

Jessen, Jonas; Kluge, Jochen

Article — Accepted Manuscript (Postprint)

The effectiveness of interventions to reduce informality in low- and middle-income countries

World Development

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Jessen, Jonas; Kluge, Jochen (2021) : The effectiveness of interventions to reduce informality in low- and middle-income countries, World Development, ISSN 0305-750X, Elsevier BV, Amsterdam, Vol. 138, <https://doi.org/10.1016/j.worlddev.2020.105256> , <https://www.sciencedirect.com/science/article/abs/pii/S0305750X20303831?via%3Dihub>

This Version is available at:

<https://hdl.handle.net/10419/266399>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0>

The Effectiveness of Interventions to Reduce Informality in Low- and Middle-Income Countries

Jonas Jessen*
Jochen Kluge†

2nd November 2020

Abstract

Labor markets in low- and middle-income countries are characterized by high levels of informality. A multitude of interventions have been implemented to increase the formalization of firms and workers, including information campaigns, simplified registration procedures, reductions of payroll taxes, and interventions enforcing formalization. We compile a database of 170 impact estimates from 38 academic studies that evaluate empirically one or more of these formalization interventions. The quantitative analysis correlates the impact estimates of the studies — given as a measure of sign and statistical significance, the effect size or percent impacts — with explanatory factors such as intervention type, outcome variable, scope of the intervention, and contextual factors such as study design and economic context. Several key findings emerge: first, tax incentives are a particularly effective intervention type, and information interventions constitute an important component to combine with other interventions. Second, the outcome "worker registration" shows significantly better results than other outcomes. Third, large-scale policies are more effective on average than singular programs. Fourth, interventions tend to be more effective in a more promising (labor) market context.

Keywords: Formalization · Firm Registration · Labor Registration · Impact Evaluation

© <2022>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

*Corresponding author, DIW Berlin and Free University Berlin, jjessen@diw.de, Mohrenstr. 58, 10117 Berlin, Germany. +49-30-89789-265

†Humboldt University Berlin and KfW, jochen.kluve@kfw.de

1 Introduction

High levels of informality are a central feature of labor markets in low- and middle-income countries. Whereas a precise definition of informality is challenging to give — Maloney (2004) defines it as "[b]roadly speaking, the small-scale, semi-legal, often low-productivity, frequently family-based, perhaps pre-capitalistic enterprise" — the fact that either firms or workers are not registered with the tax or social security system, and therefore not formal participants of the labor market, is regarded as a key concern.¹ Specifically, there are several reasons why policymakers would worry about informality: first, it reduces the tax base, thus negatively impacting the provision of public and private goods as well as reducing the state capacity for other redistributive policies. Secondly, it may lead to an inefficient allocation of resources, as informal firms are commonly much less productive than formal ones despite small differences in the human capital of their workers (La Porta and Shleifer, 2014). Third, informal workers are not covered by any of the institutions — such as pension systems, health insurance, etc. — that protect formal employees.²

Figure 1 displays shares of informal employment (as a percentage of non-agricultural employment) in a series of countries in Sub-Saharan Africa (left hand side, blue color), in Latin America (center, gray color), and Asia (right hand side, red color). The figure illustrates that levels of informality are generally pervasive, and differ somewhat between major regions: in Latin America, these shares are in the range of 23.6 (Uruguay, 2017) and 77.3% (Bolivia, 2015), and thus lower than in Sub-Saharan Africa — where the majority of countries have shares higher than 70% — and Asia. Notwithstanding these regional differences and the general downward trend in informality since the early 2000s, the figure shows that overall levels of informal employment have remained high across low- and middle-income countries.

In order to address this persistent challenge, a multitude of policies and programs have been implemented in many countries with the aim to increase the formalization of firms or workers, or both. Among initiatives to formalize businesses have been, for instance, the implementation of one-stop shops for business registration and the simplification of payroll taxes and social security contributions (e.g. Bruhn, 2011; Fajnzylber et al., 2011, for Mexico). Other approaches concern information interventions, e.g. in-

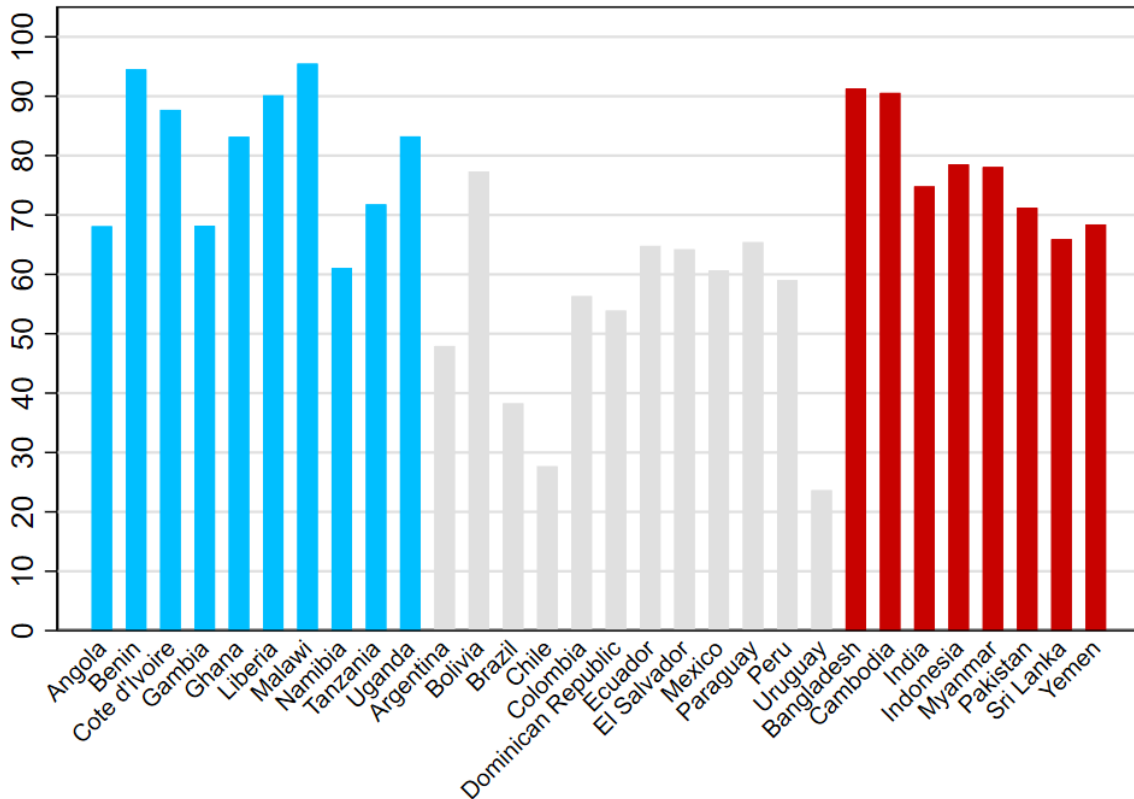


Figure 1: Informal employment as % of non-agricultural employment

Note: Informal employment relative to all non-agricultural employment in a given country. Informal employed includes all jobs in unregistered and/or unincorporated businesses. Percentages reported are from 2018 or latest year available. Data from The World Bank.

formation campaigns that explicate the step-by-step procedures and potential benefits of business registration (De Giorgi and Rahman, 2013, for Bangladesh). Also programs that reduce the costs of business registration have been considered and put into practice (Alcázar and Jaramillo, 2016, for Peru), as are financial mechanisms in which a bonus payment is given to firms who are willing to register (de Mel et al., 2013, for Sri Lanka). Finally, a potential policy alternative to incentive-based approaches are interventions that enforce business formalization (e.g. De Giorgi et al., 2018, for Bangladesh).

The second type of approaches targets the formalization of labor, such as the registration of workers. This has included, e.g., tax reduction and bureaucracy simplification policies such as SIMPLES in Brazil (Monteiro and Assunção, 2012). Other cost-reducing approaches include reductions in payroll taxes (e.g. Bernal et al., 2017, for Colombia) or the simplification of labor registration (e.g. Ronconi and Colina, 2011, for Argentina). Finally, also when targeting workers the enforcement of formalization legislation through

labor inspections is a policy option (e.g. Pignatti, 2018, for Colombia). Clearly, several of the approaches mentioned here and in the above paragraph are potentially combinable into a multi-component approach.

Given the policy relevance of labor market informality and this large spectrum of interventions aiming to increase formality, the empirical evaluation of these interventions is of key interest to policy makers, to learn about the effectiveness of formalization initiatives. Some of the earlier evidence specifically on the formalization of small firms is reviewed in Bruhn and McKenzie (2014) and Khamis (2014), who find that approaches focusing on the ease of formalization alone will not induce most informal firms to become formal, while increased enforcement of rules can increase formality. They also find that entry-reform type of interventions result in only a modest increase in the number of formal firms (at best), and, when looking at the "limited existing evidence base" at the time, call for further efforts to evaluate formalization policies and a broader perspective on intervention types, including credit and labor market policies (Bruhn and McKenzie, 2014).

Therefore, several additional policy questions are of key concern: what is known about which type of "formalization" intervention works? Is there a difference between the short-run and long-run effects? Are different outcomes affected differentially? Is it easier to "formalize" firms or labor? Does the scope of the intervention play a role? That is, are interventions more effective when implemented at scale ("policies", typically nationwide and permanent) or when implemented for narrowly defined groups, regions, or sectors ("programs", small-scale and often one-off)? In this paper, we address these questions and through this contribute to the academic and policy debate on formalization interventions. By providing a systematic overview of the existing evidence and analyzing patterns of effectiveness by various explanatory factors, the findings can be of tentative guidance for policy makers when designing such interventions.

We start with the compilation of a database of available empirical studies that assess the effects of interventions that reduce informality, in low- and middle-income countries. We then first provide a summary overview of these studies and a descriptive analysis of their findings. In a second step, we implement an empirical analysis of this database of "formalization interventions" in the spirit of a meta analysis (e.g. Stanley and Doucouli-

agos, 2012; Grimm and Paffhausen, 2015; Card et al., 2018) — that is, we quantitatively analyze the patterns of intervention effectiveness by intervention type, outcome, time horizon, and scope of the intervention, as well as contextual factors including study design features and the economic context. This quantitative approach — while of explorative nature given the state of the literature in this field — provides new evidence on patterns of effectiveness of formalization interventions across low- and middle-income countries.

The following Section 2 delineates a framework distinguishing five categories of formalization interventions that have been implemented in practice. The section also explicates the compilation of the database. Section 3 presents a descriptive statistical analysis of the data, while Section 4 explains and implements the quantitative approach to investigate correlates of the effectiveness of formalization interventions. We implement sign/significance and effect size meta regression models, and we discuss and correct for publication bias. Section 5 concludes.

2 A typology and database of formalization interventions

2.1 Typology

Formality is a multi-dimensional concept and can rarely be regarded as a simple dichotomous yes / no measure. From the firm perspective, it may involve to have a business registration number, being registered for taxes or being registered at the local council; aspects which are separated in many countries (see, e.g., Campos et al., 2018). For workers, formality can mean simply being registered as workers by the formal firm, or the firm actually making required social security contributions. To account for the various aspects of formality, which differ by country and are often hard to address holistically, in this paper we consider any program or policy ("formalization intervention") that is aimed at increasing an aspect of formality. To adequately distinguish between different forms of formalization interventions that have been implemented and evaluated in practice, we identify five main types classifiable as follows:

- (i) information interventions

- (ii) simplification / registration interventions
- (iii) tax incentives / social security reduction
- (iv) labor inspection / enforcement interventions
- (v) financial incentives

First, *information interventions* provide informal firms — and/or would-be entrepreneurs — with information regarding (a) the registration process, and (b) the benefits of registration. The latter includes e.g. protection of the business name, (better) access to bank loans, limited liabilities, greater ownership rights, and the enhancement of social status. Information interventions are typically not provided as stand-alone interventions, but are often combined with simplification / registration interventions. The studies in our data (see below) that analyze this intervention type primarily stem from experimental settings. De Giorgi and Rahman (2013) use a randomized controlled trial (RCT) in which informal firms were informed about a recent registration reform in Bangladesh that effectively reduced the duration of the registration of a business from 42 days to one day. The study focuses on randomized exposure to the information about the reform (not on the reform itself) and finds that one year after the treatment the self-reported knowledge about the reform increased among small and medium enterprises, but actual registration did not. The authors conjecture that lack of information may not be the main constraint, but rather higher taxes and regulation, and the (perceived) low benefits of registration.

Second, the purpose of *simplification / registration interventions* is to simplify business entry regulations, or business registration procedures. These reforms typically lead to a sizable decrease in the number of days required for registration. One example is the study by Bruhn (2011) who analyzes the Rapid Business Opening System (SARE) in Mexico implemented in the years 2002-2006. The reform reduced the average number of days for a business registration from 30.1 to 1.4. The study uses variation in roll-out across time and municipalities to identify the reform effects, and the results indicate that the total number of firms increased by 5% in eligible industries. The author identifies former wage workers opening businesses as the main channel for the reform effects.

Third, *tax incentive and social security reduction interventions* pursue the objective

to reduce the “costs of being formal” by reducing the tax burden and/or social security contributions. This is intended to make the registration of firms and workers more attractive by reducing the direct costs of being formal. An example is the SIMPLES reform in Brazil 1996 analyzed by Monteiro and Assunção (2012). The reform combined six different federal taxes and social security contributions into one monthly-based rate. The study uses a difference-in-differences approach with sectors affected and not affected by the reform, and finds a statistically significant and large positive effect in one eligible sector (retail), while the estimated effects are insignificant in four other sectors (construction, manufacturing, transportation, services).

The fourth category, *labor inspection and enforcement interventions*, intends to increase compliance with firm and/or labor registration regulations through enforcement. Such enforcement can take place e.g. through labor inspector visits, or through official letters from the tax authorities. E.g. De Andrade et al. (2013) conduct a field experiment in Brazil where municipal inspectors are randomly assigned to visit informal firms. The authors find that while assistance in the registration process and eliminating the initial cost of registration had little effect on registration rates, visiting firms resulted in a 21 to 27 percentage point increase in registration.

Finally, the fifth intervention category provides a *financial incentive* to enhance the potential effect of (i) information or (b) simplification / registration interventions. That is, it is typically not provided as a stand-alone intervention, but in combination with either of the two. For instance, de Mel et al. (2013) implement an RCT in Sri Lanka with informal firms and four treatment arms T1 through T4: (T1) provides an information intervention plus the reimbursement of the (modest) direct registration costs. (T2) through (T4) each provide the information treatment, too, plus an additional payment of the local equivalent of USD 88, 175, and 350, respectively. The empirical analysis finds that T1 did not show any effect; for T2, 17-22% of firms registered, for T3, 48%. There was no additional impact for T4.

2.2 Compilation of the database

The objective of the data compilation is to construct a database of the universe of impact evaluations and quantitative assessments of formalization interventions worldwide, focusing on low- and middle-income countries. This systematic process proceeds in three main steps (e.g. Kluve et al., 2019): 1) The first step is to search for relevant studies that analyze one or more of the intervention types defined in Section 2.1; 2) the second step is to verify a set of inclusion criteria to arrive at the final set of relevant studies; and 3) the third step is to systematically extract information from these primary studies and code it into the database.

The first step uses a broad set of search terms³ and applies them to a title and keywords search in a series of websites and research databases in which relevant studies would be contained, specifically: IDEAS/RePEc, SSRN, IZA, NBER, 3ie Impact Evaluation Repository and World Bank Open Knowledge Repository. In addition, we conduct a forward citation search in Google Scholar using five highly cited articles on formalization / informality (Kugler and Kugler, 2009; McKenzie and Sakho, 2010; de Mel et al., 2013; Bruhn and McKenzie, 2014; La Porta and Shleifer, 2014) and screen all papers that cited at least one of those articles.⁴ For all papers fulfilling our inclusion criteria (see below) we also conduct a backward citation search to see if they cite any other paper not yet identified in the search. Finally, in the process of developing and presenting this paper at conferences and seminars we received additional paper recommendations from experts in the field, which we also include in the pool of potential studies.

All studies identified through the first step are given a full-text assessment in the second step, in which the following inclusion criteria are considered:

- We only include empirical studies with a quantitative assessment of the effect or impact of a formalization program or policy using some version of a selection correction (counterfactual impact assessment, i.e. estimation of a causal treatment effect). In general, this can include studies based on experimental designs (RCTs) or quasi-experimental methods
- Distinguishable estimate of the effect or impact of the formalization program or

policy, with an indication of the statistical significance of the estimate

- Distinguishable formalization program or policy that can be categorized into at least one of the five intervention types
- Studies that assess the impact on at least one of six relevant outcome variables (specified and discussed in detail below)
- In line with the objectives of the study (see introduction above), and because in general symmetry of effects cannot be assumed, only studies with a *switch-on* type of intervention that target improved formalization outcomes are included. That is, for instance, a study that looks at how tax increases may lead to a reduction in formal employment would not be in-scope
- Study available in English
- Search hits of studies focusing on the "hidden economy" or "shadow economy" are not in-scope⁵

The search process and identification of studies for the database is illustrated in Figure 2. A large number of studies were initially found in the databases. After title-screening and removing duplicates—i.e. studies responding to multiple keywords, available in several versions and/or appearing in more than one database—91 studies were full-text screened, of which 23 passed the inclusion criteria. A further 8 studies were identified through the forward citation search using prominent studies on formalization interventions or informality in general. Finally, 7 additional studies were identified through backward citation search and expert recommendations.⁶

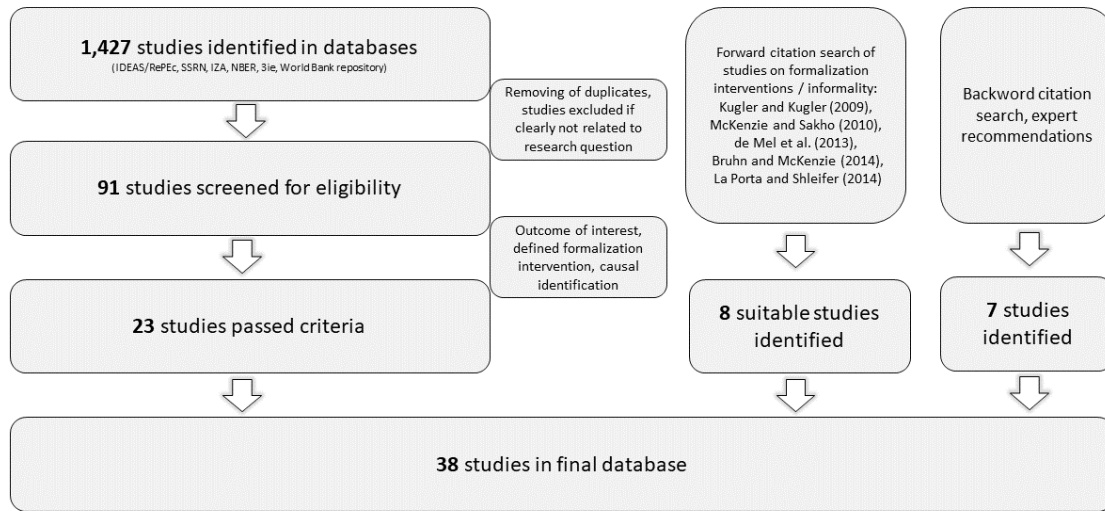


Figure 2: Search process and identification of studies

All studies fulfilling these criteria are then used to extract information into a database. In our case, 38 primary studies were identified by the search process that also fulfilled the inclusion criteria. The main information to be extracted from the primary studies concerns, first, the intervention analyzed (as discussed in Section 2.1), secondly, the outcome used to measure intervention effectiveness, and thirdly the estimate of the intervention impact.

To measure the effects of interventions, we consider six outcomes; a) firm registration, b) worker registration, c) wages, d) firm profitability, e) tax revenue, and f) investment. More than 70% of observations in our data stem from the first two categories. The first outcome looks at the number of registered firms or the probability of a set of firms to register. The second outcome looks at the same outcome at the level of the worker and considers the number of formal jobs, individual registration or the share of formally employed workers in an economy (or local labor market). Besides those two *primary* outcomes of formalization interventions, we also consider four *secondary* outcomes which are both prominent in the academic debate and of key interest to policy makers.

When examining the impact on wages we seek to identify whether the formalization interventions have led to increased wages for workers. This could be the case if, e.g.

employing registered workers has become cheaper for firms (intervention type (iii) and (v)) which increases workers' bargaining power, or if being registered leads to increased firm performance. The latter point is also directly considered by the outcomes d) and f), which both look at firm outcomes. A reason why formalization interventions have become increasingly popular is that it is widely believed that formality can improve firm performance, e.g. by giving firms access to credit markets and by making it easier for firms to grow.⁷ A major goal of governments aiming at increasing formality is to boost tax revenue (outcome e)). Large informal economies are a main reason why many low-and middle-income countries have a low tax base. By making it cheaper for firms or workers to become formal, it is hoped that the increase in formality can increase tax revenue despite reducing marginal tax rates.

In addition to intervention type and outcome, the coding process needs to extract a measure of the intervention effect. This measure can then be correlated in the empirical analysis with other variables from the primary studies, to investigate whether the estimated effectiveness shows systematic features by intervention characteristics. As a first measure, we use a trinomial indicator of sign and statistical significance (convention: at the 5% level) of the estimated intervention effect or impact: (i) negative and statistically significant, (ii) not statistically different from zero, and (iii) positive and statistically significant.

Ideally, in a second step, one would like to extract a measure of the size of the estimated impact, i.e. the coefficient of the estimated treatment effect (see e.g. Card et al., 2018). Doing this in a comparable way in the given context of formalization interventions is challenging because of the heterogeneity of the outcomes considered in the primary studies. Eventually, however, it was possible to code an effect size measure for more than two-thirds of our total sample (68%, see below) using estimates evaluating the *percent impact* of the intervention. This measure is defined as

$$Effect\ size = \frac{\hat{\beta}}{\bar{\mu}_c} \quad (1)$$

where $\hat{\beta}$ is the estimated intervention effect coefficient and $\bar{\mu}_c$ refers to the mean outcome of the control group or the pre-intervention value of the treatment group. Our effect

size measure deviates somewhat from the commonly used standardized mean difference (Borenstein et al., 2011), as in the quasi-experimental studies $\hat{\beta}$ is often not the mean difference between two groups, and the standard deviation required for the denominator is often not reported in the primary studies. Note that whereas this measure cannot be interpreted in units of specific outcomes, the standardization of $\hat{\beta}$ with the baseline value $\bar{\mu}_c$ has the advantage of making the effect size dimensionless and thus comparable across heterogeneous outcomes. This is a desirable property (see Duvendack et al., 2012), which is analogously achieved in more classical meta analysis settings by the conversion from a raw to a standardized mean difference.

Complementing this first effect size measure, we also code the *percentage point impact* of the interventions for our two main outcomes — worker and firm registration. This measure is directly interpretable and of relevance for decision makers: One policy goal, for instance, may be to increase firm registration by ten percentage points. Other related relevant effects for public policy, such as an increase in the tax base, can also be approximated well using this measure. A disadvantage of the percentage point measure is that it says little about the distribution of the dependent variable: The success of an intervention that increased worker registration by one percentage point may be interpreted quite differently depending on whether the baseline value was close to zero (a large relative effect) or close to, say, 80% (a small relative effect). This would be captured by the effect size measure explained above.

The coding process, in addition, includes several more variables in three main groups:

a) Study characteristics:

- Country
- Authors; Title; Publication status (year; journal, if applicable)
- Bibliometric information (citations, impact factor of journal)

b) Intervention characteristics:

- Target of the formalization intervention: firms, workers, or both
- Scope of the intervention: program or policy

- Year of the policy change or implementation of intervention (if applicable)
- c) Empirical analysis:
- Time horizon of the study: start and end date
 - Unit of observation: (i) firm, (ii) worker, (iii) linked, (iv) other
 - Data source and size of the estimation sample
 - Time horizon of the outcome measurement: coded in months since reform date / start of the intervention, then categorized as short-run (up to 12 months), medium-run (13-24 months), and long-run (more than 24 months)⁸
 - Identification strategy and empirical method
 - Coefficients, standard errors, statistical significance of the estimates (5% level), mean value of the control and effect size with confidence intervals
 - Macroeconomic indicators: GDP growth, poverty index, unemployment rate, share in self-employment

When a study reports estimated impacts for (a) separate interventions, (b) separate outcomes, (c) separate groups of firms or workers, or (d) at separate time horizons, then these estimates are coded separately; that is, one study typically yields more than one observation in the data. Overall, it was thus possible to extract 170 impact estimates from the 38 primary studies. Table 7 in the appendix gives the list of studies and key variables coded. The trinomial measure of sign/significance is available for all estimates, and the effect size is available for 115 estimates. Note that for each of the relevant categories we code only the "best possible" estimate, i.e. either the one explicitly highlighted by the authors as the preferred estimate, or the one on which the interpretation of the paper's findings are focused.⁹ That is, we do not code treatment coefficients from slightly varying specifications, or from robustness checks.¹⁰

Table 1 presents an overview of the countries in the data. It can be seen that a set of countries with specific reforms, some of which were analyzed in more than one paper, is prominently represented in the data (e.g. Brazil with 38 estimates from 12 studies, and Colombia with 33 estimates from 6 studies). Overall, the majority of analyses of formalization policies and programs originates in countries in Latin America (102 impact

Table 1: Distribution of countries in database on formalization interventions

Country	Observations		Studies	
	Freq. (1)	Percent (2)	Freq. (3)	Percent (4)
Algeria	4	2	1	3
Argentina	9	5	3	8
Bangladesh	3	2	2	5
Benin	21	12	1	3
Brazil	38	22	12	32
Colombia	33	19	6	16
Georgia	10	6	1	3
Indonesia	6	4	1	3
Malawi	12	7	1	3
Mexico	10	6	4	11
Peru	12	7	2	5
Russia	2	1	1	3
Serbia	1	1	1	3
Sri Lanka	5	3	1	3
Turkey	4	3	1	3
	170		38	

estimates = 60%), but still 40% of estimates (68) are from other countries. The overall number of countries in the sample (15) is not very large, which indicates that perhaps the quantitative analysis of these interventions is not very widespread yet.¹¹ Evidently, at the outset of this research we would have hoped for primary studies from many of the countries featured e.g. in Figure 1.

Figure 3 illustrates how the rigorous assessment of their effectiveness is, in fact, a very recent phenomenon. The figure shows the distribution of the years in which the respective reform or intervention was implemented, and indicates that more than two-thirds of the impact estimates for which this information was available (98 out of 157, i.e. 62%) are from interventions implemented in 2010 or later. This is despite the issue of informality having entered the academic and policy debate almost half a century ago: (Hart, 1973) defined many of the key concepts, and especially the book *The Other Path* by de Soto (1989) further popularized the topic.

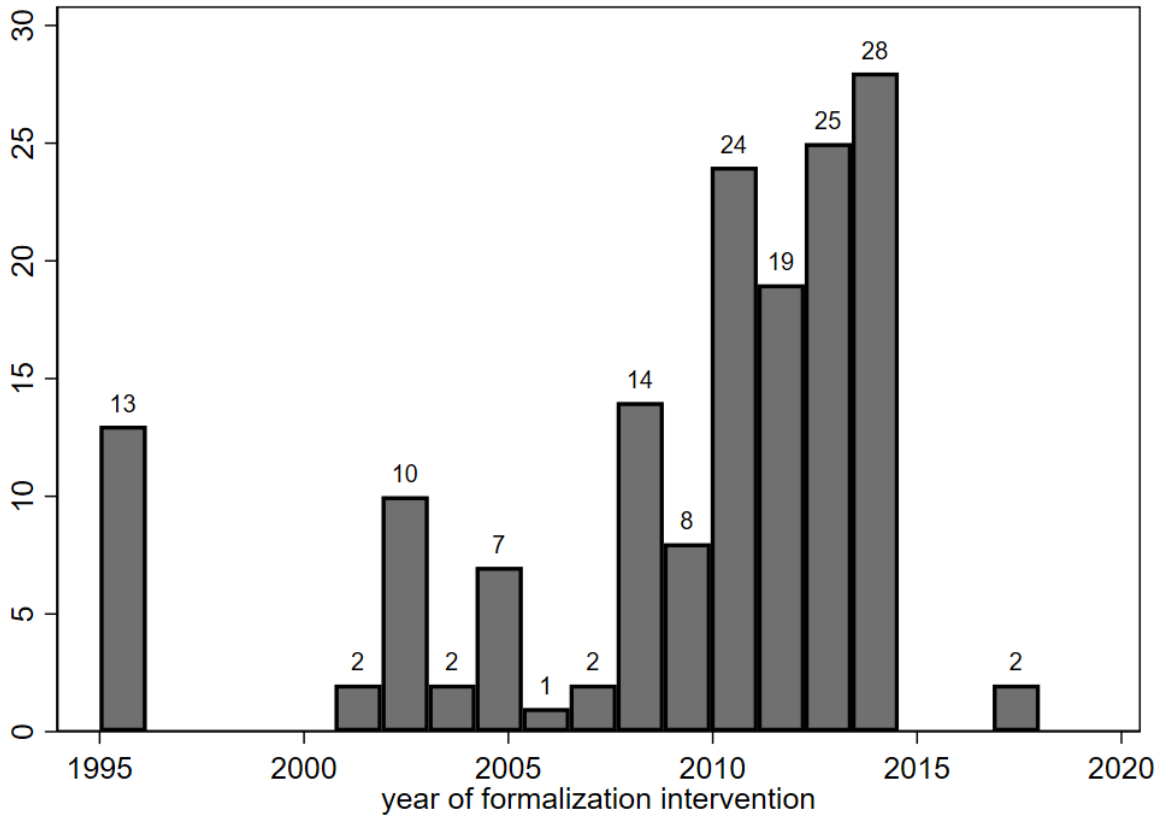


Figure 3: Year in which the formalization intervention was implemented

Note: Figure shows the distribution of estimates of formalization interventions by the year in which interventions were implemented (if applicable).

3 Descriptive analysis

This section presents a descriptive statistical analysis of the main patterns in the database. Table 2 gives an overview of the features of the impact evaluations of formalization interventions. The first two columns give a summary of the full sample, i.e. all estimates. Looking at the "intervention type", estimates of simplification / registration interventions represent the largest share in the data (92 estimates, 54%), followed by tax incentives (39%) and information approaches (32%). Financial incentives also cover almost one quarter of estimates in the data, and 15 estimates describe impacts of labor inspection interventions (9%).

As the next panel indicates, the majority of estimates originates from single-component intervention and evaluations (88 estimates, or 52%). 38% of estimates are from interven-

Table 2: Features of impact evaluations of formalization interventions

	Full sample		Experimental		Quasi-exp.	
	Freq. (1)	Perc. (2)	Freq. (3)	Perc. (4)	Freq. (5)	Perc. (6)
Number of estimates	170		61		109	
Number of studies	38		9		29	
<u>Intervention type (number of studies in parentheses)</u>						
Information intervention	55 (9)	32	49(7)	80	6(2)	7
Simplification / registration	92(18)	54	47(4)	77	45(12)	48
Tax incentives	67(15)	39	0	0	67(15)	52
Financial incentives	41(7)	24	29(3)	56	13(4)	12
Labor Inspection	15(5)	9	4(2)	7	11(3)	10
<u>Combination of interventions</u>						
Single interventions	88	52	6	10	82	75
Two combined	64	38	43	70	21	20
Three combined	18	11	12	20	6	6
<u>Formalization target</u>						
Firm	101	59	61	100	40	37
Worker	63	37	0	0	40	37
Both	6	4	0	0	6	6
<u>Formalization scope</u>						
Program	61	36	61	100	0	0
Policy	109	64	0	0	109	100
<u>Time horizon</u>						
Short-term (0-12 months)	78	46	23	38	55	50
Medium-term (13-24 months)	70	41	33	54	37	34
Long-term (>24 months)	22	13	5	8	17	16
<u>Outcome</u>						
Registered firms	58	34	31	51	27	25
Formal jobs	63	37	8	13	27	25
Wages	15	9	2	3	13	12
Firm profitability	16	9	12	20	4	4
Tax revenue	16	9	6	10	10	9
Investment	2	1	2	3	0	0

Note: Registered firms denote the number of formally registered firms or registration probability. Formal jobs denote number of formal jobs, worker registration or probability to register. Evidently it varies from country to country what precisely "registration" entails. Columns (1) and (2) show an overview for all estimates, columns (3) and (4) show estimates from experimental studies and the last two columns those from quasi-experimental settings.

tions that combine two different approaches, and 11% of estimates are from interventions that combine three. That is, almost half of the sample covers multi-component interventions.

Looking at the "formalization target" there are slightly more interventions targeting firms than workers (59% and 37%, respectively). A residual 4% of impact estimates is from interventions that target both. In terms of the intervention scope, 109 of the impact estimates (64%) refer to „policy“-type interventions, while 61 (36%) refer to „program“-type interventions. This relation differs slightly when looking at the primary study level (not shown in the table): 9 of the 38 primary studies (24%) analyze singular „programs“, while 29 primary studies (76%) analyze „policies“ at scale. That is, the typical study analyzing a program produces on average fewer impact estimates than the typical study analyzing a policy.

The large majority of impact estimates are available for the short-term (78, i.e. 46%) and for the medium-term (70, i.e. 41%), while for the long-term time horizon only a limited number of estimates has been produced (22, i.e. 13%).

Finally, and perhaps unsurprisingly, in terms of the outcomes the majority of impact estimates investigate either impacts of formalization interventions on (a) the number of registered firms / firms' registration probability, or (b) the number of formal jobs / formal employment / worker registration, with more than a third in the sample each. Wages, firm profitability and tax revenue are analyzed in about 9% of cases each, while formalization impacts on investment remain the exception.

Columns (3) to (6) differentiate by whether the estimates are derived from experimental or quasi-experimental settings. Some substantial differences are apparent; almost all experiments involve some form of information and simplification / registration intervention, and more than half include a financial incentive. Quasi-experimental estimates on the other hand mostly stem from tax incentives and simplification / registration. Almost all experiments combine several types of interventions, and all are targeted at firms. The majority of quasi-experimental estimates consist of single interventions and most of them target labor registration, which is also reflected in "formal jobs" and "wages" being the predominant outcome measures. As, almost by definition, all *programs* are analyzed

in an experimental settings and all *policies* by using quasi-experimental methods, the relationship between identification approach and intervention scope is deterministic. In this paper, we distinguish by program / policy when having the scale of the intervention itself in mind, and between experiments / quasi-experimental settings when the focus of interpretation relates to the underlying empirical approach.

3.1 Sign and statistical significance of estimated intervention effects

In Table 3 we begin investigating patterns of effectiveness by looking at sign and significance of the estimates. The table shows that just below half (75) of impact estimates in the full sample are positive and statistically significant (44%), giving a first measure of the probability with which formalization interventions can be expected to be successful. At the same time, only 10 impact estimates (6%) are negative and statistically significant;¹² this means that 50% of impact estimates (85) are not statistically different from zero.

Distinguishing between experimental and quasi-experimental designs in columns (3) to (6) of Table 3 shows that the share of positive significant estimates is larger for the quasi-experimental sample (54% vs. 26% in the experimental sample). At the same time, the share of insignificant estimates is larger in the experimental sample (66% vs. 41% in the quasi-experimental sample). Since the experimental studies are generally based on much smaller sample sizes, this latter pattern suggests that some part of the insignificant estimates may be due to research designs based on low statistical power.

Table 3: Distribution of estimated intervention impacts by sign and significance

	Full sample		Experimental		Quasi-exp.	
	Freq. (1)	Percent (2)	Freq. (3)	Percent (4)	Freq. (5)	Percent (6)
Negative and statistically significant	10	6	5	8	5	5
Insignificant	85	50	40	66	45	41
Positive and statistically significant	75	44	16	26	59	54

Note: Statistical significance is determined at a 5% level.

Figure 4 stratifies the distribution of intervention effect estimates that are negative

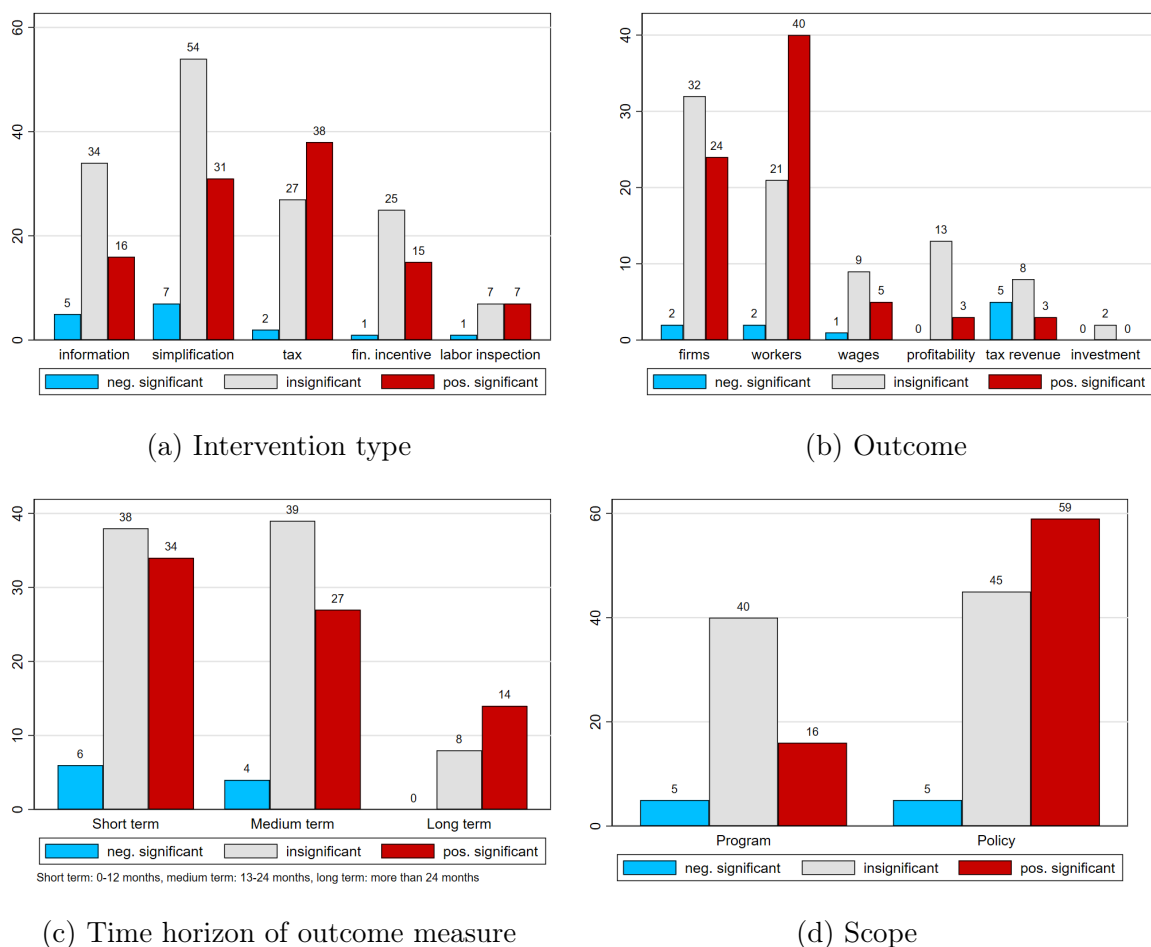


Figure 4: Sign and significance of estimates

Note: The number of estimates for the differentiation by intervention type add up to more than 170 estimates as almost half of the estimates combine more than one intervention type (see Table 2). Statistical significance determined at the 5% level.

significant, insignificant, or positive significant by intervention type, by outcome, by the time horizon of the outcome measure and by program / policy. The distribution by intervention type in panel (a) shows that only for "tax incentive" intervention types the share of positive significant estimates is larger than for insignificant estimates, indicating that these interventions, on average, display more positive results. The second main finding of the graph is that there is no pronounced pattern by intervention type; the relative share of positive significant impacts is highest for tax incentives (56%), followed by labor inspection (47%), financial incentives (37%), simplification/ registration (34%) and information interventions (29%).

The patterns by outcome in panel (b) on the other hand are rather pronounced. First, the shares of positive significant impacts differ strongly: the number of formal jobs

and worker registration have by far the highest probability of showing positive impacts (63%). For the second main outcome, firm registration, this probability of a positive impact is still 42%. For wages, with much fewer estimates, 33% are significantly positive estimates. All estimates for investment are insignificant, and tax revenue is the only outcome containing a larger number of negative than positive estimates.¹³

Panel (c) of Figure 4 distinguishes the sign/significance pattern by time horizon. Given the large shares of insignificant estimates at the short-term and medium-term time horizons, no conclusive dynamic pattern can be identified between these two. At the long-term time horizon, statistically significant estimates clearly dominate — given the small number of observations, however, this can at best give a tentative indication that formalization interventions may have more positive effects in the longer run. Finally, mirroring columns (3)-(6) of Table 3, panel (d) shows that policies implemented at scale are much more likely to yield positive effects than singular programs.

3.2 Effect sizes

In addition to these main patterns by sign and significance, we can look at more detailed patterns of intervention effectiveness in Figures 5 and 6, which investigate results by the effect sizes — the estimated percent impacts — of the formalization interventions. Note that for presentation purposes we censor the very large percent impacts at 50 percent and adjust the confidence intervals accordingly: several of the percent impact estimates are very large, attaining values of more than a 50 percent impact (18 out of 115). These originate in the field experiments in which firms are offered/exposed to treatments. Given the design of these studies, the mean value of the outcome in the control group will be relatively low compared to the treatment arms — because the treatment arms are typically offered different formalization incentives, whereas the control group is not — hence generating very large percent impacts (to be clear, this is not a shortcoming of these studies, but instead a shortcoming of the percent impact measure).

To illustrate this, consider the study from de Mel et al. (2013): in the time period of analysis two out of 105 control firms registered (0.019 registration probability), whereas in one of the treatment groups 30 firms did. Controlling for covariates, the treatment-

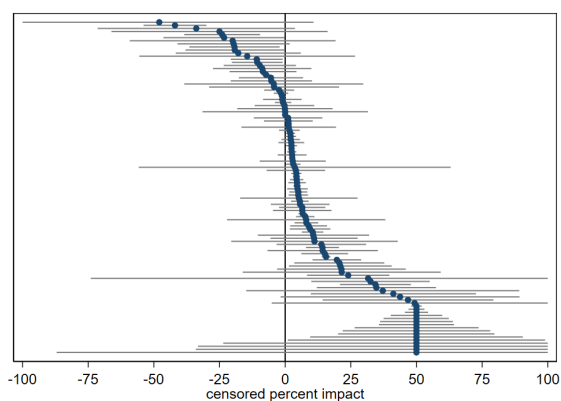
on-the-treated point estimate on the registration probability is 0.471, which implies an estimated percent impact of 2,477 percent (this is the largest in our sample). In order to make the presentation of our findings accessible in the figures, we therefore censor all percent impacts at 50.

Figure 5 displays the distribution of estimated effect sizes and confidence intervals for all estimates and by intervention category. The graph for all estimates shows a wide range of effect sizes ranging from sizable negative to very large positive, also with widely varying — and frequently very large — confidence bands. Looking specifically at intervention types, the forest plots look quite different: Information interventions (panel (b)) show a broad range of effect sizes, with the largest number across interventions of very large point estimates at the censoring value, about half of which are quite precisely estimated, and the other half display broad confidence bands. For simplification / registration interventions (panel (c)) we observe the largest number of available effect sizes, again covering the full range of effect sizes, and with strongly varying confidence bands.

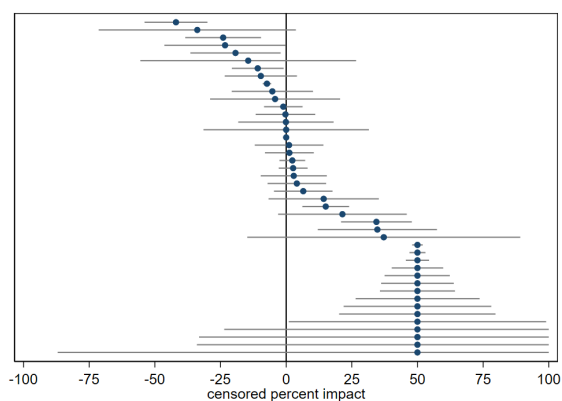
The effect size estimates for tax incentives (panel (d)) have a different pattern: There are fewer overall, and a much larger share are small effect sizes estimated with high precision. This likely reflects that often large administrative data sets analyzing policies at scale constitute the origin of these estimates. This pattern is similar for the labor inspection interventions (panel (e)), only with fewer available effect sizes, while financial incentives interventions (panel (f)) again display a more steady distribution across a larger range of effect sizes.

Figure 6 displays additional dimensions of heterogeneity by outcome type and scope of the intervention.¹⁴ Specifically, the top two panels distinguish the outcomes worker registration and firm registration. Whereas the former shows a continuous distribution across the full effect size range, with most weight of the distribution on the small, positive effect sizes (most of which precisely estimated), the latter has most estimated effect sizes at the censoring value with strongly varying degrees of precision.

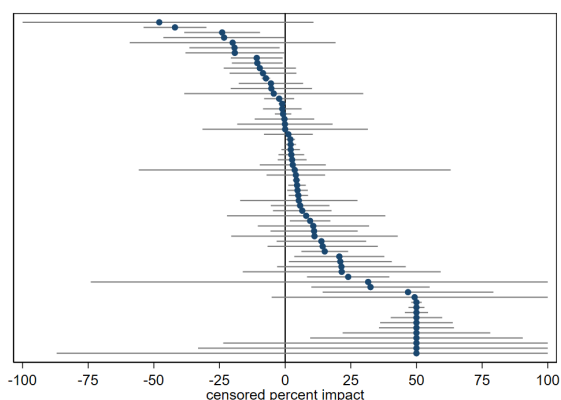
Distinguishing policies from programs also produces two very diverse patterns: specifically, the estimated effect sizes for policies show a smooth distribution from small negative to sizable positive, with the large majority of effect sizes clustering in the small-



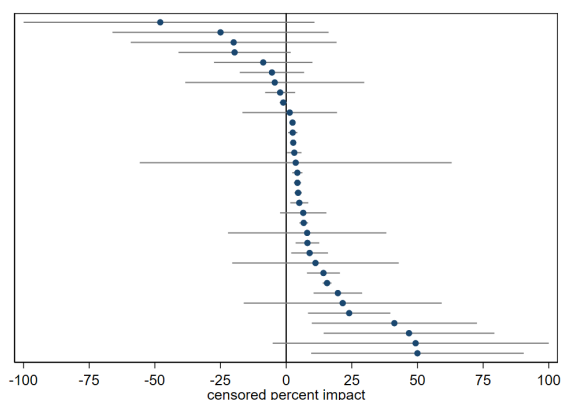
(a) All estimates



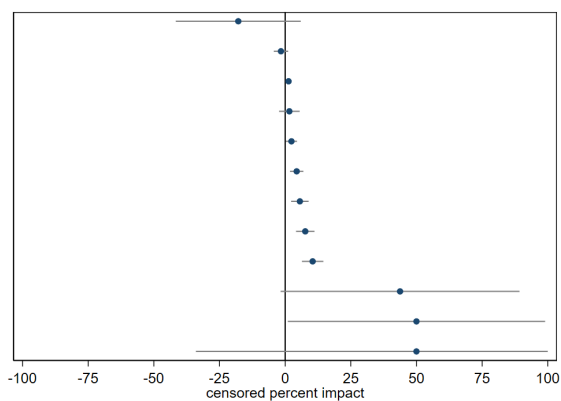
(b) Information intervention



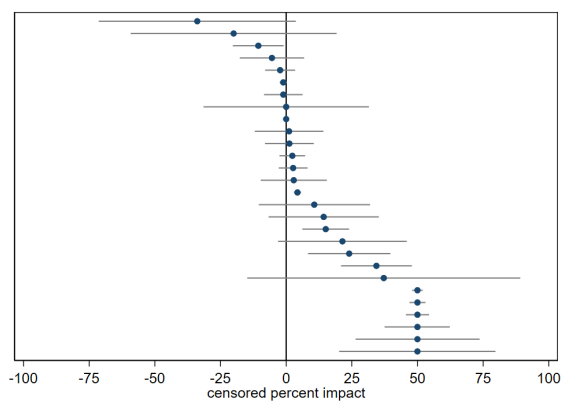
(c) Simplification / registration



(d) Tax incentives



(e) Labor inspections



(f) Financial incentive

Figure 5: Forest plots of effect sizes I

Note: Forest plots of percent impacts and corresponding 95% confidence intervals for all estimates and by intervention type. For illustrative purposes percent impacts larger than 50 percent are censored at 50 percents and the confidence intervals are scaled accordingly. Similarly, lower and upper bounds of the confidence intervals are censored at -100 and +100 respectively to ensure similar scaling. Confidence intervals are of course symmetric. In Tables 5 and 6 the uncensored estimates are used.

positive area. Confidence bands are mostly narrow. Most estimated effect sizes for programs, on the other hand, are very large and positive (at the censoring value), and

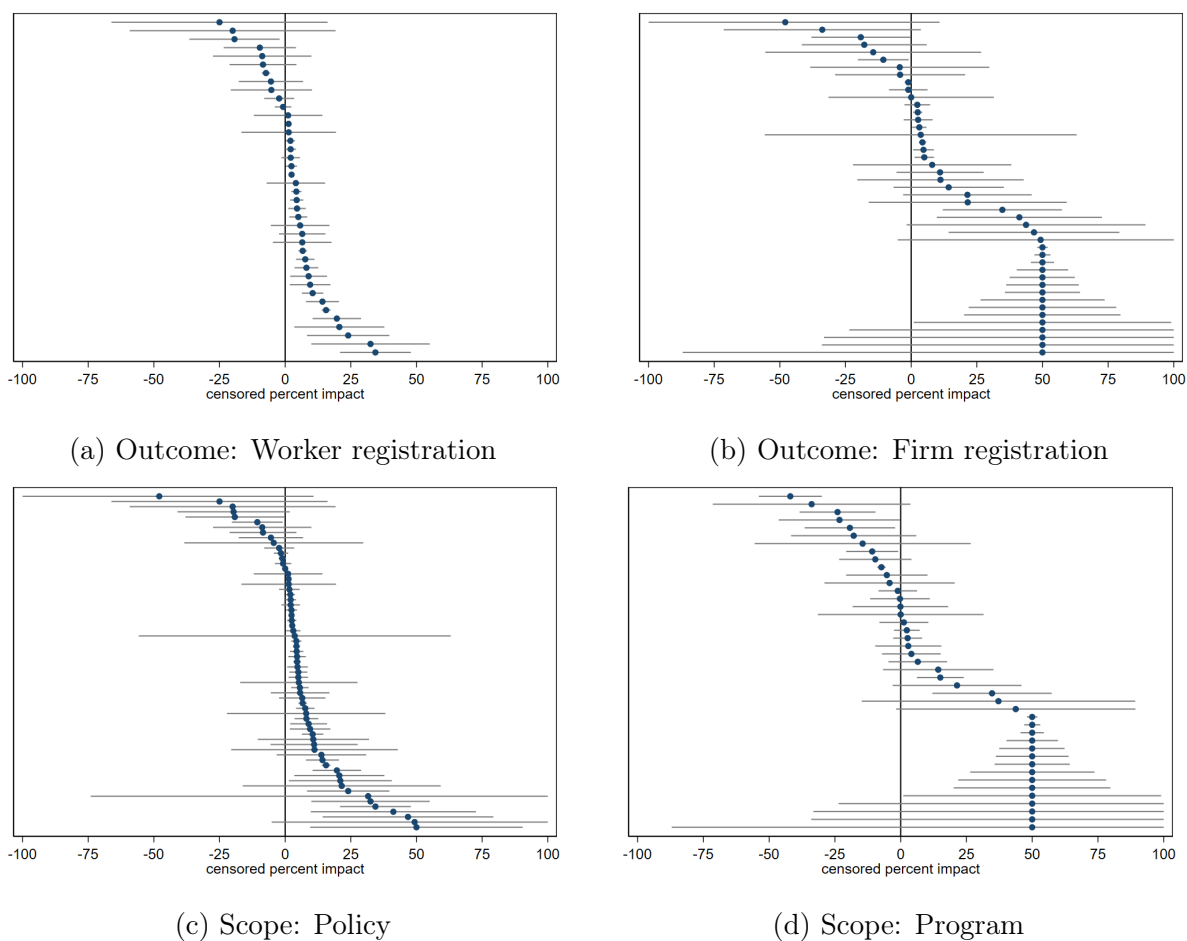


Figure 6: Forest plots of effect sizes II

Note: Forest plots of percent impacts and 95% confidence intervals for two main outcomes and by scope of the intervention. See Figure 5 for other notes.

confidence bands are typically wide. Moreover, because all *policy* estimates (c) are based on quasi-experimental designs and all *program* estimates (d) are based on experiments, the latter panel (d) highlights again the conjecture that several experiments with low statistical power contribute to the large number of insignificant intervention effects in the sample.

Figure 7 complements the descriptive analysis by plotting the distributions of the percentage point impacts. These distributions in Figure 7 are narrower than those for the percent impacts in the previous two figures, which indicates that the estimated percentage point treatment effects generally relate to small baseline formalization probabilities. As panel (a) shows, the percentage point impacts for worker registration, which primarily stem from interventions at scale, range from -19.3 to 25, with a mean of 2.38. For

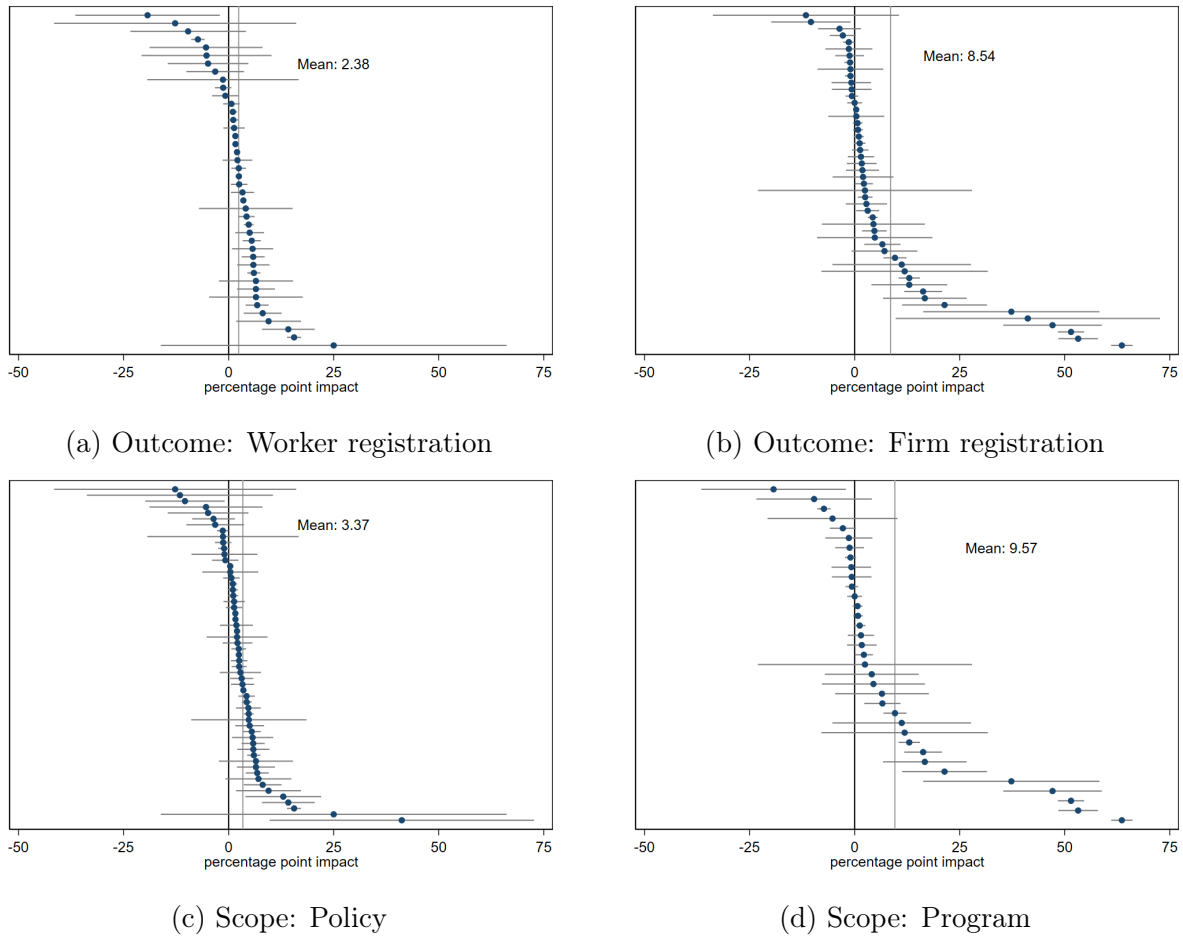


Figure 7: Percentage point impacts

Note: Forest plots of percentage point impacts and 95% confidence intervals. In contrast to Figure 5, coefficients are not standardized by the mean of the control group / baseline mean. Coefficients are only coded for the outcomes worker and firm registration, for which the percentage points increases are straightforward to interpret. The mean value of the coefficients is indicated by the thin vertical gray line.

firm registration (panel (b)) the distribution of coefficients is somewhat wider, and the majority of coefficients is in the interval between -11.6 and 21.4 percentage points, with a mean of 8.53.¹⁵

When distinguishing by the scope of the intervention in panels (c) and (d) of Figure 7, one sees a similar pattern as for the percent impacts above, where the *policy* coefficients in (c) are usually quite small and precisely estimated and clustering between -3.6 (10th percentile) and 9.5 (90th percentile) percentage points, with a mean of 3.37. The *program* estimates (d) cover a notably wider range, with an average percentage point impact of 9.57.

4 Quantitative analysis

4.1 Conceptual framework

The multivariate analysis is based on a conceptual approach used in related research designs, such as e.g. the meta analysis of the effects of active labor market policies by Card et al. (2018). Specifically, consider a formalization intervention that models an outcome y — worker registration, firm registration — observed for members of both a treatment group and a comparison group. Let b represent the estimated impact of the intervention on the outcomes of the treated units from a given evaluation design, and let β represent the probability limit of b (i.e., the estimate that would be obtained if the primary study sample size were infinite). Under standard conditions the estimate b will be approximately normally distributed with mean β and some precision P that depends on both the primary study sample size and the design features of the study. This leads to:

$$b = \beta + P^{-\frac{1}{2}}z \tag{2}$$

where z is a realization from a distribution that will be close to $N(0, 1)$ if the sample size is large enough. The term $P^{-\frac{1}{2}}z$ has the interpretation of the realized sampling error that is incorporated in b . In the next step, assume that the limiting intervention effect associated with a given study (β) can be decomposed as:

$$\beta = X\alpha + \epsilon \tag{3}$$

where α is a vector of coefficients and X captures the observed sources of heterogeneity in β , arising for example from differences in the type of intervention, characteristics of the target group or contextual factors. The term ϵ represents fundamental heterogeneity in the limiting intervention effect arising from the particular way it was implemented, specific features of the intervention or its target group, or the nature of the (labor) market environment. Equations 2 and 3 lead to a model for the observed intervention effect

estimates of the form:

$$b = X\alpha + u \tag{4}$$

where the error $u = \epsilon + P^{-\frac{1}{2}}z$ includes both the sampling error in the estimate b and the unobserved determinants of the limiting intervention effect for a given primary study.

To investigate whether there are systematic patterns in the effectiveness of formalization interventions, we implement (simple) regression models based on Equation 4 to correlate intervention effects with the set of explanatory variables capturing intervention type, outcome, primary study features, and economic context. In our case these models can be interpreted as providing descriptive summaries of the variation in average intervention effects due to differences in the observed features of a given formalization intervention and target group, and contextual factors (including methodological study features and bibliometric information). Recall that we perceive this analysis to be of explorative nature, since the available meta database remains limited in size. Future research with a larger database on hand—recall that the implementation and assessment of formalization interventions is quite a recent phenomenon—will be able to build on this and use more rigorous meta-analytical tools.

The empirical analysis proceeds in two steps. In the first step, in order to use the full sample size, we implement basic "sign and significance" models. That is, on the basis of having extracted for each estimate information related to whether it is "statistically significant negative", "statistically significant positive", or "not statistically significant from zero", the quantitative analysis estimates an ordered probit model for this trinomial classification of sign/significance of intervention effects and presents marginal effects. In a second step, we implement the main meta regression using effect size models, employing the percent impact as dependent variable. This part of the analysis also comprises testing (and correcting) for publication bias. We implement all regressions using precision weights (inverse of the variance, cf. Stanley and Doucouliagos, 2012), and cluster standard errors at the study level.

4.2 Basic models for sign and significance of intervention effects

Our basic meta regression model is implemented using an ordered probit specification for the trinomial outcome given by sign and significance of the reported intervention effect. The range of explanatory variables covers i) the intervention type (because many studies contain multiple-component interventions, no base category is required); ii) outcomes, where we contrast the two main outcomes in the data (firms, workers) with other outcomes pooled in the base category; iii) intervention and study characteristics given by the effect horizon, data type, scope of the intervention and the empirical strategy; and iv) bibliometric information. Additionally, we include v) four macroeconomic context variables: GDP growth, poverty index, unemployment rate and the share in self-employment. All context variables are measured in the particular country over the specific observation period covered in the primary study.¹⁶ Conditional correlations identified through the regressions indicate patterns in the effectiveness of interventions.

Table 4 starts with the basic model and reports marginal effects from an ordered probit model in which the sign and significance of the estimated impact is correlated with the set of explanatory variables. For this analysis, almost the full database can be used, as sign and significance was codable for all intervention effect coefficients.¹⁷ While the effect size models are preferable in most aspects of meta-regression analysis, one advantage of the sign and significance model is that it abstracts from the magnitude of the coefficient and more easily incorporates the very large impact estimates as well as heterogeneous outcomes. The table step-by-step expands the specification for all estimates and shows marginal effects for negative significant estimates—columns (1) to (3)—and positive significant estimates—columns (4)-(6)—with statistically insignificant estimates functioning as the baseline.

A first result from the top panel (i) is that information interventions and tax incentives are consistently correlated with a higher probability of a positive and significant outcome. Whereas the former is mostly combined with other intervention types, thus making the interpretation less clear, tax incentives are commonly stand-alone interventions, and appear to be relatively more effective on average. There is also some indication mainly from the full specification (columns 3 and 6) that financial incentives have a lower

Table 4: Sign/significance of estimated program effect - marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Negative significant estimate			Positive significant estimate		
<i>i) Intervention type</i>						
Information intervention	-0.021*** (0.007)	-0.025** (0.011)	-0.025*** (0.006)	0.059*** (0.019)	0.077* (0.040)	0.076*** (0.017)
Simplification / registration	-0.003 (0.010)	0.000 (0.011)	-0.003 (0.004)	0.008 (0.027)	-0.001 (0.035)	0.010 (0.013)
Tax incentives	-0.054** (0.025)	-0.048** (0.022)	-0.033*** (0.008)	0.154*** (0.049)	0.146*** (0.048)	0.098*** (0.020)
Financial incentives	-0.035* (0.021)	-0.016 (0.011)	0.011*** (0.004)	0.099** (0.042)	0.047** (0.022)	-0.034*** (0.013)
Labor inspection	-0.025 (0.022)	-0.014 (0.020)	-0.017* (0.009)	0.070 (0.057)	0.042 (0.061)	0.052* (0.027)
<i>ii) Outcome (base = other outcomes)</i>						
Registered firms		-0.033 (0.026)	-0.031*** (0.009)		0.101* (0.060)	0.092*** (0.019)
Formal jobs		-0.017* (0.010)	-0.016*** (0.003)		0.052* (0.026)	0.049*** (0.011)
<i>iii) Study characteristics (base = short-term, survey data, policy, other identification)</i>						
Medium-term		0.012 (0.012)	-0.000 (0.003)		-0.036 (0.035)	0.001 (0.009)
Long-term		-0.023 (0.014)	-0.030*** (0.006)		0.069* (0.040)	0.090*** (0.015)
Administrative data			-0.009** (0.004)			0.028** (0.013)
Program / experiment			-0.004 (0.011)			0.012 (0.034)
DD			-0.004 (0.005)			0.012 (0.014)
Square root of sample / 100			-0.002*** (0.001)			0.007*** (0.003)
<i>iv) Bibliometric characteristics</i>						
Published			0.037*** (0.008)			-0.111*** (0.016)
Citations per year			-0.001*** (0.000)			0.004*** (0.001)
<i>v) Macroeconomic indicators</i>						
GDP growth			-0.001 (0.001)			0.004 (0.003)
Poverty index			-0.000 (0.000)			0.000 (0.001)
Unemployment rate			0.000 (0.001)			-0.001 (0.002)
Share in self-employment			0.000 (0.000)			-0.000 (0.001)
Clusters	37	37	37	37	37	37
Estimates	168	168	168	168	168	168

Note: Table entries show marginal effects from an ordered probit model, in which the dependent variable takes on the values of +1, 0, and -1 for an estimated program effect being positive statistical significant, insignificant, and negative statistical significant, respectively. Regressions are inverse variance weighted. Standard errors in (parentheses) are clustered at the study level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

probability of positive outcomes, and that labor inspection interventions have a higher probability of positive outcomes.

Looking at the second panel (ii), the two main formalization outcomes, firm and worker registration, yield relatively more positive results compared to the other outcomes, as is the case when effects are measured in the long term (panel (iii)). These findings are consistent with the evidence of the descriptive approach taken in subsection 3.1. Other study characteristics results in panel (iii) indicate that studies based on administrative data are associated with a higher probability of finding significant positive outcomes (relative to studies based on survey data), as are studies with a larger sample size.

The regression results for bibliometric characteristics (panel (iv)) indicate that estimates from studies published in peer-reviewed journals (75% of estimates in the sample) are significantly correlated with a lower probability of reporting positive significant outcomes, while highly cited papers tend to contain more positive effects in the multivariate analysis. The latter variable is of course endogenously determined, and it could be that studies reporting positive significant effects tend to be cited more often. Table 4 does not indicate that the macroeconomic context is a strong predictor for the effectiveness of interventions as measured by the sign/significance of estimates (panel (v)).

4.3 Effect size models and publication bias

In any quantitative empirical exercise that compiles a meta database using a set of primary studies, a potential concern is that of publication bias (e.g. Rothstein et al., 2005) — i.e., the concern that the set of estimated intervention effects in the available literature may contain a systematic bias. In this context, it might imply that reported coefficients are biased towards being positive and statistically significant, perhaps as analysts believe for these to be more easily publishable. This could either arise because analysts only write up and circulate studies that show a positive effect ("file drawer bias") or because they choose specifications that tend to yield positive and significant effects ("p-hacking").

We can investigate this issue when moving to the next step of our analysis using the

effect size sample, i.e. the sample of 115 estimates from 30 primary studies for which it was possible to extract and code percent impact intervention effects. To examine publication bias visually, Figure 8 shows funnel plots of the relationship between the effect size and the square root of the primary study sample size, separately for the quasi-experimental and experimental estimates (because of the fundamental differences in sample size between this two types of studies). The effect sizes of the quasi-experimental estimates (panel (a)) display an inverted funnel shape in which especially the very precisely estimated impacts center close around zero. This pattern indicates a "well-behaved" sample with no visual evidence for publication bias. For the effect sizes from experimental studies (panel (b)) a different picture emerges: more precise estimates tend to be smaller or even negative, and large and positive estimates stem from studies with smaller sample sizes.

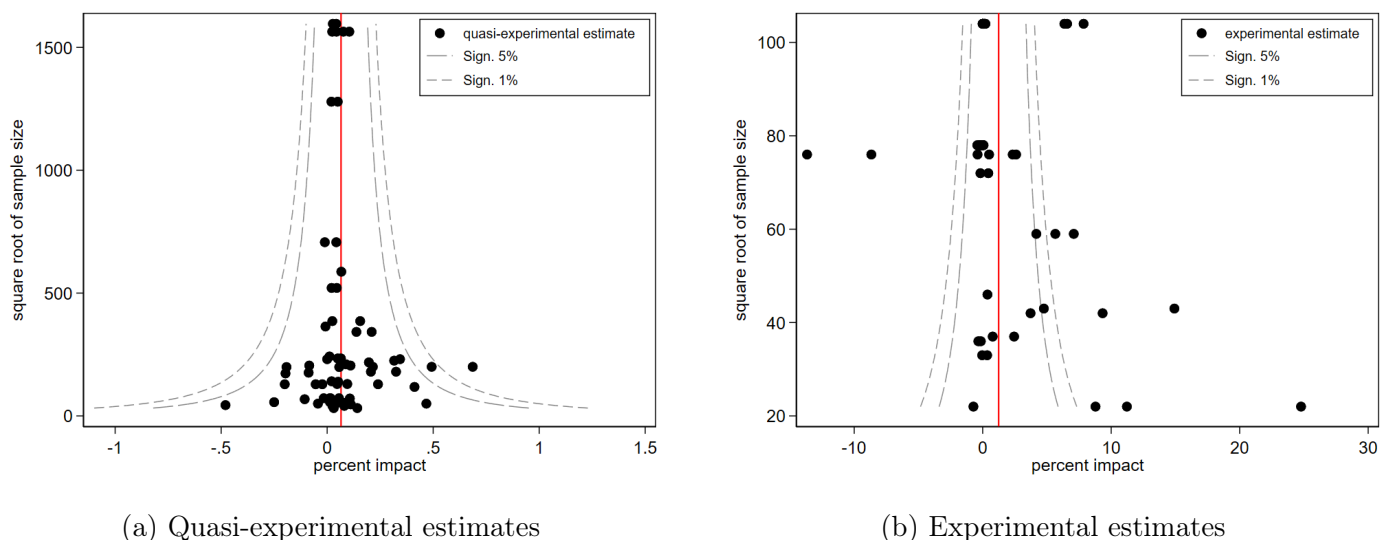


Figure 8: Funnel plots for percent impacts

Note: Figures plot the percent impacts against the square root of the sample size. A percent impact of 1 indicates a relative impact of 100 percent. The vertical line indicates the mean value of percent estimates (0.077 and 2.003, respectively). The number of estimates in the figures are 65 and 50, respectively.

A more formal test for publication bias is to regress the estimated intervention effect from a given study and specification on the associated sampling error of the estimate and other potential control variables. Using the notation of subsection 4.1, the regression model is

$$b = X\alpha + \theta P^{-\frac{1}{2}} + \nu \quad (5)$$

where ν represents a residual. The estimate of θ is interpreted as a test for asymmetry in the funnel plot relationship between the estimated intervention effects and their precision. If the sample contains more imprecisely estimated large positive effects than large negative effects, θ will be positive.

Estimation results for this model are presented in Table 5. The results for the quasi-experimental sample (panel (b)) generally confirm the visual finding from the funnel plot, giving no statistically significant indication of publication bias. Also for the experimental sample (panel (c)) the coefficients are not statistically significant – however, the coefficients are large in magnitude and mirror the visual results from the funnel plot, and the standard errors in this small subsample become very large. The suspicion that this may hide publication bias in the experimental estimates is confirmed when looking at the pooled results in panel (a): the coefficients are a weighted average of the ones in panels (b) and (c), and due to the larger sample size they are marginally significant, driven by the experimental estimates.

Based on these findings, in Table 6 we estimate the meta regression model for the

Table 5: Tests for publication bias

	(1)	(2)	(3)
a) All estimates			
$\hat{\theta}$	1.2751*	1.2767*	1.3092*
	(0.6683)	(0.6822)	(0.7555)
Estimates	115	115	115
b) Quasi-experimental			
$\hat{\theta}$	0.5960	0.6092	0.6746
	(0.3527)	(0.3832)	(0.4017)
Estimates	65	65	65
c) Experimental			
$\hat{\theta}$	2.6096	2.6954	2.4373
	(2.1546)	(2.3193)	(3.1911)
Estimates	50	50	50

Note: Entries correspond to estimated coefficients of the error term of the percent impact. Precision weighted estimates use the inverse of the variance, winsorized at the 10th and 90th percentiles of the respective samples to stabilize the estimates. Control variables in columns (1) to (3) correspond to the controls used in the same columns of Table 4. Standard errors (in parentheses) are clustered at the study level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Effect size model for estimated percent impact

	All			Quasi-exp.		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>i) Intervention type</i>						
Information intervention	-0.069** (0.032)	-0.054* (0.030)	0.020 (0.105)	0.151** (0.060)	0.170** (0.065)	0.122** (0.044)
Simplification / registration	-0.060*** (0.021)	-0.072*** (0.023)	-0.037 (0.044)	0.061* (0.032)	0.066* (0.036)	0.081** (0.039)
Tax incentives	-0.023 (0.026)	-0.025 (0.024)	0.065 (0.051)	0.094*** (0.030)	0.101*** (0.032)	0.140*** (0.019)
Financial incentives	0.037 (0.023)	0.047* (0.026)	0.008 (0.053)	-0.102*** (0.030)	-0.095** (0.036)	-0.065* (0.038)
Labor inspection	-0.050* (0.027)	-0.051** (0.024)	-0.027 (0.048)	0.065* (0.034)	0.070* (0.035)	0.067** (0.029)
<i>ii) Outcome (base = other outcomes)</i>						
Registered firms		0.042* (0.024)	0.026 (0.020)		0.020 (0.016)	-0.020 (0.025)
Formal jobs		0.043*** (0.015)	0.027* (0.013)		0.033*** (0.010)	0.012* (0.006)
<i>iii) Study characteristics (base = short-term, survey data, policy, other identification)</i>						
Medium-term		-0.016 (0.016)	-0.014 (0.017)		-0.024 (0.016)	-0.023 (0.015)
Long-term		0.018 (0.015)	0.039* (0.021)		0.009 (0.016)	0.030 (0.023)
Administrative data			0.002 (0.017)			0.001 (0.014)
Program / experiment			-0.125 (0.171)			
DD			-0.016 (0.029)			-0.051* (0.025)
Square root of sample / 100			0.002** (0.001)			-0.000 (0.001)
Publication bias correction	1.275* (0.668)	1.277* (0.682)	1.309* (0.756)	0.756** (0.340)	0.754** (0.362)	0.887** (0.372)
<i>iv) Bibliometric characteristics</i>						
Published			-0.064** (0.024)			-0.054** (0.020)
Citations per year			0.003* (0.002)			0.004** (0.002)
<i>v) Macroeconomic indicators</i>						
GDP growth			0.002 (0.002)			0.001 (0.001)
Poverty index			0.003 (0.003)			-0.012** (0.005)
Unemployment rate			-0.002 (0.003)			-0.006*** (0.002)
Share in self-employment			-0.001 (0.001)			0.001 (0.001)
Constant	0.063** (0.030)	0.033 (0.028)	0.010 (0.108)	-0.045 (0.032)	-0.072** (0.033)	0.023 (0.061)
Clusters	30	30	30	23	23	23
Estimates	115	115	115	65	65	65

Note: Table entries are coefficients from a weighted linear regression, in which the dependent variable is the percent impact of the estimate. Regressions are precision weighted with the inverse of the variance, the standard error of the estimates is added as an explanatory variable ('publication bias correction'). Standard errors in (parentheses) are clustered at the study level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

effect size sample including the standard error of the estimated intervention effect, which is the preferred correction for publication bias as recommended e.g. in Stanley and Doucouliagos (2012). We present results for the full sample (columns 1-3) and for the quasi-experimental subsample (columns 4-6). We refrain from fitting this model to the experimental subsample as it stems from too few studies (7).

Looking at the intervention type in panel (i), the results confirm the indicative pattern found for the sign/significance model above: On average, information interventions and tax incentives are significantly correlated with larger intervention effect sizes. This pattern is particularly pronounced for the sample of quasi-experimental estimates. And as for the sign/significance model financial incentive interventions are correlated with relatively smaller (or negative) effect sizes. Finally, for the two intervention types "simplification / registration" and "labor inspection" no clear picture emerges: results from the full sample indicate no or a slightly negative correlation with effect sizes, while results from the quasi-experimental sample point towards a slightly positive correlation.

Panel (ii) of Table 6 considers the outcome variables. The regression results indicate that the worker outcome ("formal jobs") has a statistically significant positive correlation with larger effect sizes, relative to other outcomes, while no such correlation is found for the "registered firms" outcome. The results in panel (iii) suggest that there are no strong patterns of effect sizes by program characteristics: some marginally significant correlations indicate that effect sizes may be larger in the long-term and when estimated in larger samples (column 3), and that effect sizes may be smaller from difference-in-differences designs (column 6).

Results for the bibliometric covariates reflect the same pattern found in Table 4 for the sign/significance model: the primary studies published in peer-reviewed journals seem to be those studies reporting smaller effect sizes, on average. This is a reassuring result against the indication of some degree of publication bias in this literature found above.

Finally, panel (v) examines the role of contextual economic factors. While no significant correlations are found for the full sample (column 3), column 6 shows some interesting results for the quasi-experimental sample: specifically, intervention effect sizes are significantly larger in a better (labor) market environment, as measured by a lower poverty

index and a lower unemployment rate. Given that the quasi-experimental sample comprises precisely the policy-type interventions - i.e. the interventions implemented at scale in the given country - this is the adequate sample to examine this relationship; and it is likely relevant for policy makers to know that policy efforts to reduce informality may be more effective in a more promising labor market context (with a lower unemployment rate) or that countries must have passed a certain level of economic development (as measured by the poverty index) for such interventions to be effective.

5 Conclusion

Against the background of high levels of informality in labor markets in low-and middle-income countries many interventions have been implemented worldwide in an effort to increase the formalization of workers and firms. As it is of key importance for policy makers to know which intervention works under what circumstances, this paper has analyzed patterns of the effectiveness of formalization interventions. The analysis has proceeded in a systematic and quantitative way: we first compile a database of primary studies of impact evaluations and quantitative assessments of formalization interventions, covering 170 impact estimates from 38 studies originating in 15 countries. We then investigate patterns of the estimated impacts in the data, both descriptively and quantitatively, using meta regression techniques. For this, we consider the sign and significance of the estimated impact and the effect size.

Several key patterns emerge from our analysis. First, while half of the estimates are not statistically different from zero, a large share of the impact estimates is statistically significant and positive (44%), indicating a generally sizable success probability of formalization interventions. Insignificant estimates are partially determined by the experimental studies in the sample, some of which generate relatively imprecisely estimated intervention effects, likely due to low statistical power. Second, the evidence suggests that particularly tax incentives are more likely to generate positive significant impacts. Third, in terms of outcomes we observe that interventions targeting the formalization of workers are more effective than those targeting firms, which may not be too surprising as the stakes of formalizing workers are lower and arguably also more easily reversible.

Furthermore, one should take into account that formal firms are often a prerequisite for their workers to be registered (Ulyssea, 2018, refers to this as the *intensive margin* of informality), so in countries with few registered firms, this should be the first focus of policy makers. Fourth, policies implemented at scale have a substantially larger share of positive significant estimates than singular programs. Fifth, estimates measured with a longer time horizon (> 24 months) are more effective, suggesting that reaping the benefits of efforts to increase formalization may require patience by policy makers, especially as some short-run effects of, say, tax incentives may lead to temporary reductions in tax revenue. Sixth, the multi-variate analysis of the effect size indicates that interventions are more effective in better (labor) market conditions. This is especially relevant as formalization often comes with a short-term cost for firms and during difficult economic circumstances they may be more reluctant to incur such costs. At the same time from a governmental perspective it may not be wise to implement such reforms / interventions during an economic downturn as interventions like tax breaks temporarily reduce revenue and if they prove to be ineffective in the short-run, this could be a strain on public funds.

Given the pervasiveness of the policy challenge of informality in low-and middle-income countries, one would have hoped for a larger database — but the meta data we were eventually able to compile arguably reflects the limited number of thorough quantitative evaluations of formalization interventions that are available to date. The fact that the majority of these studies has been produced very recently suggests a surge in interest in this topic, and makes our analysis a timely exercise. Future research will undoubtedly fruitfully build on this, and systematic assessments of formalization interventions will be able to focus on a more homogeneous set of outcomes (say, only worker registration or tax revenue), as the meta sample size increases, allowing for a more thorough implementation of meta-analytical tools as is possible to date. At the same time, this also calls on researchers to rigorously evaluate policy interventions aimed at decreasing informality throughout the world and present results in a transparent way to allow the construction of an extensive database on formalization impact estimates. Systematically identifying patterns that make such interventions more effective will be of great value to policy makers in informing them which interventions and under what circumstances are particularly effective, thus helping them design better policies. Some first guidance arises from our explorative meta analysis: do consider tax incentives, combine your intervention with an

information component, measure effects in the long-term, expect larger effects on workers than on firms, and during times when the labor market is in better shape. Given the sustained levels of informality in labor markets of low-and middle-income countries, interventions targeted at reducing informality will doubtlessly remain high on the political agenda of those countries for years to come.

Notes

1. Ulyssea (2018) makes the distinction between firm and worker informality explicit by contrasting the *extensive margin* of informality—firms not being registered— with the *intensive margin*—formal firms not registering their workers. These two margins do not have to move in the same direction.
2. In describing and analyzing informal labor markets there have been debates on whether they are best described by workers being in the informal sector voluntarily or involuntarily, and thus whether formal and informal labor markets are segmented or not; or perhaps a mixture of the two. E.g. Khamis (2012) provides a concise overview of these issues. Our starting point is less about what precisely defines informal labor markets and how they are best characterized, but the fact that there is an economic rationale to address informality through a set of interventions.
3. The search terms are "formalize", "formalization", "formality", "formal employment", "firm registration", "business registration", "worker registration", "labor registration", "registration simplification", "bureaucracy simplification", "market entry regulation" and "labor inspection".
4. For both the website / database search and the forward citation search, the reference date is June 11, 2020. The articles used in the forward citation search had been cited 201, 272, 267, 141 and 690 times, respectively, on the reference date.
5. Whereas both concepts are clearly related to informality in the labor market, they are typically used to describe unregulated labor activity in high-income countries, but less so the informal labor markets in low- and middle-income countries on which we focus here.
6. One study (Feld, 2020) was discovered as it cites an earlier version of this article.
7. However, some research also finds that immediately registering when starting an enterprise can be harmful for firm performance as it binds scarce resources at an early stage (Williams et al., 2017).

8. In some studies outcomes over several years are pooled. In this case, we code the average time between the intervention and measurement of the outcomes.
9. Some papers present different definitions of formality. If these can be separated and have different interpretations—as in Campos et al. (2018) where business and tax registration can occur independently—we code the coefficient of each estimate. If, in contrast, definitions are nested—as in De Giorgi et al. (2018) where their most 'trusted' estimate implies that the tax identification was verified rather than that registration was simply self-reported—we only code the authors' preferred definition of formality.
10. This approach distinguishes our work e.g. from the paper by Floridi et al. (2019) who code every single estimate (568) available from a set of 18 primary studies analyzing the impact of formalization interventions specifically on firms.
11. In contrast, The World Bank counts in its *Doing Business* report that since 2005 alone 722 reforms to starting a business have been implemented worldwide. See <https://www.doingbusiness.org/en/reforms/reforms-count> for continuous updates [last accessed June 11, 2020].
12. While negative and statistically significant effects may be counterintuitive at first, we note that half of those come from the outcome 'tax revenue' which is in line with the incentive structure of some interventions in the short-term. Additionally, the share of negative and significant estimates (6%) is broadly in line with what one would expect to find under the null given the pre-determined significance level of 5%.
13. Four of these negative coefficients stem from one study (Benhassine et al., 2018), where tax exemptions were applied to all newly formalized firms. After expiration of the exemptions, it is likely that more taxes will be paid in the long-run.
14. Additional forest plots distinguishing by sign and significance of the estimates and by the time horizon can be found in the appendix (Figure 9).
15. All estimates larger than 50 percentage points stem from the same study where a large share of firms obtained a business registration certificate in response to the intervention (Campos et al., 2018).
16. On average in the analysis sample, GDP growth is 3.1 percent, ranging from a minimum of -4.0 (Argentina, 1995) to a maximum of 10.5 (Serbia, 2005). Poverty gap at USD 3.20 per day; mean 14.0, minimum 0.8 (Argentina, 2013), maximum 52.7 (Malawi, 2012). Unemployment rate; mean 7.7, minimum 2.6 (Benin, 2014), maximum 20.9 (Serbia, 2005). Self-employment share; mean 50.6, minimum 9.2

(Russia, 2001), maximum 89.7 (Benin, 2014). All indicators are obtained from World Bank Open Data.

17. Only one study (Mullainathan and Schnabl, 2010) is excluded because the inverse variance weights could not be constructed.

References

- Alcázar, L. and Jaramillo, M. (2016). The Impact of Formality on Microenterprise Performance: A Case Study in Downtown Lima. *GRADE research paper*.
- Benhassine, N., McKenzie, D., Pouliquen, V., and Santini, M. (2018). Does inducing informal firms to formalize make sense? Experimental evidence from Benin. *Journal of Public Economics*, 157:1–14.
- Bernal, R., Meléndez, M., Eslava, M., and Pinzon, A. (2017). Switching from Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the 2012 Colombian Tax Reform. *Economía*, 18(1):41–74.
- Borenstein, M., Hedges, L. V., Higgins, J. P., and Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.
- Bruhn, M. (2011). License to sell: the effect of business registration reform on entrepreneurial activity in Mexico. *The Review of Economics and Statistics*, 93(1):382–386.
- Bruhn, M. and McKenzie, D. (2014). Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, 29(2):186–201.
- Campos, F., Goldstein, M., and McKenzie, D. (2018). How Should the Government Bring Small Firms into the Formal System? *World Bank Policy Research Working Paper 8601*.
- Card, D., Kluve, J., and Weber, A. (2018). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.
- De Andrade, G. H., Bruhn, M., and McKenzie, D. J. (2013). A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *The World Bank Economic Review*, 30:24–54.
- De Giorgi, G., Ploenzke, M., and Rahman, A. (2018). Small Firms' Formalisation: The Stick Treatment. *The Journal of Development Studies*, 54(6):983–1001.

- De Giorgi, G. and Rahman, A. (2013). SME's Registration: Evidence from an RCT in Bangladesh. *Economics Letters*, 120(3):573–578.
- de Mel, S., McKenzie, D., and Woodruff, C. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5(2):122–50.
- de Soto, H. (1989). *The other path*. Harper & Row New York.
- Duvendack, M., Hombrados, J. G., Palmer-Jones, R., and Waddington, H. (2012). Assessing 'what works' in international development: meta-analysis for sophisticated dummies. *Journal of Development Effectiveness*, 4(3):456–471.
- Fajnzylber, P., Maloney, W. F., and Montes-Rojas, G. V. (2011). Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program. *Journal of Development Economics*, 94(2):262–276.
- Feld, B. (2020). Direct and Spillover Effects of Enforcing Labor Standards: Evidence from Argentina. *mimeo*.
- Floridi, A., Demena, B., and Wagner, N. (2019). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *ISS Working Paper Series/General Series*, 642(642):1–37.
- Grimm, M. and Paffhausen, A. L. (2015). Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries. *Labour Economics*, 32:67–85.
- Hart, K. (1973). Informal income opportunities and urban employment in Ghana. *The Journal of Modern African Studies*, 11(1):61–89.
- Khamis, M. (2012). A note on informality in the labour market. *Journal of International Development*, 24(7):894–908.
- Khamis, M. (2014). Formalization of jobs and firms in emerging market economies through registration reform. *IZA World of Labor*.

- Kluge, J., Puerto, S., Robalino, D., Romero, J. M., Rother, F., Stöterau, J., Weidenkaff, F., and Witte, M. (2019). Do youth employment programs improve labor market outcomes? A quantitative review. *World Development*, 114:237–253.
- Kugler, A. and Kugler, M. (2009). Labor market effects of payroll taxes in developing countries: evidence from Colombia. *Economic Development and Cultural Change*, 57(2):335–358.
- La Porta, R. and Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3):109–26.
- Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7):1159–1178.
- McKenzie, D. and Sakho, Y. S. (2010). Does it pay firms to register for taxes? The impact of formality on firm profitability. *Journal of Development Economics*, 91:15–24.
- Monteiro, J. C. and Assunção, J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. *Journal of Development Economics*, 99(1):105–115.
- Mullainathan, S. and Schnabl, P. (2010). Does less market entry regulation generate more entrepreneurs? Evidence from a regulatory reform in Peru. In *International Differences in Entrepreneurship*, pages 159–177. University of Chicago Press.
- Pignatti, C. (2018). Enforcement and Compliance with Labour Legislation: The Effect of the Colombian “Action Plan”. *mimeo*.
- Ronconi, L. and Colina, J. (2011). Simplification of Labor Registration in Argentina: Achievements and Pending Issues. *IDB Working Paper Series*, (277).
- Rothstein, H. R., Sutton, A. J., and Borenstein, M. (2005). Publication bias in meta-analysis. *Publication bias in meta-analysis: Prevention, assessment and adjustments*, pages 1–7.
- Stanley, T. D. and Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. Routledge.
- Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8):2015–47.

Williams, C. C., Martinez-Perez, A., and Kedir, A. M. (2017). Informal entrepreneurship in developing economies: The impacts of starting up unregistered on firm performance. *Entrepreneurship Theory and Practice*, 41(5):773–799.

A Studies in the database

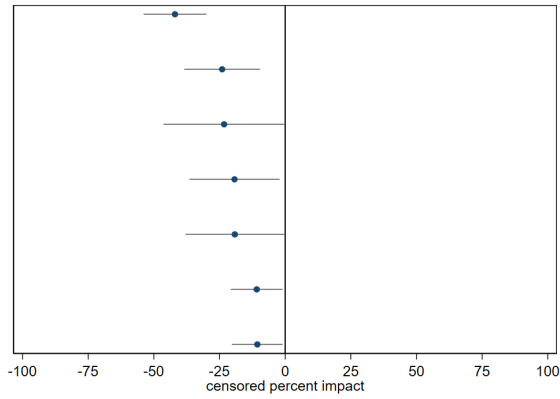
- Alcázar, L. and M. Jaramillo (2016). The Impact of Formality on Microenterprise Performance: A Case Study in Downtown Lima. *GRADE research paper*
- Almeida, R. and P. Carneiro (2009). Enforcement of Labor Regulation and Firm Size. *Journal of Comparative Economics*, 37, 28-46.
- Almeida, R. and P. Carneiro (2012). Enforcement of Labor Regulation and Informality. *American Economic Journal: Applied Economics*, 4, 64-89.
- Aparicio, G. (2014). Does Formality Improve Firm Performance? Evidence from a Quasi-experiment in Mexico. *mimeo*
- Benhassine, N., McKenzie, D., Pouliquen, V., and Santini, M. (2018). Does inducing informal firms to formalize make sense? Experimental evidence from Benin. *Journal of Public Economics*, 157, 1-14.
- Bernal, R., M. Eslava and M. Meléndez (2017). Switching from Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the 2012 Colombian Tax Reform. *Economía*, 1, 41-47.
- Betcherman, G., N. Meltem Daysal and C. Pagés (2010). Do Employment Subsidies Work? Evidence from Regionally Targeted Subsidies in Turkey. *Labour Economics*, 17, 710-722.
- Bosch, M., D. Fernandes and J. M. Villa (2015). Nudging the Self-employed into Contributing to Social Security: Evidence from a Nationwide Quasi Experiment in Brazil. *IDB Working Paper Series, WP 633*.
- Bruhn, M. (2011). License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico. *The Review of Economics and Statistics*, 93, 382-386.
- 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, 275-283.

- Bruhn, M., Cho, C., Marusic, A., Nguyen, H., Reyes, J.D. and Tran, T. (2018). Courts and Business Registration: Evidence from Serbia. *The World Bank Policy Research Working Paper*, 8611.
- Bruhn, M., and McKenzie, D. (2013). Using administrative data to evaluate municipal reforms: an evaluation of the impact of Minas Fácil Expresso. *Journal of Development Effectiveness*, 5(3), 319-338.
- Bruhn, M. and J. Loeprick (2016). Small Business Tax policy and Informality: Evidence from Georgia. *International Tax and Public Finance*, 23, 834-853.
- Campos, F., M. Goldstein and D. McKenzie (2018). How Should the Government Bring Small Firms into the Formal System? *The World Bank Policy Research Working Paper*, 8601.
- Cruces, G., Galiani, S., and Kidyba, S. (2010). Payroll taxes, wages and employment: Identification through policy changes. *Labour Economics*, 17(4), 743-749.
- de Andrade, G. H., M. Bruhn and D. McKenzie (2016). A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms. *The World Bank Economic Review*, 30 (1), 24-54.
- De Giorgi, G., and Rahman, A. (2013). SME's Registration: Evidence from an RCT in Bangladesh. *Economics Letters*, 120(3), 573-578.
- De Giorgi, G., M. Ploenzke and A. Rahman (2018). Small Firms' Formalisation: The Stick Treatment. *The Journal of Development Studies*, 54(6), 983-1001.
- de Mel, S., D. McKenzie and C. Woodruff (2013). The Demand for, and Consequence of, Formalization among Informal Firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5, 122-150.
- Fajnzylber, P., W. F. Maloney and G. V. Montes-Rojas (2011). Does Formality Improve Micro-firm Performance? Evidence from the Brazilian SIMPLES Program. *Journal of Development Economics*, 94, 262-276.
- Feld, B. (2020). Direct and Spillover Effects of Enforcing Labor Standards: Evidence from Argentina. *mimeo*.

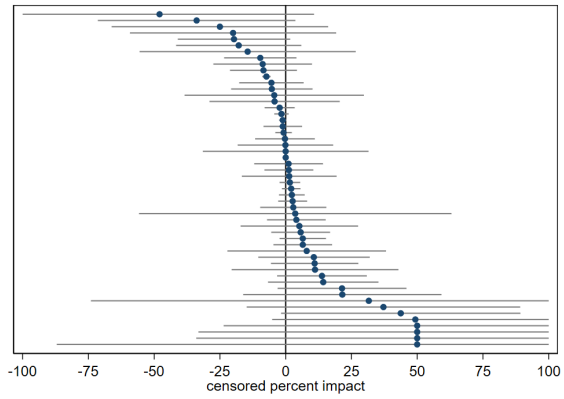
- Fernández, C., and Villar, L. (2017). The impact of lowering the payroll tax on informality in Colombia. *Economía*, 18(1), 125-155.
- Galiani, S., Meléndez, M., and Ahumada, C. N. (2017). On the effect of the costs of operating formally: New experimental evidence. *Labour Economics*, 45, 143-157.
- Kaplan, D. S., E. Piedra and E. Seira (2011). Entry Regulation and Business Start-ups: Evidence from Mexico. *Journal of Public Economics*, 95, 1501-1515.
- Kugler, A., M. Kugler and L. O. Herrera Prada (2017). Do Payroll Tax Breaks Stimulate Formality? Evidence from Colombia’s Reform. *Economía*, 1, 3-40.
- Madalozzo, R. and A. Bruscato Bortoluzzo (2011). The Impact of Tax Exemptions on Labor Registration: The Case of Brazilian Domestic Workers. *Inspere Working Paper*, 232.
- Monteiro, J. C., and Assunção, J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. *Journal of Development Economics*, 99(1), 105-115.
- Morales, L. F., and Medina, C. (2017). Assessing the effect of payroll taxes on formal employment: The case of the 2012 tax reform in Colombia. *Economía*, 18(1), 75-124.
- Mullainathan, S. and P. Schnabl (2010). Does Less Market Entry Regulation Generate More Entrepreneurs? Evidence from a Regulatory Reform in Peru. *International Differences in Entrepreneurship, NBER*, 159-177.
- Pignatti, C. (2018). Enforcement and Compliance with Labour Legislation: The Effect of the Colombian “Action Plan”. *mimeo*.
- Piza, C. (2018). Out of the Shadows? Revisiting the impact of the Brazilian SIMPLES program on firms’ formalization rates. *Journal of Development Economics*, 134, 125-132.
- Rocha, R., Ulyssea, G., and Rachter, L. (2018). Do lower taxes reduce informality? Evidence from Brazil. *Journal of Development Economics*, 134, 28-49.

- Ronconi, L. and J. Colina (2011). Simplification of Labor Registration in Argentina: Achievements and Pending Issues. *IDB Working Paper Series*, 277.
- Rothenberg, A. D., Gaduh, A., Burger, N. E., Chazali, C., Tjandraningsih, I., Radikun, R., Cole, S., and Weiland, S. (2016), Rethinking Indonesia's informal sector. *World Development*, 80, 96-113.
- Scherer, C. (2015). Payroll tax reduction Brazil: Effects on employment and wages. *ISS Working Paper*, 602.
- Slonimczyk, F. (2012). The Effect of taxation on Informal Employment: Evidence from the Russian Flat Tax Reform. *Informal Employment in Emerging and Transition Economies*, Chapter 2, 55-99.
- Souag, A., and Assaad, R. (2018). The impact of the action plan for promoting employment and combating unemployment on employment informality in Algeria. *Middle East Development Journal*, 10(2), 272-298.
- Zucco, C., Lenz, A.-K. Goldszmidt, R. and Valdivia, M. (2020), Face-to-face vs. virtual assistance to entrepreneurs: Evidence from a field experiment in Brazil. *Economics Letters*, 188.

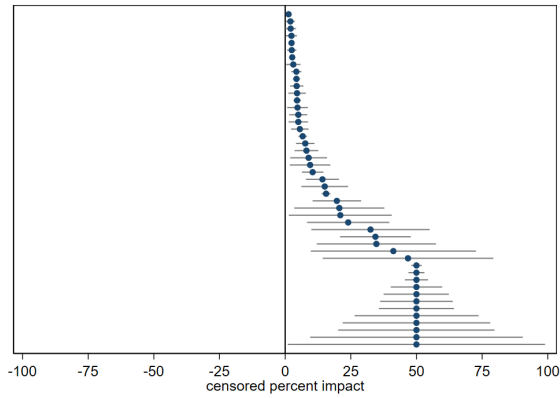
B Appendix figures and tables



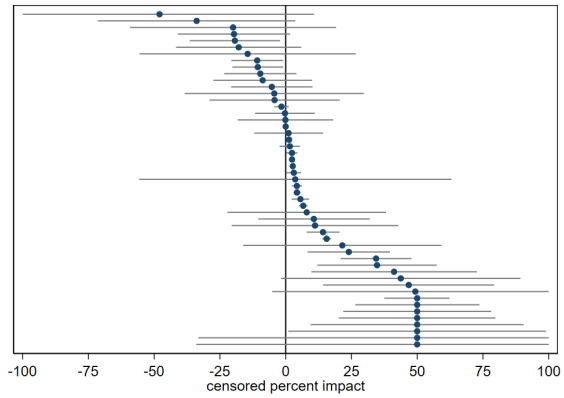
(a) Negative and statistically significant



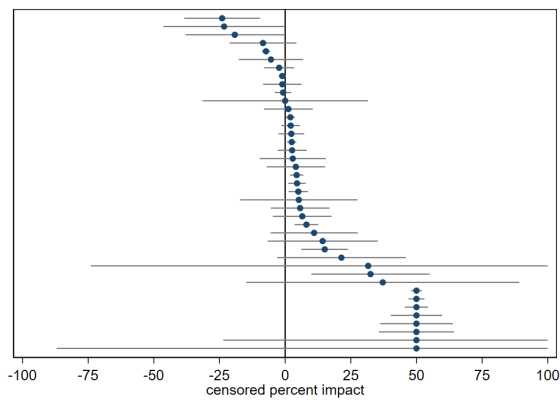
(b) Insignificant



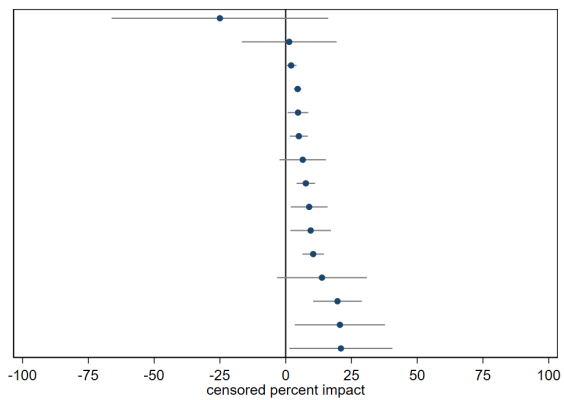
(c) Positive and statistically significant



(d) Time horizon: 0-12 months



(e) Time horizon: 13-24 months



(f) Time horizon: > 24 months

Figure 9: Forest plots of percent impacts

Note: Forest plots of percent impacts and 95% confidence intervals by statistical significance (at the 5 percent level) and by time horizon. See Figure 5 for other notes.

Table 7: Overview of study characteristics

Study	Country	Intervention type	Sample size	Identification	Effect horizon (months)	Outcomes	Number of estimates
Alcázar and Jaramillo (2016)	Peru	Simplification, financial incentive	339-572	RCT	18-30	Firms, workers, wages, profitability, investment	10
Almeida and Carneiro (2009)	Brazil	Labor inspection	1509-3074	IV	12	Firms, profitability	3
Almeida and Carneiro (2012)	Brazil	Labor inspection	5232-5242	Cross-state / city variation	12	Workers, wages	4
Apiricio (2014)	Mexico	Simplification	117259	DD	60	Profitability	2
Benhassine et al. (2018)	Benin	Information, simplification	3596-6206	RCT	12-24	Firms, workers, profitability, investment	21
Bernal et al. (2015)	Colombia	Tax incentive	2548240	DD	5	Workers, wages	2
Betcherman et al. (2010)	Turkey	Tax incentive	1056-1702	DD	8-14	Firms, workers	4
Bosch et al. (2015)	Brazil	Information, simplification	172640	DD	6	Workers, tax revenue	3
Bruhn (2011)	Mexico	Simplification	1636250	DD	18	Firms, workers	2
Bruhn (2013)	Mexico	Simplification	39855-42125	DD	18	Firms, workers	4
Bruhn et al. (2018)	Serbia	Simplification	108	DD	42	Firms	1
Bruhn and McKenzie (2013)	Brazil	Simplification, financial incentive	4704-4875	DD	10-12	Firms, tax revenue	2
Bruhn and Loeprick (2016)	Georgia	Tax incentive	14000-30000	DD and RDD	1-24	Firms, tax revenue	10
Campos et al. (2018)	Malawi	Information, simplification, financial incentive	10900	RCT	20.75	Firms, profitability	12
Cruces et al. (2010)	Argentina	Tax incentive	4018	FE	30	Workers, wages	2
de Andrade et al. (2016)	Brazil	Information, financial incentive, labor inspection	1346-5186	RCT	6-10	Firms	4
De Giorgi and Rahman (2013)	Bangladesh	Information	1133	RCT	15.5	Firms	1
De Giorgi et al. (2017)	Bangladesh	Information, labor inspection	1410	RCT	3	Firms	2
de Mel et al. (2013)	Sri Lanka	Information, financial incentive	520-2181	RCT	1-20	Firms, profitability	5
Fajnzylber et al. (2011)	Brazil	Simplification, tax incentive	80000	DD	5	Firms, workers, profitability, tax revenue	5
Feld (2020)	Argentina	Information, financial incentive	53691-58828	DD	12	Workers, wages	3
Fernández and Villar (2017)	Colombia	Tax incentive	345729	DD	12	Workers	1
Galiani et al. (2017)	Colombia	Information, simplification	1817-1881	RCT	6-18	Firms	4
Kaplan et al. (2011)	Mexico	Simplification	17028	DDD	36	Firms, workers	2
Kugler et al. (2017)	Colombia	Tax incentive	120770-78200257	DD	18	Workers	12
Madalozzo and Boroluzzo (2011)	Brazil	Tax incentive	31174	DD	12	Workers	1
Monteiro and Assunção (2012)	Brazil	Simplification, tax incentive	1988-2871	DD	10	Firms	5
Morales and Medina (2017)	Colombia	Tax incentive	80089-3003746	IV	12	Workers, wages	10
Mullainathan and Schnabl (2010)	Peru	Simplification, financial incentive	5258-10275	Before-after	12-24	Firms	2
Pignatti (2018)	Colombia	Labor inspection	2447341	IV	12-48	Workers	4
Piza (2018)	Brazil	Simplification, tax incentive	2597	RDD-DD	6	Firms	1
Rocha et al. (2018)	Brazil	Simplification, tax incentive, financial incentive	16676-500472	DD	6-18	Firms, workers	6
Ronconi and Colina (2011)	Argentina	Simplification	32413-272132	DD	21-48	Workers	4
Rothenberg et al. (2016)	Indonesia	Simplification	19521-133021	FE	18	Firms, workers, wages	6
Scherer (2015)	Brazil	Tax incentive	149472	DD	7	Firms	2
Slonimczyk (2012)	Russia	Tax incentive	44452-47718	DD	48	Firms	2
Souag and Assaad (2018)	Algeria	Tax incentive	3163-55079	DD	30	Workers	4
Zucco et al. (2020)	Brazil	Information	1141	RCT	10	Firms	2

Note: Table entries are the 38 primary studies contained in the meta database. The six intervention types and the six outcomes are explained in the main text. "Sample size", "identification" and "effect horizon" refer to the coded estimates. "Identification" categorizes Randomized Controlled Trial (RCT), Difference-in-Differences (DD, DDD), Regression Discontinuity Design (RDD), Instrumental Variables (IV), Fixed Effects (FE), and special cases explained in the table. "Effect horizon" specifies the months after the intervention during which effects are assessed. "Number of estimates" is the number of coefficients coded from the respective studies.