Compliance with Labour Legislation: Evidence from a Natural Experiment

May, 2018

Preliminary, please do not cite

Clemente Pignatti*
Graduate Institute of International and Development Studies
Email: clemente.pignatti@graduateinstitute.ch

Abstract

As part of the negotiations for the "US-Colombia Trade Promotion Agreement", the US Administration required the implementation of a set of measures to reinforce compliance with labour legislation. This led to the ratification of the "Colombian Action Plan Related to Labour Rights" in 2011, following which the number of labour inspectors operating in Colombia more than doubled in four years. Combining geographical differences across departments in the intensity of the reform with time differences in the exposition to the policy for subsequent cohorts of employed individuals, the analysis finds that an additional labour inspector per 100,000 employed individuals increases formal employment by around 2 percentage points. This effect partially results from a shift from self- to dependent employment; while no significant effects are found on overall labour market status (i.e. employment, unemployment, inactivity). However, the positive effect on formal employment materialises only in the formal economy and in urban areas. Due to potential endogeneity in the implementation of the policy across departments, I instrument the actual change in the number of labour inspectors with its planned change according to the reform.

* I would like to thank Martina Viarengo and Jean-Louis Arcand (Graduate Institute of Geneva) for their excellent supervision as well as Verónica Escudero and Santo Milasi (ILO), Jochen Kluve (Humboldt University), Andrea Garnero (OECD) for feedback and discussions and conference participants in Paris (Sorbonne University), Davis (PacDev 2018), Geneva (Graduate Institute BBL seminar and PhD Development Seminar) and Bogotá (Universidad del Rosario and WB-IZA-NJD Conference on Jobs and Development) for their insightful comments. I am also grateful to the Colombian Ministry of Labour for providing access to data on labour inspection. The views expressed herein are those of the author and do not necessarily reflect the views of the International Labour Organization.

I. Introduction

Developing economies are often characterised by restrictive systems of labour legislation, which formally guarantee standards of protection comparable to those granted to workers in developed economies (ILO, 2015). However, the level of protection actually enjoyed by workers depends on how effectively legislation is implemented and in developing economies non-compliance is often pervasive. This can be ascribed to a variety of reasons; ranging from labour market characteristics (e.g. prevalence of small businesses), workers' bargaining power (e.g. limited presence of trade unions) and broader institutional factors (e.g. rule of law). In this context, labour inspection represents the main policy tool used by governments to guarantee compliance. However, very little is known with respect to the effectiveness of this intervention. The absence of research is mostly connected with the lack of adequate data (on both the treatment indicator and the outcome of interest) as well as problems of econometric identification. On this latter aspect, enforcement and compliance can be stronger in areas generally characterised by more stable political and economic institutions. This is likely to generate problems of omitted variable bias, which will most likely result into an overestimation of the causal impact of inspection on compliance. At the same time, governments may want to reinforce labour inspection in areas characterised by a higher risk of non-compliance. In this case, the presence of a simultaneous relationship between the outcome of interest and the treatment indicator would result into an underestimation of the effect of the policy.

An exogenous change in treatment intensity (i.e. in the intensity or scale of labour inspection) is therefore needed to solve the identification problem and several options have been proposed in the literature in this respect. Ronconi (2010) uses electoral years as instruments for the number of inspectors in Argentina, under the assumption that in the proximity of an election the government is more willing to protect labour rights. The results show how enforcement increases compliance with mandated benefits such as minimum wage, maximum hours and health insurance. Almeida and Carneiro (2009, 2012) use distance to the labour office as an instrument for enforcement in Brazil, under the assumption that this exogenously reduces the probability that an employer is visited by labour inspectors. They find that stricter enforcement makes formal jobs more attractive at the expense of firms' size. Bhorat et al. (2012) use the number of officials working in South African regional labour centres in units different from the labour inspection as an instrument for the number of inspectors working in the same labour centre, under the assumption that non-inspectors will not affect labour law compliance. The results do not show any impact of enforcement on compliance with minimum wage. In the absence of major policy changes in labour inspection, all these studies exploit indirect shocks that generate presumably exogenous variations in enforcement levels occurring either over time or space. Of course, in an instrumental variable framework this requires assuming that these indirect shocks affect the outcomes of interest only through the hypothesised mechanism.

In this paper, I am able to partially relax this assumption by exploiting a drastic and externally induced reform of labour inspection in Colombia. In 2006, Colombia and the US signed a trade agreement (the "US-Colombia Trade Promotion Agreement", CTPA) aimed at reducing trade tariffs between the two countries. The agreement was immediately ratified by the Colombian Parliament in 2007, while the US Congress did not succeed in ratifying the CTPA by the end of the Congressional session in December 2008. A new bargaining round started in 2009, during which the newly elected US Government requested to include in the agreement measures to reinforce compliance with labour legislation in Colombia.[1] The negotiations led to the ratification in 2011 of the "Colombian Action Plan Related to Labour Rights" (in short, the Action Plan) which became an integral part of the CTPA. One of the main interventions included in the Action Plan was the commitment by the Colombian Government to more than doubling the number of labour inspectors within four years.[2]

Despite the drastic and exogenous nature of the reform, this variation alone would not be sufficient for identification purposes. Indeed, there might be multiple reasons why compliance with labour legislation has varied in Colombia in the period under consideration (i.e. independently from the role played by labour inspection).[3] Instead, central to the identification strategy of this paper is the fact that the rise in the number of inspectors was not uniform across departments. Rather, the central government set separate targets for the different departments. This generates differences in programme intensity that can be exploited to estimate causal effects under the (weaker) assumption that trends in the outcomes of interest would have not systematically differed between departments in the absence of the programme. Supportive evidence in this sense is provided by (i) checking the evolution in the outcomes of interest before the implementation of the policy; and (ii) analysing the determinants of the allocation of inspectors across departments. The presence of a policy change substantially facilitates the identification of a causal effect compared to previous contributions, since working in differences (rather than levels) eliminates the need to control for a series of factors jointly affecting enforcement and compliance (e.g. rule of law, level of corruption) that are generally regarded as unobservable but time-invariant (i.e. they can be accounted for by department fixed effects). Similar difference-in-difference approaches have been used to estimate the impact of schooling on labour market outcomes (Duflo, 2001) and the effect of female labour supply on the wage structure (Acemoglu et al., 2004).

Compared to these studies, I am able to complement the difference-in-difference framework with an instrumental variable approach that takes into account the possible endogeneity in the implementation of the policy. Indeed, the drastic nature of the reform required the different departments to rapidly scale up their enforcement systems and hire a number of new inspectors following a centrally established timeline. This contrasted with the lack of public investment in labour inspection in the previous years and the general absence of trained candidates for the positions. As a result, the yearly recruiting targets were systematically missed and the final target of 904 inspectors (originally set for 2014) was still to be met at the end of 2016. This could generate concerns over the exogeneity of the treatment indicator (i.e. total change in the number of inspectors in a department normalised by employment levels) in case that differences across departments in implementing the policy reflect broader (unobservable and time varying) differences (e.g. commitment of the labour office). This would cause an upward (downward) bias if the implementation gap reflects a weaker (stronger) institutional environment in a given department. For these reasons, I instrument the treatment indicator with the number of inspectors that was planned to work in that department according to the legislation. This information is obtained from combining together information from the calls for applications that were issued by the Colombian Ministry of Labour. In this respect, the study follows an instrumental variable approach as implemented by De Giorgi et al. (2015) among others.

[1] This request can be connected to the increasing policy trend of including labour provisions in trade agreements (ILO, 2016) and was specifically motivated by reports of systematic violations of labour rights in Colombia.
[2] Colombia had one of the lowest levels of coverage of labour inspection internationally (OECD, 2016).
[3] For instance, in 2012 a major tax reform was implemented nationwide with the aim of promoting formalization.

The results indicate that an additional inspector per 100,000 employed individuals increases the probability of being in a formal job by around 2 percentage point. This effect partially results from a shift from self- to dependent employment; while no significant effects are found on overall labour market status (i.e. employment, unemployment, inactivity). However, the positive effect on formal employment materialises only in the formal economy and in urban areas. The remainder of the paper is organised as follows. Section II reviews the literature, Section III presents the policy change and the system of labour inspection in Colombia, section IV presents the dataset to be used and the descriptive statistics; section V discusses the validity of the estimation strategy; section VI presents the main empirical results and robustness tests; section VII explores how the effect on formal employment materialises; section VIII concludes.

II. Literature review

The empirical economic literature (and the policy debate) has been traditionally focused on the effects of the strictness of formal employment legislation on labour market outcomes (Botero et al. 2004); while very few studies have examined the role played by the level of legislation effectively enforced. This research gap is particularly significant given the high levels of non-compliance with the legislation (especially in developing countries), which make the formal level of legislation only an imperfect proxy of the degree of regulation faced by workers and enterprises. As stated in Ronconi (2010), there are two key challenges in assessing empirically the effectiveness of labour inspection. First, there is a problem of endogeneity due to the possible presence of omitted variables as well as the potentially simultaneous relationship between enforcement and compliance. Secondly, it is problematic to adequately measure both the treatment indicator (strength of labour law enforcement) and the outcome of interest (compliance with labour legislation). This section will briefly review how previous contributions have dealt with these different identification problems. All the papers reviewed in this section have some commonalities, such as the focus on a single country and the use of an instrumental variable approach. However, they differ (substantially) on the choice of the instrument and (to a lesser extent) on the measures of enforcement and compliance used.

A. Econometric identification

Two main sources of endogeneity could bias simple estimations of the relationship between enforcement and compliance – assuming the absence of measurement errors. First, omitted variables could bias the results if the relationship between enforcement and compliance is affected by an unobservable characteristics that influence at the same time the treatment indicator and the outcome of interest. The direction of this source of bias is unclear, even though it is more common to imagine a positive relationship between the omitted variable (e.g. general rule of law) and the strength of enforcement – thus leading to an upward bias of simple impact estimates. Secondly, there could be a simultaneous relationship between the treatment indicator and the outcome of interest. This source of bias generates from the fact that enforcement efforts are likely to be directed towards areas with a higher perceived risk of non-compliance – thus leading to an underestimation of the causal effect in a simple OLS framework. The literature has pretty consistently adopted an instrumental variable approach in order to solve these identification problems. For instance, Ronconi (2010) exploits the fact that in the proximity of an election, governments might be more willing to strengthen labour law enforcement in order to gain political support. Under the assumption that election years are not otherwise correlated with compliance with the legislation, the results show that an additional inspector per 100,000 people increases compliance with mandated benefits in Argentina by 1.4 percentage points. Almeida and Carneiro (2012) exploit the fact that inspectors in Brazil are assigned to labour centres and then need to drive by cars to visit firms. As a result, distance to the labour centre will decrease the probability of a firm being inspected. At the same time, the effect of distance will be greater the fewer inspectors are based in a labour centre. Using this indicator of enforcement, the authors find that labour inspection increases formal employment and reduces wage dispersion in the formal sector. However, stronger enforcement increases wages in the informal sector as a result of lower mandated benefits. Bhorat et al. (2012) exploit the institutional organization of labour centres in South Africa, where inspectors and non-inspectors are based in the same centres while having different tasks. For this reason, they use the number of non-inspectors in a labour centre as an instrument for the number of inspectors in that same centre – which constitute their proxy for enforcement. Their results do not reveal any significant effect of enforcement on non-compliance with the minimum wage. Viollaz (2016a, 2016b) uses the arrival cost of labour inspectors (measured as the logarithm of per capita crossing vehicles per kilometre) as an instrument for enforcement in Peru – under the assumption that this will exogenously decrease the probability of being visited by labour inspectors. The results reveal limited effects of inspection on compliance.[4]

B. Measurement error

Challenges related to measurement errors can concern both the treatment indicator (i.e. how to measure enforcement) as well as the outcome of interest (i.e. compliance with the labour legislation).[5] In theory, firms respond to both the probability of being sanctioned as well as the amount of the sanction to be expected in case of non-compliance (Ronconi, 2010). Ideally, a measure of labour law enforcement should therefore include both the threat of the sanction (e.g. as captured by the number of inspectors and/or the number of inspections) as well as the size of the sanction (e.g. as captured by the amount of the fine). In practice, data on labour inspection is often very scant and only one of the two dimensions of enforcement is used in practice – generally covering the threat of being caught. In particular, most papers use the number of inspectors (Bhorat et al. 2012, Ronconi 2010) or the number of inspections (Almeida and Carneiro 2012; Almeida and Poole 2017) as a proxy for enforcement. This is normalised by the employment levels or the number of firms in the area of interest (i.e. city, region) in order to take into account of size effects. Bhorat et al. (2012) complement data on the number of inspectors with information on the allocated budget by the Ministry of Labour and the number of labour centres in the province; while Almeida and Carneiro (2009) use as treatment indicator an interaction term between distance to the labour office and the number of inspectors.[6]

Turning to the measurement of the outcomes of interest, it is generally agreed to use information on compliance reported by the workers – generally as collected in labour force surveys or censuses (Almeida and Carneiro 2009, Bhorat et al. 2012, Ronconi 2010). Indeed, firms are likely to under-report non-compliance with the legislation and the measurement error is also expected to be correlated with the level of enforcement (Ronconi 2010). A second problem related to the measurement of the outcome of interest concerns the choice of the indicator of labour law

[4] Other studies use similar methodologies to assess the impact of enforcement on other outcomes of interest (i.e. differently from law compliance). For instance, Almeida and Carneiro (2009) analyse the impact of enforcement on firms' size; while Almeida and Poole (2017) analyse the impact on hiring decisions following a monetary devaluation.
[5] From an econometric point of view, measurement error in the treatment indicator is a more serious concern since it will automatically generate inconsistent estimates.
[6] Other measures of enforcement used in the literature include the inverse of seignorage (Ihrig and Moe 2004) and the workforce average education (Botero et al. 2004).

compliance. Indeed, labour legislation is generally very complex and several rules regulate working conditions for different groups of workers. Additionally, labour inspectors can be assigned priority areas in certain specific types of violations of the legislation. Accordingly, most papers focus on few measures of labour law compliance based on the content of the labour code in the country and the institutional characteristics of labour inspection. These measures have included the incidence and depth of non-compliance with the minimum wage (Bhorat et al. 2012), the coverage of formal employment (Almeida and Carneiro 2009), the presence of an employment contract and pension contributions (Viollaz 2016), the right to paid holidays and compliance with the legislation over working hours' (Ronconi 2010).

III. Programme

A. Action Plan

The present analysis exploits a drastic and externally induced reform of the system of labour inspection implemented in Colombia to identify a causal effect of enforcement on compliance. In particular, the United States and the Colombian governments signed in November 2006 a comprehensive trade agreement aimed at substantially reducing administrative and fiscal barriers between the two countries (the "United States-Colombian Trade Promotion Agreement", CTPA) The strategic importance of this agreement for Colombia can hardly be overstated: the US is Colombian main trade partners, accounting for 39 per cent of the country's exports and 29 per cent of its imports. On the other hand, Colombia accounts for only around one per cent of US imports and exports. This generated an asymmetry in the importance of the agreement for the two countries, which – combined with existing geopolitical considerations – resulted in a substantial imbalance in the bargaining power during the negotiations. The ratification process required the approval of the trade agreement by both countries. Given the strategic importance of the agreement, the Colombian Congress approved the bill already in June 2007 with a large majority (84 yes and 3 no in the House Floor). In April 2007, the agreement was also sent for approval to the US Congress (Stenzet, 2008). However and despite the US Administration support for the CTPA, the debate never led to a vote amid concerns raised by the US Congress over violations of labour rights in Colombia.[7] A legislative rule change to indefinitely delay action on the CTPA was instead approved by the US Congress, after which the US Administration committed to include "worker protections in several pending trade accords".[8] As a result, the CTPA was not approved by the end of the Congressional session in December 2008.

The newly elected US Administration immediately specified its opposition to the approval of the CTPA in its current form and requested additional guarantees with respect to compliance with labour rights in Colombia.[9] This led to the definition in April 2011 of a set of measures aimed at reinforcing compliance with labour law in the country (the "Colombian Action Plan Related to Labour Rights", or Action Plan) to be implemented in collaboration with the ILO.[10] In particular,

[7] It is important to note that since January 2007 the US had a divided Government, with a Republican Administration (in favour of the agreement) and a Democratic Congress (opposing it).

[8] This was meant to apply to pending trade agreements with Colombia, Panama, Peru and South Korea.

[9] In June 2009, the US President specified there was not a "strict timetable" to the agreement amid controversies for the violation of labour rights in Colombia (e.g. the major US trade union federation was also opposed to the CTPA). This led to delays in the beginning of the new bargaining round and the final definition of the Action Plan.

[10] It is difficult to assess the extent to which economic actors could have anticipated the effects of the Action Plan (which would partially invalidate the estimation strategy). Few details of the policy could however alleviate this concern, such as (i) the fact that the policy was not debated in the national political arena (but rather between

the Action Plan established the creation of the Ministry of Labour (that resulted from the separation from the Ministry of Social Protection) as the appropriate institutional vehicle to promote labour rights and employment policies. Additionally, reforms were agreed to strengthen collective bargaining and prevent the misuse of temporary agency workers. For the purpose of this analysis, the Action Plan also established that the Ministry of Labour would have hired 480 new inspectors within four years (i.e. by the end of 2014, 100 new inspectors to be hired already in 2011) in a move that would have more than doubled the total number of inspectors in the country. Additionally, the Action Plan included measures aimed at reinforcing citizens' ability to file complaints to labour inspectors (either anonymously or otherwise). This included the institution of a free telephone line as well as the possibility to file complaints online from the website of the Ministry. Awareness campaigns were also launched in order to increase citizens' use of these tools.[11] After the Action Plan was agreed, the US Congress approved the CTPA in October 2011 and the trade agreement went into force in May 2012.

As a result of the Action Plan, the number of labour inspectors in Colombia increased from 353 in 2010 to 726 in 2014.[12] This represents a substantial improvement in international terms, as the number of inspectors per 100,000 employed individuals has evolved from around 2 to 4 between 2011 and 2015 (ILO Statistics). This value is comparable to the one observed in other countries in the region (e.g. 4 inspectors per 100,000 employed persons in Argentina, 3 in Brazil, Ecuador and Peru, 1 in Paraguay – although Chile and Uruguay have a much higher level), but is still considerably below the value registered in most developed economies – despite the lower perceived risk of non-compliance with labour law in these economies (see Figure 1 in the Appendix).[13] Importantly for the identification strategy of this paper, the increase in the number of inspectors was not homogeneous across departments. In particular, the central government set different targets at the department level and provided the necessary financial resources for the hiring process. Additionally and despite the commitment of the Government to rapidly raise the number of inspectors, the hiring process took longer than expected. This was mostly due to organizational delays as well as the lack of enough qualified candidates for the new positions. As a result, the planned target of 904 inspectors (originally set for 2014) was still to be met in 2016.[14]

B. Labour inspection

Labour inspectors in Colombia are in charge of securing law enforcement and guaranteeing the respect of collective bargaining rights. Candidates need to hold at least a Bachelor's Degree in either law, public administration or medicine and have seven months of relevant work experience. In their operations, inspectors are asked to follow ILO guidelines as summarised in a manual prepared by the Ministry of Labour. This contains the professional and personal requirements to be met by labour inspectors and details the administrative procedures and the code of conduct. Additionally, training initiatives are regularly organised by the Ministry of Labour. Labour inspectors in Colombia have the authority to verify compliance with a number of labour norms;

governments of two different countries); and (ii) its swift implementation (first call for applications was released already in April, with the aim of hiring 100 inspectors already in 2011).

[11] However, all other policy changes contained in the Action Plan (i.e. apart from the hiring of new inspectors) were ruled out nationally and they did not set targets that differed across departments.

[12] In particular, the number of inspectors increased from 353 in 2010 to 412 in 2011, 467 in 2012, 575 in 2013 and 726 in 2014. As a reference, the number of inspectors was equal to 280 in 2008.

[13] For instance, the number of inspectors per 100,000 employed persons is equal to 15 in Germany, 8 in France and 6 in Japan. However, Anglo-Saxon countries all have a lower number of labour inspectors.

[14] Apart from hiring new inspectors, the Government also committed to fill existing vacancies. This explains why the final target (904 inspectors) was higher than the existing level in 2010 (353) plus the new hiring (480).

including minimum wage, social security contribution, working hours, leave days, employment contract and collective bargaining rights. Priority areas have been identified in terms of labour formalization and labour intermediation. Inspectors can either undertake preventive actions or respond to complaints. Based on the situation with which they are confronted, inspectors can either (i) conduct a preventive and informative function (e.g. inform the employer about a breach in the legislation); (ii) act as mediators between the employer and the workers (e.g. help defining an agreement in case of collective dismissals); or (iii) impose fines and sanctions. Sanctions can be of financial nature (going from 1 to 50,000 times the minimum wage) and/or involve the temporary closure of the enterprise (from 3 to 10 days, up to 30 in case of recidivism) or its immediate shut-down in case of health or security risks. To prevent corruption, inspectors are hired with an open-ended contract and can be dismissed only for disciplinary reasons

IV. Data

A. Database

The present analysis draws from three different sources of information. First, the population of interest comes from repeated cross sections of the Colombian Integrated Household Survey (GEIH) conducted by the National Administrative Department of Statistics (DANE). The current version of the GEIH is operational since 2008; when (i) the sample size and coverage has been increased, (ii) electronic devices have been introduced for data collection, and (iii) the scope of the analysis has been extended. The GEIH has a two-stage stratified sample and interviews every year around 250,000 households. The sample that will be used in this study corresponds to the entire working age population between 2009 and 2014.[15] The current version of the database has been made available to the author by the ILO Department of Statistics, which reviews the survey in order to construct internationally comparable labour market indicators.

Using information on the department of residence and the year of the interview, I match the survey data with department-level administrative information on the number of inspectors provided by the Colombian Ministry of Labour.[16] Some remarks concern the matching process between the administrative and survey data. In particular, the Ministry of Labour follows the traditional division of the country into 32 departments to organize the system of labour inspection – while the GEIH covers only the 24 main departments in Colombia. Since it is impossible to connect observations in the GEIH to the 32 departments, I conduct the analysis for the 24 departments for which I have both survey and administrative data.[17] In practice, the problem is of limited empirical importance given the limited size of the population living in the excluded departments (representing less than 5 per cent of the Colombian population).[18]

[15] Starting the analysis from 2009 is motivated by the fact that we have a new consistent measure of formal employment from that year (i.e. due to changes in the questionnaire).

[16] Ideally, I would need information on the department where the individual works – rather than the department of residence. Unfortunately, this information is reported in the GEIH only from 2012 onwards. However and given the large size of the departments, this is unlikely to create major issues. In particular, in 2014 only 0.8 per cent of the sample reported living and working in two different departments.

[17] Starting from 2014, the GEIH covers five additional departments (Arauca, Casanare, Putumayo, San Andrés and Vichada). However, the department variable still reports only 24 values.

[18] Another codification problem concerns the departments of Cundinamarca and Bogota. In particular, the system of labour inspection within the Ministry of Labour initially considered the two departments as part of the same administrative unit (named Cundinamarca), so that data on labour inspectors is reported jointly until 2012 and then separately only from 2013 onwards. In order to have the entire series (2009-2014) for both departments, I compute the ratio of inspectors working in Cundinamarca (Bogota) using the two separate series from 2013 to 2015. This ratio is relatively constant for the three years with available information (from 77 to 81 per cent of the total number of

Finally, I obtained information on the quality of the public administration from the Departmental Institutional Environment and Performance Survey (EDID). This is a survey conducted by DANE that interviews randomly selected public officials working in the central administration in the different departments,[19] with the aim of obtaining information on their perception of the institutional environment. In particular, the survey asks detailed information on the official's perception of the management, efficiency and transparency of the public administration – including how tasks and responsibilities are assigned; according to which principle officials are recruited; whether the institution has the necessary resources. Answers to these questions are grouped to compute intermediary indicators (credibility in the rules, credibility in the policies, adequacy of resources, result-based management, accountability, labour welfare, prevention of irregular practices and citizen participation) and then overall indicators (institutional environment and institutional performance). This information is available at the department level from 2009 onwards.[20]

B. Outcome of interest and treatment indicators

The main outcome of interest in the analysis corresponds to formal employment as reported by employed persons in the GEIH. The definition of formal employment follows ILO guidelines and the variable has been coded by the ILO Department of Statistics. In particular, employees are defined as formal workers if their employer contributes (either partially or entirely) to their pension scheme. Employers and own-account workers are instead included in formal employment if they are registered at the relevant public authority or keep track of the accounting activities of the business. Finally, contributing family workers are always presumed in informal employment. The decision to focus on formal employment is motivated by the explicit policy priority of inspectors in terms of labour formalization (see section III above) as well as the overarching nature of this concept as an indicator of employment quality independently from the status in employment. In particular, the alternative strategy of selecting different indicators of labour law compliance (e.g. minimum wages, maternity leave, paid holidays) would have limited the analysis to specific sub-groups of the working population that are not necessarily representative of the overall labour market (around 43.5 per cent of the employed population was estimated being contributing family workers in Colombia in 2017).[21] Additionally, this would have increased the risk of not taking into account general equilibrium effects (e.g. increase in mandated benefits for employees paralleled by a decrease in the share of employees in the labour force).

As stated in Ronconi (2010), an ideal measure of enforcement would cover both the threat of being caught (as proxied for instance by the number of inspectors or inspections) as well as the size of the sanction (as proxied by the amount of the sanctions). In practice, data on labour inspection is generally very scant and the existing contributions included information only on the threat of being sanctioned (i.e. number of inspectors or number of inspections). I follow these contributions and use as main treatment indicator the number of inspectors operating in each

inspectors for the region works in Bogota) and its average is used to obtain separate series for those years when only aggregate information is available (2009-2012).

[19] In this sense, the sample does not include officials working in local offices or individuals hired under alternative work arrangements (e.g. external collaborators)

[20] Unfortunately, it is not possible to trace to which entity of the public administration in a given department the respondents belong to (e.g. Ministry of Labour, anti-corruption authority). This information is recorded, but not made available by DANE due to confidentiality issues.

[21] The analysis of wage aspects is anyway limited by the absence of adequate information in the GEIH on the number of weeks and hours worked within the month.

department between 2010 and 2014.[22] This is normalised by the number of employed individuals in a given department in order to take account of the different size of each department (Bhorat et al. 2012, Ronconi 2010).[23] Following discussions with policy experts, using the number of inspectors appears to be a more sensitive approach than relying on the number of inspections conducted (or the amount of the fines imposed).[24] Indeed, the new system of labour inspection in Colombia – as reformed in collaboration with the ILO – has moved away from the traditional enforcement model based exclusively on sanctioning strategies (i.e. maximise the number of inspections or the amount of fines) towards a strategic compliance model. This new model starts from the recognition that not all employers are motivated to comply with the legislation exclusively on the grounds of a cost-benefit analysis (i.e. comparing the economic benefits of non-compliance with the possible costs of being caught); but they rather respond to a variety of factors such as habits, peer pressure and civic motivations. Under this assumption, sanctioning tools represent only one of the options to be used by inspectors to promote compliance.[25] More broadly, under this scenario inspectors aim to (i) understand the causes of non-compliance; (ii) engage with relevant stakeholders; and (iii) design tailored interventions with multiple arms (ILO, 2017).[26]

C. Descriptive statistics

For ease of exposition, I divided the 24 departments with available information into quartiles according to the intensity of the policy in that given department as a result of the Action Plan (i.e. low, medium-low, medium-high and high).[27] Descriptive statistics for these groups at baseline are presented in Table 1. An analysis of the data reveals the absence of notable differences in individual characteristics (e.g. age, share of men and average years of education) across departments characterised by different levels of policy intensity (Panel A). Some differences emerge with respect to the share of individuals in employment (higher in medium-high policy intensity departments), the share of individuals outside the labour market (lower in medium-high policy intensity departments) and the share of employed individuals in a formal job (lower in high policy intensity departments) (Panel B). At the same time, the quality of the public sector (as measured by the EDID) does not systematically differ (Panel C). Finally, the last panel shows how the intensity of the policy was substantially different across departments (the change in the number of inspectors per 100,000 employed ranges from 1.1. to 2.5) and it did not necessarily match initial differences in the number of inspectors. Additionally, the implementation gap (i.e. difference between target and actual number of inspectors in 2014) is higher in high policy intensity departments (Panel D).

[22] I do not include in the analysis all labour inspectors operating in Colombia – but only those assigned to a specific department and that can thus be matched to individuals in the GEIH. This leaves out inspectors working in the newly created directorial offices as well as those assigned to special offices in charge of a specific town (i.e. overall corresponding to around 10 per cent of inspectors).

[23] I normalise the treatment indicator by the number of employed individuals – rather than by the number of firms, as done for instance in Almeida and Carneiro, 2012 – as the literature has shown how inspectors tend to disproportionately target large establishments.

[24] Additionally, the target as part of the Action Plan was set in terms of inspectors (rather than inspections or fines).

[25] This is confirmed by an analysis of the data on the number of labour inspections. Indeed and while inspections have also increased during the period under analysis (OECD, 2015), their number is much more volatile across years and this volatility is unlikely to reflect actual variations in the level of enforcement. For instance, the number of procedures started has increased from 16,546 in 2014 to 21,055 in 2015 but then decreased to 4,966 in 2016.

[26] In particular, interventions undertaken by labour inspectors to promote compliance could include (i) enforcement activities (e.g. inspections and sanctions); (ii) education activities (e.g. training and recommendations); (iii) communication strategies (e.g. name and shame campaigns); (iv) political actions (e.g. proposing to set up a parliamentary commission); and (v) systemic interventions (e.g. certifications, grants) (ILO, 2017).

[27] This is done according to the total change in the number of inspectors (per 100,000 employed individuals) between 2010 and 2014 (first variable in panel D of Table 1)

Table 1: Descriptive statistics at baseline (2010)

	Low	Medium-low	Medium-high	High
	Panel A: Personal characteristics			
Age	39.06	39.02	38.72	38.12
	(0.56)	(1.01)	(1.73)	(1.48)
Male	0.49	0.49	0.49	0.49
	(0.01)	(0.01)	(0.01)	(0.01)
Years of education	7.63	7.74	7.56	7.01
	(0.92)	(1.26)	(0.55)	(0.59)
	Panel B: Labour market			
Employed	0.59	0.58	0.62	0.56
	(0.05)	(0.05)	(0.03)	(0.03)
Unemployed	0.08	0.07	0.07	0.08
	(0.02)	(0.01)	(0.01)	(0.02)
Inactive	0.34	0.35	0.31	0.36
	(0.06)	(0.05)	(0.04)	(0.04)
Formal employment	0.27	0.27	0.25	0.23
	(0.11)	(0.13)	(0.06)	(0.08)
	Panel C: Quality of public sector			
Institutional environment	3.60	3.62	3.65	3.61
	(0.20)	(0.14)	(0.15)	(0.29)
Institutional performance	3.62	3.64	3.71	3.68
	(0.17)	(0.07)	(0.21)	(0.18)
	Panel D: Policy intensity (per 100,000 employed)			
Total change in inspectors (2014-2010)	1.09	1.36	1.90	2.49
	(0.40)	(0.13)	(0.21)	(0.37)
Gap with respect to target (in 2014)	0.02	0.14	0.03	0.39
	(0.32)	(0.12)	(0.08)	(0.58)
Inspectors in 2010	1.78	2.17	2.25	2.03
	(0.63)	(0.83)	(0.56)	(0.67)

Note: Departments are divided into quartiles (low, medium-low, medium-high and high) according to the intensity of the policy as captured by the total change in the number of inspectors per 100,000 employed between 2010 and 2014.

V. Identification strategy

A major challenge in the identification of a causal effect of enforcement on compliance with labour legislation is that the level of enforcement is generally set at the national level and/or it rarely varies over time (Almeida and Carneiro 2009). As a result, it might be difficult to disentangle whether the (positive or negative) relationship between enforcement and compliance arises from a causal effect or is rather the result of some spurious correlation at the level of the unit of interest – generally the municipality or the district. Two main reasons of concern relate to the possible presence of omitted variable bias and the possible simultaneous relationship between enforcement and compliance (Section II). For these reasons, previous studies have exploited indirect shocks (over time or space) that generate presumably exogenous variations in the level of enforcement. Of course, these studies critically rely on the assumption that these indirect shocks affect the outcomes of interest (i.e. labour law compliance) only through the hypothesised mechanism (i.e. enforcement levels). Unfortunately, it is possible to imagine different reasons of violation of this assumption. For instance, Almeida and Carneiro (2009) use firms' distance to the labour office as an instrument for enforcement. However, the decision of being far from the labour office could correlate with other characteristics that independently determine compliance (e.g. specialization in certain sectors, remoteness to specific markets) or it could even be the result of a strategic decision of non-compliant firms (i.e. which might want to minimise the risk of being inspected by placing themselves far from a labour office). Similarly, Bhorat et al. (2012) use the number of non-inspectors operating in a labour office as an instrument for the number of inspectors (and therefore the level of enforcement) in the province where the labour office operates. However, the decision over the allocation of non-inspectors to a labour office is likely to be determined by similar considerations as the allocation of inspectors (e.g. risk of non-compliance, strength of the institutions) – thus presenting again concerns over omitted variable bias or simultaneity.[28]

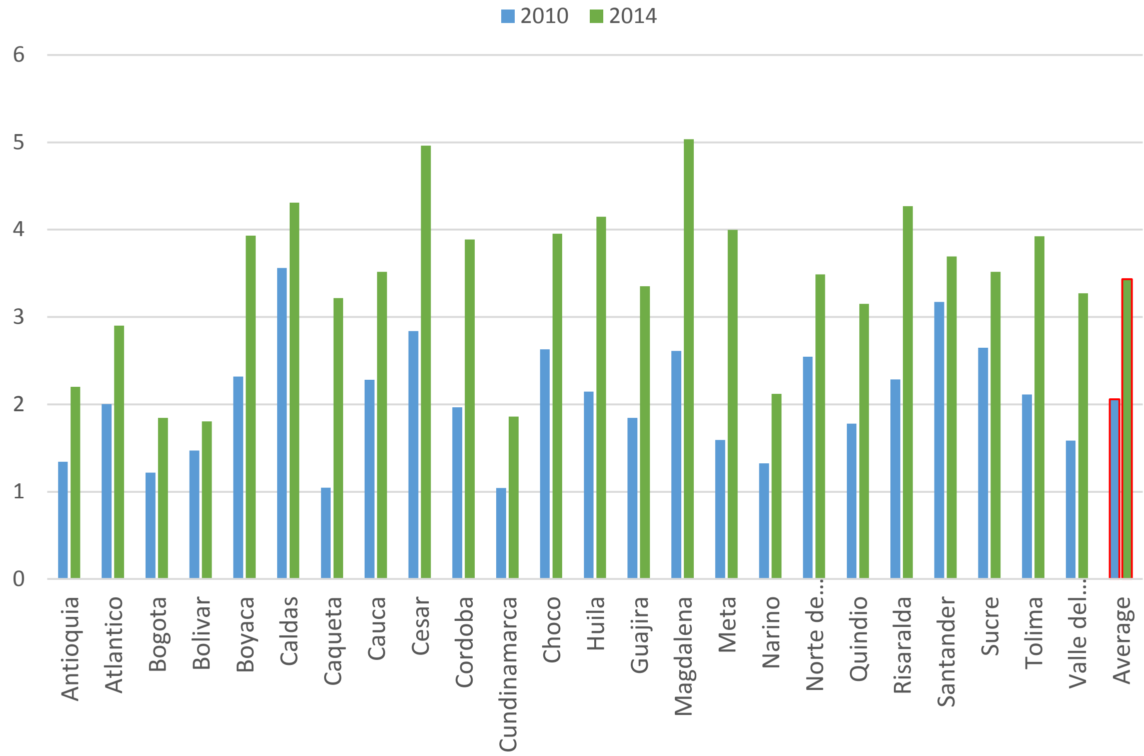
Many of these identification problems relate to the possible presence of unobserved differences in the rule of law (and therefore the risk of non-compliance with the labour legislation) across the geographical areas of interest (e.g. municipalities, regions) characterised by different levels of enforcement (i.e. omitted variable bias). These differences are likely to be determined by a multitude of cultural factors (e.g. trust in the government) and economic incentives (e.g. sectorial specializations) that are difficult to account for.[29] However, many of these factors are generally regarded as constant over time and an analysis of longitudinal nature could alleviate these identification problems. Unfortunately, most of the available studies take a cross-sectional approach due to the absence of adequate information on enforcement and compliance over successive years. The only notable exception is represented by Ronconi (2010), which constructs a panel database of Argentina's provinces between 1995 and 2002. In order to deal with the other major issue of econometric identification (i.e. the possible presence of simultaneity), the study uses electoral cycles as an instrument for inspection under the assumption that in the proximity of an election governments might be more willing to protect labour rights by reinforcing inspection. However, electoral cycles are likely to generate shifts in governments' efforts that go well beyond the role played by labour inspection (e.g. changes in the legislation). Similarly, employers' perceptions might vary in the proximity of an election (e.g. expectation that the next government will be less/more lenient on non-compliance) and workers' organizations might shift their priorities away from denouncing cases of non-compliance (e.g. due to electoral campaigning). All these dynamics would generate violations of the exogeneity assumption of the instrumental variable strategy, thus causing a bias in the impact estimates of enforcement on compliance.

The identification strategy of this paper exploits the drastic and externally induced increase in the number of inspectors following the implementation of the Colombian Action Plan. However, this variation alone would not be sufficient for identification purposes. Indeed, there might be multiple reasons why compliance with labour legislation has varied in Colombia in the period under consideration (i.e. independently from the role played by labour inspection). Instead, central to the identification strategy of this paper is the fact that the rise in the number of inspectors was not uniform across departments. While the average increase in the number of inspectors per 100,000 employed individuals was equal to 67 per cent between 2010 and 2014 (from 2.05 to 3.43 inspectors per 100,000 employed); this increase varied from a minimum of 21 per cent (from 3.56 to 4.31 in the department of Caldas) to a maximum of 208 per cent (from 1.04 to 3.21 in the department of Caquetá – see Figure 1). This generates differences in programme intensity that can be exploited to estimate causal effects under the (weaker) assumption that trends in the outcomes of interest would have not systematically differed between departments in the absence of the programme. This section will examine the plausibility of this assumption by (i) looking at trends in the outcomes of interest before the implementation of the policy (section IV.A); and (ii)

[28] Additionally, inspectors and non-inspectors are likely to cooperate (e.g. information sharing, overlap of functions) – thus blurring the difference between the endogenous regressor and the instrumental variable.
[29] Available studies address this shortcoming by controlling for a rich set of observable characteristics that could proxy for the unobservable ones (Almeida and Carneiro 2009 and 2012, Bhorat et al. 2012).

analysing the determinants of the allocation of inspectors across departments (section IV.B). This identification strategy has already been adopted in the literature[30] to study the effect of an increase in the number of schools on educational attainments (Duflo 2001), the effect of an increase in female labour supply on the subsequent wage structure (Acemoglu et al. 2004) and the effect of a shift in sex-specific income on sex-differential survival rates of children (Qian 2008).[31]

Figure 1 – Actual change in the number of inspectors (per 100,000 employed)



Source: Author's calculations based on GEIH and administrative data.

Compared to these studies, I am able to complement this difference-in-difference framework with an instrumental variable approach that takes into account the possible endogeneity in the implementation of the reform. Indeed and as discussed above, the hiring targets set by the central Government were systematically missed by the departments. As a result, the planned target of 904 inspectors (originally set for 2014) was still to be met in 2016. Since the hiring process took place at the department level, this introduces a possible source of endogeneity. In particular, the ability of the different departments to comply with the policy reform might be correlated with time-varying characteristics at the department level that are not accounted for by the department fixed effects. For instance, the motivation to implement the Action Plan might have varied according to the political stance of the local government (or its alignment with the central government). Similarly, the hiring of new inspectors might have changed the behaviour of inspectors that were already working in the labour office. As an extreme, it could be assumed that local governments

[30] This approach has also been used specifically in studies on labour legislation, adopting either a difference-in-difference framework (Micco and Pages 2007) or an event study approach (Ahsan and Pages 2007, Amin 2007, Autor et al. 2007 Besley and Burges 2004). From a methodological point of view the paper closest to this one is Almeida and Poole (2017), which exploit a policy shock for identification purposes and use individual level data on compliance and matched with administrative measures of enforcement at the city level.

[31] Compared to these studies, the advantage of the policy reform used in this paper is its presumed exogeneity (i.e. as asked by the US Administration).

received pressures from lobby groups not to hire the new inspectors. For this reason, I instrument the change in the number of inspectors working in a given department between 2010 and 2014 with the theoretical change that should have taken place during the same time in that department according to the Action Plan. This information is obtained from the calls for applications that can be consulted on the website of the Ministry of Labour. This section will also explore the determinants behind the implementation gap and discuss the validity of the instrumental variable approach (section IV.C).

A. Common trends

The main idea behind the identification strategy can be presented using a two-by-two table (Duflo, 2001). Table 2 presents differences in the rates of formal employment for the four groups of departments divided according to the intensity of the policy (low, medium,-low, medium-high and high, as defined above) before and after the implementation of the Action Plan. Panel A shows that before the policy was implemented (2009 to 2010), the rate of formal employment was essentially constant in all four groups of departments – providing suggestive evidence in favour of the common trend assumption. At the same time, Panel B shows that in the period after the policy change (i.e. 2010 to 2014) formal employment was increasing in all groups of departments. However, the rate of increase was higher in the departments characterised by higher policy intensity (i.e. 6 versus 4 percentage points). The difference in this difference can be interpreted as the causal effect of the programme. For instance, the second and fourth quartiles of departments experienced a difference in the intensity of the policy by around one inspector per 100,000 employed individuals (confront Table 1). Under the common trend assumption, this seems to have generated an increase in formal employment by around 2 percentage points – which can be ascribed to the programme. However, this is only suggestive evidence based on data at the department level and few observations. Section V will provide more robust tests for the validity of the common trend assumption as well as more precise estimates of the causal effect.

Table 2 – Means of rates of formal employment

	Panel A: Control Experiment (2009-2010)			
	Low	**Medium-low**	**Medium-high**	**High**
2009	0.28	0.27	0.24	0.23
2010	0.27	0.27	0.25	0.23
Difference	-0.01	0.00	0.01	0.00
	Panel B: Experiment of interest (2010-2014)			
	Low	**Medium-low**	**Medium-high**	**High**
2010	0.27	0.27	0.25	0.23
2014	0.31	0.31	0.29	0.29
Difference	0.04	0.04	0.05	0.06

Note: Departments are divided into quartiles (low, medium-low, medium-high and high) according to the intensity of the policy as captured by the total change in the number of inspectors per 100,000 employed between 2010 and 2014.

B. Allocation of inspectors

The next step concerns understanding the determinants of the allocation of inspectors across departments (Acemoglu et al., 2004). Unfortunately, there was no explicit rule that was defined (or made public) by the central Government to allocate the newly hired inspectors across departments. This complicates the analysis compared to a case in which a policy is implemented gradually across regions for some reasons that can be formally identified in the legislation (e.g. funding availability) and controlled for by the researcher.[32] In order to investigate the assignment mechanism, I regress the theoretical change in the number of inspectors per 100,000 employed individuals between 2010 and 2014 (the instrumental variable in the rest of the analysis) on department level characteristics at baseline in a single cross section analysis for 2011.[33] This is meant to provide suggestive evidence on whether the (future) intensity of the policy is correlated with some department level characteristics at baseline. The covariates in this regression include the number of inspectors per 100,000 employed in 2010 (a proxy for the strength of enforcement in the department before the policy came into effect), the initial rate in the outcome of interest (i.e. formal employment) as well as its one year variation and the two main indicators of the quality of the public administration from the EDID (and their one year variation). The results show how the intensity of the policy is not associated with any of the included covariates.[34] This provides suggestive evidence that the allocation rule – despite not being made public – should not represent a threat to the validity of the present estimation strategy. The same message is conveyed by Figure 2 in the Appendix, which shows the lack (presence) of a systematic relationship between the intensity of the policy and changes in the outcome of interest before (after) the policy (Panel A and B, respectively).

Table 3 – Determinants of the allocation of inspectors across departments

	Total theoretical change in the number of inspectors per 100,000 employed individuals (2010-2014)						
Rate formal employment	-0.681						0.0605
	(0.820)						(1.091)
Δ Rate formal employment		0.537					
		(8.421)					
Inspectors (per 100,000) in 2010			0.205				0.258
			(0.151)				(0.230)
Δ Inspectors (per 100,000) in 2010				-0.0681			
				(0.297)			
Institutional environment					-0.715		-1.277
					(1.206)		(1.175)
Institutional performance					0.900		1.163
					(1.351)		(1.383)
Δ Institutional environment						0.219	
						(1.498)	
Δ Institutional performance						-0.276	
						(1.461)	
Constant	1.890***	1.663***	1.294***	1.693***	1.009	1.656***	1.610
	(0.338)	(0.128)	(0.257)	(0.182)	(2.671)	(0.115)	(2.843)
Observations	24	24	24	24	24	24	24
R-squared	0.021	0.000	0.056	0.001	0.011	0.002	0.084

Note: Robust standard errors in parenthesis. The dependent variable is the change in the number of inspectors per 100,000 employed individuals between 2010 and 2014. * p<0.1, ** p<0.05, *** p<0. 01.

C. Instrumental variable

As a final step, I will explicit the rationale and validity of the instrumental variable approach that completes the identification strategy. Indeed and as discussed above, the hiring targets in terms of

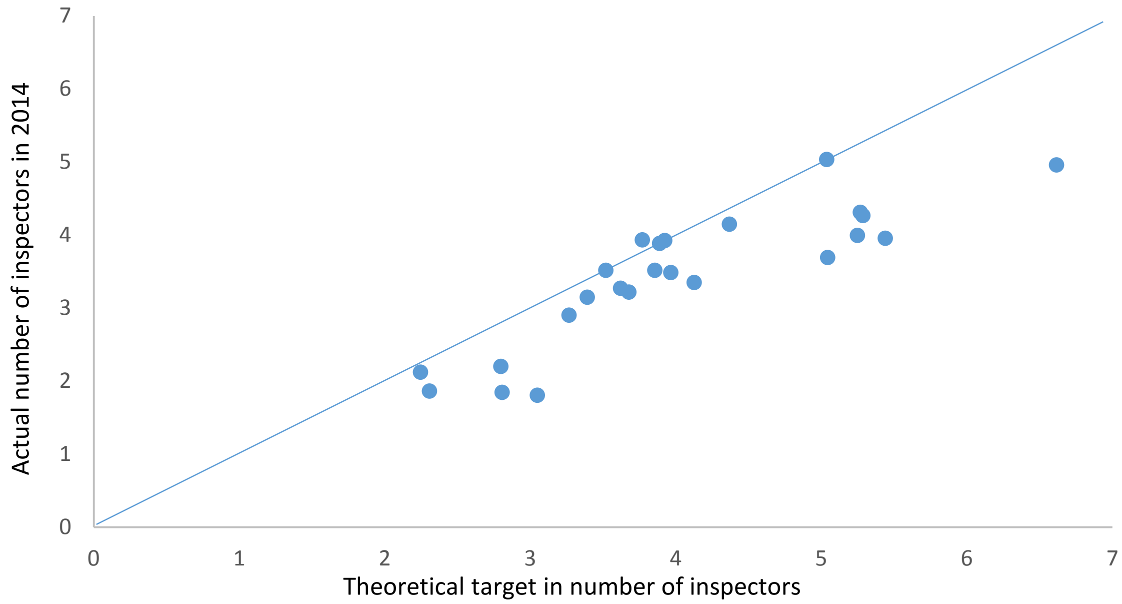
[32] However, it is important to note how we are only concerned with time-varying (but not time invariant) characteristics at the department level (i.e. the analysis will have department fixed effects).

[33] The analysis conducted using the actual (rather than the theoretical) change yields similar results. However and since the objective is to understand the determinants behind the allocation rule made by the government, the theoretical change is preferred in this step.

[34] Specifications including other department characteristics (e.g. education) yield similar results.

new inspectors set by the Action Plan were systematically missed at the department level. For that reason, the target of 904 inspectors set for 2014 was still to be met in 2016. Figure 2 plots the relationship between the total hiring target and the actual number of inspectors operating in each department in 2014 – and in this sense represents the first stage relationship in the identification strategy (although in levels rather than in changes). As shown in the figure, not all departments missed the hiring target set by the central government. Rather, some departments had met the target and are therefore on the 45 degree line; while other departments are on the right of that line. This might generate concerns if we were to directly use the number of inspectors operating in a department as the treatment indicator, insofar as missing the hiring target by a given department is unlikely to be purely random. For this reason, the paper will use the target change in the number of inspectors set by the Action Plan as an instrument for the actual change in the number of inspectors operating in the department in 2014. This is obtained from repeated calls for applications that were issued every year, detailing the number of vacancies by department.[35]

Figure 2 – Target and actual number of inspectors (per 100,000 employed)



Note: the figure shows the relationship between the target number of inspectors per 100,000 employed individuals as set by the Action Plan and the actual number of inspectors operating in a given department in 2014. The distance with the 45 degree line therefore measures the implementation gap of the Action Plan.

This instrumental variable strategy has several empirical benefits and has been already used in the literature (De Giorgi et al., 2015). First, the first stage relationship between the instrumental variable and the endogenous treatment indicator is generally very strong. This will further be proved in the regression analysis, but it can already be deduced by Figure 2 above. Secondly, the monotonicity assumption is likely to hold since a higher target has generally resulted into more inspectors *ceteris paribus*. A possible problem could be associated to the presence of the so-called always takers (i.e. departments that would have hired additional inspectors also in the absence of the policy), but this was unlikely to happen given previous trends in labour inspection. Finally, the exclusion restriction is realistic since issuing a new vacancy can be assumed to have an impact on formal employment only through the new inspector that is eventually hired. The main threat would be associated to the possible presence of a scare effect (economic agents reacting to the

[35] The hiring process, requirements and conditions of employment were instead common across the country.

announcement), but these dynamics are eventually more likely to have taken place for the Action Plan as a whole – rather than for the single new vacancy opened.[36]

Despite the fact that the instrumental variable strategy should address concerns over the possibly endogenous implementation of the policy, it is still important to predict the direction of the bias generating from a simple OLS estimation. For instance, missing the hiring target might reflect a weaker institutional environment (e.g. low ability and/or commitment of inspectors initially operating in the department), limited organizational capacities (e.g. to set up the hiring process) and even political pressures not to implement the Action Plan (e.g. through corruption) – all aspects that would lead to an overestimation of the causal effect from a simple OLS analysis. At the same time, it could be the case that the implementation gap is higher in those departments that followed closely the formal procedures for hiring the new inspectors (e.g. issued the calls for applications, interviewed pre-screened candidates). In that case, the fewer new inspectors hired in these departments are better qualified and/or operate in a labour office which is more sensitive to the respect of internal procedures – aspects pointing towards a possible underestimation of a simple OLS specification.

To investigate the determinants of the implementation delays, I regress the implementation gap (i.e. difference between theoretical and actual number of inspectors per 100,000 employed individuals in 2014) on some measures of the quality of the public administration at the department level in 2014 from the EDID.[37] The results (Table 1 in the Appendix) reveal how the implementation gap is positively associated with the quality of the institutional performance in a given department, with the relationship being driven by the capacity of the public administration in that department to prevent irregular practices. Despite this is only suggestive evidence, it points to the possible positive nature of the implementation gap (e.g. departments preventing irregular practices take more time to hire the new inspectors and therefore accumulate a delay in the implementation of the Action Plan). This would result into an underestimation of the causal impact of enforcement on compliance of a simple OLS analysis, as the fewer new inspectors would be more qualified and/or would operate within a better work environment.

VI. Estimation results

A. Basic results

The identification strategy introduced in the section above can be generalised to a regression framework (Duflo 2001). Consider the difference in the probability of being in a formal job between a cohort of employed individuals exposed to the programme (i.e. interviewed between 2011 and 2014) and a cohort of employed individuals not exposed to the programme (i.e. interviewed in 2009 and 2010). If the hiring of additional inspectors led to higher labour law compliance, this differences will be positively associated with the number of inspectors hired in a particular department. In practice, I will run the following regression:

$$FE_{i,s,t} = c + \alpha_t + \delta_s + \beta\left(X_{i,s,t} * Post\right) + \gamma(C_s * Post_t) + \ \theta(P_s * Post_t) + \varepsilon_{i,s,t} \quad (1)$$

[36] As a final point, it is worth mentioning that this instrumental variable approach allows only for unilateral deviations of the endogenous regressor with respect to the instrument. Indeed and while some departments have missed the target; there is no department that has exceeded it (i.e. all dots lie on or to the right of the 45 line in Figure 2, with the exception of one). This is a case similar to Angrist (2006), but here the treatment indicator takes a continuous (rather than binary) form.

[37] In particular, I include the two main indicators of institutional environment and institutional performance as well as the different sub-indicators that compose them.

where $FE_{i,s,t}$ is a dummy for formal employment for individual *i*, employed in department *s,* in year *t* (i.e. taking the value of one if the individual is in formal employment); c is a constant, α_t is the set of year dummies; δ_s represents the department dummies; $X_{i,s,t}$ is a vector of individual characteristics (age, gender and years of education), C_s contains department specific variables (i.e. the strength of labour enforcement and the rate of formal employment both measured at baseline); $Post_t$ is a dummy taking the value of one after the policy has been implemented (from 2011 onwards in all regions) and P_s represents the intensity of the policy (actual change in the number of inspectors per 100,000 employed individuals in a given department between 2010 and 2014). As discussed above, this last variable is instrumented by the planned change in the number of inspectors that should have taken place in that department between 2010 and 2014 according to the Action Plan. A linear probability model is preferred to discrete choice models in order to avoid the risk of inconsistent estimates in case of misclassification of the dependent variable (Hausman et al. 1998). The Heckman correction for sample selection is included to take into account that the observed outcomes of interest (i.e. formal employment) are recursively defined for individuals that are already in employment. This correction uses as exclusion restriction whether the individual is paying or not for the house where she lives (e.g. rent, mortgage). Standard errors are clustered at the department and year level, where treatment variations take place (Abadie et al. 2017).[38]

Table 4 presents the results of different estimations of equation (1), with additional sets of covariates added sequentially. The only coefficient reported refers to the effect of the policy (θ in the notation of the equation above). The estimates suggest that (in the most complete specifications), an additional inspector per 100,000 employed individuals increases the probability of being in a formal job by 1 percentage points in the linear probability model and 2.1 percentage points in the instrumental variable results. These estimates are economically significant and in line with the results obtained by previous studies.[39] First stage results (presented in Table 2 in the Appendix) show a very strong relationship between the instrument and the endogenous regressor – as expected given the nature of the instrument – and the test of exogeneity confirms the soundness of opting for an instrumental variable approach. The difference between the linear probability and instrumental variable results point towards a systematic underestimation of results obtained from the linear probability model (of around one percentage point). This is consistent with the suggestive evidence presented above (section III) that the implementation gap was higher in those departments where the public administration prevents the emergence of internal irregular practices. This would imply that the fewer new inspectors hired in these departments were more qualified and/or started operating in a better work environment, which increased their productivity in terms of ensuring labour law compliance.

[38] Instrumental variable results with bootstrapped standard errors are reported in Table 3 in the Appendix.
[39] For instance, Ronconi (2010) finds that an additional inspector per 100,000 employees increases compliance with mandated benefits by 1.4 percentage points in Argentina.

Table 4 – Treatment effects on formal employment

	Linear probability			Instrumental variable		
	(1)	(2)	(3)	(1)	(2)	(3)
θ	0.0154***	0.0153***	0.00948**	0.0259***	0.0242***	0.0211***
	(0.00440)	(0.00480)	(0.00402)	(0.00558)	(0.00512)	(0.00521)
R-squared	0.065	0.153	0.153			
First stage (F)				209.954	208.709	214.576
Test of exogeneity (p)				0.0018	0.0149	0.0066
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes
N	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the dummy taking the value of 1 after the policy change. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

B. Interaction term analysis

The results discussed above show how estimates are robust to the inclusion of additional covariates – thus potentially ruling out that they are spuriously driven by differences in (the levels of) either individual or departmental characteristics. However, they do not shed light on the possible presence of differential trends in the outcomes of interest before the implementation of the policy change (section III provided some suggestive evidence at the department level in this respect). In order to isolate the effects of the policy from other time varying trends, the continuation of the analysis will use interaction terms with single years before and after the policy change. In particular, the identification strategy presented before can be extended to a specification with multiple interactions (Duflo, 2001). This has the advantage of allowing to test the common trend assumption as well as providing more insights on the effect of the policy over the years after its implementation. In particular, I run the following specification:

$$E_{i,s,t} = c + \alpha_t + \delta_s + \beta_t \sum_{t=2010}^{2014}(X_{i,s,t} * d_t) + \gamma_t \sum_{t=2010}^{2014}(C_s * d_t) + \theta_t \sum_{t=2010}^{2014}(P_s * d_t) + \varepsilon_{i,s,t} \quad (2)$$

where d_t is a dummy that indicates the year in which the individual appears in the sample (an year dummy); while all other variables are indexed as before. Individuals interviewed in 2009 (first year in the sample) constitute the control group and that year dummy is omitted. Figure 4 plots the estimates of the interaction term between the dummy of being in a given year and the policy intensity in a particular department – while the left panel of Table 5 reports the results with the three different specifications as presented before (i.e. with no controls in the first column, adding individual controls in the second column and departmental controls in the third column). For formal employment, the coefficients remain around zero until 2011 and increase afterwards, becoming statistically significant from 2013.[40] As hypothesised, the programme had no effect on cohorts of employees not exposed to it; while it had a positive effect on subsequent cohorts. There are no notable differences between the instrumental variables and linear probability estimations

[40] The year of 2011 was a year of only partial implementation and we should expect only a minimal effect of the policy in this year. Indeed, the first call for applications for labour inspectors was issued in April 2011 and it is also reasonable to expect some delays before the first newly hired inspectors started operating (e.g. recruitment process, training).

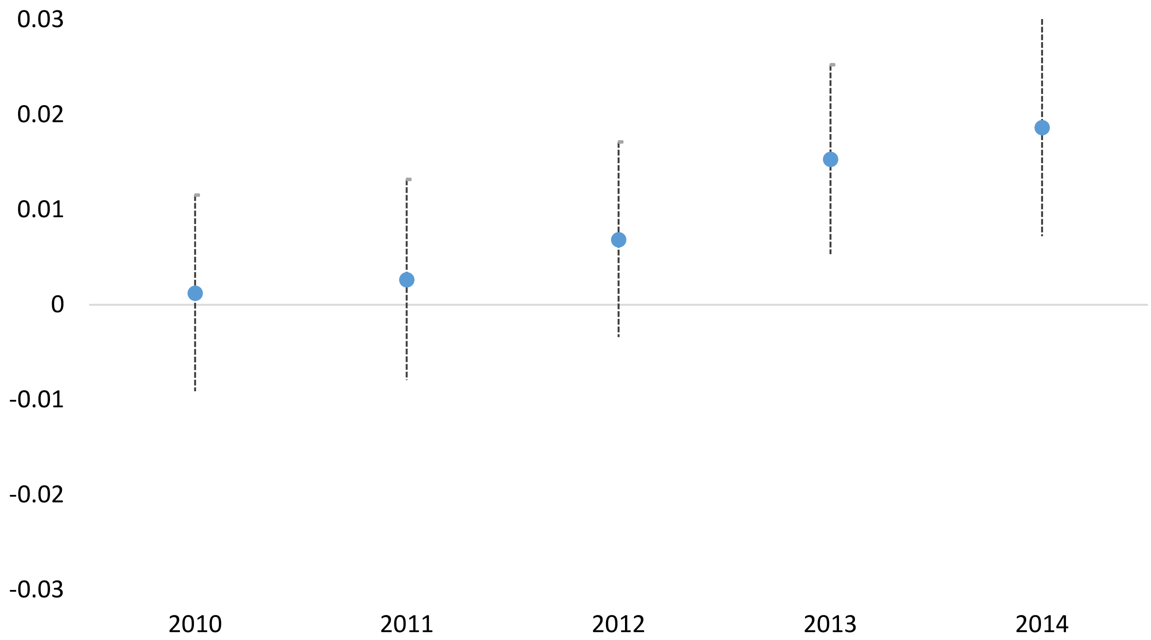
and for ease of expositions only the former are reported in Table 5. In terms of magnitude, the effect estimated with the interaction term analysis is similar to the one presented above. In particular, an additional inspector per 100,000 employed individuals increased the probability of being in a formal job by around 1.5 percentage points in 2013 and 1.8 percentage points in 2014.

Rather than testing whether $\theta_t = 0$ for the years before the implementation of the policy, I can impose this restriction and study the effects of the policy in the years after its approval. This is more efficient and allows to obtain more precise estimates of the impact of the policy (Duflo 2011). In practice, I run the following regression:

$$E_{i,s,t} = c + \alpha_t + \delta_s + \beta_t \sum_{t=2012}^{2014}(X_{i,s,t} * d_t) + \gamma_t \sum_{t=2012}^{2014}(C_s * d_t) + \theta_t \sum_{t=2012}^{2014}(P_s * d_t) + \varepsilon_{i,s,t} \quad (3)$$

where all year dummies up to 2011 are set equal to zero. The results (Table 5, right panel) show that the estimates of the interaction between the intensity of the policy in the department and the year dummies are statistically significant and do not substantially change with respect to the previous specification (i.e. they are slightly smaller in magnitude).

Figure 3 – Coefficients of the interactions between year dummies and the programme intensity in the department (and 90 per cent confidence intervals)



Note: The figure shows the estimated coefficients of interaction terms between year dummies and the intensity of the programme (θ in the equation above) for regressions having as dependent variable formal employment. The blue points correspond to point estimates and the lines are the 90 per cent confidence intervals. The coefficients correspond to the IV estimations with both individual and departmental characteristics (IV Models 3 in Table 5).

Table 5 – Treatment effects by year

	Instrumental variable			Instrumental variable		
	(1)	(2)	(3)	(1)	(2)	(3)
Θ 2010	0.000753	0.00266	0.00121			
	(0.00856)	(0.00485)	(0.00527)			
Θ 2011	0.0114	0.00422	0.00261			
	(0.00865)	(0.00499)	(0.00538)			
Θ 2012	0.0180**	0.00757	0.00685	0.0139**	0.00524	0.00554
	(0.00858)	(0.00466)	(0.00524)	(0.00585)	(0.00328)	(0.00346)
Θ 2013	0.0357***	0.0161***	0.0153***	0.0315***	0.0137***	0.0140***
	(0.00855)	(0.00455)	(0.00508)	(0.00584)	(0.00312)	(0.00324)
Θ 2014	0.0407***	0.0183***	0.0186***	0.0365***	0.0159***	0.0173***
	(0.00951)	(0.00499)	(0.00583)	(0.00718)	(0.00373)	(0.00431)
Test of exogeneity (p)	0.000	0.000	0.000	0.000	0.000	0.000
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes
N	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the year dummies between 2010 and 2014. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

C. Robustness tests

The main issue to verify refers to whether the limited number of clusters (corresponding to 24 departments over 6 years) should represent a concern in the estimation strategy. Ideally, we would like to have a more detailed level of disaggregation that would better capture the level of enforcement to which individuals and enterprises are faced. For instance, Almeida and Carneiro (2012) use information on 5,242 Brazilian cities. However, labour inspection in Colombia is organised along the 32 departments and the GEIH provides information only for the 24 main departments (see section IV for details).[41] For this reason, more disaggregated information is unavailable neither for the outcome of interest nor for the treatment indicator. However, other studies have faced similar issues – generally owing to the traditional scarcity of information on labour law enforcement. In particular, Ronconi (2010) conducts the analysis at the provincial level using information on the 24 Argentinian provinces over successive years. Similarly, Bhorat et al. (2010) have data on enforcement disaggregated only at the level of the nine South African provinces – which they match with individual data from the labour force survey.

In order to deal with this issue, I conduct two separate exercises. First, I re-run the baseline specification presented above but now with bootstrapped standard errors following the procedure proposed by Camero et al. (2008). Indeed, the traditional adjustment of clustering standard errors assumes that the number of clusters tends to infinity. However, with few clusters (as in the present case) the standard errors are downward biased. In these cases, Camero et al. (2008) propose to bootstrap clustered standard errors in order to reduce the bias. Table 4 in the Appendix presents

[41] The GEIH formally collects more detailed information on the place of residence of the individual, but only information on the department is publicly released.

the results of the baseline specification (i.e. Table 4 above) with this adjustment and shows how adjusting the estimates for the presence of few clusters does not change the conclusions of the analysis. As a second exercise, I conduct the analysis at the departmental level – where variations in treatment intensity actually occur. This is also useful to benchmark the results to those obtained by previous studies, which mostly conduct the analysis at the macroeconomic level – the province in the case of Ronconi (2010) and the city in Almeida and Carneiro (2012). For doing that, I construct a database at the department level between 2009 and 2014 and run similar regressions as those presented above (e.g. formal employment of an individual *i* in department *s* in year *t* becomes the rate of formal employment in department *s* and year *t*). In particular, I present three different specifications (only with time and department dummies, including individual controls and with department controls) both with a simple post-policy dummy and in an interaction term setting (corresponding to Tables 5 and 6 above). Even in this case, the treatment indicator is instrumented with its planned change according to the Action Plan.[42] The results (available in Table 6) are extremely encouraging. In particular and despite the limited sample size and the substantially different methodological approach, the results of the post-policy dummy (Panel A) are still statistically significant and with a magnitude similar to the one presented above. Additionally, the interaction term analysis (Panel B, left side) confirms the absence of treatment effects until 2011 and a positive treatment effect afterwards. Similar results are also obtained by setting all year dummies before 2012 equal to zero (Panel B, right side). This provides strong evidence in support of the fact that the analysis presented above is not sensitive to the presence of few clusters.

[42] Only the instrumental variable results are presented, while the linear probability estimates are available upon request.

Table 6 – Macroeconomic analysis

	Panel A: Post-policy dummy					
	(1)	(2)	(3)			
θ	0.0133***	0.0108**	0.0143***			
	(0.00407)	(0.00467)	(0.00454)			
First stage (F)	0.994	0.995	0.995			
Test of exogeneity (p)	163.231	143.549	216.098			
	Panel B: Interaction term analysis					
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Θ 2010	0.00658	0.00515	0.00518			
	(0.00657)	(0.00782)	(0.00782)			
Θ 2011	0.00702	0.00215	0.00559			
	(0.00748)	(0.00867)	(0.00828)			
Θ 2012	0.0178***	0.0140*	0.0166**	0.0133***	0.0116***	0.0131***
	(0.00651)	(0.00768)	(0.00705)	(0.00352)	(0.00409)	(0.00390)
Θ 2013	0.0141*	0.0115	0.0152**	0.00951**	0.00906*	0.0116**
	(0.00718)	(0.00798)	(0.00753)	(0.00464)	(0.00464)	(0.00453)
Θ 2014	0.0269***	0.0251***	0.0280***	0.0223***	0.0227***	0.0244***
	(0.00723)	(0.00855)	(0.00782)	(0.00472)	(0.00556)	(0.00526)
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes
N	144	144	144	144	144	144

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the year dummies between 2010 and 2014. Individual characteristics the share of men and the average age and average years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

VII. A theory of labour inspectors

The key result that emerges from the analysis presented above is that labour inspectors are effective in promoting compliance with the labour legislation (in terms of formal employment) and this result is both economically significant (around 2 percentage points for any additional inspector per 100,000 employed individuals) and in line with previous studies. In this section, I try to understand the mechanisms through which this effect operates and the groups of employed individuals that are more likely to be affected by the policy intervention. The objective is to acquire a better understanding of how labour inspection operates in Colombia and in which circumstances it can be more (or less) effective.

A. General equilibrium effects

The discussion in Section IV has shown how increased enforcement of labour legislation has promoted formal employment. At the same time, it is important to analyse whether increasing labour law enforcement generates any general equilibrium effects on labour market outcomes. For instance, it could be hypothesised that better employment conditions might come at the expense of fewer employment opportunities. This trade-off between employment quality and quantity has been largely studied for both advanced and developing economies and it has also been explicitly taken into consideration in the area of labour inspection (Almeida and Carneiro 2012; Almeida and Poole 2017). In order to shed light on the possible presence of general equilibrium effects, I run the analysis discussed before but now using as outcome variables the binary variables for status in the labour market (i.e. employment, unemployment and inactivity). For ease of exposition, here and in the continuation of the analysis I present only the instrumental variable results (with the three specifications adding one set of controls after the other, as explained above) and only the results of the interaction between the treatment indicator and the post-policy dummy – corresponding to Table 4 in the main analysis.[43] The results do not report any evidence of an effect of labour inspectors on employment, unemployment or inactivity (Table 7). The impact estimates are not only statistically non-significant despite the large sample size (which now includes the entire working age population, compared to the employed subsample included before), but they are also very low in magnitude – suggesting the absence of any general equilibrium effects.

Table 7 – Treatment effects on employment indicators

	Employment			**Unemployment**			**Inactivity**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
θ	-0.0001	0.0002	0.0003	0.0045	0.003	0.0027	-0.0044	-0.0032	-0.0029
	(0.0003)	(0.0004)	(0.0004)	(0.003)	(0.003)	(0.0029)	(0.0029)	(0.0027)	(0.0030)
First stage (F)	207.335	206.214	213.159	207.335	206.214	213.159	207.335	206.214	213.159
Test of exogeneity (p)	0.4472	0.8363	0.6619	0.1653	0.3177	0.3796	0.1773	0.3268	0.4087
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes	No	No	Yes
N	3,550,881	3,550,881	3,550,881	3,550,881	3,550,881	3,550,881	3,550,881	3,550,881	3,550,881

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the dummy taking the value of 1 after the policy change. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

As a second step, I want to understand whether the policy had any impact on the composition of the labour force in terms of status in employment (i.e. employee, self-employed). Indeed, it could be the case that employers reacted to the increased cost of dependent employment (materialised through higher risk of being caught in case of non-compliance) by changing the nature of the employment relationship (e.g. from dependent employment to self-employment) in order to lower labour costs (Almeida and Carneiro 2012). A similar mechanism has been observed in advanced economies as a result of increased employment protection on open-ended contracts, resulting into a more frequent use of fixed-term employment relations. To this end, I replicate the analysis discussed above but using as outcomes of interest binary indicators for the employment status.[44] The results can be consulted in Table 8 and indicate that the increase in labour law enforcement has marginally increased the probability of being in dependent employment and (in parallel)

[43] Of course, even in this case the results are valid only if the parallel trend assumption is satisfied. For this reason, Table 5 in the Appendix presents the interaction term analysis for the instrumental variable results of the full specification (corresponding to column 3 in Table 5) for the different outcomes of interest that will be introduced in this part of the analysis.

[44] According to the ILO statistical classification, this corresponds to the categories of (i) employees, (ii) employers, (iii) own-account workers, and (iv) contributing family workers. Categories from (ii) to (iv) can be grouped into self-employment. For ease of exposition, only the two main categories of employees and self-employed are used here.

decreased the probability of being in self-employment.[45] However, the impact estimates are lower in magnitude compared to the overall effect obtained for formal employment. Overall, this leads to the possible interpretation that the increase in formal employment documented above has partially resulted from a shift in the composition of the labour force from self- to dependent employment – where key aspects of the employment relationship (e.g. contribution to social protection, presence of an employment contract) are more likely to be respected.[46]

Table 8 – Treatment effects on employment by status

	Employee			Self-employed		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
θ	0.0140**	0.0148**	0.0156**	-0.0137*	-0.0145*	-0.0152**
	(0.00703)	(0.00743)	(0.00735)	(0.00704)	(0.00746)	(0.00739)
First stage (F)	209.954	208.709	214.576	209.954	208.709	214.576
Test of exogeneity (p)	0.8406	0.7013	0.9094	0.8769	0.6689	0.8743
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes
N	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the dummy taking the value of 1 after the policy change. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

B. What do labour inspectors target?

In this sub-section, I try to analyse whether the positive treatment effect is homogenous across the economy – or if rather concerns only specific sectors. To this end, I conduct two simple exercises. First, I split the sample between employed individuals in the formal and informal economy.[47] Indeed, it has been argued that labour inspectors tend to target formal firms that are easier to find and reach (Almeida and Carneiro 2012). This might generate concerns that they miss the most severe forms of non-compliance with the labour legislation, which might occur in the informal economy (e.g. agricultural businesses are overly represented in the informal economy). Similarly, I divide the sample between individuals resident in urban as compared to rural areas. Even in this case, previous studies have shown how travel distance represents one of the main constraints to the activities of labour inspectors (Almeida and Carneiro 2009 and 2012, Viollaz 2016a and 2016b), which might generate concerns over their ability to reach scarcely populated areas.[48] The results of this exercise can be seen in Table 9 below (Panel A for the formal versus informal economy and Panel B for rural versus urban areas), which confirms many of the

[45] Disaggregating the results along the categories of self-employment, the effect arises from a decrease in the probability of being contributing family workers.

[46] It should also be noted that contributing family workers are by definition always considered as informal workers.

[47] The informal economy is defined by the ILO recommendation 204 of 2015 as all economic activities that "are – in law or in practice – not covered on insufficiently covered by formal arrangements". In this sense, it refers to the condition of the economic activity rather than the condition of the employment relationship – to which the definition of formal employment refers. For this reason, an individual can be in (in)formal employment either in the formal or the informal economy.

[48] The definition of urban and rural areas is directly taken from the GEIH.

conclusions of previous studies. In particular, the effect of labour inspectors in increasing formal employment is visible only in the formal economy – the effect in the informal economy is negative and even statistically significant in the most parsimonious specification. At the same time, the positive treatment effects materialises mostly in urban areas – while the coefficient is substantially smaller and of only marginally significnat in rural areas.

Table 9 – Treatment effects on different segments of the economy

	Panel A: Formal and informal economy					
	Formal economy			**Informal economy**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
θ	0.0145***	0.0139***	0.0130**	-0.000602**	-0.000392	-0.000374
	(0.00531)	(0.00500)	(0.00524)	(0.000269)	(0.000239)	(0.000241)
First stage (F)	201.715	200.972	190.267	210.423	208.944	226.697
Test of exogeneity (p)	0.000	0.000	0.000	0.6685	0.4604	0.4446
N	808,183	808,183	808,183	1,278,494	1,278,494	1,278,494
	Panel B: Urban and rural areas					
	Urban areas			**Rural areas**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
θ	0.0277***	0.0274***	0.0254***	0.0119	0.0139*	0.0128*
	(0.00554)	(0.00509)	(0.00547)	(0.00745)	(0.00749)	(0.00701)
First stage (F)	208.259	207.121	213.58	186.317	184.446	188.82
Test of exogeneity (p)	0.0022	0.001	0.0005	0.8402	0.389	0.1887
N	1,881,747	1,881,747	1,881,747	204,930	204,930	204,930
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the dummy taking the value of 1 after the policy change. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

C. Who benefits from labour inspection?

As a final step, I want to understand whether the operations of labour inspectors have a differential effect on various groups of employed individuals. This is of particular importance given that the analysis just discussed has revealed how inspectors target (or are effective) mostly in urban areas and in the formal economy. This might translate into a differential treatment impact for different categories of workers that are over/under-represented in those segments of the economy. Additionally, analysing differential effects by individual characteristics might be important if there is any expectation that particular groups should be prioritised in terms of ensuring labour law compliance (e.g. due to generally worse employment conditions). Similarly as before, I split the sample in two dimensions (by gender and educational level) and analyse whether the effect on formal employment is homogeneous across these groups. The results can be consulted in Table 10. They show how treatment effects are similar (in both magnitude and statistical significance) between men and women. This provides reassuring evidence, since a differential gender effect would be difficult to explain from a theoretical standpoint. At the same time, dividing the sample between low- and high-educated individuals (i.e. high-school degree or above) shows how the positive treatment effect is stronger for low-educated individuals (high-school or less). In particular, the effect for high-educated individuals is smaller in magnitude and statistical significant only at the ten per cent in the most complete specification. This can probably be explained by the higher initial levels of informal employment among low-educated individuals.

Table 10 – Treatment effects by societal groups

	Panel A: By gender					
	Men			**Women**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
θ	0.0427***	0.0244***	0.0214***	0.0519***	0.0230***	0.0204***
	(0.00972)	(0.00516)	(0.00531)	(0.0111)	(0.00595)	(0.00594)
First stage (F)	215.52	214.241	223.664	202.101	200.769	200.761
Test of exogeneity (p)	0.021	0.003	0.000	0.0565	0.1203	0.0231
N	1,150,680	1,150,680	1,150,680	935,997	935,997	935,997
	Panel B: By education					
	Low-educated			**High-educated**		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
θ	0.0251***	0.0286**	0.0229*	0.0234***	0.0134*	0.0140*
	(0.00688)	(0.0119)	(0.0119)	(0.00694)	(0.00780)	(0.00794)
First stage (F)	212.321	212.499	218.663	204.167	202.738	201.664
Test of exogeneity (p)	0.3506	0.4392	0.9848	0.0243	0.1633	0.1787
N	537,810	537,810	537,810	640,119	640,119	640,119
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the dummy taking the value of 1 after the policy change. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

VIII. Conclusions

Labour inspection represents one of the main policy instruments available to governments in order to ensure compliance with labour legislation and the respect of rights at work. This is especially the case in developing economies, where mostly informal labour markets are characterised by systematic breaches of the terms of the employment relation (e.g. lack of social protection coverage, non-respect of minimum wage legislation) while other institutional characteristics (e.g. prevalence of small businesses, absence of trade unions) limit workers' capacity to voice their concerns. Despite the widespread reliance on labour inspection around the world, few available studies exist to assess its effectiveness. This can be at least partially connected to econometric identification challenges related to simultaneity between enforcement and compliance with labour legislation in a simple cross-country analysis. Micro-econometric approaches represent a valid alternative in these cases, provided that an exogenous change can be identified.

This paper uses a natural experiment generated by the "Colombian Action Plan Related to Labour Rights". This represents the single most drastic change in labour inspection policy in Colombia, following which the number of inspectors more than doubled in four years. This contrasts with the absence of investment in labour inspection in the previous period. Critical to the identification strategy of this paper, the intensity of the policy differed quite markedly across departments. Combining this geographical variation with time differences in the exposition to the policy for subsequent cohorts of employees, this paper estimates that an additional inspector per 100,000 employed individuals increases formal employment by around 2 percentage points. In order to take into account the possible endogeneity between the implementation of the policy, I instrument the actual treatment (i.e. change in the number of inspectors) with its planned change as obtained from the legislation. I also find that the positive effect on formal employment arises at least partially from a shift in employment status from self- to dependent employment; while not generating any general equilibrium effect on employment levels. However, the positive treatment effect is confined to the formal economy and urban areas.

References

Abadie, A; Athey, S.; Imbens, G.W.; and Wooldridge, J. (2017) When Should You Adjust Standard Errors for Clustering?, Arvix Working Paper

Acemoglu, D.; Autor, D.H.; Lyle, D. (2004) "Women, War and Wages: The Effect of Female Labour Supply on the Wage Structure at Midcentury", *Journal of Political Economy*, 112(3), 497–551

Ahsan, A. and Pages, C. (2007) Are All Labor Regulations Equal? Assessing the Effects of Job Security, Labor Dispute, and Contract Labor Laws in India, World Bank Policy Research Working Paper No. 4259

Almeida, R. and Carneiro, P. (2009) "Enforcement of labor regulation and firm size", *Journal of Comparative Economics*, 2009, vol. 37, issue 1, 28-46

–. (2012) "Enforcement of Labor Regulation and Informality", *American Economic Journal: Applied Economics*, 4(3), 64–89

Almeida, R. and Poole, J. (2017) "Trade and labor reallocation with heterogeneous enforcement of labor regulations", *Journal of Development Economics*, 2017, vol. 126, issue C, 154-166

Amin, M. (2007) Labor Regulation and Employment in India's Retail Stores. Policy Research Working Paper; No. 4314.

Angrist, J.D. (2006) "Instrumental variables methods in experimental criminological research: what, why and how", Journal of Experimental Criminology, 2,23–44

Autor, D.H; Kerr, W.R and Kugler, A.D. (2007) "Does Employment Protection Reduce Productivity? Evidence from U.S. States", in *Economic Journal*, Vol. 117, No. 521, pp. 189-217, 2007

Bertrand, M.; Duflo, E.; and Mullainathan, S. (2004) "How much should we trust differences-in-differences estimates?", *The Quarterly Journal of Economics*, 119(1):249–275.

Besley, T. and Burgess, R. (2004) "Can Labor Regulation Hinder Economic Performance? Evidence from India", *The Quarterly Journal of Economics*, Volume 119, Issue 1, 1 February 2004, Pages 91–134,

Bhorat, H.; Kanbur, R. and Mayet, N. (2012) "Estimating the Causal Effect of Enforcement on Minimum Wage Compliance: The Case of South Africa", *Review of Development Economics*, Volume16, Issue4

Botero, J.C; Djankov, S. La Porta, R.; Lopez-de-Silanes, F. and Shleifer, a. (2004) "The Regulation of Labor", *The Quarterly Journal of Economics*, Volume 119, Issue 4, 1 November 2004, Pages 1339–1382

Cameron, Colin A., Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Standard Errors." Review of Economics and Statistics 90 (3):414–427.

De Giorgi, G.; Paccagnella, M.; Pellizzari, M. (2015) "Gender Complementarities in the Labor Market", *Research in Labor Economics*, 41(8), 277–298

Duflo, E. (2001) "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *The American Economic Review*, 91(4), 795–813

Ihrig, J. and Moe, K. (2004) "Lurking in the shadows: the informal sector and government policy", *Journal of Development Economics*, 2004, vol. 73, issue 2, 541-557

Appendix

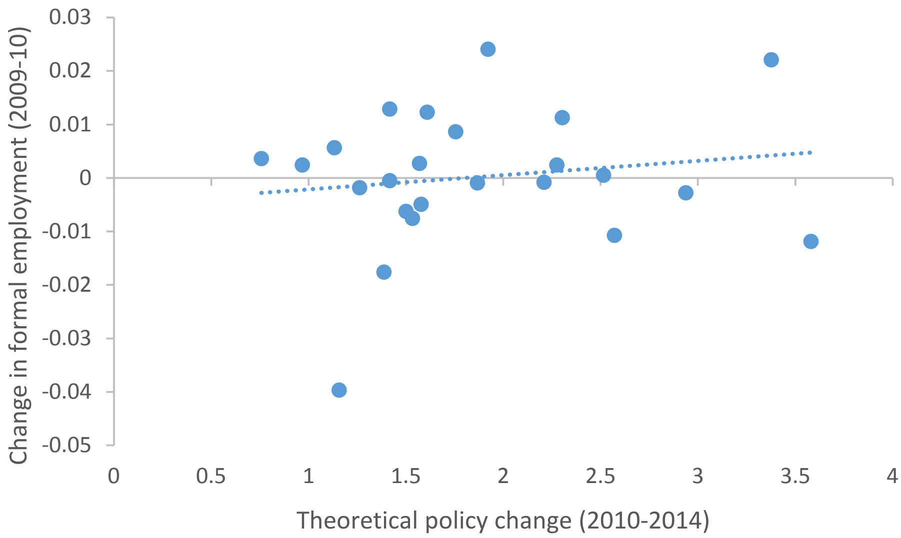
A. Additional figures

Figure 1: Number of workers per labour inspectors

Source: OECD (2016)

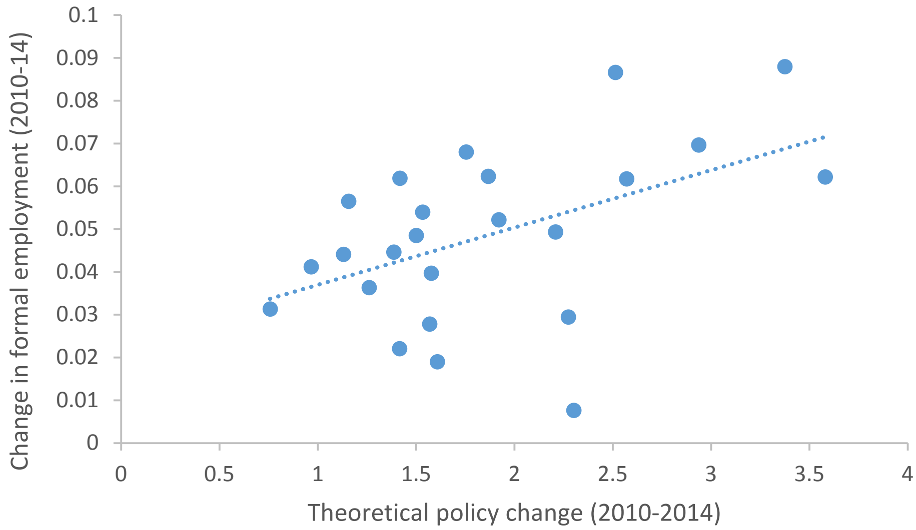
Figure 2 – Formal employment and labour inspection

Panel A: Relation between policy intensity (2010 to 2014) and change in formal employment (2009 to 2010)



Note: The figure present the relationship between the policy intensity in a given department (total planned change in the number of inspectors per 100,000 employed individuals between 2010 and 2014) and the change in the rate of formal employment between 2009 and 2010.

Panel B: Relation between policy intensity (2010 to 2014) and change in formal employment (2010 to 20104



Note: The figure present the relationship between the policy intensity in a given department (total planned change in the number of inspectors per 100,000 employed individuals between 2010 and 2014) and the change in the rate of formal employment between 2010 and 2014.

B. Additional tables

Table 1 – Determinants of the implementation gap

	Gap between theoretical and actual number in inspectors per 100,000 employed individuals in 2014			
	(1)	(3)	(4)	(6)
Institutional environment	0.289			
	(0.341)			
Credibility in the rules		0.637		
		(0.757)		
Credibility in the policies		-0.483		
		(0.836)		
Adequacy of resources and predictability		0.237		
		(0.552)		
Institutional performance			0.378**	
			(0.177)	
Result based management				-0.548
				(1.009)
Accountability				0.708
				(0.934)
Labour welfare				0.0921
				(1.200)
Prevention of irregular practices				0.695**
				(0.261)
Development planning and citizen participation				-0.677
				(0.741)
Constant	-0.926	-1.254	-1.186*	-0.711
	(1.257)	(1.306)	(0.603)	(2.059)
Observations	24	24	24	24
R-squared	0.032	0.056	0.114	0.221

Note: Robust standard errors in parenthesis. The dependent variable is the difference between the theoretical and actual number of inspectors per 100,000 employed individuals in 2014. * p<0.1, ** p<0.05, *** p<0. 01.

Table 2 – First stage relationship

	Actual number of inspectors
Rate formal employment in 2010	-1.731***
	(0.3632)
Level of labour inspection in 2010	-0.1218*
	(0.0665)
Age	0.0003***
	(0.0001)
Male	0.0676***
	(0.0184)
Years of education	0.0047***
	(0.0013)
Target number of inspectors	0.5846***
	(0.0399)
Department dummies	Yes
Year dummies	Yes
R-squared	0.9415
N	2,086,677

Note: Standard errors clustered at the department and year level are in parenthesis. The first stage relationship corresponds to the instrumental variable model presented in table 4 (column 3 of the IV panel). * p<0.1, ** p<0.05, *** p<0. 01.

Table 3 – Instrumental variable analysis with bootstrapped standard errors

	Panel A: Post-policy dummy					
	(1)	**(2)**	**(3)**			
θ	0.0259***	0.0242***	0.0211***			
	(0.00739)	(0.00717)	(0.00638)			
First stage (F)	209.954	208.709	214.576			
Test of exogeneity (p)	0.0018	0.0149	0.0066			
	Panel B: Interaction term analysis					
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Θ 2010	0.000753	0.00266	0.00121			
	(0.0109)	(0.00554)	(0.00788)			
Θ 2011	0.0114	0.00422	0.00261			
	(0.0120)	(0.00524)	(0.00938)			
Θ 2012	0.0180	0.00757	0.00685	0.0139	0.00524	0.00554
	(0.0119)	(0.00556)	(0.00867)	(0.00896)	(0.00460)	(0.00590)
Θ 2013	0.0357***	0.0161***	0.0153*	0.0315***	0.0137***	0.0140***
	(0.0124)	(0.00549)	(0.00855)	(0.00867)	(0.00417)	(0.00535)
Θ 2014	0.0407***	0.0183***	0.0186*	0.0365***	0.0159***	0.0173**
	(0.0136)	(0.00606)	(0.0113)	(0.00953)	(0.00475)	(0.00705)
Department dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual covariates	No	Yes	Yes	No	Yes	Yes
Department covariates	No	No	Yes	No	No	Yes
N	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677	2,086,677

Note: Bootstrapped standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the year dummies between 2010 and 2014. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.

Table 4 – Treatment effects on formal employment

	LPM with adjustment for few clusters (Cameron et al 2008)		
	(1)	(2)	(3)
θ	0.0200***	0.0145***	0.0147***
	(0.00509)	(0.00461)	(0.00433)
R-squared	0.065	0.153	0.153
Department dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Individual covariates	No	Yes	Yes
Department covariates	No	No	Yes
N	2,086,677	2,086,677	2,086,677

Note: Bootstrapped standard errors corrected with the procedure proposed by Camero et al (2008) for few clusters are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the dummy taking the value of 1 after the policy change. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. * p<0.1, ** p<0.05, *** p<0. 01.

Table 5 – Parallel trends for other outcomes of interest

	Employed	Unemployed	Inactive	Employee	Self-employed
Θ 2010	-0.000736	-0.000686	0.00142	-0.00966	0.00947
	(0.00282)	(0.00552)	(0.00648)	(0.0116)	(0.0117)
Θ 2011	-0.00589**	0.00246	0.00343	-0.00315	0.00338
	(0.00232)	(0.00397)	(0.00436)	(0.0107)	(0.0107)
Θ 2012	-0.00579**	0.00429	0.00149	0.00207	-0.00173
	(0.00228)	(0.00387)	(0.00438)	(0.0103)	(0.0104)
Θ 2013	-0.00529**	0.00120	0.00409	0.0180	-0.0179
	(0.00232)	(0.00484)	(0.00549)	(0.0127)	(0.0127)
Θ 2014	-0.00532**	0.00362	0.00170	0.0219**	-0.0217**
	(0.00235)	(0.00489)	(0.00560)	(0.0108)	(0.0109)
Department dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Individual covariates	Yes	Yes	Yes	Yes	Yes
Department covariates	Yes	Yes	Yes	Yes	Yes
N	3,550,811	3,550,811	3,550,811	2,086,677	2,086,677

Note: Standard errors clustered at the department and year level are in parenthesis. The coefficients reported are those of an interaction between the treatment indicator (i.e. change in the number of inspectors per 100,000 employed in a given department between 2010 and 2014) and the year dummies between 2010 and 2014. Individual characteristics include gender, age and years of education. Department characteristics include the (normalised) number of inspectors and the rate of formal employment in 2010. The IV analysis instruments the change in inspectors per thousands of employed individuals with its planned change. * p<0.1, ** p<0.05, *** p<0. 01.