017)	(0.015)					
Step 3: Business is formal acording to the CGA$^{\beta}$	0	0.002	0.146***	0.171***	3,596	0.000***	0.000***	0.051*	0.000***
	[0]	(0.02)	(0.013)	(0.011)					
Step 4: Additional services:									
Business registered to CGAs	0	0.009	0.129***	0.154***	3,596	0.000***	0.000***	0.041**	0.000***
	[0]	(0.02)	(0.012)	(0.011)					
Business attended to at least one group training at CGAs	0	0.006	0.113***	0.141***	3,596	0.000***	0.000***	0.016**	0.000***
	[0]	(0.019)	(0.012)	(0.01)					
PANEL B: Endline survey data									
Formalization process: (only formal businesses)									
Number of days it took to formalize$^{\eta}$	22.9	-14.6	-21.6**	-15.6**	329	0.557	0.929	0.322	0.11
	[27.3]	(11.4)	(8.6)	(7.5)					
Amount paid for formalization	66,931	-73,289***	-57,145***	-56,552***	332	0.421	0.333	0.957	0.001***
	[57,036]	(19,449)	(15,343)	(13,302)					
Share of business who paid something to formalize	1	-0.788***	-0.693***	-0.773***	332	0.706	0.945	0.563	0.001***
	[0]	(0.246)	(0.194)	(0.168)					
Program Knowledge:									
Ever heard of the Entreprenant status	0.131	0.187***	0.207***	0.252***	2,582	0.668	0.114	0.083*	0.000***
	[0.338]	(0.041)	(0.026)	(0.023)					
Was able to explain what is the Entreprenant status	0.055	0.174***	0.148***	0.198***	2,582	0.508	0.476	0.023**	0.000***
	[0.228]	(0.034)	(0.022)	(0.02)					

Notes: Column 1: Standard deviations presented in brackets. Columns 2-4: coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on treatment dummies, controlling for strata dummies (dummies for each triplet). ***, **, * indicate statistical significance at 1, 5 and 10%. β: For the control group and group 1, CGA did not have any information as they are not following up with these businesses. η: Top-coded at the 99th percentile.

Table 3: Impact on Formalization

Dependent variables:	(1) Admin. Data (GUFE)	(2) Declared that the business is formal	(3) Showed a document	(4) formality or found in admin. data	(5) Showed a document or found in admin. data
Group 1	0.096***	0.066**	0.069***	0.107***	0.130***
	(0.023)	(0.026)	(0.024)	(0.029)	(0.029)
Group 2	0.130***	0.108***	0.093***	0.143***	0.146***
	(0.014)	(0.017)	(0.015)	(0.018)	(0.018)
Group 3	0.163***	0.128***	0.120***	0.176***	0.181***
	(0.013)	(0.015)	(0.013)	(0.016)	(0.016)
Observations	3,596	3,061	2,929	3,061	2,929
R-squared	0.392	0.436	0.453	0.446	0.464
Adjusted R-squared	0.086	0.072	0.075	0.090	0.094
Mean dependent variable in Control	0.023	0.052	0.026	0.059	0.040
Pvalue Test Group1=Group2	0.175	0.153	0.353	0.257	0.602
Pvalue Test Group1=Group3	0.003	0.017	0.028	0.015	0.075
Pvalue Test Group2=Group3	0.022	0.211	0.066	0.068	0.057
Pvalue Test Group1=Group2=Group3	0.002	0.037	0.026	0.016	0.049
Pvalue Test Group1=Group2=Group3=0	0.000	0.000	0.000	0.000	0.000

Note: Administrative data from GUFE and survey data May 2016. OLS regression of the outcome variable on treatment dummies, controlling for strata dummies (dummies for each triplet). ***, **, * indicate statistical significance at 1, 5 and 10%.

Table 4: Heterogeneous Impact on Formalization by Baseline Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variables:	**Formalized: GUFE data**							
Variable for heterogeneous analysis:	Female owner	Operates in Dantokpa market	Trader	Doesn't look like formal species	Index of business size below median	Does not have secondary education	One visit or fewer from tax inspectors	Female owner (sample restricted$^{\alpha}$)
Impact in group […] for heterogeneous variable=0								
Group1	0.130***	0.102***	0.142***	0.119**	0.077**	0.150***	0.123**	0.175***
	(0.035)	(0.026)	(0.033)	(0.054)	(0.032)	(0.035)	(0.050)	(0.045)
Group2	0.188***	0.151***	0.178***	0.207***	0.140***	0.182***	0.164***	0.228***
	(0.024)	(0.016)	(0.022)	(0.031)	(0.020)	(0.021)	(0.030)	(0.031)
Group3	0.198***	0.174***	0.186***	0.197***	0.146***	0.222***	0.196***	0.209***
	(0.021)	(0.014)	(0.019)	(0.027)	(0.018)	(0.018)	(0.026)	(0.027)
Additional impact in group […] for heterogeneous variable=1								
Group1 x Heterogenous variable (int1)	-0.064	-0.047	-0.099**	-0.032	0.029	-0.098**	-0.039	-0.078
	(0.047)	(0.055)	(0.046)	(0.059)	(0.046)	(0.045)	(0.055)	(0.073)
Group2 x Heterogenous variable (int2)	-0.091***	-0.099***	-0.087***	-0.094***	-0.022	-0.086***	-0.042	-0.108**
	(0.030)	(0.035)	(0.029)	(0.034)	(0.028)	(0.026)	(0.032)	(0.047)
Group3 x Heterogenous variable (int3)	-0.066**	-0.079**	-0.053**	-0.049	0.022	-0.112***	-0.049*	-0.049
	(0.026)	(0.031)	(0.026)	(0.030)	(0.025)	(0.022)	(0.028)	(0.042)
Observations	3,596	3,596	3,596	3,596	3,596	3,596	3,596	1,619
R-squared	0.398	0.398	0.398	0.398	0.396	0.405	0.397	0.404
Adjusted R-squared	0.094	0.094	0.095	0.093	0.091	0.105	0.092	0.096
Mean heterogenous variable	0.629	0.217	0.550	0.818	0.500	0.591	0.804	0.415
Mean dep. var. in Control heterogenous=0	0.047	0.030	0.028	0.072	0.043	0.055	0.050	0.038
Mean dep. var. in Control heterogenous=1	0.021	0.034	0.033	0.022	0.018	0.014	0.026	0.013
Pvalues of test: Heterogeneous=0								
Group1=Group2=Group3	0.158	0.012	0.403	0.334	0.110	0.071	0.343	0.584
Group1=Group2=Group=0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pvalues of test: Heterogeneous=1								
Group1+int1=Group2+int2=Group3+int3	0.020	0.303	0.004	0.010	0.017	0.157	0.024	0.347
Group1+int1=Group2+int2=Group3+int3=0	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000

Note: Administrative data from GUFE and survey data March 2015. OLS regression of the outcome variable on treatment dummies and interaction terms (treatment dummies X variable for heterogeneous analysis), controlling for strata dummies (dummies for each triplet). α : sample restricted to non-traders outside Tokpa market.***, **, * indicate statistical significance at 1, 5 and 10%.

Table 5: Cost Effectiveness Analysis

	In CFAF			In USD		
	Group 1	**Group 2**	**Group 3**	**Group 1**	**Group 2**	**Group 3**
Program costs:						
Total Program costs	**21 304 850**	**154 397 653**	**195 493 401**	**35 746**	**259 056**	**328 009**
Costs by intervention:						
One-stop-shop for formalization	6 325 293	18 975 879	25 301 172	10 613	31 839	42 452
Interventions to increase take up	14 979 557	135 421 774	170 192 229	25 133	227 218	285 557
Costs by types:						
Total set up costs	5 728 222	36 001 489	45 733 290	9 611	60 405	76 734
Total variable costs	15 576 628	118 396 164	149 760 111	26 135	198 651	251 275
Cost per formalization						
Number of businesses	301	899	1199	301	899	1199
Program impact:						
Impact on formalization (in pp)	9,6%	13,0%	16,3%	9,6%	13,0%	16,3%
Number of firms which formalized because of the program	29	117	195	29	117	195
Total costs…						
… per business included in treatment	70 780	171 744	163 047	119	288	274
… per formalization	737 294	1 321 106	1 000 289	1 237	2 217	1 678
Variable costs…						
… per business included in treatment	51 750	131 698	124 904	87	221	210
… per formalization	539 058	1 013 059	766 283	904	1 700	1 286
Cost per formalization with targetting (see Table 4)						
Targeting firms that looked more like formal firms before head (18% of firms)						
Impact on formalization (in pp)	11,9%	20,7%	19,7%	11,9%	20,7%	19,7%
Variable costs per formalization:	434 871	636 220	634 031	730	1 067	1 064
Targeting firm owners with secondary education (41% of firms)						
Impact on formalization (in pp)	15,0%	18,2%	22,2%	15,0%	18,2%	22,2%
Variable costs per formalization:	344 997	723 613	562 631	579	1 214	944
Targeting firms outside Dantokpa, with a male owner, with secondary education. (16% of firms)						
Impact on formalization (in pp)	17,5%	27,7%	29,1%	17,5%	27,7%	29,1%
Variable costs per formalization:	295 712	475 443	429 224	496	798	720
Targeting firms with a bank account. (22% of firms)						
Impact on formalization (in pp)	20,0%	23,7%	22,3%	20,0%	23,7%	22,3%
Variable costs per formalization:	258 748	555 686	560 108	434	932	940

Notes: 1 USD = 596 CFAF (exchange rate on June 1st, 2016). See Table A6 for more details on program costs.

Table 6 : Impact on intermediate outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Has a bank account[B]	Loan contracted in 2014-16[B] (bank or MFI)	Attended business training in the past year	The firm does any form of account-ing[B]	Share of business practises implement-ed (26 questions)	Index of tax harass-ment[λ]	Has sold goods to the public administration or to a large company (last 3 months)	In the last month a client asked for a receipt[B]
1st stage: impact of treatment allocation:								
Group1 X year1 (b1)	0.028	-0.030	0.008	-0.099***	-0.046***	-0.057	0.014	-0.030
	(0.031)	(0.024)	(0.018)	(0.026)	(0.012)	(0.040)	(0.023)	(0.029)
Group2 X year1 (b2)	-0.008	-0.018	0.081***	0.007	-0.004	-0.053**	0.008	-0.027
	(0.018)	(0.014)	(0.014)	(0.017)	(0.008)	(0.024)	(0.014)	(0.018)
Group3 X year1 (b3)	0.017	-0.009	0.112***	0.023	0.004	-0.030	0.026**	0.004
	(0.016)	(0.013)	(0.012)	(0.015)	(0.008)	(0.022)	(0.013)	(0.016)
Group1 X year2 (c1)	0.054*	0.006	0.023	-0.052*	-0.018	-0.067*	-0.002	-0.026
	(0.031)	(0.025)	(0.019)	(0.028)	(0.015)	(0.039)	(0.022)	(0.031)
Group2 X year2 (c2)	0.011	0.051***	0.113***	0.020	-0.005	-0.031	-0.006	-0.017
	(0.019)	(0.016)	(0.015)	(0.018)	(0.009)	(0.025)	(0.013)	(0.019)
Group3 X year2 (c3)	0.003	0.015	0.145***	0.047***	0.007	-0.066***	0.008	-0.014
	(0.016)	(0.014)	(0.013)	(0.016)	(0.008)	(0.021)	(0.013)	(0.016)
Observations	6,211	6,215	5,949	6,166	6,169	5,217	5,361	5,394
Mean Dep. var in control year1	0.249	0.13	0.033	0.198	0.262	0.008	0.093	0.234
Mean Dep. var in control year2	0.257	0.173	0.056	0.234	0.273	00	0.083	0.25
Adjusted R-squared	0.147	0.191	0.140	0.155	0.099	0.133	0.089	0.252
Test for impact constant…								
…accross treatments, year1 (b1=b2=b3)	0.378	0.610	0.000	0.000	0.000	0.604	0.515	0.191
…accross treatments, year 2 (c1=c2=c3)	0.285	0.085	0.000	0.003	0.178	0.398	0.606	0.923
Coef. are jointly 0 (b1=b2=b3=c1=c2=c3=0)	0.561	0.000	0.000	0.000	0.003	0.037	0.388	0.538
(IV) impact of Formalization:								
Formalization instrumented by 1st stage	0.053	0.031	0.669***	0.152**	0.009	-0.255***	0.068	-0.059
treatment variables	(0.074)	(0.061)	(0.048)	(0.066)	(0.034)	(0.091)	(0.050)	(0.067)
P-values	0.469	0.613	0.000	0.022	0.782	0.005	0.172	0.377
Sharpened two-stage q-values[μ]	0.755	1	0.001	0.071	1	0.023	0.402	0.755

Note : Panel data from midline and endline surveys in 2015 and 2016. All regressions are controlling for strata dummies (dummies for each triplet). Standard errors (in parentheses) are clustered at the firm level. α: truncated at the 99th percentile. B: controling for baseline value. μ: Sharpened two-stage q-values as described in Anderson (2008) using P-values in table 6 and 7. λ: summary index of the following questions: "Was asked to pay a bribe by a tax inspector in the last 6 months"; "Received a sexual suggestion or other inappropriate request from a tax inspector in the last 6 months"; "Was threatened with business closure by a tax inspector in the last 6 months"; " Received more than 1 visit by a tax inspector in the last 6 months"; "Received at least one visit by a labour or hygiena inspector "; "Feel that he/she paid more taxes than he/she should have paid according to the law"; "Thinks that tax officials override their duty and ask firms to pay too much taxes". ***, **, * indicate statistical significance at 1, 5 and 10%

Table 7 : Impact on firm performances

	(1) Total sales in the last day$^{\alpha\beta}$ (CFAF)	(2) Total sales in the last week$^{\alpha\beta}$ (CFAF)	(3) Last month profit$^{\alpha\beta}$ (CFAF)	(4) Summary index of sales and profit$^{\alpha\beta}$	(5) Total number of employees$^{\alpha}$	(6) Any tax paid for business activity in 2015^{B}	(7) Sum of all taxes paid in 2015^{B} (CFAF)
1st stage: impact of treatment allocation:							
Group1 X year1 (b1)	2,228	12,496	-8,053*	0.008	-0.22**	0.013	-19
	(2,754)	(14,029)	(4,798)	(0.057)	(0.10)	(0.030)	(1,747)
Group2 X year1 (b2)	540	-7,376	-3,016	-0.052*	-0.06	0.048***	-51
	(1,451)	(7,312)	(3,021)	(0.031)	(0.09)	(0.018)	(1,091)
Group3 X year1 (b3)	-114	-1,224	-3,106	-0.010	-0.11	0.005	-2,041**
	(1,384)	(6,399)	(2,858)	(0.030)	(0.08)	(0.016)	(949)
Group1 X year2 (c1)	602	12,192	470	0.041	-0.09	-0.066**	-3,308**
	(2,930)	(14,243)	(5,742)	(0.060)	(0.10)	(0.030)	(1,678)
Group2 X year2 (c2)	1,246	-5,235	-874	-0.007	0.05	-0.055***	-3,413***
	(1,832)	(8,010)	(3,377)	(0.036)	(0.07)	(0.018)	(1,047)
Group3 X year2 (c3)	1,847	3,998	242	0.026	0.08	-0.067***	-5,967***
	(1,669)	(7,911)	(3,233)	(0.035)	(0.07)	(0.017)	(869)
Observations	5,918	6,043	5,874	5,926	6,206	6,163	6,096
Mean Dep. var in control year1	17,373	99,984	53,313	-0.02	1.14	0.507	18,856
Mean Dep. var in control year2	17,882	106,803	54,536	-0.003	1.23	0.413	14,221
Adjusted R-squared	0.227	0.260	0.159	0.254	0.350	0.356	0.257
Test for impact constant…							
…accross treatments, year1 (b1=b2=b3)	0.684	0.401	0.593	0.406	0.457	0.085	0.168
…accross treatments, year 2 (c1=c2=c3)	0.908	0.449	0.957	0.670	0.303	0.826	0.024
Coef. are jointly 0 (b1=b2=b3=c1=c2=c3=0)	0.873	0.796	0.438	0.447	0.111	0.000	0.000
(IV) impact of Formalization:							
Formalization instrumented by 1st stage treatment variables	4,718	-1,877	-10,235	-0.008	-0.12	-0.127*	-18,789***
	(6,511)	(31,925)	(13,388)	(0.143)	(0.30)	(0.075)	(4,463)
P-values	0.469	0.953	0.445	0.957	0.687	0.091	0.000
Sharpened two-stage q-values$^{\mu}$	0.755	1	0.755	1	1	0.251	0.001

Note : Panel data from midline and endline surveys in 2015 and 2016. All regressions are controlling for strata dummies (dummies for each triplet). Standard errors (in parentheses) are clustered at the firm level. α: truncated at the 99th percentile. B: controling for baseline value. μ: Sharpened two-stage q-values as described in Anderson (2008) using P-values in table 6 and 7. ***, **, * indicate statistical significance at 1, 5 and 10%

APPENDICES

Figure A1: Sampling strategy and Survey completion rates

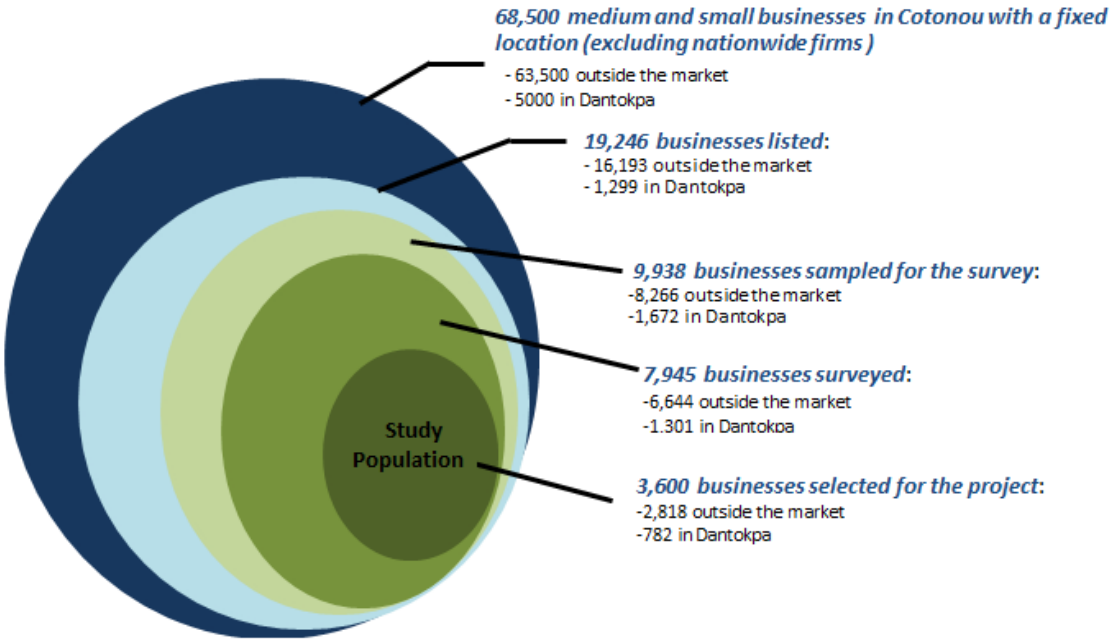


Figure A2: Evaluation Design

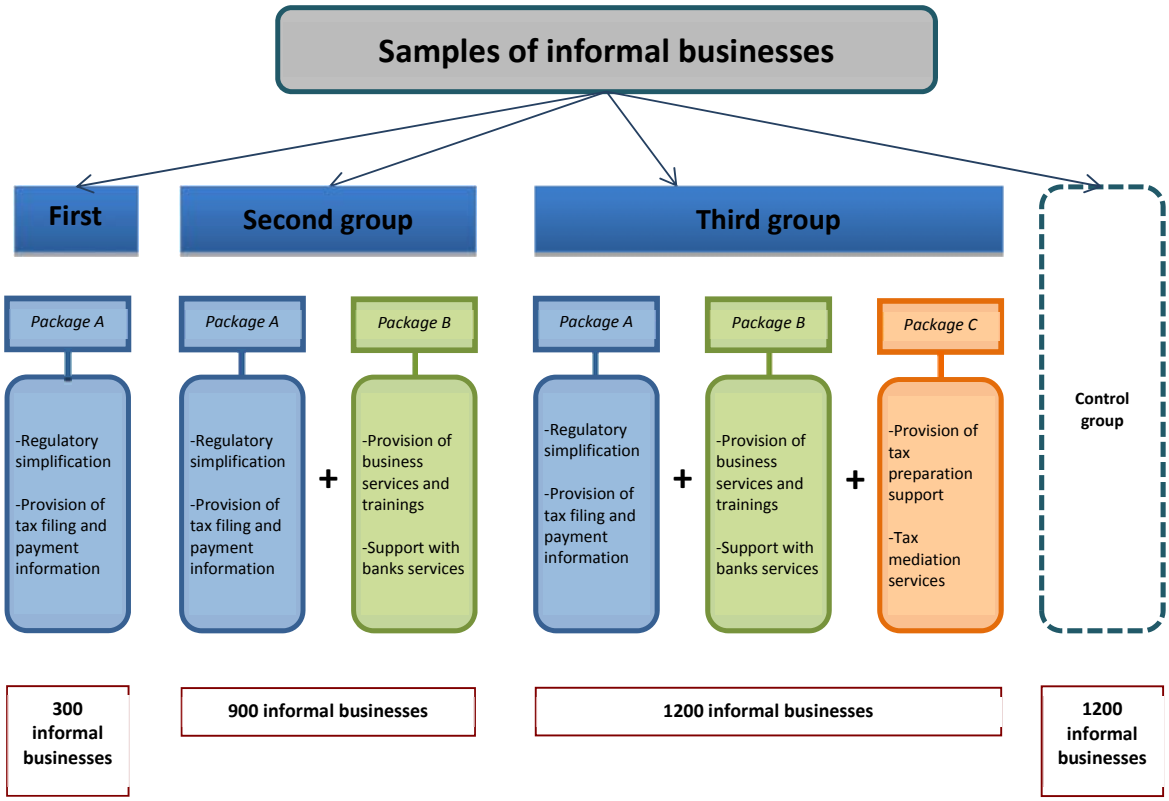


Table A1 : Potential Benefits and Costs of Formalization in Benin

	(1)	(2)	(3)	(4)	(5)
Type of status	**Informal (fixed location)**	**Entreprenant status**	**Individual enterprises**	**Limited liability company**	**Trader card[1]**
Cost of the status	n.a.	Free of charge	CFAF 10,000 (USD 17)	CFAF 17,000 (USD 29)	Extra CFAF 5,000 (USD 8)
Time to register	n.a.	1 day	1 day	1 day	1 day
Maximum Turnover	n.a.	CFAF 30 million for traders. CFAF 20 million for craftmen, CFAF 10 million for services	No	No	No
Needs to pay taxes	55% pay taxes	Yes with tax exemptions after formalization[2]	Yes with tax exemptions after formalization[2]	Yes with tax exemptions after formalization[2]	Yes with tax exemptions after formalization[2]
Open business bank account	No (or difficult)	Yes	Yes	Yes	Yes
Apply to a bank loan	Requires collateral	Yes	Yes	Yes	Yes
Export license	No	No	No	Yes (need to get an export card)	Yes (need to get an export card)
Can work with large private companies	Possible but complicate	Yes	Yes	Yes	Yes
Access to large public contract	No	No	No	Yes	Yes
Access to small public contract	No	Yes	Yes	Yes	Yes
Registered at the chamber of Commerce	No	Yes	Yes	Yes	Yes
Provide invoices to customers for tax purposes	No	Yes	Yes	Yes	Yes
Register more than one activity for the firm	n.a.	No	Yes	Yes	Yes

Notes: [1]: For the trader card, businesses also need to get the individual enterprise status. [2]: Businesses that formalized, registered with CGA, and that had not paid taxes before, have a tax exemption for the first year after formalization, in addition to a reduction of 40% in the amount of taxes due for the following 3 years. 1 USD = 596 CFAF (exchange rate on

Table A2: Balance Checks among study population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Difference between […] and Control group				P-value for difference…			P-values joint tests
	Mean [SD] in Control Group	Group 1	Group 2	Group 3	N	Group 1 and 2	Group 1 and 3	Group 2 and 3	G1=G2=G3 =0
Firm owner characteristics									
Female owner	0.63	0.001	0.000	-0.002	3,596	0.717	0.363	0.415	0.63
	[0.483]	(0.003)	(0.002)	(0.002)					
Age of the owner	39.25	1.13	0.56	-0.2	3,557	0.482	0.068*	0.094*	0.113
	[10.75]	(0.73)	(0.46)	(0.41)					
Firm owner has at least some formal education	0.707	-0.034	-0.004	0.026	3,591	0.396	0.064*	0.143	0.158
	[0.456]	(0.032)	(0.02)	(0.018)					
Firm characteristics									
Trade	0.551	-0.007*	0.000	0.000	3,596	0.104	0.071*	1	0.316
	[0.498]	(0.004)	(0.002)	(0.002)					
Services	0.259	0.007	-0.001	0.007	3,596	0.76	0.999	0.585	0.936
	[0.438]	(0.023)	(0.014)	(0.013)					
Craft	0.165	-0.008	-0.01	-0.008	3,596	0.928	0.97	0.921	0.883
	[0.371]	(0.023)	(0.014)	(0.013)					
Firm area in m²	19.02	5.99*	-0.21	-2.31	3,590	0.076*	0.008***	0.282	0.055*
	[42.13]	(3.15)	(1.96)	(1.75)					
Business connected to electricity network	0.617	-0.001	0.008	0.002	3,594	0.811	0.941	0.759	0.984
	[0.486]	(0.034)	(0.021)	(0.019)					
Number of employee	1.18	-0.04	-0.02	0.02	3,596	0.824	0.518	0.521	0.861
	[1.68]	(0.09)	(0.06)	(0.05)					
The firm does any form of accounting	0.173	0.013	-0.001	0.015	3,594	0.632	0.917	0.306	0.678
	[0.378]	(0.027)	(0.017)	(0.015)					
Amount of sales in an average week	60,828	73	281	-1,052	3,596	0.934	0.618	0.342	0.755
	[57,039]	(2,257)	(1,405)	(1,255)					
Amount of profit in the last month	46,249	139	316	1,156	3,596	0.941	0.634	0.527	0.791
	[44,867]	(2,135)	(1,329)	(1,188)					
Firm owner owns a bank account	0.22	0.004	0.004**	0.005***	3,514	0.985	0.771	0.665	0.016**
	[0.414]	(0.003)	(0.002)	(0.002)					
Taxes									
Firm pays taxes	0.561	-0.013	-0.017	-0.022	3,560	0.927	0.8	0.807	0.687
	[0.496]	(0.033)	(0.021)	(0.018)					
Amount of taxes paid in the previous year	19,450	-3,867**	-306	-657	3,482	0.1	0.101	0.77	0.269
	[28,146]	(1,952)	(1,211)	(1,077)					

Notes: Baseline survey data (March 2014). Column 1: Standard deviations presented in brackets. Columns 2-4: coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on treatment dummies, controlling for strata dummies (dummies for each triplet). ***, **, * indicate statistical significance at 1, 5 and 10%.

Table A3: Follow-up Surveys Attrition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mean [SD] in Control Group	Difference between […] and Control group				P-value for difference…			P-values joint tests G1=G2=G3 =0
		Group 1	Group 2	Group 3	N	Group 1 and 2	Group 1 and 3	Group 2 and 3	
Two years follow-up survey results (May-June 2016):									
Surveyed and business still operating	0.712	0.006	0.012	0.009	3,596	0.878	0.928	0.896	0.937
	[0.453]	(0.033)	(0.02)	(0.018)					
Surveyed with short phone survey and business still operating	0.037	0.005	-0.007	0.004	3,596	0.404	0.929	0.179	0.599
	[0.188]	(0.014)	(0.009)	(0.008)					
Surveyed and business shut down	0.081	-0.006	0.015	0.008	3,596	0.359	0.495	0.588	0.635
	[0.273]	(0.021)	(0.013)	(0.012)					
Surveyed and business owner deceased	0.012	0.015**	-0.006	-0.004	3,596	0.012**	0.011**	0.691	0.058*
	[0.108]	(0.007)	(0.005)	(0.004)					
Survey attrition (refused, not found, sickness, traveling, maternity leave..)	0.159	-0.02	-0.014	-0.018	3,596	0.819	0.912	0.817	0.61
	[0.366]	(0.026)	(0.016)	(0.014)					
Including refused to answer	0.083	0.001	-0.013	-0.005	3,596	0.531	0.758	0.533	0.759
	[0.276]	(0.02)	(0.012)	(0.011)					
One year follow-up survey results (April-May 2015):									
Surveyed and business still operating	0.811	0.008	-0.023	-0.013	3,596	0.328	0.458	0.579	0.575
	[0.392]	(0.029)	(0.018)	(0.016)					
Surveyed with short phone survey and business still operating	0.01	0.015	0.002	0.014***	3,596	0.184	0.898	0.031**	0.025**
	[0.1]	(0.009)	(0.006)	(0.005)					
Surveyed and business shut down	0.057	-0.002	0.008	0.007	3,596	0.607	0.635	0.877	0.844
	[0.232]	(0.018)	(0.011)	(0.01)					
Surveyed and business owner deceased	0.004	0	0.001	-0.003	3,596	0.908	0.51	0.206	0.57
	[0.065]	(0.004)	(0.003)	(0.002)					
Survey attrition (refused, not found, sickness, traveling, maternity leave..)	0.118	-0.023	0.012	-0.005	3,596	0.179	0.459	0.227	0.517
	[0.322]	(0.023)	(0.014)	(0.013)					
Including refused to answer	0.066	-0.007	0.006	0.009	3,596	0.504	0.382	0.831	0.749
	[0.248]	(0.018)	(0.011)	(0.01)					

Notes: Column 1: Standard deviations presented in brackets. Columns 2-4: coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on treatment dummies, controlling for strata dummies (dummies for each triplet). ***, **, * indicate statistical significance at 1, 5 and 10%. Sample sizes by group are the following: control group: 1,197, group1: 301, group 2: 899, group 3: 1,199.

Table A4: Balance Checks among Businesses Surveyed at second follow-up survey

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Difference between […] and Control group				P-value for difference…			P-values joint tests
	Mean [SD] in Control Group	Group 1	Group 2	Group 3	N	Group 1 and 2	Group 1 and 3	Group 2 and 3	G1=G2=G3 =0
Firm owner characteristics									
Female owner	0.627	0.002	-0.001	-0.001	3,064	0.55	0.403	0.787	0.836
	[0.484]	(0.003)	(0.002)	(0.002)					
Age of the owner	39.37	0.93	0.55	-0.01	3,035	0.671	0.246	0.276	0.407
	[10.58]	(0.81)	(0.51)	(0.46)					
Firm owner has at least some formal education	0.724	-0.049	-0.009	0.015	3,061	0.309	0.069*	0.28	0.256
	[0.447]	(0.035)	(0.023)	(0.02)					
Firm characteristics									
Trade	0.536	-0.006	0.001	0.002	3,064	0.085*	0.036**	0.779	0.206
	[0.499]	(0.004)	(0.003)	(0.002)					
Services	0.262	0.017	0.006	0.005	3,064	0.712	0.655	0.956	0.915
	[0.44]	(0.026)	(0.017)	(0.015)					
Craft	0.176	-0.015	-0.023	-0.005	3,064	0.781	0.707	0.278	0.523
	[0.381]	(0.026)	(0.017)	(0.015)					
Firm area in m²	19.4	6.39*	0.5	-2.56	3,059	0.16	0.017**	0.199	0.084*
	[43.52]	(3.77)	(2.39)	(2.14)					
Business connected to electricity network	0.617	-0.016	0.006	0.003	3,063	0.602	0.616	0.9	0.961
	[0.486]	(0.038)	(0.024)	(0.021)					
Number of employee	1.22	-0.07	-0.05	0.07	3,064	0.849	0.177	0.072*	0.21
	[1.7]	(0.11)	(0.07)	(0.06)					
The firm does any form of accounting	0.174	-0.007	0.002	0.005	3,062	0.804	0.688	0.843	0.976
	[0.379]	(0.03)	(0.019)	(0.017)					
Amount of sales in an average week	59,792	174	968	-1,237	3,064	0.776	0.572	0.163	0.543
	[56,781]	(2,500)	(1,586)	(1,414)					
Amount of profit in the last month	46,563	648	-47	-979	3,064	0.79	0.487	0.529	0.821
	[45,839]	(2,345)	(1,488)	(1,327)					
Firm owner owns a bank account	0.227	0.005	0.004*	0.005**	2,998	0.805	0.915	0.545	0.066*
	[0.419]	(0.003)	(0.002)	(0.002)					
Taxes									
Firm pays taxes	0.555	-0.015	-0.027	-0.033	3,037	0.776	0.642	0.816	0.454
	[0.497]	(0.038)	(0.024)	(0.021)					
Amount of taxes paid in the previous year	19,779	-4,285**	-657	-1,413	2,976	0.131	0.184	0.575	0.227
	[28,779]	(2,160)	(1,369)	(1,214)					

Notes: Baseline survey data (March 2014). Only businesses surveyed at second follow-up are included. Column 1: Standard deviations presented in brackets. Columns 2-4: coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on treatment dummies, controlling for strata dummies (dummies for each triplet). ***, **, * indicate statistical significance at 1, 5 and 10%.

Table A5: Impact of the "leaflets intervention"

	(1) (IV) Formalized after May 2016 (Admin Data)
Received a leaflet	0.010
(Instrumented by assignment to leaflet group)	(0.006)
Observations	1,197
R-squared	0.050
Share of firm selected in Leaflet group that received it	0.706
Formalization rate in control (no leaflet) group	0.003

Note: Administrative data from GUFE: May, June and July 2016) and survey folow up data 2016. IV regressions, controlling for strata dummies (used to randomize the "leaflets intervention"). ***, **, * indicate statistical significance at 1, 5 and 10%.

Table A6: Implementation costs

	(1)	(2)	(3)	(4)	(5)
	TOTAL COSTS		**Share**	**Share**	**Share**
	In CFAF	**In USD**	**Group 1**	**Group 2**	**Group 3**
Costs associates to GUFE (one-stop-shop for formalization)					
Set up costs:					
-Set up costs of Hardware and Sofware: 3 computers including softwares and buying and setting up the serveur	4 500 000	7 550	12,5%	37,5%	50%
-Investments required to edit Entreprenant cards: printer and software	12 500 000	20 973	12,5%	37,5%	50,0%
Variable costs:					
-Salary of GUFE entreprenant staffs : Two full-time staffs for 18 months, and one for 12 months	21 600 000	36 242	12,5%	37,5%	50%
-Hardware and sofwares: Maintenance for 2 years	6 812 344	11 430	12,5%	37,5%	50%
-Office supplies	5 190 000	8 708	12,5%	37,5%	50%
Costs associates to the CGA (implementing agency)					
Set up costs:					
- Initial training of CGA advisors	13 450 000	22 567	5,1%	42,0%	52,8%
- Motobikes for CGA advisors (including insurance)	19 200 000	32 215	5,1%	42,0%	52,8%
- Office and mobile phones (one phone for each advisor)	1 250 000	2 097	5,1%	42,0%	52,8%
- Hardware and sofware: computers, serveurs and	35 553 000	59 653	5,1%	42,0%	52,8%
- Set up costs of upgrading CGA office	1 010 000	1 695	5,1%	42,0%	52,8%
Variable costs:					
- Salary of CGA staffs for 2 years: CGA supervisor, 24 advisors (10 of them only for 14 months) and one hotline	175 144 000	293 866	5,1%	42,0%	52,8%
- CGA overheads: printing, administrative fees, office supplies, water and electricity	12 979 560	21 778	5,1%	42,0%	52,8%
- Transportation costs: gazoline	18 600 000	31 208	5,1%	42,0%	52,8%
- Communication	11 160 000	18 725	0,0%	44,3%	55,7%
- Maintenance of hardware and sofware:	6 147 000	10 314	5,1%	42,0%	52,8%
- Office rent	9 600 000	16 107	5,1%	42,0%	52,8%
- Group trainings organization: office supplies and coffee	16 500 000	27 685	0,0%	44,3%	55,7%
Total Set up costs	87 463 000	146 750	6,5%	41,2%	52,3%
Total Variable costs	283 732 904	476 062	5,5%	41,7%	52,8%
Total costs of program implementation	371 195 904	622 812	5,7%	41,6%	52,7%

Notes: Data on costs and group allocation from CGA and GUFE. 1 USD = 596 CFAF (exchange rate on June 1st, 2016)

Table A7: Baseline characteristics of formal businesses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	firms at listing survey	**Newly formalized firms in…** Control group	Group 1	Group 2	Group 3	col (1)=col (3),(4),(5)	col (2)=col (3),(4),(5)
Number of observations	608	27	33	140	222		
Firm owner characteristics							
Female owner	0.419	0.444	0.333	0.493	0.518	0.000***	0.622
	[0.494]	[0.506]	[0.479]	[0.502]	[0.501]		
Age of the owner	43.6	38.8	40	38.5	38.9	0.000***	0.959
	[10.5]	[9.2]	[9.8]	[8.7]	[8.9]		
Business owner has some formal education	0.884	0.889	0.818	0.836	0.856	0.000***	0.545
	[0.32]	[0.32]	[0.392]	[0.372]	[0.352]		
Business owner has some secondary education	0.74	0.704	0.576	0.529	0.581	0.000***	0.151
	[0.439]	[0.465]	[0.502]	[0.501]	[0.494]		
Firm characteristics							
Trade	0.584	0.556	0.273	0.436	0.473	0.161	0.257
	[0.493]	[0.506]	[0.452]	[0.498]	[0.5]		
Services	0.26	0.296	0.394	0.279	0.275	0.959	0.91
	[0.439]	[0.465]	[0.496]	[0.45]	[0.447]		
Craft	0.09	0.074	0.212	0.257	0.216	0.000***	0.058*
	[0.287]	[0.267]	[0.415]	[0.439]	[0.413]		
Firm area in m²	52.5	41.3	42.9	23.5	17.8	0.000***	0.125
	[106.5]	[103.1]	[154.2]	[47.4]	[34.8]		
Business connected to electricity network	0.898	0.704	0.758	0.75	0.743	0.000***	0.62
	[0.303]	[0.465]	[0.435]	[0.435]	[0.438]		
Number of employee	2.961	1.185	1.455	1.379	1.423	0.000***	0.559
	[4.59]	[1.388]	[1.416]	[1.668]	[2.196]		
The firm does any form of accounting	0.642	0.222	0.303	0.257	0.243	0.000***	0.721
	[0.48]	[0.424]	[0.467]	[0.439]	[0.43]		
Amount of sales in an average week	542,167	89,500	61,930	56,768	54,001	0.000***	0.001***
	[4,434,990]	[61,392]	[57,387]	[46,359]	[53,366]		
Amount of profit in the last month	223,041	56,185	55,145	50,719	46,702	0.000***	0.4
	[726,068]	[45,206]	[46,173]	[42,550]	[44,246]		
Firm owner owns a bank account	0.789	0.423	0.406	0.418	0.33	0.000***	0.569
	[0.409]	[0.504]	[0.499]	[0.495]	[0.471]		
Firm pays taxes	0.836	0.667	0.531	0.584	0.534	0.000***	0.244
	[0.371]	[0.48]	[0.507]	[0.495]	[0.5]		
Amount of taxes paid in the previous year	316,636	31,262	20,358	26,795	22,841	0.000***	0.28
	[2,591,065]	[33,019]	[36,866]	[34,633]	[31,662]		
Thinks that it's difficult to know in advance how much taxes she will have to pay	0.725	0.708	0.708	0.673	0.651	0.66	0.654
	[0.447]	[0.464]	[0.464]	[0.471]	[0.478]		
Ratio tax/ annual profit for all businesses	0.128	0.076	0.047	0.066	0.058	0.000***	0.39
	[0.221]	[0.105]	[0.083]	[0.104]	[0.088]		
Ratio tax/ annual profit for businesses paying taxes	0.169	0.117	0.091	0.116	0.11	0.000***	0.813
	[0.313]	[0.11]	[0.097]	[0.115]	[0.095]		

Notes: sources: listing-baseline survey March 2014

Table A8: Mechanisms explaining take up on formalization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Mean [SD] Control Group	Difference between [...] and Control group				P-value for difference...			P-values of joint test	P-values of joint test
		Group 1	Group 2	Group 3	N	G1 and G2	G1 and G3	G2 and G3	G1=G2=G3	G1=G2=G3=0
PANEL A: formal businesses (according to survey): two main reasons why registered:										
Being able to open a bank account/ It is easier to get a loan	0.196	-0.14	0.189	-0.037	367	0.179	0.621	0.091*	0.203	0.358
	[0.401]	(0.24)	(0.19)	(0.164)						
To comply with the law/get access to the legal system/not been fined or asked for bribes	0.478	-0.118	-0.273	-0.199	367	0.567	0.726	0.615	0.822	0.635
	[0.505]	(0.268)	(0.212)	(0.183)						
It gives access to new markets (public administration and large companies)	0.413	-0.503**	-0.223	-0.313*	367	0.239	0.35	0.481	0.492	0.147
	[0.498]	(0.233)	(0.185)	(0.16)						
Better reputation/social acceptance for the business	0.217	0.069	-0.057	-0.102	367	0.582	0.388	0.721	0.649	0.783
	[0.417]	(0.227)	(0.18)	(0.156)						
It gives access to government or NGO program (including the CGA)	0.043	0.545**	0.31	0.415**	367	0.346	0.543	0.435	0.601	0.079*
	[0.206]	(0.246)	(0.195)	(0.168)						
Including:										
It gives access to CGA benefits	0	0.328	0.19	0.326*	367	0.568	0.994	0.3	0.58	0.223
	[0]	(0.239)	(0.189)	(0.164)						
Other	0.152	0.331**	0.244*	0.235**	367	0.595	0.492	0.913	0.784	0.131
	[0.363]	(0.161)	(0.127)	(0.11)						
PANEL B: informal businesses: two main reasons why not registered:										
Answered that she is going to formalize soon	0.176	0.003	0.003	0.007	2,217	0.998	0.919	0.876	0.985	0.991
	[0.381]	(0.039)	(0.025)	(0.023)						
High registration costs / registration process complicate or time consuming	0.309	0.065	-0.005	0.014	2,217	0.196	0.293	0.558	0.433	0.571
	[0.462]	(0.048)	(0.031)	(0.028)						
Doesn't see any benefits of formalization / the business is too small	0.319	-0.104**	-0.016	0.016	2,217	0.106	0.014**	0.321	0.039**	0.079*
	[0.466]	(0.048)	(0.031)	(0.028)						
It increases the amount of taxes to be paid/ risk of tax inspection	0.284	-0.052	-0.004	-0.027	2,217	0.352	0.585	0.461	0.611	0.599
	[0.451]	(0.046)	(0.03)	(0.027)						
Doesn't have enough information on formalization	0.089	0.015	-0.018	-0.047***	2,217	0.295	0.03**	0.136	0.048**	0.02**
	[0.285]	(0.028)	(0.018)	(0.017)						
Doesn't have a legal ID	0.006	0.021*	0.000	0.014*	2,217	0.121	0.55	0.09*	0.163	0.104
	[0.079]	(0.012)	(0.008)	(0.007)						
More paperwork / it requires to do accounting	0.014	0.021	0.015	0.008	2,217	0.745	0.41	0.49	0.604	0.244
	[0.116]	(0.014)	(0.009)	(0.009)						
More corruption	0.005	-0.011	-0.002	-0.002	2,217	0.257	0.187	0.937	0.416	0.459
	[0.07]	(0.007)	(0.004)	(0.004)						
Husband forbid it	0	0.001	0.003	0.003	2,217	0.711	0.749	0.889	0.933	0.667
	[0]	(0.004)	(0.003)	(0.003)						
Other	0.015	0.033**	-0.003	0.014	2,217	0.035**	0.217	0.095*	0.08*	0.077*
	[0.121]	(0.015)	(0.01)	(0.009)						

Notes: Data from second follow-up survey (June 2016). Column 1: Standard deviations presented in brackets. Columns 2-4: coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on treatment dummies, controlling for strata dummies (dummies for each triplet). ***, **, * indicate statistical significance at 1, 5 and 10%.

Table A9 : Impact on other intermediate outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Has done any type of advertising in the last 6 months[B]	Standardized index of business presentation[η]	Value of inventories and raw materials[α]	Actual value of all investments done in the firm[α]	Number of customers in a typical week[α]	Has gained a new regular customer in the past 3 months	Introduced a new product or services in the last 12 months	Standard-ized index of trust in institutions	Subjective standard of living on a Cantril ladder[λ]	Anticipated Subjective standard of living in 5 years on a Cantril ladder[λ]
1st stage: impact of treatment allocation:										
Group1 X year1 (b1)	-0.021	-0.016	-81,855	-492,949***	-3.33	-0.034	0	0.077	-0.046	-0.033
	(0.026)	(0.053)	(118,108)	(175,080)	(4.46)	(0.031)	0	(0.081)	(0.147)	(0.139)
Group2 X year1 (b2)	-0.037**	-0.025	-141,562*	-116,823	-3.63	-0.015	0	-0.001	-0.177*	-0.034
	(0.016)	(0.033)	(72,254)	(114,176)	(2.98)	(0.019)	0	(0.049)	(0.094)	(0.082)
Group3 X year1 (b3)	-0.031**	-0.008	-20,556	-102,287	-1.74	-0.020	0	0.068	-0.051	-0.080
	(0.014)	(0.029)	(66,740)	(101,130)	(2.62)	(0.017)	0	(0.043)	(0.086)	(0.076)
Group1 X year2 (c1)	0.037	-0.020	507,258**	306,979	-1.36	0.002	0.025	0	0.170	0.239*
	(0.029)	(0.050)	(243,549)	(256,853)	(4.74)	(0.031)	(0.035)	0	(0.173)	(0.145)
Group2 X year2 (c2)	0.016	-0.059*	364,556***	206,022	-6.07**	0.008	-0.012	0	0.189*	0.285***
	(0.018)	(0.031)	(122,084)	(139,484)	(2.96)	(0.020)	(0.023)	0	(0.099)	(0.080)
Group3 X year2 (c3)	0.002	-0.026	449,110***	143,176	-1.98	-0.003	0.003	0	0.279***	0.171**
	(0.015)	(0.027)	(127,183)	(127,976)	(2.59)	(0.018)	(0.021)	0	(0.091)	(0.074)
Observations	5,390	4,367	4,503	5,102	5,071	5,357	2,561	2,623	5,294	5,109
Mean Dep. var in control year1	0.134	00	572,740	1,693,253	46.46	0.799	.	-0.002	4.602	8.638
Mean Dep. var in control year2	0.153	-0.012	1,151,506	1,971,251	45.23	0.805	0.162	.	4.845	8.950
Adjusted R-squared	0.067	0.092	0.293	0.281	0.133	0.059	0.051	0.033	0.092	0.041
Test for impact constant…										
…accross treatments, year1 (b1=b2=b3)	0.849	0.900	0.232	0.088	0.820	0.854	0.000	0.369	0.461	0.869
…accross treatments, year 2 (c1=c2=c3)	0.455	0.588	0.820	0.805	0.421	0.877	0.621	0.000	0.660	0.442
Coef. are jointly 0 (b1=b2=b3=c1=c2=c3=0)	0.007	0.658	0.000	0.000	0.566	0.732	0.812	0.336	0.000	0.000
(IV) impact of Formalization:										
Formalization instrumented by 1st stage	-0.078	-0.150	679,408*	52,453	-15.20	-0.055	-0.000	0.243	0.352	0.323
treatment variables	(0.062)	(0.117)	(362,873)	(478,538)	(10.78)	(0.070)	(0.099)	(0.196)	(0.356)	(0.286)
P-values	0.213	0.200	0.061	0.913	0.158	0.437	0.999	0.214	0.323	0.259
Sharpened two-stage q-values[μ]	0.749	0.749	0.749	0.76	0.749	0.749	0.76	0.749	0.749	0.749

Note : Panel data from midline and endline surveys in 2015 and 2016. All regressions are controlling for strata dummies (dummies for each triplet). Standard errors (in parentheses) are clustered at the firm level. **α:** top-coded at the 99th percentile. **B:** controling for baseline value. **η:** the standardized summary index includes the following questions: "Is the business premise generally well organized?", "Is the premise generally clean and in good shape?", Are there posters or pictures advertizing some products or services in particular?", "Are prices of merchandizes visible inside the premise?", "Are commodities grouped by type?", "Are commodities globally clean and in good shape?" (last three questions were only asked to traders). **λ**: The Cantril ladder goes from 0 to 10 with 10 for the best situation possible.**μ**: Sharpened two-stage q-values as described in Anderson (2008) ***, **, * indicate statistical significance at 1, 5 and 10%

Table A10 : Impact on other measures of business performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Inverse hyperbolic of Total sales in the last day$^{\alpha\beta}$	**Inverse hyperbolic of Total sales in the last week$^{\alpha\beta}$**	**Inverse hyperbolic of Last month profit$^{\alpha\beta}$**	**Above the 95th percentile of the control group weekly sales distribution**	**Above the 95th percentile of the control group profit distribution**	**Hired someone in the last 6 months**	**Number of hours worked in the business last week by the owner**
1st stage: impact of treatment allocation:							
Group1 X year1 (b1)	0.298	-0.073	0.135	-0.041***	-0.002	0.000	-7.30***
	(0.345)	(0.327)	(0.302)	(0.015)	(0.015)	(0.044)	(1.98)
Group2 X year1 (b2)	0.454**	0.160	-0.073	0.003	0.008	0.004	-2.85**
	(0.210)	(0.190)	(0.189)	(0.009)	(0.009)	(0.027)	(1.14)
Group3 X year1 (b3)	0.023	0.018	-0.038	0.000	0.005	0.055**	-4.12***
	(0.188)	(0.175)	(0.172)	(0.009)	(0.009)	(0.024)	(1.06)
Group1 X year2 (c1)	-0.972***	-0.294	-0.243	0.009	0.006	0.017	1.17
	(0.368)	(0.326)	(0.310)	(0.019)	(0.018)	(0.028)	(1.75)
Group2 X year2 (c2)	-0.212	-0.423**	-0.492**	0.017*	-0.001	-0.016	0.89
	(0.218)	(0.204)	(0.199)	(0.010)	(0.010)	(0.017)	(1.06)
Group3 X year2 (c3)	-0.165	-0.193	-0.258	0.008	0.001	0.009	1.84*
	(0.192)	(0.178)	(0.176)	(0.009)	(0.010)	(0.016)	(0.98)
Observations	5,918	6,043	5,874	5,408	5,410	4,081	5,422
Mean Dep. var in control year1	6.646	9.753	9.727	0.046	0.049	0.226	61.75
Mean Dep. var in control year2	6.292	9.509	9.303	0.056	0.053	0.118	66.7
Adjusted R-squared	0.125	0.138	0.123	0.131	0.148	0.077	0.112
Test for impact constant…							
…accross treatments, year1 (b1=b2=b3)	0.166	0.728	0.825	0.024	0.839	0.232	0.129
…accross treatments, year 2 (c1=c2=c3)	0.106	0.585	0.538	0.701	0.946	0.344	0.704
Coef. are jointly 0 (b1=b2=b3=c1=c2=c3=0)	0.002	0.052	0.105	0.038	0.974	0.226	0.000
(IV) impact of Formalization:							
Formalization instrumented by 1st stage	-0.156	-0.545	-1.027	0.028	0.019	0.081	-7.35*
treatment variables	(0.832)	(0.803)	(0.814)	(0.035)	(0.036)	(0.066)	(4.36)
P-values	0.851	0.498	0.208	0.411	0.596	0.220	0.092
Sharpened two-stage q-values$^{\mu}$	1	1	1	1	1	1	1

Note : Panel data from midline and endline surveys in 2015 and 2016. All regressions are controlling for strata dummies (dummies for each triplet). Standard errors (in parentheses) are clustered at the firm level. **α:** top-coded at the 99th percentile. **B:** controling for baseline value. **μ**: Sharpened two-stage q-values as described in Anderson (2008) ***, **, * indicate statistical significance at 1, 5 and 10%

Table A11: Panel data: heterogeneous impact of formalization on firm outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Heterogenous variables:	**Female owner**			**Operating in Dantokpa market**			**Doesn't look like formal species**			**Does not have secondary education**			**Index of business size below median**		
Dependent variables:	Profits$^{\alpha}$	Index of profits and sales	Number of emplo-yees$^{\alpha}$	Profits$^{\alpha}$	Index of profits and sales	Number of emplo-yees$^{\alpha}$	Profits$^{\alpha}$	Index of profits and sales	Number of emplo-yees$^{\alpha}$	Profits$^{\alpha}$	Index of profits and sales	Number of emplo-yees$^{\alpha}$	Profits$^{\alpha}$	Index of profits and sales	Number of emplo-yees$^{\alpha}$
Impact of Formalization on dep. var. for heterogeneous variable=0															
Formalized (Gufe data)	-16,140	0.039	0.29	-15,696	-0.225*	-0.09	-47,233	0.015	-0.46	-22,447	0.111	-0.44	-11,342	-0.097	-0.05
(Instrumented by treatment assignment)	(19,827)	(0.177)	(0.49)	(12,898)	(0.124)	(0.34)	(29,063)	(0.295)	(0.77)	(18,721)	(0.189)	(0.37)	(22,347)	(0.228)	(0.54)
Sharpened two-stage q-values$^{\mu}$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Impact of Formalization on dep. var. for heterogeneous variable=1															
Formalized (Gufe data) x Heterog. var.	10,419	-0.100	-0.86	66,261	1.978**	-0.05	53,158	-0.009	0.46	27,451	-0.277	0.73	6,628	0.154	-0.16
(Instrumented by $Group_i$ x $year_j$ x Heterog.	(26,919)	(0.280)	(0.62)	(59,162)	(0.781)	(0.63)	(33,261)	(0.340)	(0.85)	(31,897)	(0.347)	(0.71)	(26,586)	(0.284)	(0.62)
Sharpened two-stage q-values$^{\mu}$	1	1	1	1	0.52	1	1	1	1	1	0.982	1	1	1	1
Observations	5,874	5,926	6,206	5,874	5,926	6,206	5,874	5,926	6,206	5,874	5,926	6,206	5,874	5,926	6,206
R-squared	0.328	0.406	0.471	0.326	0.370	0.474	0.326	0.407	0.474	0.327	0.406	0.471	0.329	0.404	0.474
Mean heterogenous variable	0.621	0.622	0.624	0.193	0.196	0.198	0.823	0.824	0.820	0.594	0.593	0.59	0.506	0.505	0.502
Formalization rate in Control hetero=0	0.037	0.037	0.035	0.025	0.024	0.024	0.053	0.052	0.049	0.039	0.038	0.036	0.031	0.03	0.03
Formalization rate in Control hetero=1	0.014	0.013	0.014	0.015	0.013	0.014	0.016	0.015	0.016	0.011	0.01	0.012	0.014	0.014	0.014
Mean Outcome control year1 hetero=0	60,623	-0.067	1.606	45,481	-0.124	1.316	81,683	0.241	1.744	62,789	0.023	1.074	68,219	0.159	1.669
Mean Outcome control year2 hetero=0	62,985	-0.037	1.646	50,227	-0.085	1.347	83,543	0.232	1.792	62,399	0.011	1.248	68,881	0.16	1.703
Mean Outcome control year1 hetero=1	48,918	0.009	0.856	83,291	0.393	0.433	46,930	-0.078	0.996	46,554	-0.05	1.179	39,120	-0.188	0.617
Mean Outcome control year2 hetero=1	49,424	0.018	0.986	72,117	0.326	0.771	47,799	-0.057	1.101	48,966	-0.012	1.221	40,281	-0.166	0.759

Note : Panel data from midline and endline surveys in 2015 and 2016. All regressions including control for baseline values of the dependent variable (if available) and strata dummies (dummies for each triplet). α: truncated at the 99th percentile. μ: Sharpened two-stage q-values as described in Anderson (2008). ***, **, * indicate statistical significance at 1, 5 and 10%

Appendix 1: Details of Intervention Implementation

The advisors from CGA delivered the program to each business owner following four main steps:

(1) First visit: A CGA advisor conducted a first visit to each business to explain the benefits of becoming an *entreprenant*, specific by group, and to distribute informational leaflets. If a business owner was not present on the day of the visit, the CGA advisor attempted to call the owner on the phone. If the owner could not be reached, the CGA advisor made another attempt by trying a visit or a call in different moments of the day. After four attempts (visits or calls), the business was considered as not interested.

(2) Second visit: For businesses receiving package B, the same CGA advisor called, arranged, and confirmed a meeting, which took place approximately two weeks after the first visit, and provided 1-2 hours of personalized training. If a business owner was not present on the planned day of the second visit or could not be reached, the CGA advisor made another attempt by trying a visit or call in different moments of the day. After 3 attempts (visits or calls), the business was considered not interested. Registration at GUFE was not mandatory to be eligible to this second visit. For those also receiving package C, the CGA advisor devoted additional time in reviewing the procedures to calculate the taxes to pay, and the option of receiving tax mediation help, if necessary.

(3) Formalization decision: After having received the first and/or the second visit, business owners decided whether or not to register as *entreprenants* at GUFE.

(4) Provision of additional benefits: Businesses in treatment groups 2 and 3 could also register with CGA, and receive counselling and business training (group sessions). Businesses in group 3 could benefit from tax mediation services with CGA, if needed. Finally, businesses could open a bank account with specific conditions at BoA or Orabank.

Appendix 2: Details of Sampling Procedure

Sampling protocols for inside and outside the market were different:

- For Danktopa market, we used a precise map of the market made by the public company managing markets in Benin (SOGEMA). This map allowed to divide geographically the market in small areas. We then randomly selected areas in the markets in which 50% of the businesses (with fixed location) where sampled for the survey.[16]
- For other neighborhoods of Cotonou, we were able to obtain detailed maps of each of the 144 neighborhoods in Cotonou. Those maps allowed the easy identification of *ilots* (blocks), the official administrative unit within a neighborhood. We used this administrative unit as a reference for the listing survey sampling. We then used information given by the tax administration (and confirmed by the survey company) in order to characterize neighborhoods as high or low firm density areas. We randomly sampled 38% of *ilots* in high density neighborhoods and 10% of the *ilots* in low density neighborhoods. In each *ilots* 68% of businesses where sampled for the survey in average.

Overall, 19,246 businesses were listed. The listing survey allowed us to estimate the total number of businesses operating in Cotonou (with a fixed location, excluding international and nationwide businesses and liberal professions) to approximately 68,500, including around 5,000 in Dantokpa market.[17] Among those 19,246 businesses, 9,938 businesses were randomly selected to be surveyed. 7,945 (80%) businesses were successfully surveyed, 1,000 (10%) businesses refused to be surveyed, and 995 (10%) businesses were dropped because the business owner was not available or not reached after 4 attempts. Figure A1 details the listing survey results inside and outside the market.

From the 7,945 businesses surveyed, a population of 3,596 businesses was then selected to participate in the study based on the following goals:

- Drop businesses already formal
- Drop businesses that will probably not cooperate in the future or which will be probably difficult to find (i.e. businesses that refused to provide information on profits or turnover during baseline survey)

[16] Few areas were excluded from the sampling frame because they almost exclusively included businesses selling illegal products (i.e. taint oil, medicine, and voodoo products) or by large formal businesses.
[17] Some sections of Dantokpa market were not included in the listing survey. Therefore, the total number of firms in Dantokpa is probably significantly higher.

- Trim the database from (a) businesses very close to formalization who would have formalized anyway and (b) businesses very far from formalization which would not be interested by the program
- Remove businesses that ever got a loan from a commercial bank that will most probably not been interested by the program (less than 3% of informal businesses)
- Reduce the standard deviation of the main outcomes (profit and turnover)[18]
- Include a sufficient number of businesses in Dantokpa market.[19]

[18] Outside Dantokpa market we excluded businesses with sales or profit lower than CFAF 12,000 (USD 20), profit greater than CFAF 150,000 (USD 252) or sales greater than CFAF 400,000 (USD 671). In Dantokpa market, we excluded businesses with sales or profit lower than CFAF 10,000 (USD 17), profit greater than CFAF 200,000 (USD 336) or sales greater than CFAF 500,000 (USD 839).

[19] We choose to take 22% of the total study population from Dantokpa market to have the same share of businesses from the market as in the 2008 firm census (INSEA, 2008).

Appendix 3: Matching program data to administrative data on formalization

This appendix describes the protocol to match businesses in the administrative database on formalization provided by the GUFE (around 550 businesses every month) with the program data (3,596 informal businesses prior to the program start).

Information available:

We had the following information in both databases:

- Surname of the business owner. It can be written with different spellings in each database.
- Between 1 and 5 first names. In GUFE data we usually have more than one first name. In the program data we only have one first name in most of the cases.
- The business activity as described by the owner (no codes). The business activity is missing for 30% of businesses in the GUFE data.
- Business addresses. In the GUFE data addresses were given by the business owner whereas in the program data, we are using "official" addresses used by the tax administration (there are 144 neighborhoods in Cotonou). In practice only neighborhoods can be matched. In GUFE data, there are few missing variables and some cases for which the neighborhood does not belong to the official list of neighborhoods.
- Gender of the business owner.
- Phone number of the business owner.

Definition of a match:

We consider it to be a match if: (i) the surname, at least one first name, the activity, and the neighborhood match or (ii) if surname and at least one first name match and either the activity, the neighborhood or the phone number also match, and the others are missing (or does not exist for the neighborhood or the phone number).

Method of matching:

We used first the STATA command "reclinck" designed for fuzzy name matching. This command uses record linkage methods to generate matching scores. For this first step, we used tree variables: surname, first name and gender. As a second step, we looked manually (in an

Excel file) to all matches and validate each match only if names, activity and neighborhood were consistent in both databases.

The "reclink" command allows inputting different weights to match on each of the three variables used (surname, first name and gender). In order not to rely on the weights used, we reiterate the process with different weights until no additional matches were found.

Since it is possible that the first name in the program data corresponds to the second name in the GUFE data, we also reiterate the whole process for all combinations of first to fifth names. Surname and first name were inverted in one of the two databases. So we also reiterate the process with other combination of surname and first names.

Checking that the matching method is working:

To assess whether our matching method is working efficiently, we used the following methods:

1. First we looked at whether we could find additional matches using a more usual method of matching. That is looking at the two lists (sorted by surnames) and trying to find each business of the GUFE data in the program database. So it means looking mainly at businesses with surnames starting with the same letters. We were not able to find any additional matches.
2. Secondly, we looked at the proportion of business which formalized with the *entreprenant* status during the first 3 months after program launch. Indeed, most businesses which formalized with the *entreprenant* status should be also in the program data (in theory they are the only businesses aware of this new status). We matched 78% (119/153) of the newly registered *entreprenants*.
3. We then took the 34 businesses which formalized with the *entreprenant* status and were not matched with the program data and we tried harder to match these businesses. We took the program data and looked at all the businesses in the same neighborhood as the unmatched businesses. We were able to find 6 new matches. These 6 matches are very imperfect matches with surnames somewhat different and first name sometimes different. For two cases, the match was not done with our main method because the surname is missing in the program data.

As a result of this process, we use a conservative definition of a match between the business names in the two databases as "two businesses with a close surname, and at least one close

first name, and either the same phone number, or the same sector of activity and an address in the same neighborhood." Using this definition, the likelihood that a business in the program database was considered as formal, whereas the business was in reality not formal, was very low. The opposite case (i.e. a business was considered as informal, whereas it is in reality formal) is however possible, so this measure of formalization may underestimate the actual number of businesses which formalized in all groups. We therefore also supplement the administrative measure of formalization with survey measures.