Banco de México

Documentos de Investigación

Banco de México

Working Papers

N° 2020-19

Increasing the Cost of Informal Workers: Evidence from Mexico

Brenda Samaniego de la Parra
University of California, Santa Cruz

León Fernández Bujanda
Banco de México

December 2020

Documento de Investigación
2020-19

Working Paper
2020-19

Increasing the Cost of Informal Workers: Evidence from Mexico*

Brenda Samaniego de la Parra†
University of California, Santa Cruz

León Fernández Bujanda‡
Banco de México

Abstract: This paper estimates the effects of increasing the cost of informal jobs on formal firms' and workers' outcomes. We create novel datasets combining administrative records and household surveys data, and exploit exogenous variation in this cost generated by over 480,000 random work-site inspections in Mexico. Increasing the cost of informal jobs at formal firms leads to lower employment growth, lower formal job creation, and higher formal and informal job destruction. For informal workers, inspections increase the probability of being formalized at the inspected firm, but also increase the probability of dissolving the informal match. Transitioning to a formal job due to an inspection increases the probability of being poached to a new, formal job.

Keywords: Firm Behavior, Informal Economy, Informal Labor Market
JEL Classification: D22, E26, J46

Resumen: Este documento estima los efectos del aumento del costo de los empleos informales sobre el desempeño de las empresas formales y sus trabajadores. Creamos nuevas bases de datos que combinan registros administrativos y datos de encuestas de hogares, y aprovechamos la variación exógena en dichos costos generados por más de 480,000 inspecciones aleatorias a establecimientos en México. El aumento del costo de este tipo de empleo en las empresas formales lleva a un crecimiento más bajo del empleo, una menor creación de puestos de trabajo formales y una mayor destrucción de puestos de trabajo formal e informal. Para los trabajadores informales, las inspecciones aumentan la probabilidad de formalizarse en la empresa inspeccionada, pero también incrementa la probabilidad de disolver el vínculo laboral informal. La transición a un trabajo formal debido a una inspección aumenta la probabilidad de ser contratado por otra empresa en un nuevo empleo formal.

Palabras Clave: Comportamiento de la Empresa, Economía Informal, Mercados Laborales Informales

---

*IMSS data was accessed through the Econlab at Banco de México. Inquiries regarding the terms under which the data can be accessed should be directed to: econlab@banxico.org.mx. This project received financial support from the Research Center for the Americas at the University of California, Santa Cruz and from the Institute for Mexico (RCA UCSC) and the United States (UC MEXUS).

† Economics Department. Email: bsamanie@ucsc.edu.

‡ Dirección General de Investigación Económica. Email: lfernandezb@banxico.org.mx.

1 Introduction

61% of the world's employed population works in the informal economy and 93% of all informality is concentrated in emerging and developing economies (ILO, 2018[30]). Informal jobs coexist with formal jobs, not just within narrowly defined industries and occupations, but even within the same firm. In this paper, we first document new facts about the prevalence and dynamics of informal jobs within formal firms in Mexico. We then examine the effects from increasing the cost of informal employment on firm and worker outcomes using inspections at randomly selected formal establishments as a positive shock to the cost of employing workers informally.

There are two main challenges to estimating the effects of increasing the cost of informality. First, informality is hard to measure and even more so to track informal workers' labor market trajectories and their outcomes. Informally employed workers are, by definition, excluded from administrative records. Second, the estimation requires credible, exogenous variation in the cost of informal employment. Without this variation, it is difficult to disentangle the effects of an increase in formal jobs from the many unobserved factors that endogenously determine formality status, wages, hours worked, and other observed outcomes of interest.

We create a novel dataset compiling administrative employer-employee matched records and a rotating panel household survey to address the challenge of measuring informality. The administrative data allows us to observe the complete formal employment histories for the universe of workers' in the economy. The survey data includes both formal and informal employment spells for a large, representative sample of households in Mexico. These data allow us to track workers' transitions between informal and formal jobs. Further, we can distinguish changes in formality status within the same firm from those that occur when a worker switches employers.

We use over 480,000 random work-site inspections by the Ministry of Labor (STPS) as an exogenous, firm-level shock to the cost of employing workers informally.[1] Two characteristics of these inspections are crucial for our estimation strategy. First, every year, STPS randomly chooses a large, fixed number of establishments to inspect from a directory of formal establishments. Second, STPS inspectors do not have the ability to directly fine or punish employers that have informal workers. Instead, they report any informality detected during

[1]We use the terms "work-site" and "establishment" interchangeably. Similarly, we use "employer" and "firm" interchangeably.

an inspection to the Social Security Institute (IMSS) who then decides whether to conduct a follow-up visit which can result in a large fine. A random inspection by STPS therefore can increase the probability of a future enforcement action by IMSS which in turn raises the expected cost of hiring or continuing to employ informal workers at those firms that were randomly selected for a STPS inspection.

We first estimate the effect of increasing the cost of informal employment on informal workers' flows and wages. We find that informally employed workers at inspected establishments experience a 7 percentage point increase in the quarterly probability of being formalized by the same employer, relative to similar workers at non-inspected establishments. We find no evidence that this higher formalization probability is accompanied by a change in wages up to three quarters after the inspection. Consistent with inspections increasing the cost of informal employment, we also find that the quarterly probability of separating from the inspected employer to unemployment increases from 2.9% to 4.1% on the quarter of inspection.

When offering informal jobs, employers avoid payroll taxes, severance payments, and other costs, but face a positive probability of getting caught. We argue that asymmetrical information about workers' productivity between current and prospective employers also contributes to firms' decisions to hire, and keep, workers in informal jobs. Relative to their co-workers that were formalized without an inspection ("organically formalized workers"), employees that start a formal job immediately after an inspection have shorter formal tenures with the inspected employer. Conditional on separating, the probability of having a formal job at the next employer is higher than that of informal workers at non-inspected establishments and similar to that of "organically formalized" co-workers. Moreover, while wages at the inspected employer are 1.4% higher for organically formalized workers, this wage gap disappears after being poached by a new, formal firm. These findings are consistent with employers using formality status as a signal of worker productivity.

We next focus on formal firms' outcomes using administrative data. We find that increasing the expected cost of informal employment (through random inspections) has a negative and significant effect on total formal employment. The quarterly formal employment growth rate permanently decreases by 2 percentage points. Although the worker-level results from survey data indicate that within-firm informal to formal jobs transitions increase, this rise is dominated by a persistent decline in the formal worker poaching rate (2 percent decline) and a 3.7 percent increase in the formal job destruction rate. We also find that small firms are most likely to respond to higher informality costs by increasing formal job creation.[2]

We contribute to the literature on the effects of policies aimed at reducing informality. Most of the literature analyzes firms' decision to register with the government, rather than the formality status of the jobs within a firm. A common approach is to develop a model where firms balance the benefits of participating in the formal sector[3] against the lower regulatory costs and the probability of punishment in the informal sector. Ulyssea (2018) and Bobba et al. (2020) extend this literature by introducing informal jobs within formal firms. Their estimation approach is similar to the prior literature, however, in that their measures of the effects of increasing the cost of informality are based on policy counterfactuals based on the calibrated model. We contribute to this literature by directly estimating firms' responses to firm-level, exogenous variation in the cost of informal employment. We focus on informality within formal firms. Furthermore, we also estimate the effects on workers' outcomes.

Our paper is closely related to the empirical strand of the literature that relies on identifying cross-sectional variation in the cost or benefits of informality. Using survey data coupled with a policy that simplified the tax regime in Brazil, Fajnzylber, Maloney and Montes-Rojas (2011) finds that formalization significantly increases firm size, profits, and revenue. Exploiting cross-industry variation in most-favorable nation tariffs in Peru, Cisneros-Acevedo (2020) argues that trade liberalization increases the incentives to hire workers informally. Almeida and Poole (2017) and Ponczek and Ulyssea (2018) examine how variation in the enforcement of labor regulation across cities in Brazil, measured by the number of inspectors per firm, mediates the employment effect of a currency and a trade shock, respectively. Other papers exploiting geographical variation in labor inspections to analyze the effect of enforcement include McKenzie and Sakho (2010) in Bolivia, Ronconi (2010) in Argentina and Almeida and Carneiro (2005), (2012) in Brazil.

We differ from these papers in three important ways. First, our identification strategy relies on a large set of inspections at randomly selected establishments by design. STPS inspections do not result in the formalization or termination of informal matches. Instead, they increase the cost of maintaining informal jobs and creating new informal matches by

[2] This includes both within-firm informal to formal transitions, across firm informal to formal transitions and transitions to formality from non-employment.

[3] The specific benefits of formality highlighted in the model vary across papers and include, for example, lower borrowing costs (D'Erasmo and Moscoso Boedo (2012)), lower search frictions (Zenou (2008)), higher productivity (Albrecht, Navarro and Vroman (2009)) or productivity dynamics, Bosch and Esteban-Pretel (2012)). Leal Ordóñez (2014) and Meghir, Narita and Robin (2015) instead assume that formal and informal firms have access to the same technology and markets, and the decision instead hinges on the costs of formality (taxes) versus those of informality (probability of getting caught).

raising the probability of a fine in the future. Second, we combine a rotating panel survey and administrative employer-employee matched data with a very large set of random work-site inspections over a long period of time. The combined data allow us to implement a difference-in-difference estimation strategy at the firm and worker levels, and to analyze a wide range of response margins. Third, our estimates do not rely on structural assumptions about the trade-offs between formal and informal employment, but we nonetheless propose a theoretical framework to interpret our empirical results and improve our understanding of this trade-off.

We provide new facts on the prevalence of informal employment within formal firms, how it varies by firm size, and its dynamics.[4] We find that most informal to formal job transitions occur within the same firm while the majority of formal to informal flows are due to separations from the current employer. The probability of transitioning to a formal job at the same employer is highest in the first quarter of employment and decreases with on-the-job tenure and worker's age. Distinguishing between within and across firm changes in formality status is a relevant addition to the literature using microdata to analyze informal labor market dynamics such as Levy (2008), Bosch and Maloney (2010), Bosch and Esteban-Pretel (2012), and Gallardo del Angel (2013). The distinction is relevant for the literature using labor flows to assess the degree of labor market segmentation including Maloney (1999), Pratap and Quintin (2006), Ulyssea (2010), among others.

Our results also contribute to the theoretical literature. A standard modeling assumption is that the cost of informality is increasing in firm size (either in terms of productivity, capital, or number of employees). This assumption is based on the fact that larger firms are more likely to be formal (i.e. registered with the government), and rationalized by arguing that the probability of being detected increases with firm size.[5] Ulyssea (2018) and Bobba et al. (2020) expanded this literature by explicitly modeling firms' decision to formalize some of their workers (i.e. informality within formal firms). In so doing, they kept the previous literature's assumption that the cost of informality increases with firm size. However, in the case of informality within formal firms, detection need not depend on the firm's visibility. A firm with many employees may be better able to hide a given number of informal workers than a smaller firm for whom even a single employee missing would be easily detectable.

[4]In a concurrent paper, Bobba et al. (2020) also decompose within and across firms informal-to-formal job transitions. They identify these transitions using worker self-reported tenure data. Tenure data is not collected in ENOE on every quarter and can be subject to bunching so we instead use direct information on whether the worker is with the same employer as in the previous quarter.

[5]Rauch (1991), Fortin et al. (1997), De Paula and Scheinkman (2010), Leal Ordóñez (2014).

Alternatively, very small firms may be able to easily relocate to avoid fines even if detected. Hence, the cost of the intensive margin of informality may be increasing or decreasing in firm size. We find that the average share of informal workers is strictly decreasing in firm size ranging from 73 percent at establishments with 5 workers or less down to 8 percent at worksites with over 100 employees.[6] We also find that firms are more likely to exit the formal sector when the cost of hiring informal workers at formal firms increases. The magnitude of this effect is decreasing in firm size.

The paper is organized as follows. In section 2, we provide a description of the requirement to register workers with IMSS and the accompanying payroll taxes. We then describe IMSS's and STPS's enforcement tools. Section 3 describes the data. In section 4 we present our empirical findings on the effects of increasing the cost of informal employment on workers and firms using survey and administrative data. Section 5 develops a model with search frictions and asymmetrical information about workers' productivity to discuss our empirical findings. Section 6 concludes.

2 Formal Jobs, Payroll Taxes, and Monitoring in Mexico

2.1 Registration requirements

In Mexico, employers must register all wage-earning employees with the Mexican Social Security Institute (IMSS) within 5 business days of hiring.[14] Registration with IMSS gives workers access to a set of benefits including public health care and day care services, maternity leave, sick leave, disability insurance, and a retirement fund, among others. IMSS provides these services and collects payroll taxes to finance them. Employers must deduct these taxes from each enrolled worker's paycheck.

Federal Labor Law levies the majority of the payroll tax on the employer.[13] Employers' contribution has a fixed and a variable component. [7] Due to the fixed fee, employers' con-

[6]This is consistent with Perry et al. (2007) findings that the share of informal workers is highest at firms with 2-5 workers, and rapidly declines with firm size.

[7]The fixed cost is equal to 20.4% of the daily minimum wage (MW), times the number of days the employee worked in the period. Employers' contributions have an upper bound at a daily wage equivalent to 25 minimum wages (USD$101.45). On January 2016, a constitutional amendment instructed that all references to the minimum wage in all federal and state laws and regulations should be understood as instead referring to the newly created Measurement and Revaluation Unit (*UMA*). When it was first introduced, 1 *UMA* was equal to 1 minimum wage, therefore, the amendment has no effect during our sample period.

tributions as a percent of the worker's after tax wage ranges from 17% for an employee that earns 25 minimum wages to 35% for a minimum wage earner. Firms must also face other costs for their formal workers including minimum wages, severance and overtime pay, paid leave and minimum vacation days, workers' right to training, safety and health conditions in the workplace, among others. The Federal Labor Law also requires businesses to share a portion of their after-tax profits with their formal workers.[13]

2.2 Monitoring and the Cost of Informal Employment

According to its inspection bylaws, STPS must inspect establishments periodically by random selection from the National Firm Directory (DNE).[49] Between January 2005 and June 2016, STPS performed 487,118 inspections at establishments across Mexico. STPS's self-proclaimed objective is to incentivize compliance, not mainly through fines but by helping employers understand and abide by the law. After each visit, STPS inspectors file a report and give a copy to the employer. The report lists all detected violations to labor regulation, including whether there are any informally employed workers.[8] STPS then schedules a follow-up visit[9] and fines the establishment only if any of the originally detected violations are still occurring.[13,48] Crucially, STPS inspectors check for informal workers but do not have the authority to punish firms for this specific violation. Instead, when they detect informal workers, STPS inspectors notify IMSS.

IMSS also carries out inspections regularly. Unlike STPS inspections which cover a range of labor regulations [50] with a corrective approach, IMSS inspections focus on informality and highlight a punitive approach. Their goals are to verify that employers are registering and paying the mandated social security fees for each of their employees, to collect any missing payments and fine employers for noncompliance.[25] When determining which establishments to visit, IMSS uses a "risk-based" model that prioritizes cases where there is evidence of irregularities. To develop such evidence, IMSS takes into account firm size, industry, history of noncompliance, and notifications made to IMSS by STPS and other fiscal authorities.[26]

STPS inspections increase the cost of informal employment by raising the probability of a visit, and fine, from IMSS. The cost of a visit from IMSS is high. Data from a special

[8]STPS inspectors make note of how many workers are employed at the establishment and whether the employer provided proof of payroll tax payments to IMSS.

[9]In the case of dangerous or extreme cases (such as improper management of hazardous waste or use of child labor), employers are fined and must comply with regulation immediately or can be closed down.

establishment survey indicates that a manager at a large establishment spends on average 10 hours preparing for a visit from IMSS, not including the time the inspection itself takes. According to 25% of employers on the survey, the main reason firms incur in bribes is to avoid the inspection itself, not the corresponding fine.[27] Further, if IMSS does confirm the presence of informality, employers must pay back-due payroll taxes and fines ranging from 20 to 250 daily minimum wages per worker. Employers can also be charged with providing false information which carries an extra fine of similar magnitude per worker. Fraud against IMSS is punishable with up to nine years in jail.[14]

Employers have incentives to formalize (or terminate) their workers promptly after receiving a visit by STPS. If an employer admits to having informal workers before being visited by IMSS, fines are partially, and sometimes entirely, waived if promise to formalize. Moreover, employers who come forward with IMSS about their informal employment can request extensions and installment payment plans for their back-due payroll taxes.[10]

3 Data Description

This paper combines 3 datasets in a novel way: the quarterly National Employment and Occupation Surveys (ENOE), the Mexican Social Security Institute (IMSS) administrative employer-employee matched administrative-records, and the Ministry of Labor's Directory of Firms and Inspections Logs.[11] We create three baseline samples combining these various data sources. We highlight their contribution below and provide additional information on each of the underlying data sources in the following subsections.

- All formal firms that have at least one establishment registered in the DNE. This dataset is constructed by merging the DNE and its inspection logs with IMSS administrative data on the universe of formal employers. This allows us to track formal firms' life-cycles.[12]

[10]Garcia, Kaplan and Sadka (2012) show that when workers include IMSS as a co-defendant in a labor dispute, firms are more likely to settle. The authors argue that notifying IMSS about the labor dispute raises the stakes for firms, since IMSS is the only entity with the authority to fine firms for not registering their workers. Our hypothesis about STPS inspections increasing the likelihood that firms register or fire workers, before a potential visit from IMSS, is consistent with their findings.

[11]This is, to our knowledge, the first paper to exploit these three sources of data. See Kumler, Verhoogen and Frías (2020) for an example comparing wage distributions from IMSS and ENOE to investigate the extent of tax-evasion via wage under-reporting.

[12]The sample begins in 2005. We do not observe the creation of firms before this date, but can track them

- All the formal employment spells for workers who are ever formally employed by a firm with at least one establishment registered in the DNE. This dataset allows us to track the formal employment trajectories for the universe of employees by merging employee-level data from IMSS with the DNE.
- All individuals in ENOE's survey who during any of the survey's waves are employed at a firm with at least one establishment registered in the DNE. This dataset is a random sample of both informal and formal employees at each of the firms with establishments registered in the DNE. ENOE's panel structure allows us to directly observe each of these individuals' labor market transitions, including those across formal and informal jobs within the same firm.

The DNE is the connecting thread across all samples since only those establishments registered in the DNE have non-zero probability of being randomly inspected. The main upsides of focusing on the first two samples are that they cover the universe of formal workers and firms, and we can follow them for as long as they remain formal. The downside is that since their source is administrative data they lack information about all informally employed workers. The third sample fills this gap. However, unlike the samples based on IMSS data where we could observe the universe of formal workers and their trajectories, this sample is constrained to ENOE's rotating panel structure so we can at most follow each individual for 5 quarters and only observe a non-representative sample of each firms' workforce. Appendix B provides details on the matching process across all datasets.

3.1 The Ministry of Labor and Social Welfare's National Firm Directory (DNE) and Inspection Logs

The National Firm Directory (DNE) is a list of formal firms' establishments compiled by the Ministry of Labor and Social Welfare (STPS). It includes information on firm's name, the date on which the firm was registered in the DNE, and its establishments' addresses. Firms' unique tax identifier (RFC) and IMSS employer ID (*registro patronal*) are optional fields in the DNE. STPS uses the DNE to randomly select establishments for inspection.

Between January 2005 and June 2016, there were 248,548 firms with 423,129 establish-

after if they continue to exist. We end our sample in 2016 to match the final period of the DNE's sample, so we can also not observe firms' exits beyond this year.

ments in the DNE.[13] Firms do not have a legal obligation[14] to register in the DNE, hence, it is not an exhaustive list of all formal work-sites operating in Mexico. By comparison, there were 833,716 formal employers in the country on average during the same time period.[15] However, firms are required by law to offer training to their employees. Training programs must be registered with STPS and firms who need help setting them up can receive direct assistance from STPS.[16] STPS then uses this training registry[17] as the main source of information for the DNE. Another channel through which firms enter the DNE is through STPS non-random inspections programs or when a complaint is filed against them. In recent years, STPS has made additional efforts to exchange information with other government entities and private sector institutions that have data on establishments operating in Mexico.

Every year, STPS establishes a goal for the number of inspections. It then selects which work-sites to visit via random draws from the DNE. We first test for evidence of "haphazard" instead of random sampling. To do so, we compare the distribution of the universe of firms in the DNE to the group that was selected to be inspected each year along various variables that inspectors could have potentially used to sort the DNE. As highlighted by Hall et al. (2012), if auditors attempt to minimize the effort in selecting audit subjects from an ordered list, the selected sample may not be evenly or independently spread throughout the population of eligible firms.

We consider two possible orderings for DNE establishments to test for haphazard selection: alphabetical and postal code. Under a null hypothesis of random selection, the order in the alphabetical list or location is uncorrelated with the probability of being inspected. Instead, each category's (defined as the first letter of the establishment's name or the first two digits of its location's zip code) representation within the set of inspected establishments is similar to its share in the universe of active establishments in the DNE. We do this test separately for each year.[18] Most categories are represented in the inspected sample according to what we would expect under random selection. For each year between 2006 and 2016, we fail to reject the null of randomness at a 10% significance level for at least 74% of categories

[13] Firms are identified as unique firm names (*razon social*) or tax ID. Establishments are unique combinations of addresses and firm names/tax IDs.

[14] Article 132 frac. XV and article 153, Federal Labor Law[13].

[15] IMSS identifies employers using a uniquely assigned ID called *registro patronal*. A firm can apply for more than one *registro patronal*.

[16] STPS programs offer firms training that helps them comply with regulation and bid for government contracts. STPS also provides registered firms free courses and worker training.

[17] The system that tracks these registered training programs is known as SICAPE or *Sistema de Capacitacion de las Empresas*.

[18] We exclude 2005 and 2016 to have complete annual data for all establishments.

and up to 96%. We do not find evidence of consistent selection of establishments across time based on establishment name or location. Tables C.2 and C.3 present the haphazard test for 2015.[19,20]

In the next sections, we examine whether the distribution of inspected establishments across industries, size (measured by number of formal employees) and firm age is consistent with random selection across these characteristics.[21]

3.2 The National Employment and Occupation Survey (ENOE)

ENOE's counterpart in the United States is the Current Population Survey (CPS). Like the CPS, ENOE is the main source of labor market statistics in Mexico and, since INEGI allows public access to its micro-data, it has been extensively used in research papers. However, the information that allows identifying firms, including the firm's name, is removed from the publicly available data. This paper uses ENOE's employer-side data, which has been so far unexploited due to its confidentiality. Using firms' names and locations, we match ENOE to STPS's Firm Directory and Inspections Logs.

The dwelling is the sampling unit in ENOE. It gathers information regarding households' composition and dwelling characteristics, as well as extensive data on each household member such as age, education, gender, labor market participation, and job characteristics. Each quarterly sample is representative of the national labor market and includes 120,260 households and 420,000 individuals on average. Like the CPS, ENOE is a rotating panel: households are interviewed for five consecutive quarters and then replaced.[22]

We focus on wage-earning employees (as opposed as self-employed workers or independent contractors) since these are the workers that employers must register with IMSS. We

[19]We also consider geographical coordinates as an ordering category for haphazard selection. These results, as well as those from other years, are available upon request from the authors.

[20]Randomly selected establishments can be assigned non-random dates for inspection within the year. For example, inspectors may which to minimize travel costs across the randomly selected establishments. As detailed in the appendix, we find evidence consistent with such geographical clustering on inspection dates after the number of inspections more than doubled in 2012 (see tables C.4 and C.5). This does not pose a threat to our identification strategy. If establishments are aware of the clustering, both treated and control groups would be expected to have similar anticipatory behavior. This would, if anything, bias our results against finding an effect from inspections.

[21]Since the DNE and inspection logs lacks information on firms' characteristics, we first need to merge the data with IMSS or ENOE to perform these tests.

[22]The attrition rate is 3% for the first quarter in the sample. 85% of households stay in the sample for a full year.

drop individuals employed in the agriculture sector or as domestic workers during any of ENOE's waves. We exclude individuals younger than 15 and those older than 80 on the first survey wave. We also exclude individuals who did not fully complete the survey. We further restrict the sample to individuals employed at firms included in the National Directory of Firms (DNE) for at least one of the quarters when they participate in ENOE's survey. Our baseline sample includes only individuals who, for at least one of the quarters of observation in ENOE, are employed (informally or formally) at formal firms included in the DNE which can therefore be randomly selected for a STPS inspection. On average, we match 11,000 employees per quarter. Table B.1 presents additional information on the matching rate between workers in ENOE and their employers in the DNE.

Table B.3 shows the distribution of employees in ENOE by industry (as reported by the workers). Column (1) presents workers' distribution for all individuals employed at formal firms (as defined by INEGI). Column (2) is the distribution for the subset of these workers that we can match to an employer in the DNE. Column (3) shows the industry distribution of workers at firms that received at least one inspection. Despite the differences in the industry distribution of all employees at formal firms in ENOE and the subset of matched workers (column (1) vs. column (2)), the various industries are proportionally represented within the set of inspected establishments (column (2) vs. column (3)). Consistent with random allocation of inspections, this indicates that industry is not correlated with the probability of receiving an inspection within the sample of matched ENOE-DNE establishments.

In table C.6 we further examine whether the probability of receiving an inspection is correlated with workers' or establishments' characteristics. Let $Z_{i,t}$ be an indicator variable equal to 1 if individual i was employed at an establishment that received an inspection in quarter t and zero otherwise. Column (1) in table C.6 tests this using a linear probability model in ENOE data. Jointly, worker's and establishment's characteristics explain less than 0.1 percent of the variation in inspection probability across establishments within the DNE (joint p-value of 0.595). Columns (2) and (3) show that these covariates do affect workers transition probabilities out of informal employment. In particular, more educated workers and larger employers are more likely to transition to a formal job or separate to unemployment. Thus, these characteristics are relevant for workers' outcomes, but uncorrelated with the probability of an inspection.

3.3 IMSS Administrative Records

For all formal workers, firms report to IMSS the day the formal employment relation started and the prevailing wage. Employers must notify IMSS whenever the wage changes and when the match is terminated. IMSS data is therefore a complete record of the wage and tenure history for the universe of formal employment relationships. For each employer, we observe an IMSS employer-specific identifier[23] and their tax ID[24], industry, and location.

Each observation in IMSS data is an employer-worker-wage match. While ENOE only tracks workers for at most 5 quarters, using IMSS data we can track all wage changes within a formal match and observe all future formal matches with any other firm, too. This makes IMSS data suitable for longer term analyses. The downside is that while in ENOE we can directly observe within-firm informal-to-formal transitions, in IMSS's records, we only start tracking a worker after they are formalized. When a new formal match begins, we can directly observe whether the worker previously had a formal job or not. But for those workers who did not have a formal job before the new match, we cannot distinguish whether they were unemployed or informally employed (whether at the same firm or elsewhere) beforehand.

For our empirical analyses, we construct two datasets from IMSS administrative records each ranging from January 2005 to April 2019: an employer-level dataset and a worker-level one. The employer data is a quarterly panel that tracks employers' size (measured as the number of formal workers), formal wage distribution, formal job creation and formal job destruction. For job creation, we distinguish between jobs filled by workers previously employed with a different formal firm ("poaching") and jobs filled by workers that were not formally employed in the previous quarter ("new formalizations"). These "new formalizations" include transitions to formality from three different initial states: an informal job at the same firm, an informal job at a different firm, and non-employment. We make an analogous distinction between job destruction where workers separate to another formal job versus all other separations. The latter category includes separations to an informal job (either within the same firm or with a new employer), to unemployment or leaving the labor force.

Table B.4 presents the matching rate between employers in IMSS administrative records and the DNE, along with descriptive statistics for the matched sample. During an average year, there were 1 million different active employers registered with IMSS. We focus on the

[23] IMSS assigns an identifier, known as *Registro Patronal*, to each employer. As we discuss in more detail in Appendix B, each firm (or Tax ID) can have more than one employer ID (*Registro Patronal).*

[24] *Registro Federal de Contribuyentes* or RFC.

set of employers that we can match to a firm in the DNE[25] and all of their employees. The matched DNE-IMSS employer data includes 24,251 distinct employers during an average year, 14,386 of which were randomly selected for an inspection by STPS.[26]

Our employee-level dataset tracks the formal labor market trajectories of all workers who were ever formally employed at any of the firms that we can match to the DNE in IMSS records. In this dataset, we record the worker's entry and exit date from the DNE firm, as well as her starting and final wages at that firm. We also identify the previous and next firm of formal employment and the wages at these prior and future jobs. We will use these data to analyze the persistence of transitions into formal employment around the dates of random inspections, and the potential effects into future formal jobs.

4 Empirical Findings

4.1 Informality within Formal Firms

A well established fact in the labor informality literature is that firm size is negatively correlated with the firms' formality status. Smaller firms are less likely to be registered or pay taxes. Using the ENOE-DNE matched sample, we can provide new stylized facts about the prevalence and dynamics of informal employment within formal firms. Our measure of size is the mode of the number of individuals employed at the establishment as reported by all ENOE respondents employed at the work-site for at least one of ENOE's waves. Table 2 shows that the share of informal workers at formal establishments is strictly decreasing in size. On average, 73% of all workers at establishments with 2 to 5 employees are informal. This share drops to 45% for establishments with 6 to 10 workers and 13.6% for those with 51 to 100 employees. Services (excluding social, financial, professional business services) have the highest share of informal employment within formal firms at 56.7% followed by restaurants and lodging with 4 informal employees out of every 9 workers.

[25] Firms are identified using firm name and tax ID's. A firm can have more than one employer ID. Appendix B discusses matching employers to firms, and then firms in the IMSS data to firms in the DNE.

[26] Since IMSS's records include all formal employers we should in theory be able to match all DNE firms' to an employer record. However, IMSS data only included firms' tax ID for recent years so our matching process additionally relies on fuzzy name matching for years prior to 2017. The low matching rate and under-representation of certain sectors in the matched dataset, relative to the universe of firms in IMSS, is in part due to the high turn-over of firms in IMSS records which makes it less likely for firms to exists long enough to register in the DNE.

Table 3 shows the predicted quarterly transition rates in formality status within and across employers matched to the DNE after controlling for workers' age, tenure, education, gender, firm size, occupation and sector fixed effects.[27] On average, half of all informal employees at a formal firm remain employed at the same establishment the following quarter; 73% of this half remain informally employed while the rest are "promoted" to a formal job. Meanwhile, 81.8% of all formal employees stay at the same establishment with the same formality status. The unconditional probability of separating to a different formal firm is similar for formal and informal employees (7.5% vs. 7.7%). However, workers who leave an informal job at a formal firm are 3 times more likely to again be hired as an informal worker by the new employer than workers that separate from a formal job.

Frequent transitions across formality status are not a new finding in the literature. Bosch and Maloney (2010) find high transitions across informal and formal jobs for Argentina, Brazil, and Mexico. Bobba et al. (2020) find that 20 percent of workers transition to a formal job within a year of informal employment in Mexico. We contribute to this literature by distinguishing between transitions that occur without a change in the current employer. We find that most of the transitions from informal to formal jobs occur with the same firm. Further, we find that within-firm formalizations occur shortly after the informal match begins. Figure 1 plots the predicted probability of within-firm formalizations by worker's tenure and age. The probability of transitioning to a formal job at the same employer is highest in the first quarter of employment and decreases with match duration.

4.2 Increasing the Cost of Informality via Inspections

4.2.1 Transitions out of Informality

We begin by estimating the effect of inspections on informal workers' labor market transitions. Between any two consecutive quarters, an informal employee at a formal work-site can either remain informally employed with the same employer or switch to any of six mutually exclusive labor market states: (1) unemployment, (2) formal job with the current employer, (3) formal job with a new formal employer or (4) informal job with a new formal employer, (5) informal job at an informal firm, and (6) out of the labor force.

We model the probability of transitions out of informal employment into each of these six possible labor market states using the multinomial logit specification in equation 1. $\Pi\left(TI_{i,j,t} = x\right)$

[27] Table B.2 in Appendix B presents descriptive statistics for these control variables.

is the probability that worker i, who is informally employed at firm j in period t, transitions to market state x in period $t+1$ where x is each of the 6 possible transitions listed above, conditional on worker and firm characteristics $X_{i,j,t}$. $V_{j,t}$ is a dummy variable indicating whether STPS detected a regulatory violation during the inspection and $Fine_{j,t}$ is the sum of fines received by j so far (including fines imposed on quarter t). $Inspected^{q}_{i,t}$ is an indicator variable equal to 1 if in period t individual i was employed at a firm that received an inspection on period $t-q$, $q \in [-3,3]$ and 0 otherwise. The coefficients of interest, β^{x}_{q}, capture the time-varying effects of inspections.

$$\Pi(TI_{i,j,t} = x) = \sum_{q=-3}^{3} \beta^{x}_{q} Inspected^{q}_{i,t} + X'_{i,j,t}\eta^{x} + \alpha^{x} V_{j,t} + \gamma^{x} Fine_{j,t} + \theta_t + \lambda_j + \varepsilon_{i,j,t} \tag{1}$$

$\forall x \in$ Unemployed, Formal with same employer, Formal and Informal with new formal employer, Informal employer, Out of the labor force.

Panels A to D in Figure 2 below show the dynamic effect of inspections on transitions before and after the inspection occurs (β^{x}_{q}). Each panel plots the average quarterly transition probability for informal workers at inspected work-sites (the treatment group) into a different labor market state before and after the inspection. The red line shows the average transition probability for the control (informal workers at work-sites included in the DNE that did not receive an inspection within a three-quarter rolling window).

Informal employees at inspected and non-inspected work-sites exhibit similar average quarterly transition rates out of informality before an inspection occurs, consistent with random work-site selection. This random assignment allows for consistent reduced form estimation of the effect of inspections on labor market flows.[28]

Consistent with our interpretation of inspections as an increase to the expected cost of maintaining informal employees, on the quarter of inspection, the quarterly probability of remaining informally employed at the same work-site decreases from 35% to 26% (Panel A). This drop is due to a 50 percent increase in the quarterly within work-site formalization rate (from 14% to 21%) on the quarter of inspection (Panel B) and an increase in the average quarterly probability becoming unemployed from 2.9% to 4.1% (Panel C).

Panels D presents the effect of inspections on informal workers' quarterly transition probabilities to a formal job with a new employer. Quarterly transition rates to formal jobs with a new employer decrease (from 8 to 6.5 percent) on the quarter of inspection. A possible explanation for this is that inspected workers are more likely to experience an unemployment spell

[28] We present additional evidence on random allocation of inspections in Appendix C.

between quarterly job transitions that is unobserved to us, but observed by the new employer. Unemployed workers have lower probabilities of receiving formal job offers than informal workers.

4.2.2 Short Term Wage Effects

We estimate the short-term change in informal workers' wages at inspected firms using ENOE's panel. ENOE allows us to see the worker's wage while informally employed and directly measure whether it experiences any changes immediately after the inspection. Inspections could change firms' wage setting policies and hence affect formal and informal wages for all workers even if they were not yet employed at the firm at the time of the inspection. We include in our baseline sample only workers employed at firms that either had not been inspected or received their first inspection during the time period when the worker was interviewed for ENOE.[29] For every quarter, the control group is the subset of workers whose work-sites have not yet been inspected. Workers are in the treatment group on the quarter their work-site is inspected and on every subsequent quarter.

ENOE is a 5-quarter rotating panel, so we observe each worker at most 5 times. We will refer to each of these observations as a wave, denoted as $v \in$[1,2,3,4,5]. A worker's firm can receive an inspection in any of these waves so the number of pre and post inspection observations varies depending on the time when the worker entered ENOE's sample. To account for this, we gather workers into cohorts, c, based on the quarter when they enter ENOE's sample. Let $datefirst_j$ be the date of firm j's first inspection. The treatment path for each of j's workers in cohort c, $\{Inspected^c_{j,v}\}^5_{v=1}$, is a non-decreasing sequence of 0's and 1's where the first 1 in the sequence corresponds to the ENOE wave v for cohort c when j receives an inspection.[30]

Figure 3 shows the dynamic effect of inspections on the after-tax wage for informal workers estimated using the log-linear regression model in equation 2. $\ln\left(w^c_{i,j,v}\right)$ is the natural log of the hourly after-tax wage for worker i, in cohort c on quarter $t = v + c$ employed at firm j, $X_{j,t}$ is a vector of time-varying work-site characteristics. i and t are sets of individual

[29] We do not observe inspections before January 2005. Therefore, to limit the possibility of left-side censoring on first inspection dates, we further restrict the analysis to workers that enter ENOE's sample on or after 2007.

[30] Note that calendar time, t, is equal to the cohort's entry date c plus ENOE's wave, v, $t = c + v$.

and time fixed-effects.[31].

$$\ln\left(w^{c}_{i,j,v}\right) = \sum_{v=1}^{5} \zeta_v Inspected_{j,c} + \gamma X'_{j,t} + \lambda_i + \theta_t + \varepsilon_{i,j,t} \quad (2)$$

Average net of tax wages for informal workers at inspected work-sites are not significantly different from those of their counterparts at non-inspected places of work, before or after an inspection occurs. Although inspections increase the probability of transitioning to a formal job, which represents a 30% increase in labor costs for an informal worker earning the average wage, there is no evidence of an effect in short-term wages.

4.3 Long-term labor market trajectories

The previous sections provide evidence that increasing the cost of informal employment through random inspections leads to higher within-firm formalization rates for informal workers with no changes in short-term wages. We next explore whether labor market trajectories are similar for individuals formalized through inspections relative to their co-workers who were already formal. To do so, we rely on IMSS records and compare co-workers who started a formal match immediately before an inspection to those who were formalized right after the inspection increased the cost of informality at the firm. We refer to the former as "inspection-driven" and the latter as "organic" formalizations.

It is important to clarify some aspects of the analyses that follow. First, we focus on workers starting a formal spell at an inspected firm around the date of the firm's first inspection. We exclude workers who were formally employed elsewhere prior to starting a formal job at the inspected firm. In so doing, we seek to focus on jobs filled by newly formalized workers. The sample will include both within and across firms transitions from informal to formal jobs, as well as transitions to formality from non-employment. Second, we do not attempt to argue that these formalizations are exogenous or that co-workers formalized before an inspection are an appropriate control group for individuals formalized after an inspection. Instead, the goal of the comparison is to assess whether their initially (and endogenously) different labor market outcomes tend to converge or further diverge through time. In other words, we seek to answer the question of whether increasing the cost of informality at a firm allows previously

[31] Borusyak and Jaravel (2017) show that event studies with two-way (unit and time) fixed-effects and dynamic treatment effects are underidentified. However, in this case, for every cohort entering ENOE's sample there is a well specified control group which can be used to independently identify the time fixed-effects: the set of informal employees at work-sites that have not been selected for a random inspection.

informal workers to access labor market trajectories similar to those of their already formal co-workers.

Equation 3 presents our baseline specification. The sample includes all workers at any of an inspected firm's establishments who started a formal job within a 6 month window of the firm's first inspection and who were not recently[32] employed at a different formal firm. $Start_{i,j} - FirstInspection_j \in [-6,6]$ is a dummy variable for the number of months elapsed from the start of a formal match between worker i and firm j to j's first inspection.

$$log\left(Y_{i,j,t}\right) = \sum_{m=-6}^{6} \alpha^m \left(Start_{i,j,t} - FirstInspection_j\right) + \gamma X'_{i,t} + \lambda_j + \theta_t + \varepsilon_{i,j,t} \tag{3}$$

Table 4 shows the estimated α^m coefficients. Each column refers to one of the following dependent variables: natural log of the starting wage of the formal spell at the inspected firm, natural log of the last observed wage of the formal spell at the inspected firm, natural log of the duration of the formal spell (tenure) at the inspected firm, and an indicator variable equal to 1 if the worker separated to a formal job with a different employer. The omitted category are workers who start a formal job a month before the employer is inspected.

Starting wages[33] are similar among workers formalized up to 3 months before an inspection. Meanwhile, starting wages for workers that start a formal job up to a month after the first inspection are 0.9% lower. Workers formalized more than 1 and up to 3 months after the first inspection still have lower wages than their counterparts hired prior to STPS's visit. Consistent with our previous findings that the increase of within-firm formalizations is concentrated in the first 3 months after an inspection (Figure 2 Panel C), the negative difference in starting wages between workers hired before the inspection is reversed after the first quarter.

The wage gap between workers hired up to a month before the inspection and those hired after is higher at the end of their formal spell at the firm. This is seen by comparing the coefficients between in the "Starting Wage" and "Final Wage" columns. For workers hired up to 1 month after the inspection, the gap increases by half a percentage point, from 0.9% to 1.4%. The "Tenure" column shows the differences in the length of a formal spell for workers by the timing of inspections relative to their formal start date. Workers who start after an

[32]We consider only workers who were not at another formal firm in the previous 6 months. Our sample, therefore, includes workers hired from non-employment or from an informal job at a different firm directly into a formal job, as well as workers who were already at the firm but in an informal job.

[33]It is important to note that "starting" wage here refers to the first observed wage for a worker as a formal employee with a given employer. If a worker was first informally employed, we do not observe their first wage at the firm but rather the first wage as a registered, formal worker.

inspection have shorter tenures than their workers hired before the inspection.

The last column in table 4 uses a linear probability model, with the same specification as in equation 3 to compare workers' probability of having a formal job with the next employer, conditional on separating from the inspected firm. We find that the probability of being poached into a formal job is between 0.3 and 0.4 percentage points lower for "inspection-driven" formal workers (i.e. those hired up to a quarter after an inspection) relative to their "organically formalized" co-workers.

While this difference in the likelihood of the next job being formal is statistically significant, it is important to put the magnitude of the effect in perspective. Table 3 shows that, conditional on separating to a new formal employer, informal workers at formal firms have a 2-to-1 probability of getting an informal job with their new employer (64.5% vs. 35.1%). The odds are reversed for formal employees (30.7% vs. 69.3%). Therefore, informal workers formalized after an inspection experience an important improvement in the odds of future formal employment (relative to the probability of a transition to a formal job at a new firm for the average informal worker.)

Finally, we use the specification in equation 4 to compare starting wages at the new firm, denoted as k. $Post\_Inspection_{i,j,t}$ is an indicator variable equal to 1 if i's formal match started on or after the date of j's inspection and zero otherwise, t and k are time and next-employer fixed effects, respectively. We find an insignificant coefficient of $\beta = -0.007$[34]

$$log\left(w_{i,j,k,t}\right) = \beta Post\_Inspection_{i,j,t} + \gamma X'_{i,t} + \lambda_k + \theta_t + \varepsilon_{i,j,k,t} \tag{4}$$

All these results indicate workers embark on an improved labor market trajectory after being formalized due to an inspection. Inspection-driven formalizations lead to more than a label change at their current employer. They get access to a set of social security benefits, and their future labor path resembles that of other formal employees.

4.4 Formal job creation and job destruction

In this section, we use IMSS administrative records to analyze the effect of increasing the cost of informal employment on firms' outcomes. We consider the effects on formal job creation,

[34] $\sigma\left(\hat{\beta}\right) = 0.005$ The sample size (N = 414,886) is smaller as we condition on separating from the inspected firm. There are 16,831 inspected employers in IMSS data that experienced at least one separation of a worker hired within 6 months of the firm's first inspection. The adjusted R-squared is 0.344.

formal job destruction, and formal employment growth. We consider two categories of job creation depending on the type of job previously held by the worker that fills the vacancy. The first type of job creation, that we label as "formal job creation from outside the formal sector," refers to new formal jobs that are filled by workers who were previously not formally employed. This category includes individuals that start a formal job after being out of the labor force, unemployed, informally employed at a different firm as well as those who were formalized by the same employer. We separately consider new formal jobs filled by workers who were previously at a formal job with a different employer (i.e. across-firm formal job to formal job transitions). We refer to these type of transitions as "job creation from within the formal sector" or "formal poaching."[35]

Equation 5 is our main specification. The outcome variable, $Y_{j,t}$, is expressed in either growth rates or inverse hyperbolic sine. We interpret α_q as the percentage point (for dependent variables expressed as growth rates) or percent change (for inverse hyperbolic sine transformations) in the outcome variable, $Y_{j,t}$, q quarters after the first inspection.[36] It is important to note that for each period t, the control group includes all the firms that were never randomly selected for inspection (approximately 79% of the matched sample) as well as those firms whose first inspection occurs after period t.

$$Y_{j,t} = \sum_{q=-6}^{6} \alpha_q Inspection^q_{j,t} + \lambda_j + \theta_{s_j,t} + \varepsilon_{j,t} \tag{5}$$

where $Inspection^q_{j,t}$ is an indicator variable equal to 1 if firm j received its first inspection in period $t-q$.[37] λ_j and $\theta_{s_j,t}$ are firm and industry-time fixed effects.[38]

Formal employment growth (Panel (d) in Figure 4), defined as the quarterly change in the number of formal workers divided by the average formal employment in the previous and current quarter, declines by 2.4 percentage points and remains depressed as far as 18 months after the inspection. This decline is due, on the one hand, to an increase in formal

[35]We use the term "formal poaching" to refer to formally employed workers poached to a new formal job. Formal firms can also poach workers who were previously informally employed elsewhere. Since we cannot distinguish these transitions from workers who are hired from non-employment, we reserve the term poaching to refer to formal-to-formal job transitions across employers.

[36]We follow Bellemare and Wichman (2020) when interpreting coefficients using the inverse hyperbolic sine transformation.

[37]Negative values for q mean that the inspection will occur in q quarters. Positive values mean that the inspection happened q quarters ago.

[38]We use IMSS's industry classification which includes over 300 industry categories.

job destruction (panel (c). A decrease in formal job creation reinforces this effect. Formal job creation from within the formal sector (panel (b)) declines by 3%. That is, after an inspection, firms are less likely to create formal jobs that are filled by workers who were previously employed by other formal firms. Meanwhile, inspections do not appear to have a significant effect on job creation from outside the formal sector, which includes within-firm formalizations.

The worker-level analyses in Figure 2 presented evidence of an increase in outflows from informal to formal jobs with the same employer after an inspection. However, panel (a) in Figure 4 does not show an increase in job creation from outside the formal sector which would include the within-firm informal-to-formal job flows found in Figure 2. As explained earlier, IMSS data does not allow us to distinguish between within-firm formalizations (which ENOE's findings indicate rise after an inspection) and new formal matches filled by non-employed workers or informally employed individuals from other firms. A decline in these across-firm formalizations, or in formal hiring from unemployment, after an inspection could counter the rise in within-firm formalization observed in ENOE's data.

The effect of an inspection likely differs across firms depending on the extent to which they rely on informal workers. Table 2 shows a wide range of variation in the use of informal employment across firms of different size. For small firms, an increase in the cost of informal jobs represents an increase in the majority of its labor force. We examine heterogeneous firm responses using the specification in equation 6.

We group firms into size categories using the mean number of formal workers at the establishment during the first year of activity.[39] We then estimate the change in formal job creation, formal job destruction and exit from the formal sector after an inspection separately for each of the following seven size categories, z: 1 to 5 formal employees, 6 to 10, 11 to 25, 26 to 50, 51 to 100, 101 to 250, and more than 251 formal workers. $Post_{j,t}$ is equal to 1 after firm j's first inspection, and zero otherwise.

$$Y^z_{j,t} = \beta^z Post_{j,t} + \lambda_j + \theta_{s_j,t} + \varepsilon_{j,t} \forall z \tag{6}$$

Figure 5 plots the values of β^z. Panel (a) shows that for employers with less than 6

[39] For firms that receive the first inspection during their first year of operation, we define their size as the average during the first quarters of operation up until the quarter before the inspection.

workers, formal job creation from outside of formality (i.e. formal jobs filled by workers who were previously not formal, including formalizing their own informal workers) increases by 5% after the first inspection. The effect declines with firm size and becomes negative, albeit not statistically significant, for employers with 26 to 250 formal workers. Job creation from outside the formal sector declines by 6.5% at the largest employers. There is no change in formal job creation from outside of the formal sector by employers with between 11 and 50 workers. Finally, new formal jobs filled by workers from outside the formal sector decline at employers with more than 50 formal workers after the first inspection and the magnitude decreases with firm size. Panel (b) plots inspection effects on total job destruction. Formal job destruction increases for establishments with 1 to 5 workers and all establishments with more than 25 workers.

Finally, panel (c) shows that an increase in the cost of informal employment increases the probability of leaving the formal sector altogether for firms with 10 workers or less by 0.6 percentage points. Some of these exits are due to firms hiding from follow-up inspections by receding into the informal sector. This behavior by small firms endogenously creates size-dependent monitoring. The government can only follow-up to enforce compliance at larger firms.[40]

5 Discussion: A Model of Asymmetric Employer Learning with Informal and Formal Jobs

We've shown that increasing the cost of informal employment via inspections leads to a rise in informal workers' probability of transitioning out of informal employment. Within-firm transitions from informal to formal jobs increase. Workers formalized after an inspection have similar probability of being formal in their next job as their already formal co-workers and similar starting wages. For firms, we find that an increase in the cost of informal employment leads to a significant and persistent decline in formal employment growth, and increase in formal job destruction and a decline in formal job creation. In this section, we develop a simple theoretical framework to discuss the mechanisms underlying these empirical results.

[40]Kaplan and Sadka (2011) find evidence of firms avoiding payments from lost labor court rulings by hiding its assets, relocating, or disappearing.

5.1 Model setup

The economy is comprised of a continuum of workers and a large mass of employers. Workers differ in terms of their productivity, θ, which is fixed. Firms use the same production technology but the costs of having informal workers varies with firm size, measured in number of formal workers. We consider three size categories: small (S), medium (M), and large (L). Let α_z, $z \in \{S, M, L\}$, denote the employment-weighted firm shares and $P(\theta)$ the distribution of workers' productivity in the economy. Both $P(\theta)$ and α_z are common knowledge.

Workers are equally productive in all firms and all jobs, but employers do not observe this productivity before hiring. Upon meeting a worker, employers unilaterally choose whether to offer a formal or an informal job,[41] or to keep searching, maximizing their expected profits.

We consider a two period model. Figure 6 depicts the timing of events. In the first period, all workers are unemployed and all firms have a vacancy to fill. At the end of the first period, production occurs and firms that matched with a worker observe her productivity. Matches are destroyed with exogenous probability δ after production occurs. At the start of the second period, incumbents (i.e. firms that matched with a worker on the first period and did not experience an exogenous separation shock) choose whether to endogenously terminate matches, formalize informal workers, or switch formal workers to an informal job, and whether to maintain or increase wages.[42]

Poachers (i.e. unmatched firms) meet an employed worker with probability λ_E and an unemployed one with probability λ_U. Poachers who meet a worker cannot observe the worker's wage or productivity, but do observe a worker's current employment status. For the employed workers they meet, poachers can also observe their formality status and current employer's size.[43] Poachers post contracts conditional on workers' observable characteristics at the time of hiring (i.e. employment/formality status and incumbent size). A contract is a one-period

[41] In a related paper, Bobba et al. (2020) also assume there is no bargaining over formality status. In their model, workers learn on the job and endogenous transitions from informal to formal jobs are due to human capital accumulation. In our framework, workers' ability is fixed and transitions instead occur due to employer learning about worker's type, θ. While both types of learning likely play a role in transitions, our finding that within-firm transition from informal to formal jobs are decreasing with workers' age and tenure (see Figure 1) is at odds with workers on-the-job learning as the main driver of within-firm formalization.

[42] In Michaud (2018), downward wage rigidity prevents employers from offering an initial high wage to poach workers and reducing wages afterwards. The timing of offers in our setup, coupled with there only being one production period after wages are set, eliminates this incentive. We maintain the assumption of downward wage rigidity for simplicity. We also adopt the "incumbents vs. poachers" terminology from Michaud (2018).

[43] We believe it is reasonable to assume size is an observable firm characteristic. We can instead assume poachers only know the distribution of firm size in the economy. The main results follow through.

commitment to a wage and job type (formal or informal).

We next describe firms' expected flow profits in an informal (Π^{INF}) and a formal match (Π^{F}). The productivity of a match is equal to the workers' productivity, θ, and does not depend on the type of job, firm size, or the workers' tenure. If a firm offers a formal job they must pay a proportional pay-roll tax τ_p and a fixed-fee τ_f.[44] With informal jobs, firms avoid these taxes but instead pay an informality cost $c(\theta,z)$ that is an increasing function of worker productivity and firm size.[45]

$$\pi^{INF}\left(\theta, w^{INF}, z\right) = \theta - w^{INF} - c(\theta, z) \tag{7}$$

$$\pi^{F}\left(\theta, w^{F}\right) = \theta - w^{F}(1+\tau) - \tau_F \tag{8}$$

Unemployed workers receive flow utility $b_U > 0$. Employed workers' flow utility in any period equals their wage.[46]

Firms are risk-neutral and can update their offers for subsequent periods after observing workers' productivity. Therefore, each period can be treated as a separate maximization problem. Given the information available to them at the beginning of the period, firms offer workers the contract (job type and wage) that maximizes expected profits in that period.

Workers are also risk neutral and maximize the present value of their lifetime utility flow. In a multi-period setting, this could imply rejecting a higher wage offer in favor of a lower initial wage with better outside options in the future (as in Meghir et al. (2015) where informal to formal job transitions can be accompanied by a wage cut). In our 2-period setting, employed workers choose the highest wage offer in the second period, and accept a job in the first period if the offer is above their reservation unemployment value.

In the next sections, we solve for incumbents' and poachers' optimal offers. We then show how their decisions change when they experience a permanent increase in the cost of informal employment, and how this affects workers' outcomes.

[44] These taxes are meant to reflect the payroll tax structure presented in Table 1.

[45] In section 4, we argue that firms' responses to random monitoring are consistent with a positive and non-linear correlation between the costs of informal employment and firm size.

[46] We could instead assume that flow utility includes a "social benefit" component that differs between formal and informal jobs. This extension changes the equilibrium wage offered by employers' but does not significantly alter the main results so we omit it to maintain simplicity. The setup is similar to Meghir et al.'s (2015) [38] when modeling an economy without unemployment insurance, like Mexico's.

5.1.1 Period 1

Let $\bar{\theta}$ be the average labor productivity in the economy. In the first period, all z-sized firms that meet a worker offer the job type where the average worker yields the highest expected profit. Let $\theta_1^{F,z}$ be the value of productivity that makes a z-size firm indifferent between a formal and an informal job with wages equal to the reservation wage w^r of the average unemployed worker. Size z-firms offer formal (informal) jobs if $\bar{\theta} \geq (<)\, \theta_1^{F,z}$.

Incumbent Maximization Problem. Period 1:

$$\max\left\{ \underbrace{\mathrm{E}\left[\Pi^{INF}\left(\theta, w^r, z\right)\right]}_{\text{hire all informally}}, \underbrace{\mathrm{E}\left[\Pi^{F}\left(\theta, w^r\right)\right]}_{\text{hire all formally}}, \underbrace{0}_{\text{no hires}} \right\} \qquad \forall z \in \{S, M, L\} \tag{9}$$

$$\text{where } \mathrm{E}\left[\Pi^{INF}\left(\theta, w^r, z\right)\right] = \int_{\theta_L}^{\theta^H} \left(\theta - w^r\left(\bar{\theta}, z\right) - c\left(\theta, z\right)\right) p\left(\theta\right) d\theta$$

$$\mathrm{E}\left[\Pi^{F}\left(\theta, w^r\right)\right] = \int_{\theta_L}^{\theta^H} \left(\theta - w^r\left(\bar{\theta}\right) \times \left(1 + \tau_p\right) - \tau_F\right) p\left(\theta\right) d\theta$$

5.1.2 Period 2

Incumbents now have full information about their workers' productivity. They also know that with probability $\lambda_E \times \alpha_{z'}$ their worker will meet a size-z' poacher who posted a wage $w^{Pz'}(k, z)$ for all workers currently employed with a size-z incumbent in a k-type job. Therefore, when deciding on their workers' current contract, incumbents determine the probability that their workers' get poached, and the type of contract that they will be offered, conditional on meeting another firm.

Let $w_1^{k,z}$ and $w_2^{k,z}$ be the equilibrium wages in a k-type job, $k \in \{F, IN\}$ at size-z incumbents in the beginning of periods 1 and 2, respectively. Let $w^{Pz'}(k, z)$ be the equilibrium wage that a size z' poacher posts for workers employed in a k-type job with a size-z incumbent. In period 2, the worker accepts the poacher's offer if the wage is higher than with the current employer.[47] If the worker leaves, the incumbent's profits are zero. If the wage offered by the incumbent at the beginning of period 2 is higher than the poacher's offer (or if the worker did not meet any poachers), the worker stays with the incumbent. By increasing wages,

[47]For simplicity, we assume workers only care about wages. Alternatively, we could assume workers also differ in terms of how much they value a formal job relative to an informal one. The set of poachers an incumbent is willing to compete with would then be defined over wage levels and job types.

incumbents decrease turnover but also reduce their share of the match surplus. In equilibrium, incumbents choose to match poachers' offers for the set of workers with productivity above a threshold which we denote as $\theta\left(w^{Pz}(k,z)\right)$.[48] Equation 10 below presents the maximization problem for incumbent firms at the beginning of period 2.

Incumbents Maximization Problem. Period 2:

$$\begin{aligned}\max\Bigg\{&\max_{w_2^{IN,z}}\left(1-\lambda_E\left(1-\sum_{z'}\alpha_{z'}\mathbf{I}\left[w^{Pz'}(IN,z)<w_2^{IN,z}\right]\right)\right)\left(\theta-w_2^{IN,z}-c(\theta,z)\right),\\&\max_{w_2^{F,z}}\left(1-\lambda_E\left(1-\sum_{z'}\alpha_{z'}\mathbf{I}\left[w^{Pz'}(F,z)<w_2^{F,z}\right]\right)\right)\left(\theta-w_2^{F,z}\times(1+\tau_p)-\tau_F\right),0\Bigg\}\end{aligned} \tag{10}$$

s.t. $w_2^{F,z}, w_2^{IN,z} \geq w_1^z$ where $\mathbf{I}\left[w^{Pz'}<w_2^z\right]$ is equal to 1 if a z'-size poacher's equilibrium wage offer is lower than than the incumbent's, and w_1^z is the first period wage.

Taking incumbents' optimal responses as given, poachers update their prior on employed and unemployed workers' productivity. Let $\gamma_{k,z}(\theta)$ be the updated probability that the productivity of a worker employed at a k-job with a z-size firm equals θ. Further, let $\theta(w(k,z))$ be the highest productivity worker that poachers can successfully attract from a size-z incumbent if they post a wage equal to $w(k,z)$ for workers in k-type jobs. Equation 11 presents the size-z' poacher maximization problem conditional on meeting an employed worker.

Poachers Maximization Problem

$$\begin{aligned}\mathrm{E}[\Pi^{Pz'}(\theta)\,|k,z]=\max\Bigg\{&\max_{w^{Pz'}(k,z)}\int_{\theta^{k,z}}^{\theta\left(w^{Pz'}(k,z)\right)}\left(\theta-w^{Pz'}(k,z)(1+\tau_p)-\tau_F\right)\gamma_{k,z}(\theta)\,d\theta,\\&\max_{w^{Pz'}(k,z)}\int_{\theta^{k,z}}^{\theta\left(w^{Pz'}(k,z)\right)}\left(\theta-w^{Pz'}(k,z)-c\left(\theta,z'\right)\right)\gamma_{k,z}(\theta)\,d\theta,0\Bigg\}\quad\forall z'\in\{S,M,L\}\end{aligned} \tag{11}$$

[48] Postel-Vinay and Robin (2004) develop a model with endogenous search intensity by workers where firms optimally choose for which workers to compete. In their setting, high-productivity firms match outside offers while low-productivity firms do not. In our setup, larger firms are more likely to match poachers' offers than smaller firms even in the absence of endogenous search intensity because larger firms have more to lose from higher turnover.

5.2 Equilibrium

Strategies for each size-z incumbents and each size-z' poachers are mappings from their available information set to an offer (job type and wage), taking competing employers' actions as given. z-sized incumbents' strategies are a wage and job type (k) for each type of worker $w_2^{k,z}(\theta)$, z'-poachers' strategies are a wage and job type (k') for all workers employed in a job k at a z-size incumbent $\left(w^{Pz'}(k,z), k'\right)$, and a strategy for each type of workers is a quit decision $Q\left(w_2^{k,z}(\theta), w^{Pk'}(k,z)\right)$. In a Symmetric Perfect Bayesian Equilibrium, strategies are sequentially rational, poachers' beliefs about workers' productivity are consistent with equilibrium distributions, and the distributions of workers in each job and firm size category are consistent with optimal strategies. We next characterize such an equilibrium. We first consider incumbents' strategies. Only 4 possible wages can be part of an equilibrium strategy for incumbents: to keep the wage from period 1 unchanged, or to offer a wage equal to one of the offers expected from the three types of poachers, up to the value of the match. An incumbent that increases the wage above the first period offer but below the lowest poacher's posted wage has lower profits if the worker does not meet a poacher (with probability $1-\lambda_E$) and does not keep the worker if an outside offer arrives. Therefore, any increase in wage must at least match the lowest outside offer to be an equilibrium strategy.

Incumbents that keep first period wages receive higher profits if the worker does not meet a poacher, but lose the worker with certainty if there is an outside offer. Equations 12 and 13 present the expected change in profits from increasing a formal and informal worker's wage to compete with a size-z' poacher's offer. Incumbents will increase formal (informal) workers' wages to match z'-sized poachers' offer[49] if their productivity θ is such that equation 12 (or 13) is positive. For workers with lower productivity, incumbents will keep the first period wage unchanged. Then, they will choose the type of job that maximizes their expected profits given these equilibrium wages following a threshold rule: formalize workers with $\theta \geq \theta^{F,z}$, offer informal jobs to workers with $\theta^z \leq \theta \leq \theta^{F,z}$ and terminate matches where $\theta < \theta^z$.[50] Equation 10 determines these thresholds.

[49] Incumbents choose period 2 wages before outside offers arrive. However, employers can rationally predict the poachers' equilibrium wage offers.

[50] Incumbents follow a threshold rule because profits from a formal match are linearly increasing in productivity while the costs of an informal match are increasing and convex on worker productivity.

Change in Incumbent Profits from Competing with Poachers

$$\frac{\delta\ E[\Pi^{F,z}(\theta)]}{\delta w = w^{Pz'}(F,z) - w_1} = (1-\lambda_E)\ \underbrace{\Big(w_1 - w^{Pz'}(F,z)\Big) \times (1+\tau_p)}_{\text{Loss in Profits if there is no outside offer}}$$
$$+\lambda_E \times \underbrace{\left(\sum_{z'' s.t. w^{Pz''}(F,z) \leq w^{Pz'}(F,z)} \alpha_{z''}\right)}_{\text{Measure of poachers with} \leq \text{offers}} \underbrace{\Big(\theta - w^{P,z'}(F,z)(1+\tau_p) - \tau_f\Big)}_{\text{Gain in profits from competing with poachers}} \tag{12}$$

$$\frac{\delta\ E[\Pi^{IN,z}(\theta)]}{\delta w = w^{Pz'}(IN,z) - w_1} = (1-\lambda_E) \qquad \underbrace{\Big(w_1 - w^{Pz'}(IN,z)\Big)}_{\text{Loss in Profits if there is no outside offer}}$$
$$+\lambda_E \times \underbrace{\left(\sum_{z'' s.t. w^{Pz''}(IN,z) \leq w^{Pz}(IN,z)} \alpha_{z'}\right)}_{\text{Measure of poachers with} \leq \text{offers}} \underbrace{\Big(\theta - w^{Pz'}(IN,z) - c(\theta,z)\Big)}_{\text{Gain in profits from competing with poachers}} \tag{13}$$

A size-z' poacher updates its prior about workers they meet based on the worker's current employment situation. Higher productivity workers are more likely to be formally employed due to the cost of informal employment increasing in worker productivity. Poachers rationally expect higher wage offers to increase the average quality of the set of workers that they can successfully poach. The equilibrium poacher's offer is such that the increase in the expected productivity of the match is equal to the increase in its total cost, which we obtain from the poacher's maximization problem in equation 11. Given these equilibrium wages, poachers then choose the job type that maximizes their profits, which is determined by the cost of informally employing the average poached worker.

Firms' learning about workers' productivity, and hence about the cost of offering them an informal contract, leads to within-firm changes in formality status. The threat of outside offers is what drives within-firm wage growth. Even though workers are equally productive in formal and informal jobs, and productivity is fixed, workers experience wage growth in both types of jobs. Wage growth is higher for formal workers, both within and across firms, because poachers expect higher productivity from formal workers and incumbents realize that formal workers are more expensive to retain. This leads incumbents to formalize fewer workers than they would if there were no outside offers.

Although the purpose of the model is not intended for calibration, we do set some parameters to match the data in the sample of non-inspected DNE firms to carry out some simulations. First, we set each α_z to match small (1-3 employees), medium (4-10), and large (11+) firms' employment weighted shares in the economy.[51] For the cost of informality, we assume the functional form $c = (\theta \tilde{z})^{\gamma}$ where $\tilde{z}$ is the size ratio of each firm category relative to the smallest group. We set $\gamma = 2$. Finally, we set λ_E, λ_U, and δ to match the job-to-job finding rate, the finding rate from unemployment, and the overall separation rate. The taxes in the formal sector, τ_F and τ, are 32% and 20%, respectively. We assume that worker's productivity in the economy follows a Gamma distribution, $\Gamma(8,20)$.

We use the calibrated model to show how incumbents respond to an x% increase in the cost of informal matches such that $c'(\theta,z) = (1+x) \times c(\theta,z)$, with $x \in \{0,1\}$. We assume the shock arrives at the end of the first period. For incumbents, this is after productivity is observed but before the decision to set wages and formality status for the second period. For poachers, it is before a hiring decision is made. We assume that upon meeting an employed worker, a poacher does not know whether the current employer received the shock or not.

We start with the effects on incumbents. As shown in Figure 7, an increase in the cost of informal employment increases both endogenous separations (by raising the match termination threshold, θ_z) and formalizations (by decreasing the threshold to offer formal jobs, $\theta^{F,z}$) for all firms. The change in the share of formal workers is highest at small firms since larger employers already formalize most of their workers. Because to outside firms the workers who are formalized after the increase in costs are indistinguishable from workers that would have been formalized in the absence of the shock (i.e. workers with $\theta > \theta^{F,z}$), poacher's offers do not change after an incumbent's cost of informality increases.

Next, we consider the effect of increasing the cost of informal employment on poachers (i.e. before a match is created). The shock increases the cost of offering informal jobs but does not change poachers' expectations about the productivity of workers they meet. Therefore, hiring decisions change only when the expected surplus of the match is higher in an informal job. For these matches, the increase in the cost of informality can lead to either an increase in formal job creation or to not making an offer at all (a decrease in total job creation). Whether poachers' decisions change depends on the magnitude of the rise in the cost of informality. The rows labeled "Min. $\Delta c(\theta,z)$" in Table 5 show the minimum percent increase in the cost of informal employment required to change a poacher's optimal informal offer to either a formal offer or to not making an offer.

[51] The values of these parameters are consistent with those estimated in Levy (2018).

The change in the cost of informal employment needed in order to change poachers' hiring behavior is decreasing in firm size. A large-sized poacher that meets an informal worker at an equal-sized incumbent would offer the worker an informal job in the absence of a rise in the cost of informality. If the cost increases by at least 18%, large poachers will no longer make job offers to workers informally employed at large incumbents. While our model does not track the evolution of firm size, table 5 and figure 7 are consistent with larger firms shrinking after a rise in the cost of informal employment (due to reduced hiring by poachers and increased separations at incumbents).

Table 5 also shows that, conditional on meeting a worker currently in an informal job, the probability that the poacher offers an informal job depends on the incumbent's size relative to the poacher's. Because informal employment is "cheaper" at smaller firms than at larger ones, a formal job at a small incumbent is a stronger signal of high worker productivity than at a larger firm. Conversely, an informal job at a small firm is not necessarily reflective of low worker productivity.

6 Conclusions

In principle, formal firms must register all their employees with the Social Security Institute and pay payroll taxes. In practice, enforcement is lax and many workers employed at formal firms, of all sizes, have informal jobs. We examine the effects from increasing the cost of informal employment on firm and worker outcomes using inspections at randomly selected formal establishments as a positive shock to the cost of informality.

Using administrative employer-employee matched data, we find that even if all firms face the same probability of being monitored by the government, smaller firms are more likely to exit the formal sector after an inspection as the cost of informal employment increases. The possibility of "disappearing" when faced with a higher likelihood of being fined leads to an endogenous positive correlation between firm size and cost of informal employment, even if inspections are randomly allocated across firms. Conditional on staying in the formal sector, within-firm informal to formal job transitions increase at smaller firms, while larger firms are more likely to respond to an increase in the cost of informal jobs by decreasing hiring and increasing job destruction. Given that the latter effects dominate in the aggregate, inspections have a negative and persistent impact on firm growth due to firms' responses. These responses are consistent with the cost of informality being increasing in firm size.

We obtain consistent results using household survey data. When employers' cost of informal jobs increases due to an inspection, informal workers' quarterly probability of remaining informal with the same employer decreases by 9 percentage points. The decrease is caused by a 50% increase in the probability that their current employer formalizes them, and a 40% increase in separation rates.

We find evidence consistent with prospective employers using formality status as a signal of workers' productivity. Future employers treat formal workers similarly regardless of whether they were formalized before or after an inspection. Conditional on separating from the inspected employer, formal workers have similar probability of being poached into a formal job and equivalent starting wages at the new job than their "organically-formalized" co-workers.

References

Agin, Marilyn A., and Anant P. Godbole. 1992. "A New Exact Runs Test for Randomness." *Computing Science and Statistics*, 281–285.

Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. "The Effects of Labor Market Policies in an Economy with an Informal Sector." *The Economic Journal*, 119(539): 1105–1129.

Alcaraz, Carlo, Daniel Chiquiar, and Alejandrina Salcedo. 2015. "Informality and segmentation in the Mexican labor market." Banco de México Working Papers 25.

Almeida, Rita, and Jennifer Poole. 2017. "Trade and labor reallocation with heterogeneous enforcement of labor regulations." *Journal of Development Economics*, 126(C): 154–166.

Almeida, Rita, and Pedro Carneiro. 2005. "Enforcement of regulation, informal labor and firm performance." *IZA Discussion Paper No.1759*, 1–47.

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

Amon, Ivan, Francisco Moreno, and Jaime Echeverri. 2012. "Phonetic Algorithm to Detect Duplicate Text String in Spanish." *Revista Ingenierías Universidad de Medellin*, 11(20).

Bellemare, Marc F., and Casey J. Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation." *Oxford Bulletin of Economics and Statistics*, 82(1): 50–61.

Bobba, Mateo, Luca Flabbi, Santiago Levy, and Mauricio Tejada. 2020. "Labor Market Search, Informality, and On-The-Job Human Capital Accumulation." *Journal of Econometrics*, *Forthcoming*.

Borusyak, Kirill, and Xavier Jaravel. 2017. "Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume." *mimeo*.

Bosch, Mariano, and Julen Esteban-Pretel. 2012. "Job creation and job destruction in the presence of informal markets." *Journal of Development Economics*, 98: 270–286.

Bosch, Mariano, and William Maloney. 2010. "Comparative analysis of labor market dynamics using Markov processes: An application to informality." *Labour Economics*, 17: 621–631.

Cámara de Diputados del H. Congreso de la Unión. June, 2015*b*. "Ley Federal del Trabajo." *http://www.diputados.gob.mx/LeyesBiblio/pdf/125*$_1$20615.pdf.

Cámara de Diputados del H. Congreso de la Unión. *November, 2015*a. *"Ley del Seguro Social."* http://www.diputados.gob.mx/LeyesBiblio/pdf/92$_1$21115.pdf.

Cisneros-Acevedo, Camila. 2020. “Unfolding Trade Effect in Two Margins of Informality. The Peruvian Case.”

De Paula, Áureo, and Jose A. Scheinkman. 2010. “Value-Added Taxes, Chain Effects, and Informality.” *American Economic Journal: Macroeconomics*, 2(4): 195–221.

D’Erasmo, Pablo N., and Hernan J. Moscoso Boedo. 2012. “Financial structure, informality and development.” *Journal of Monetary Economics*, 59(3): 286 – 302.

Fajnzylber, Pablo, William F. Maloney, and Gabriel V. Montes-Rojas. 2011. “Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program.” *Journal of Development Economics*, 94(2): 262 – 276.

Fiess, Norbert M., Marco Fugazza, and William F. Maloney. 2010. “Informal self-employment and macroeconomic fluctuations.” *Journal of Development Economics*, 91(2): 211 – 226.

Fortin, Bernard, Nicolas Marceau, and Luc Savard. 1997. “Taxation, wage controls and the informal sector.” *Journal of Public Economics*, 66(2): 293 – 312.

Gallardo del Angel, Roberto. 2013. “Gross Flows of Formal and Informal Workers in the Mexican Labor Market.” *Estudios Economicos*, 28(2): 299–324.

Garcia, Emmanuel, David Kaplan, and Joyce Sadka. 2012. “Endogenous Asymmetric Stakes in Litigation: The effects of the Mexican Social Security Authority as Co-Defendant.” *mimeo*.

Hall, Thomas W., Andrew W. Higson, Bethane Jo Pierce, Kenneth H. Price, and Christopher J. Skousen. 2012. “Haphazard Sampling: Selection Biases Induced by Control Listing Properties and the Estimation Consequences of these Biases.” *Behavioral Research in Accounting*, 24(2): 101–132.

Hsieh, Chang-tai, and Benjamin A Olken. 2014. “The Missing “Missing Middle”.” *Journal of Economic Perspectives*, 28(3): 89–108.

Instituto Mexicano del Seguro Social. 2002. “Reglamento de la Ley del Seguro Social en Materia de Afiliación, Clasificación de Empresas, Recaudación y Fiscalización.” *http://www.imss.gob.mx/sites/all/statics/pdf/reglamentos/4046.pdf.*

Instituto Mexicano del Seguro Social. 2015. “Informe de Labores y Programa de Actividades 2014-2015.” *http://www.imss.gob.mx/sites/all/statics/pdf/informes/2015/*$2015_{i}nforme_{l}abores_{a}ctividades.pdf$.

Instituto Nacional de Geografía y Estadística. *2016*a. *“Encuesta Nacional de Calidad Regulatoria e Impacto Gubernamental en Empresas.”* http://www.beta.inegi.org.mx/proyectos/encestablecimientos/especiales/encrige/2016/.

Instituto Nacional de Geografía y Estadística. *2016*b. *"Encuesta Nacional de Ocupación y Empleo 2005-2016. Retrieved from http://www.beta.inegi.org.mx/proyectos/enchogares/regulares/enoe/default.html?init=2."*

International Labour Organization. *2003. "Guidelines concerning a Statistical Definition of Informal Employment."* The 17th International Conference of Labour Statisticians (ICLS).

International Labour Organization. *2018. "Women and men in the informal economy: a statistical picture (third edition)."* Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General.

Kaplan, David, and Joyce Sadka. *2011. "The Plaintiff's Role in Enforcing a Court Ruling: Evidence from a Labor Court in Mexico."* IDB Working Paper Series*, 264.*

Kumler, Todd, Eric Verhoogen, and Judith Frías. *2020. "Enlisting Employees in Improving Payroll Tax Compliance: Evidence from Mexico."* The Review of Economics and Statistics*, 1–16.*

Leal Ordóñez, Julio César. *2014. "Tax collection, the informal sector, and productivity."* Review of Economic Dynamics*, 17(2): 262 – 286.*

Levy, Santiago. *2008.* Good Intentions, Bad Outcomes. Social Policy, Informality, and Economic Growth in Mexico. *Washington, D.C.:Brookings Institution Press.*

Levy, Santiago. *2018.* Under-Rewarded Efforts The Elusive Quest for Prosperity in Mexico. *Washington, D.C.:Inter-American Development Bank.*

Maloney, William F. *1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico."* World Bank Economic Review.

McKenzie, David, and Yaye Seynabou Sakho. *2010. "Does it pay firms to register for taxes? The impact of formality on firm profitability."* Journal of Development Economics*, 91(1): 15 – 24.*

Meghir, Costas, Renata Narita, and Jean-Marc Robin. *2015. "Wages and informality in developing countries."* American Economic Review*, 105(4): 1509–1546.*

Michaud, Amanda M. *2018. "A Quantitative Theory of Information, Worker Flows, and Wage Dispersion."* American Economic Journal: Macroeconomics*, 10: 154 – 183.*

National Institute for Transparency, Access to Information and Personal Data Protection. *2016. "Information request no. 0001400017316 and no. 0001400017416."*

Perry, Guillermo E, William Maloney, Omar S Arias, Pablo Fajnzylber, Andrew D Mason, and Jaime Saavedra-chanduvi. *2007.* Informality. Exit and Exclusion. *Washington, D.C.:The World Bank.*

Ponczek, Vladimir, and Gabriel Ulyssea. *2018. "Enforcement of Labor Regulation and the Labor Market Effects of Trade: Evidence from Brazil."* IZA Working Papers*, 11783.*

Postel-Vinay, Fabien, and Jean-Marc Robin. *2004. "To match or not to match?: Optimal wage policy with endogenous worker search intensity."* Review of Economic Dynamics*, 7(2): 297 – 330.*

Poterba, James M, and Lawrence H Summers. *1986. "Reporting Errors and Labor Market Dynamics."* Econometrica*, 54(6): 1319–1338.*

Pratap, Sangeeta, and Erwan Quintin. *2006. "Are labor markets segmented in developing countries? A semiparametric approach."* European Economic Review*, 50: 1817–1841.*

Rauch, James E. *1991. "Modeling the Informal Sector Formally."* Journal of Development Economics*, 35(1): 33–47.*

Ronconi, Lucas. *2010. "Enforcement and compliance with labor regulations in Argentina."* Industrial and Labor Relations Review*, 63: 719–736.*

Secretaría del Trabajo y Previsión Social. *2011. "Manual del Inspector Federal de Trabajo."* http://inspeccion.stps.gob.mx/Ayuda/AyudaCT.pdf.

Secretaría del Trabajo y Previsión Social. *2014*a*. "Reglamento General para la Inspección y Aplicación de Sanciones por Violaciones a la Legislación Laboral."* http://www.stps.gob.mx/02$_{s}ub_{t}rabajo/01_{d}gaj/r_{i}nspeccion.pdf$.

Secretaría del Trabajo y Previsión Social. 2014*b*. "Reglamento Interior de la Secretaría del Trabajo y Previsión Social." *http://www.stps.gob.mx/bp/secciones/conoce/Reglamento%20Interior%20STPS.pdf.*

Sheskin, David. 2004. *Handbook of parametric and nonparametric statistical procedures.* . Third edition. ed., Boca Raton:Chapman & Hall/CRC.

Smeeton, Nigel, and Nicholas J. Cox. 2003. "Do-it-yourself shuffling and the number of runs under randomness." *Stata Journal*, 3(3): 270–277.

Ulyssea, Gabriel. 2010. "The formal-informal labor market segmentation hypothesis revisited." *Brazilian Review of Econometrics*, 30: 311–334.

Ulyssea, Gabriel. 2018. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review*, 108(8): 2015–47.

Zenou, Yves. 2008. "Job search and mobility in developing countries. Theory and policy implications." *Journal of Development Economics*, 86: 336–355.

7 Tables

Table 1: Contributions to Government Mandated Benefits (per worker/day)

	Employer	Worker	Government	Total
Fixed Fee	20.4% $\times$ MW	0%	13.9%$\times$ MW	34.3% $\times$ MW
Proportional to wage				
If wage$<$ 3 MW	15.15%	2.375%	0.475%	18%
Added % on wage$\geq$ 3 MW	1.10%	0.40%	0%	1.50%
Upper bound	16.394%$\times$ 25MW	2.272% $\times$ 25MW		

Source: Own calculations based on payroll tax rates established in the Social Security Law[14].
MW refers to the daily minimum wage, equal to MXN$73.04 on January 2016.

Table 2: Share of Informal Workers at Formal Establishments Matched to DNE
(January 2005-June 2016 Average)

Industry	Establishment Size										Total
	2-5	6-10	11-15	16-20	21-30	31-50	51-100	101-250	251-500	501+	
Oil & Mining	40.5	56.5	24.6	35.6	16.5	16.8	20.2	3.9	3.0	1.3	7.7
Manufact.	79.7	61.3	53.0	37.8	30.2	21.3	10.0	4.7	3.3	2.4	20.6
Construction	61.1	44.7	34.0	35.1	31.4	23.7	21.8	10.2	6.3	5.1	27.6
Retail/Wholesale	72.5	34.5	25.5	16.8	13.5	10.8	11.0	7.9	3.3	9.2	32.4
Lodging & Food	83.1	65.5	41.2	34.6	16.4	15.7	8.6	1.7	3.8	4.0	45.3
Transport. & Communic.	79.4	46.5	34.4	33.4	19.9	15.7	10.0	8.5	8.0	11.3	27.5
Finance & Prof. Bus. Serv.	67.1	40.6	25.5	25.2	23.0	20.2	14.9	11.4	10.3	6.8	26.4
Social Serv.s	45.2	26.8	13.7	11.4	13.4	11.1	10.6	6.0	9.9	7.0	14.9
Other Serv.	80.4	55.0	46.1	40.6	31.3	37.9	25.8	18.4	35.6	27.0	56.7
Government & NGOs	62.0	47.8	36.9	32.5	42.0	34.1	31.6	25.5	22.4	11.4	17.2
Total	73.1	45.0	29.8	24.9	20.6	17.8	13.6	8.7	7.2	7.3	26.0

Source: Own calculations using data from ENOE (2005-2016) and DNE.
We calculated these shares based on the sample of individuals employed at establishments matched to firms with at least one establishment registered in the DNE. We measure establishment size using the mode of the number of individuals (including the owner, formal and informal employees) working at the establishment as reported by all workers surveyed by ENOE during each calendar year.

Table 3: Predicted Quarterly Transition Probabilities

		Initial Labor Market Status	
Labor Market Status		Formal Establishment	
Next Quarter		Informal	Formal
Same Formal Firm	Informal	38.5%	1.1%
	Formal	14.2%	81.8%
Separation	New Formal Firm	7.7%	7.5%
	Other*	40.2%	9.1%
Conditional on Separating to a New Formal Firm			
New Formal Firm	Informal	64.5%	30.7%
	Formal	35.1%	69.3%

Source: Own calculations using data from ENOE (2005-2016) and DNE.

We calculate predicted probabilities using a multinomial logit. The sample consists of individuals employed at DNE firms for at least one of the quarters when they participate in ENOE's survey.

* "Other" separations include separating to a job as an employee at an informal firm, becoming self-employed, transitioning to unemployment and movements out of the labor force.

Table 4: Labor Market Trajectories: "Organic" vs "Inspection-Driven" Formal Workers

	Months from Inspection to Start Date	At Inspected Firm: Starting Wage	At Inspected Firm: Final Wage	At Inspected Firm: Tenure	At Next Formal Employer: Pr. of Next Formal Job
Organic	m = (-3,-6]	0.037***	0.041***	0.180***	0.003**
		(0.001)	(0.001)	(0.004)	(0.001)
	m = (- 1,-3]	0.009	0.009***	0.058***	0.000
		(0.001)	(0.001)	(0.004)	(0.001)
Inspection Driven	m = [0,1]	-0.009***	-0.014***	-0.075***	-0.003**
		(0.002)	(0.002)	(0.005)	(0.001)
	m = (1,3]	-0.002***	-0.007***	-0.078***	-0.004***
		(0.001)	(0.002)	(0.004)	(0.001)
	m = (3,6]	0.003**	-0.005***	-0.105***	-0.007***
		(0.001)	(0.001)	(0.004)	(0.001)
N		2,165,139	2,165,297	2,165,297	2,133,772
N firms		31,520	31,520	31,520	31,520
Adj. R^2		0.5766	0.5527	0.2889	0.0665

Notes: "Organic" refers to inflows to formal jobs that do not follow an inspection. By contrast, "inspection-driven" formalizations are those that occur immediately after an inspection.

The sample includes all workers at any of an inspected firm's establishments who start at formal spell within a 6 month window of the firm's first inspection and who were not employed at a different formal firm in the 6 months prior.

Each cell shows the estimated α^m coefficients from equation 3. The dependent variable in each column is, respectively, the natural log of the starting wage of the formal spell at the inspected firm, natural log of the last observed wage of the formal spell at the inspected firm, natural log of the duration of the formal spell (tenure) at the inspected firm, and an indicator variable equal to 1 if the worker separated to a formal job with a different employer. The omitted category are workers who start a formal job a month before the employer is inspected.

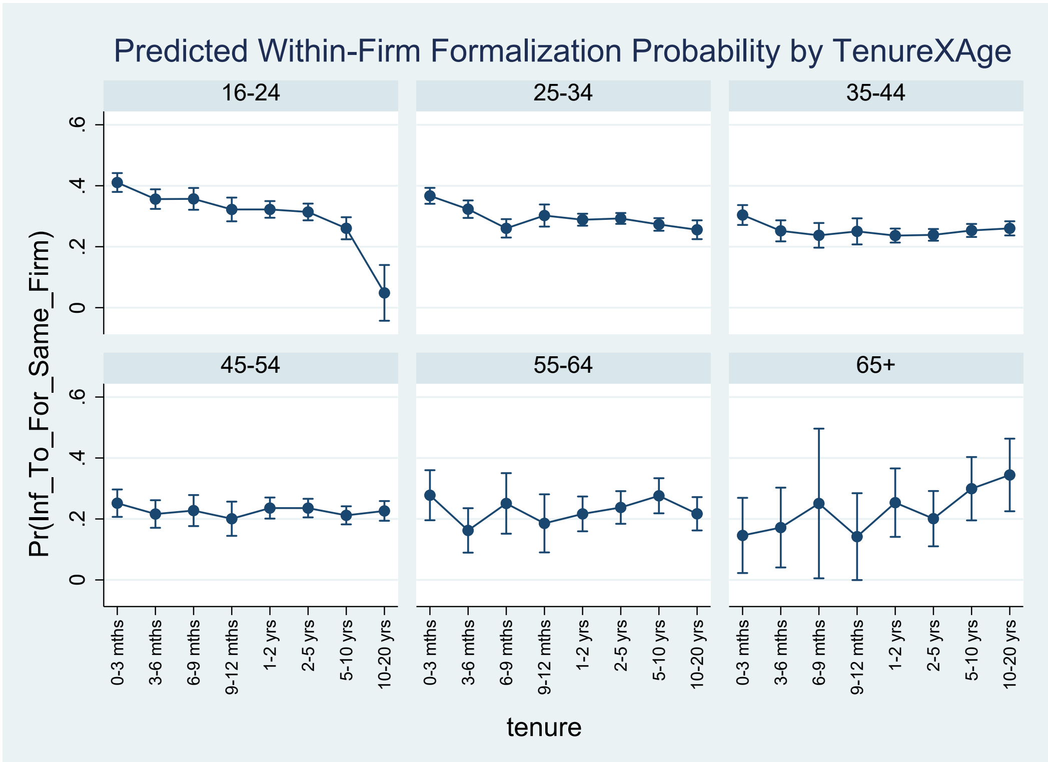
Table 5: Inspecting Poachers: Effects on Hiring and Formal Job Offers

Poacher Size and Inspection Status			Incumbent Size and Current Formality Status						
			Small		Medium		Large		Unemp.
			INF	F	INF	F	INF	F	
Small	Not Inspected		INF	F	INF	F	INF	F	INF
	Inspected	Min. $\Delta c(\theta, S)$	84%	0	233%	0	363%	0	1,139%
		New Offer	F	F	F	F	No Hire	F	F
Medium	Not Inspected		F	F	INF	F	INF	F	F
	Inspected	Min. $\Delta c(\theta, M)$	0	0	50%	0	106%	0	0
		New Offer	F	F	F	F	No Hire	F	F
Large	Not Inspected		F	F	F	F	INF	F	F
	Inspected	Min. $\Delta c(\theta, L)$	0	0	0	0	18%	0	0
		New Offer	F	F	F	F	No Hire	F	F

Notes: Each cell in the "Not Inspected" rows shows the optimal job type offered by non-inspected poachers that meet a worker with the current labor/formality status and incumbent size indicated by the column headers. The "min $\Delta c(\theta, z')$ show the minimum percent increase in the cost of informal employment that inspections need to generate for a size-z' poacher to change their offers. The "New Offer" rows indicate optimal poacher's behavior assuming that inspections increase the cost of informal employment by min. $\Delta c(\theta, z')$.

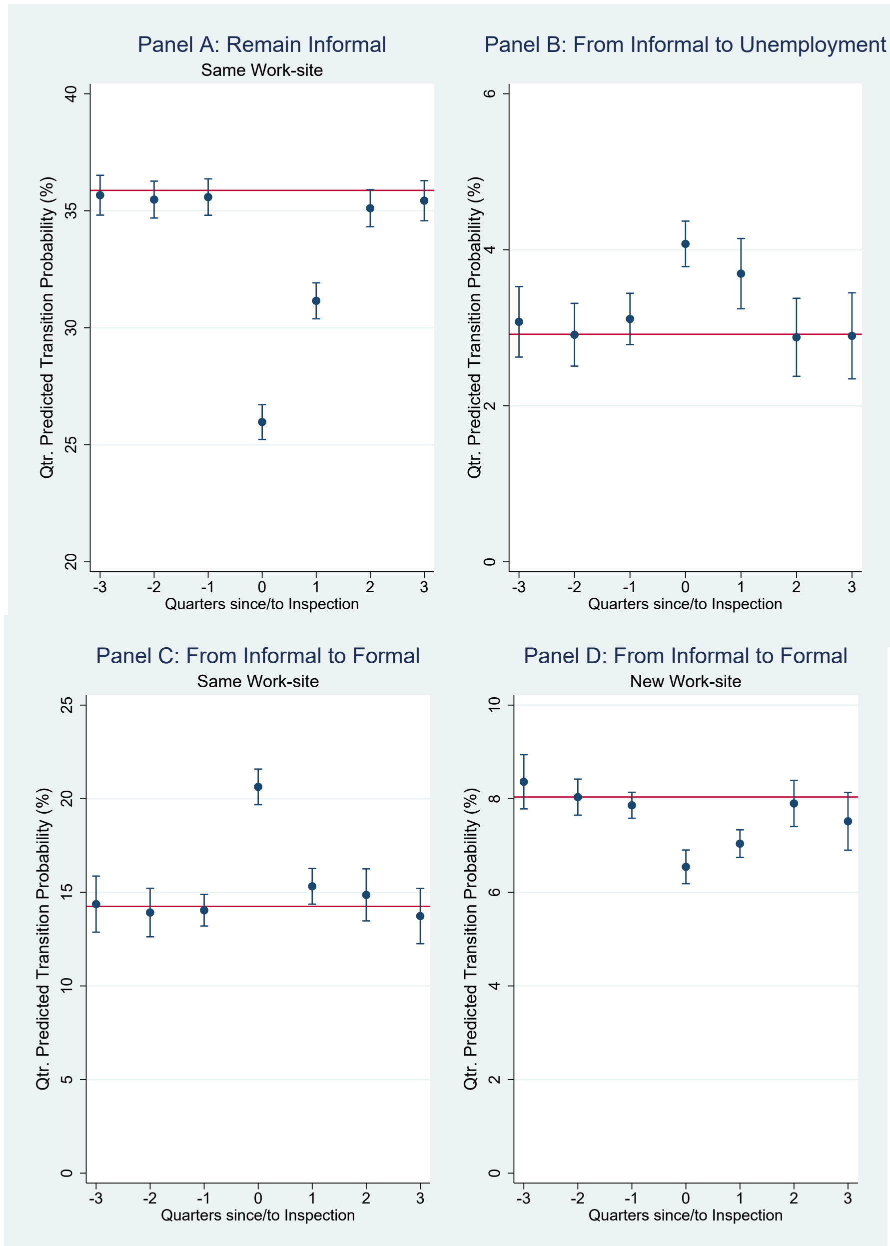
8 Figures

Figure 1: Within-Firm Formalization Probability



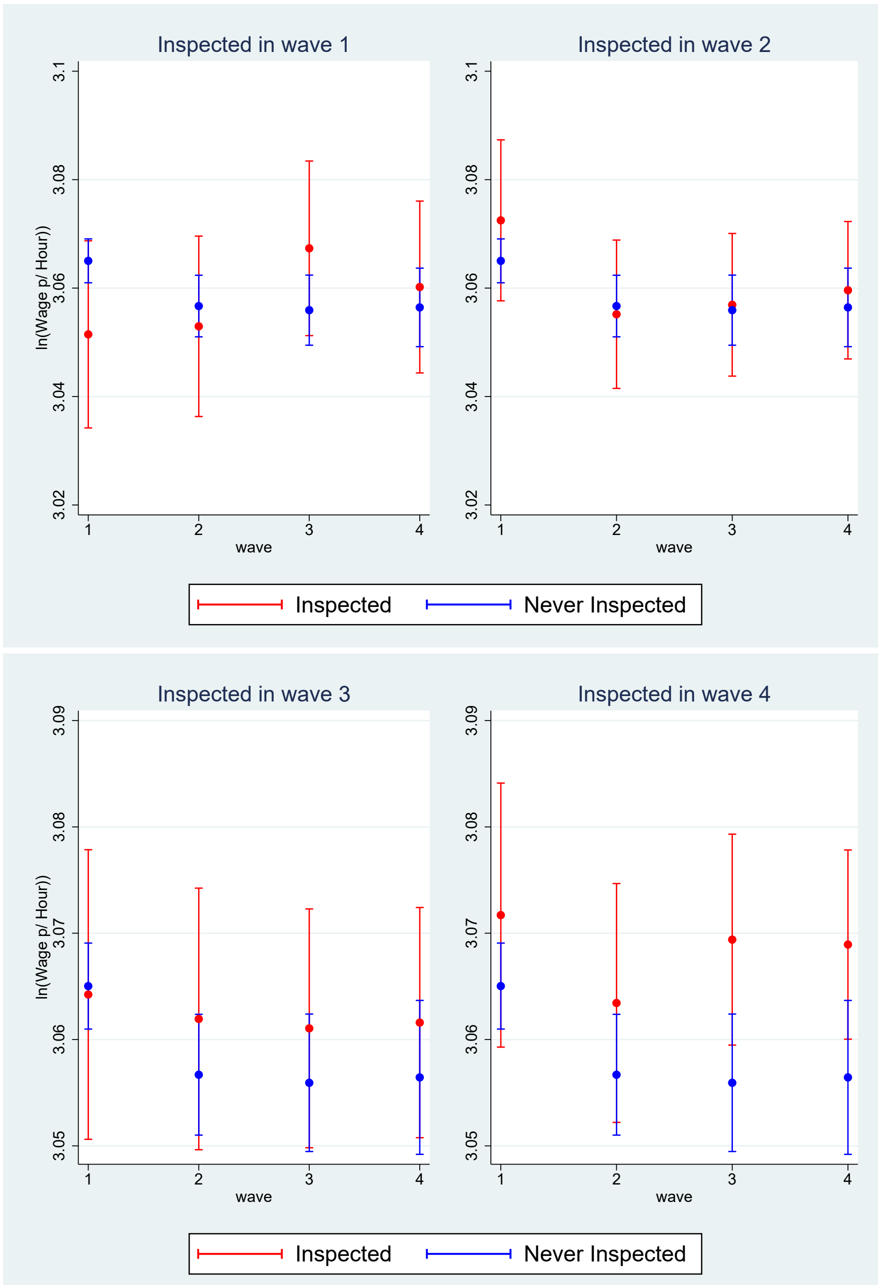
We calculate the predicted probability of "organic" within-firm informal to formal job transitions using a log-linear probability model with age and tenure groups, an interaction of these two variables, as regressors, We also control for workers' years of education, firm size as reported by the worker, firm and year fixed effects. The figure shows the marginal transition probability for each worker age and tenure groups. The sample includes all individuals, between the ages of 18 and 60, who are informally employed at a formal firm for at least one of the waves in ENOE 2005 to June 2016. We exclude domestic workers and agriculture.

Figure 2: Informal Workers' Transition Probabilities by No. of Quarters since Inspection



Notes to Figure 2 Panels A to D: These figures display the effect of inspections on the probability of transitioning out of informal employment on the treated (inspected) and control groups (not inspected) by number of quarters until and since an inspection occurs. $q = 0$ indicates the quarter of inspection. The sample includes individuals informally employed at work-sites included in the DNE between 2005-2015. For each value of q, the treatment group includes all informal workers employed at an establishment inspected q quarters ago. The control group includes informal workers at establishments in the DNE that were not inspected within a [-3,3] quarter window. The solid red line shows the quarterly transition rate dependent mean. The effect on the treated is calculated from the β_q coefficients in the system of equations, (1). Control variables include worker age, gender, education, household size, number of children in day-care age, occupation and industry dummies, i's tenure with employer j, and firm size (as reported by individual i in period t). Errors are clustered at the work-site level.

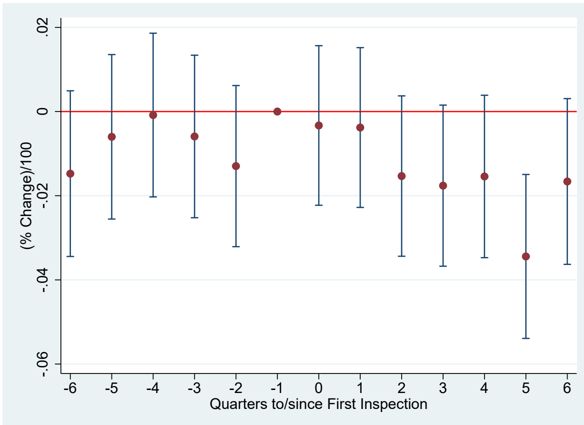
Figure 3: Average Dynamic Inspection Effect on Hourly Wage for Informal Workers



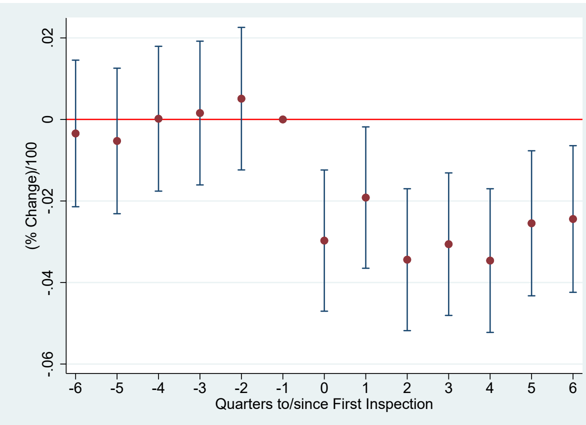
Notes to Figure 3: This figure shows the average treatment effect of inspections on wages for informal employees by timing of treatment using the log linear regression model specified in equation 2. Treatment timing, and hence the number of observed pre and post periods, varies based on the timing of the inspection relative to the quarter when the worker entered ENOE's survey.

Figure 4: First Inspection Effect on Formal Job Creation and Job Destruction

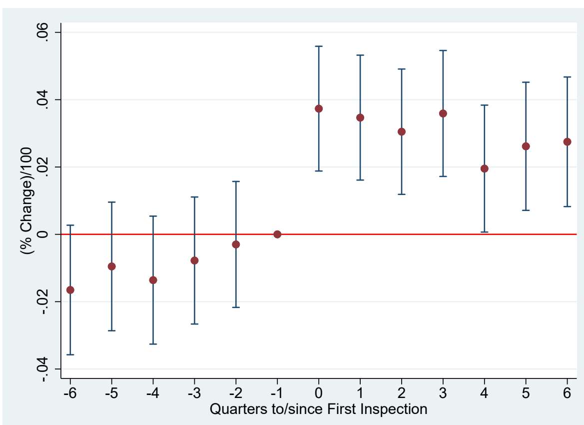
(a) Job Creation from Outside the Formal Sector ("Formalization")



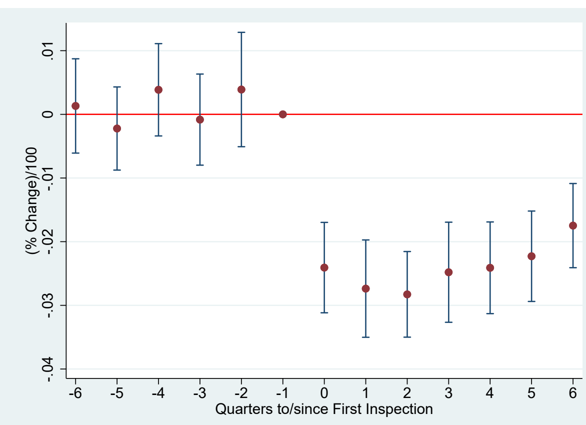
(b) Job Creation from Within the Formal Sector ("Formal Poaching")



(c) Total Job Destruction



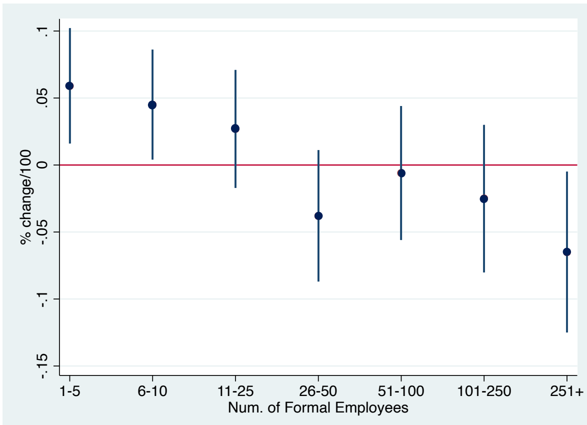
(d) Quarterly Employment Growth



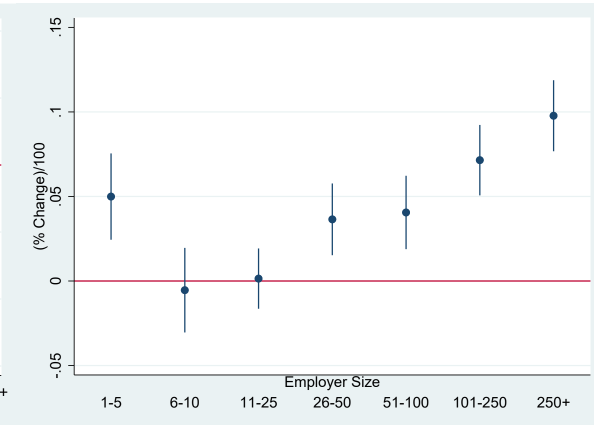
Notes to Figure 4: These figures display the effect of a firm's first inspection on its establishments' formal job creation from outside of formality (Panel (a)) and from within formality (i.e. formal workers poached from other formal firms) (Panel (b)). Formal job destruction (Panel (c)) includes both separations where the worker remains in the formal sector with a different employer, as well as transitions to informality and non-employment. Quarterly formal employment growth (Panel (d)) is equal to the change in the number of formal workers divided by the average formal employment in the previous and current quarters. $q = 0$ indicates the quarter of inspection. The sample includes formal establishments' matched in the DNE and IMSS administrative records between 2005-2016. The effect on the treated is calculated from the $\alpha^{1,q}$ coefficients in equation 5. We control for timeXindustry and firm fixed-effects. Errors are clustered at the firm level.

Figure 5: Effect of Firm's First Inspection on "Formalizations" by Firm Size

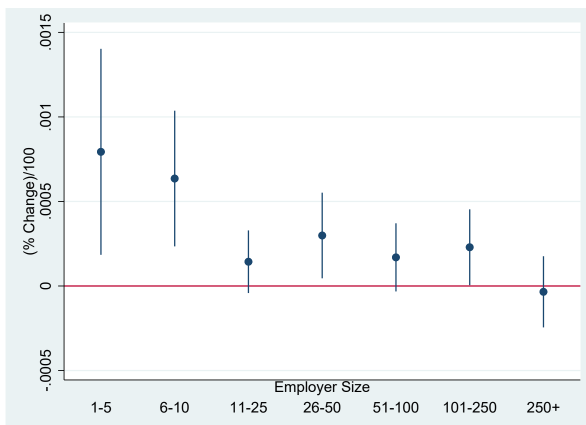
(a) Job Creation from Outside the Formal Sector ("Formalization")



(b) Total Job Destruction



(c) Firm Formal Sector Exit



Notes to Figure 5: These figures display the effect of a firm's first inspection on establishments' inflows of workers who were not employed at a different formal firm in the previous 6 months (panel (a)), total formal job destruction (panel (b)) and formal sector exit (panel (c)). The sample includes formal employers matched in the DNE and IMSS administrative records between 2005-2016. Firm size is measured as the average number of formal workers during its first year of activity in the formal sector. We include timeXindustry and firm fixed-effects, and cluster errors at the firm level.

Figure 6: Timing

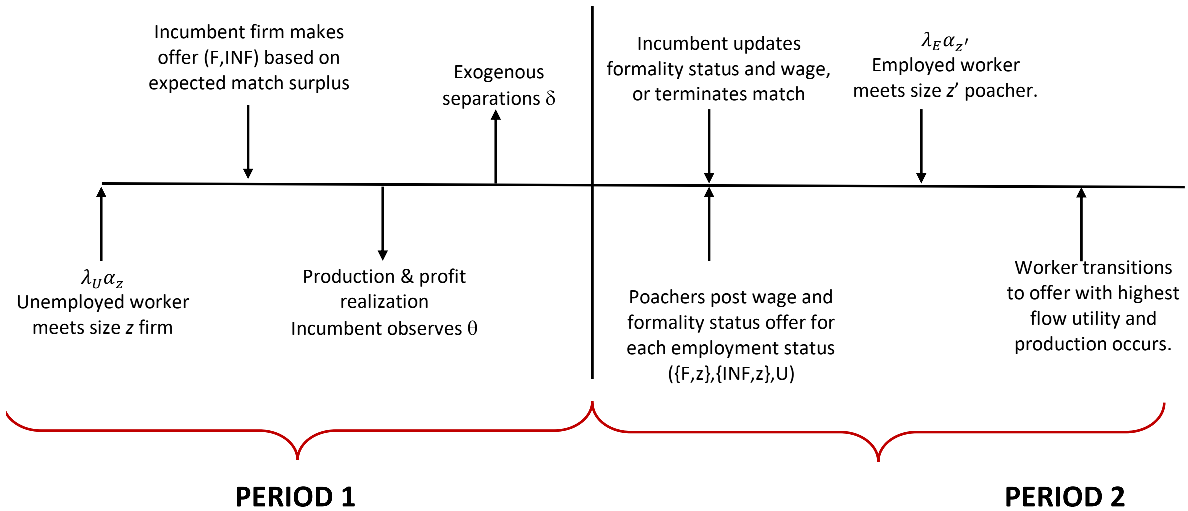
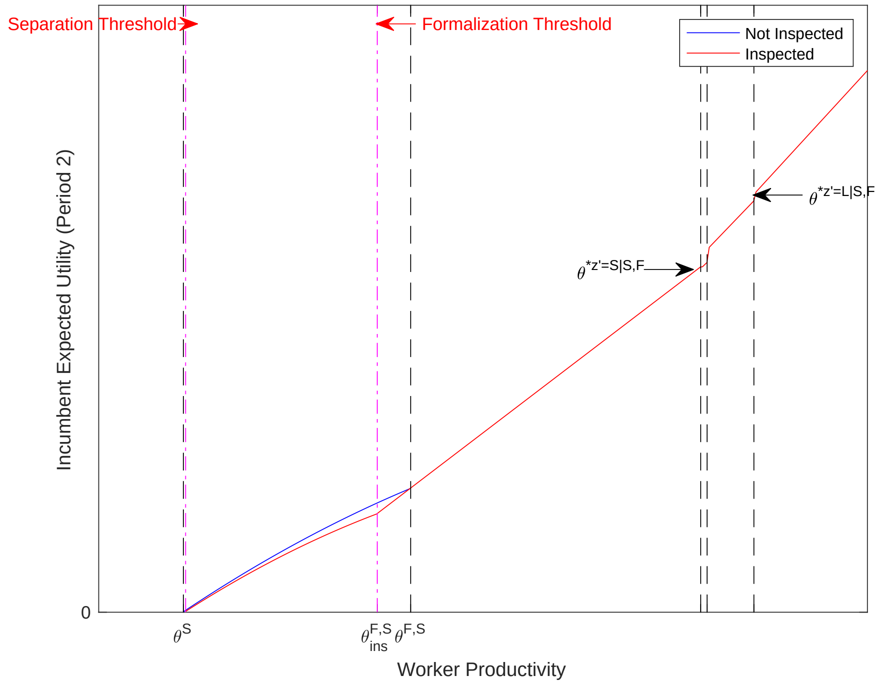


Figure 7: Increasing the Cost of Informal Workers for Incumbents: Effects on Hiring, Layoffs, Formalization, and Wages

(a) Small Incumbent. Hiring, Formalization, and Separation Thresholds



(b) Small Incumbent, Worker Wage Trajectory Trajectories

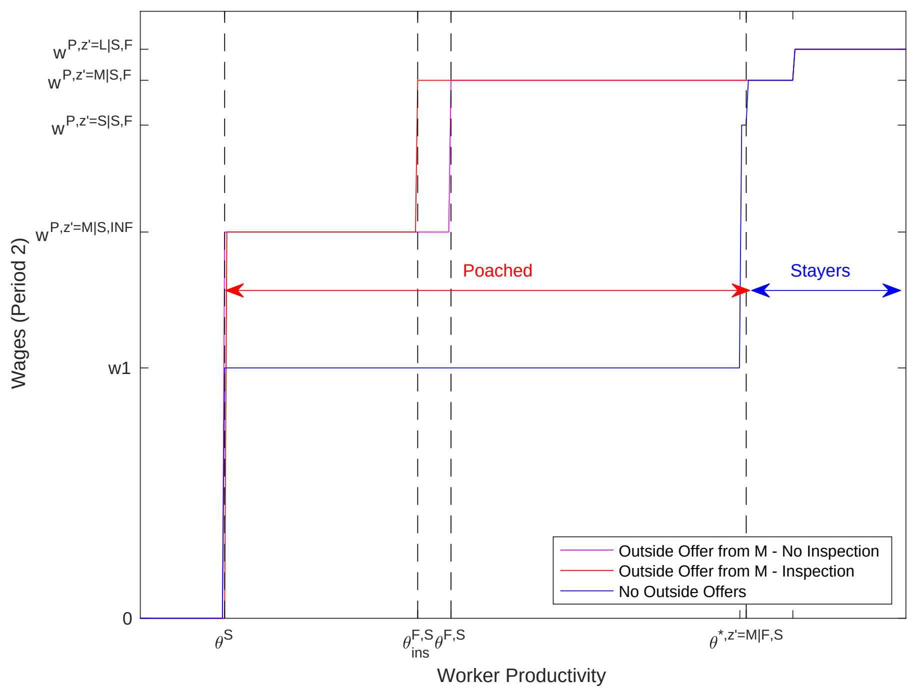
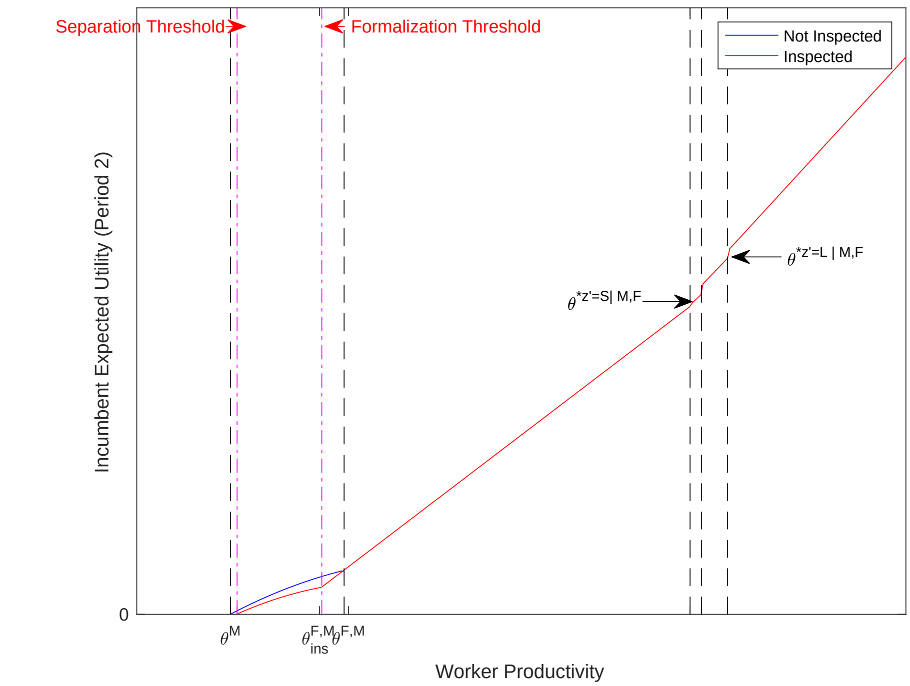


Figure 7: Continued

(c) Medium Incumbent. Hiring, Formalization, and Separation Thresholds



(d) Medium Incumbent. Worker Wage Trajectory

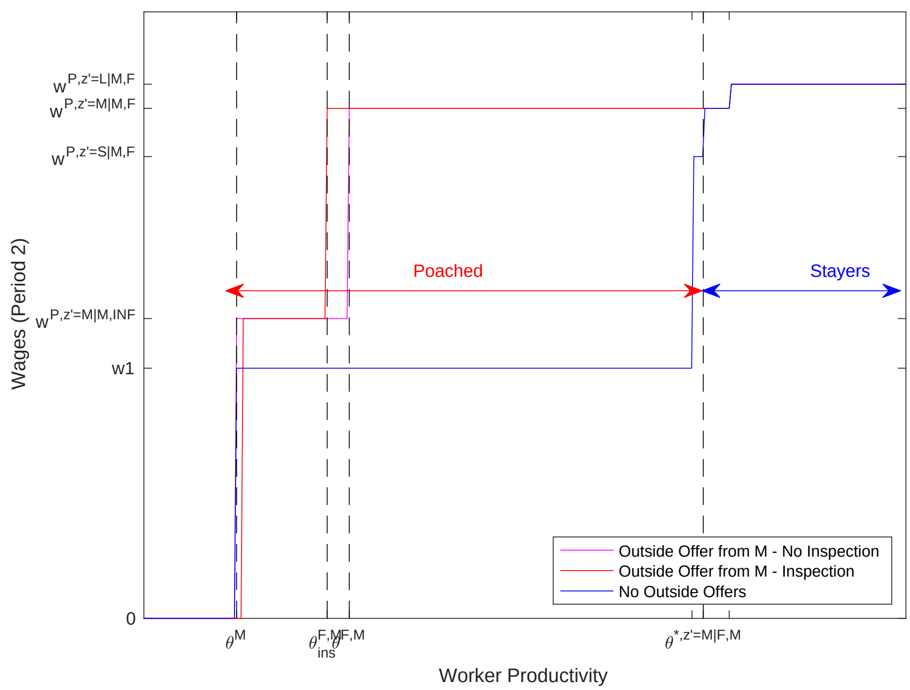
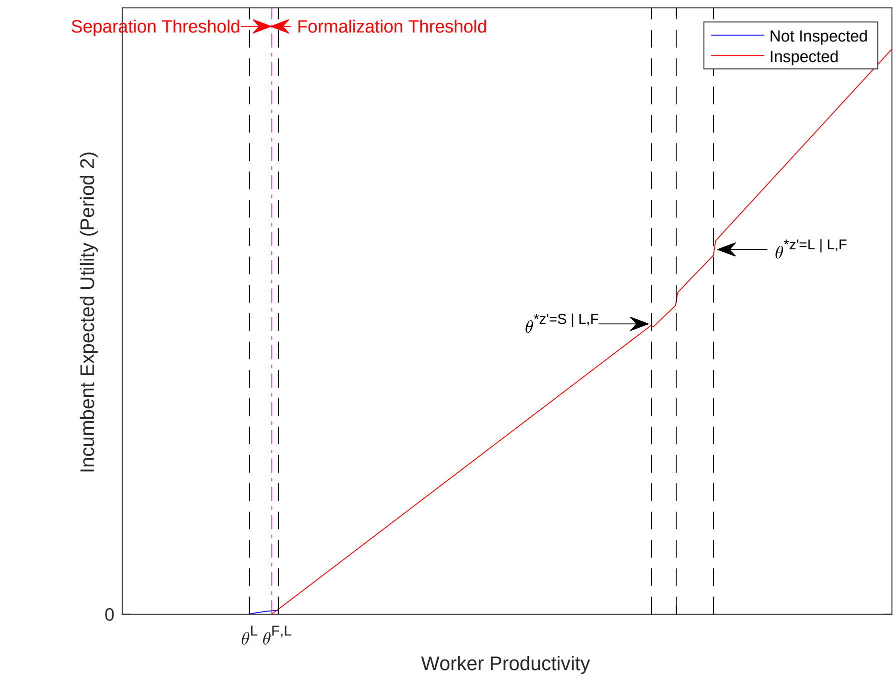
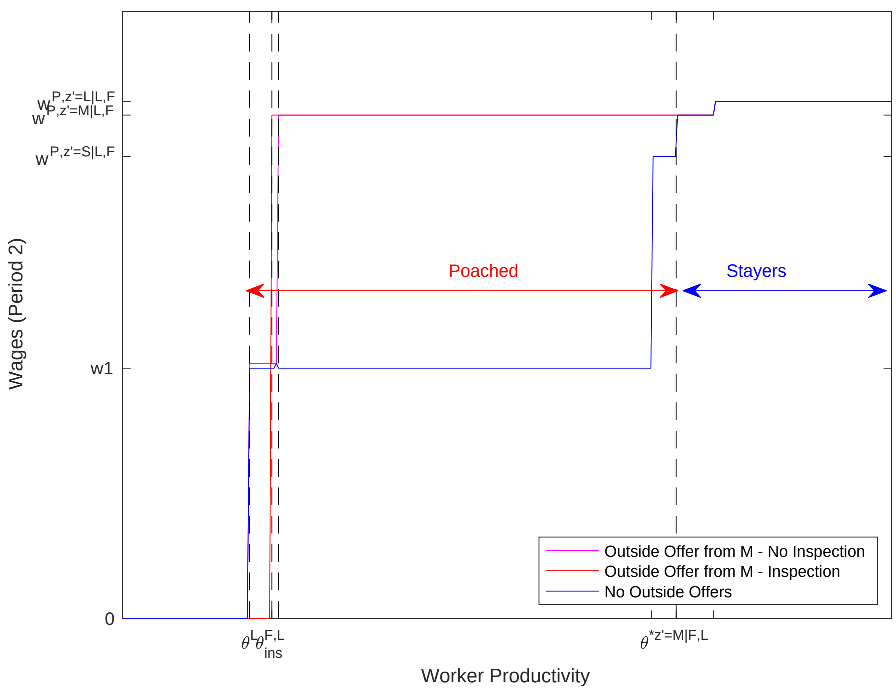


Figure 7: Continued

(e) Large Incumbent. Hiring, Formalization, and Separation Thresholds



(f) Large Incumbent. Worker Wage Trajectory



APPENDIX: FOR ONLINE PUBLICATION ONLY

A Identifying Informal Workers at Formal Firms in ENOE

We follow the 17th International Conference of Labor Statisticians[29] resolution for measuring informality. According to this Resolution, informality has two dimensions. The first dimension refers to employers' characteristics: an employer is categorized as informal when it is not a registered business with tax authorities. The second dimension refers to job characteristics. Informal jobs are those that lack the benefits and institutional protection required by the legal framework in the country. Using this definition of informality, an employee can have either a formal or an informal job at a formal firm depending on whether the employer registers the worker with IMSS or not. All jobs are informal at informal firms in Mexico because only formal employers can register their employees.

The first step to identify informal workers at formal firms is to determine which workers in ENOEs household survey are employed at formal firms.[52] INEGI, and previous research[53], uses data on firms' size and industry to determine whether a firm is formal or not. This classification strategy relies on the assumption that larger firms are more likely to be detected by authorities and hence have a higher risk of being informal. Similarly, it assumes that firms in certain industries have more incentives to register with authorities because they either require a larger scale to operate or are more likely to benefit from participating in production networks that require issuing tax deductible sale receipts which are only available to firms registered with the government.

We depart from this strategy to identify formal firms and instead rely on the Ministry of Labor's Firm Directory (DNE). The DNE is a subset of formal firms in the economy. However, our methodology exploits the Ministry's inspections and only firms in the DNE are selected for random inspections. Moreover, we argue that there are several benefits to identifying firms registered with the government directly, instead of inferring formality status from size. First, households' reporting of employers' size might be inaccurate. Using size thresholds to identify formal firms can therefore be problematic. Second, Hsieh and Olken (2014) find no evidence in the distribution of firms in Mexico to support size-based sorting into formality. Third, more than half of all employers registered at IMSS have between 2 to

[52] All employers and workers in IMSS data are by definition formal.

[53] See, for example, Maloney (1999), Fiess, Fugazza and Maloney (2010), and Alcaraz, Chiquiar and Salcedo (2015) among others.

5 employees. Using size to classify employers could therefore lead to misclassifying a large share of registered employers as informal firms. Fourth, since formal firms can hire workers off-the-books and tax authorities do not share information with IMSS, it is not clear whether the relevant measure to determine risk of getting caught is related to aggregate labor force size, share of non-registered workers or a combination of both.

For analyses based on IMSS administrative data, the baseline sample is all workers who were ever formally employed at a DNE firm. Our baseline sample in ENOE is all individuals employed in a firm included in the DNE for at least one of ENOE's waves. After identifying formal firms, we classify jobs into formal and informal jobs. For this, we follow INEGI's categories which are based on workers' reported access to mandated employer-provided social security benefits. In ENOE's data, transitions across formal and informal jobs refer to changes in workers' reported access to these benefits. In IMSS data, we do not observe these transitions directly. However, for all formal workers in IMSS data, we can observe whether the source and destination of transitions into and out of formal employment are other formal jobs or not.

When using self-reported data on access to social benefits to determine workers' formality status, time-varying misreporting and misclassification can lead us to overestimate transitions rates across formal and informal jobs. While these errors may cancel in aggregate, stock variables, the estimated flow rates between labor market states may be very sensitive to these spurious transitions. Poterba and Summers (1986) This concern might be heightened if individuals' incentives to misreport their access to social benefits is correlated with the timing of inspections.

To address this concern, we first point out that if inspections were only changing reporting behavior but not actual access to social security benefits then we would not see any changes in formal jobs in IMSS administrative data. Second, we implement a conservative correction to transitions. We identify sequential back and forth changes in reported access to social security benefits with the same employer and re-code them as misclassifications. If a worker switches between formal and informal status more than once within a three quarter period with the same employer, we consider the "true" formality status as the job in which the worker spent most time with the employer during the 5-quarter period that ENOE tracks the worker. This correction has a negligible effect on the stocks of informal and formal jobs within formal firms. However, it reduces the rate of transitions from formal to informal jobs and vice-versa by 2.5 p.p. and 2.0 p.p., respectively.[54]

[54] If establishments register their workers after an inspection to avoid being detected in a follow-up visit by

B Merging Datasets

B.1 ENOE and DNE

The National Employment and Occupation Survey (ENOE) interviews 120,260 households every quarter starting in 2005. Among other questions regarding labor market participation, it asks every household member who is employed or involved in any income generating activity the name of the firm, business or institution of employment. ENOE also includes a battery of questions regarding the type of activities performed and goods or services provided by the firm. The Mexican National Institute of Statistics and Geography (INEGI) then uses the answers provided to these questions to classify the firm into one of 178 NAICS industry codes.

DNE is a list of firms' establishments. Each establishment is identified by the firms' "official name" (*razon social*), the establishments' exact address, and for 83% of firms in the directory we also observe the firm's tax ID (*Registro Federal de Contribuyentes*). Meanwhile, in ENOE, workers self-report the name of their employer. Since the dwelling is the unit of observation in ENOE, the survey also includes information on the household's location, but not that of their place of work.[55] These differences generate two challenges when merging DNE and ENOE. First, due to spelling mistakes, abbreviations, and incomplete name reporting by the workers surveyed in ENOE, the name provided by the worker seldom is an exact match to the official name registered by the firm with STPS. Second, if a firm has more than one establishment in the workers' reported location, we need to make a decision about which establishment to match with the worker.

To match ENOE with the DNE and inspections logs, we first perfom basic name cleaning to standardize workers' reported firm names. This includes removing all punctuation, spacing and accents, eliminating articles, spelling out numbers, and replacing common abbreviations and plural forms. We then want to compare firm names in ENOE and find the closest match in the DNE. We define the closest match using a combination of a soundex algorithm and a Levenshtein distance.

IMSS but then un-register them after the verification takes place, observed informal-formal-informal transitions would not be misclassifications but rather real transitions. However, employers have incentives to avoid this "hiding" practice. Registering and unregistering workers within short periods of time can raise flags with authorities making establishments targets of directed inspection visits.

[55]These two pieces of information, the household's location and their members' employers' names, are collected and recorded by ENOE. However, due to confidentiality requirements, they are only available through INEGI's microdata lab for research purposes.

Before implementing our matching algorithm, described in more detail below, we must clean ENOE's names further. In ENOE, employers' names are often reported including the type of establishment or sector in which the firm operates. For example, the answer for a worker employed at a 7-Eleven is at times recorded as "*Autoservicio* 7-Eleven" ("Convenience Store 7-Eleven"), "*Tienda* 7-Eleven" ("Store 7-Eleven"), or with the diminutive "*Tiendita* 7-Eleven". Meanwhile the official name (*Razon Social*) for 7-Eleven, as recorded in the DNE, is "*7-Eleven Mexico, S.A. de C.V.*". In this case, the words "*Tienda*" and "*Autoservicio*" are not actual parts of the firms name so we would like to remove them. However, in other cases, these words are useful to distinguish firms with similar names in different sectors, or are part of the official name. To address this issue, we create a word cloud with the most frequently appearing words in workers' reported employer names. We then reduce these words, and all words with the same root, to the first 5 letters. This procedure reduces the weight given to these words when assessing which name is the closest match in the DNE.

Once we have standardized employers' name in both datasets, we then use a phonetic algorithm, in Spanish, to reduce mismatches from misspelling and typos.[56] Finally, for each employer name reported in ENOE, we identify the closest match in the DNE using the Levenshtein distance.[57] We consider an ENOE-DNE pair to be a match if the Levenshtein similarity ratio is at least 90% and the worker lived in the same state as the firm's location. For multi-establishment firms, we assume the effect of inspections occurs at the firm-state level. This assumption implies that a worker in state X employed at a firm that received an inspection at any of the establishments in X is coded as inspected.[58]

[56] The algorithm is our own implementation of Amon, Moreno and Echeverri (2012)

[57] For each employer name in ENOE, we actually identified the top 5 matches to conduct manual checks and robustness analysis.

[58] We could instead assume that inspections affect firms' decisions for all establishments in a given municipality, or that only the inspected establishment is affected. For the later assumption, we merge workers to the establishment that is closest to their dwelling's location. Our current matching methodology biases our results downwards since we consider workers as "treated" by the inspection even if their specific establishment was not the one targeted for the inspection.

Figure B.1: ENOE Firm Names Word Cloud

Notes: This figure shows a censored version of the word cloud used to standardized employers' names between ENOE and DNE. The goal is to decrease the weight of words frequently used by workers to describe their employers, and inaccurately reported as part of the firm's name, when calculating the Levenshtein distance between ENOE and DNE firm name pairs. Some words were censored to avoid revealing employers' identities.

On average, we match 11,000 workers per quarter (out of an average of 42,000 salaried workers at formal firms as defined by ENOE). Table B.1 shows the results of the matched inspected and non-inspected workers by formality status. Table B.3 compares the share of establishments by industry code in ENOE, the DNE and by inspection status. Consistent with random selection of establishments to inspect, conditional on being part of the DNE list, inspections are similarly distributed across establishments in different industries.

Table B.1: Matching by Workers' Inspection and Formality Status

Formal Firms Employees in ENOE's sample (Jan. 2005-Jun. 2016)					
	(1)	(2)	(3)	(4)	(5)
	Matched to DNE				Not Matched to DNE
	Timing of First Inspection Relative to Entering ENOE's sample				
	Before ENOE	During ENOE	After ENOE	Not Inspected	
Informal	16,611	2,339	39,040	48,288	284,522
Formal	57,405	8,612	139,988	126,294	944,459

Notes: This table shows the matching rate between ENOE and DNE. It includes all individuals who for at least one of the survey's waves is a wage-earning employee at a formal firm. Columns (1) through (4) ("Matched to DNE") present individuals who for at least one of ENOE's waves are employed, informally or formally, at a firm included in the DNE. Column (5) ("not matched to DNE") refers to individuals employed at establishments that we could not match to the DNE, but that are classified as formal firms based INEGI's definition.[59] Column (5) includes individuals who were unable or unwilling to provide the name of their employer. We exclude individuals employed in the agriculture sector and domestic workers during any of the waves, as well as those younger than 15 or older than 80 on the first survey wave. Individuals are classified as formal or informal based on their formality status when first observed working at a formal firm in ENOE.

Table B.2: Employees at ENOE-DNE Matched Establishments by Formality Status 2005-2016

	Informal Employees	Formal Employees
Median After-Tax Wage (2014 Pesos p/Hr.)	\$18.2	\$26.0
Median No. Hours (Weekly)	46	47
Median Tenure (Months)	24	59
% aged 15-24	34	15
% Female	39	39
% Completed 9^{th} grade	72	85
No. Workers:	106,278	332,299

Source: Own calculations based on data from the National Employment and Occupation Survey (ENOE)[28] and the National Directory of Firms (DNE)[40] Individuals are classified as formal or informal based on their formality status when first observed at the DNE-matched establishment in ENOE.

[59]See INEGI's microdata website[28] for further information on INEGI's methodology to classify formal and informal firms.

Table B.3: ENOE Workers by Industry

NAICS Sector	% of employed individuals in ENOE		
	(1)	(2)	(3)
	At formal firms in ENOE	At firms matched to DNE	At inspected firms matched to DNE
Mining & oil	0.78	1.7	1.94
Electricity	0.66	2.07	2.72
Construction	5.75	8	8.05
Manufacturing	18.23	18.76	18.12
Wholesale	3.81	6.53	6.58
Retail	18.28	16.01	15
Transport & warehouses	4.17	8.13	9.18
Information	1.34	0.64	0.56
Insurance & financial serv.	1.43	2.52	2.63
Real Estate	0.93	1	0.97
Prof. & tech. services	3.47	1.1	0.96
Corporations	0.1	0	0
Admin. & business support serv.	3.44	0.36	0.29
Education	8.53	1.76	1.4
Health & social services	4.92	1.5	1.25
Entertainment & rec. serv.	1.08	1.39	1.35
Lodging & food services	8	13.2	13.35
Other non-public services	5.57	10.92	11.74
Government & NGO's	8.6	4	3.54
NEC	0.92	0.44	0.37
Total	100%	100%	100%
No. of Workers	1,523,587	438,577	263,995

Notes for Table B.3: This table shows the distribution of workers by sector using 2-digit NAICS. Column (1) includes all individuals in ENOE, between the ages of 15 and 80, employed at formal firms, based on INEGI's definition of formal "units of production". Column (2) shows the distribution of workers employed for at least one of ENOE's waves at a firm that could be matched to the DNE. The last column shows the industry distribution of employees at inspected firms.

B.2 IMSS and DNE

IMSS employer-employee administrative data and the DNE share some variables that allow us to identify a firms' establishments in the two datasets: the firm name (*"Razon Social"*) and its tax ID (*RFC*). Due to confidentiality restrictions, we were not allowed to work directly with the non-anonymized data. Instead, that staff at Banxico's EconLab helped us to develop and implement a data cleaning and name matching algorithm to matched IMSS administrative records with the DNE. The EconLab staff is exceptionally well suited for this task as they are extremely familiar with the data and highly skilled in working with big data.

We next describe the steps the EconLab staff implemented to merge these two datasets: 1) tax ID matches, 2) name cleaning and homogeneization, and 3) (direct, phonetic, and closest distance) name matching. We start by matching establishments for which we do have tax ID information. If the tax ID matches perfectly, then we consider these employers as matched and keep them in the sample. For the remaining observations in IMSS data, we follow a similar process than the one used to match ENOE and DNE. First, we clean firms' names in both datasets removing acronyms like Corp., Ltd., Inc., etc. We homogenize capitalization and remove accents. Then we identify the employers in IMSS data that have an identical firm name (letter-by-letter match) as a firm in the DNE. We consider these as matches and continue with the rest of the non-matched employers.

We then perform a soundex algorithm (in Spanish) converting each firm name in the set of unmatched IMSS employers and in the full set of DNE firms to its phonetic equivalent. We then compare the two datasets and, for each firm in the IMSS data, we find the closest phonetic match in the DNE. If the distance between a firm in IMSS data and its closest phonetic match in the DNE is such that the probability of a true match is at least 95%, we consider it a match. Finally, for the rest of the unmatched employers in IMSS data, we calculate how different their firm's name is to each of the firm's names in the DNE using, first, the Levenshtein distance and, second, Jaro Winkler similarity measure. If the nearest match is at least a 95% match then we consider it a match.

Table B.4 describes the matched sample and compares it to the universe of firms in IMSS data.

Table B.4: IMSS Employers Matched to DNE Firm: Descriptive Statistics 2005-2016

PANEL A: IMSS Employers by Industry (2016)			
	Matched IMSS-DNE		IMSS
	(Inspected)	(All)	Universe
Manufacturing	27.00	19.52	13.42
Construction	6.44	8.02	14.84
Prof. and Bus. Services	16.96	19.31	27.12
Other Services	41.62	43.72	34.28
Other NEI	7.98	9.43	10.34
No. of Employers:	14,386	24,251	1,051,210
PANEL B: DNE-IMSS Matched Employer Characteristics			
	Mean	Median	Std. Dev.
Size (No. of Workers)	121.4	30.9	432.2
Starting Daily Wage (2018 $MXN)	$228.0	$165.6	$224.9
Employer Age (quarters)	41.1	47.0	10.5
% Female Workers	40	30	30
No. of Inspections	2	1	3

Source: Own calculations based on data from IMSS administrative records accessed through Banxico's Econ-Lab Convenio No. 45, the National Directory of Firms (DNE)[40] and STPS Inspection Logs.

Notes: Employer age is measured as the number of months elapsed from the first time an employer ID registers a worker. "Others NEI" refers to all other industries not explicitly listed including government entities, transportation, oil and mining.

C Verifying Random Selection in Inspections

As shown in table C.1, in most cases, inspections are closed without there being any reported violations for the items within STPS's enforcement responsibility. Only 10% of all inspections lead to a fine. Between 2005 and 2016, the average fine was MXN$32,194 (USD$1,740) with a maximum fine of MXN$82,569,000 (USD$4,463,189)[60] and a minimum of MXN$20.57 (USD$1.11).

[60]This fine was due to health and hygiene violations in 2013.

Table C.1: Distribution of STPS's Inspections by Result (2005-2016)

Result		No. of Inspections	% of All Inspections
Closed without report of violations		266,517	43%
Provided proof of compliance		296,367	48%
Request for time extension granted		184	0%
Fining process started	Fine imposed	23,154	4%
	Fine no yet imposed	34,620	6%

Source: Own calculations using STPS DNE and Inspections logs 2005-2016.[40]
Notes: Excludes violations beyond STPS's jurisdiction, including those related to informal employment.

Figure C.1: Yearly Inspection Probability and Distribution of Establishments by Number of Inspections Received in a Year (Conditional on being Inspected)

Source: Own calculations using the National Firm Directory (STPS) and Inspections logs 2005-2016. (INAI)[40]

The annual probability of inspections increased starting in 2012. Between 2005 to 2011, establishments' annual inspection probability was 9%. In 2013 the likelihood of being inspected increased to 17% and by 2015 it was 29%. Figure C.1 shows that the probability of receiving more than two inspections in a given year also increased.

STPS must choose establishments in the DNE at random for ordinary inspections.[48,49] Section 4, where we presented our empirical findings, shows similar trends in the outcomes of interest for inspected and non-inspected establishments in the pre-period. In this section, we conduct additional tests to assess whether establishments are randomly selected.

Table C.2: Haphazard Sampling Test 2015: Alphabetical Order

First Letter in Name	No. of Establishments	% in DNE	No. of Inspections	% of Inspections	test-stat	p_value
A	44,503	10.73%	8,941	10.55%	-1.22	0.222
B	15,690	3.78%	3,026	3.57%	-1.39	0.163
C	56,920	13.73%	11,687	13.80%	0.47	0.641
D	14,563	3.51%	2,965	3.50%	-0.08	0.935
E	18,728	4.51%	3,644	4.30%	-1.42	0.156
F	15,780	3.80%	3,126	3.69%	-0.76	0.447
G	27,110	6.54%	5,751	6.79%	1.67	0.096
H	11,755	2.83%	2,328	2.74%	-0.57	0.569
I	21,143	5.10%	3,953	4.66%	-2.86	0.004
J	5,358	1.29%	1,214	1.43%	0.91	0.362
K	2,880	.69%	517	.61%	-0.55	0.586
L	12,105	2.92%	2,353	2.77%	-0.93	0.353
M	27,912	6.73%	5,848	6.90%	1.14	0.254
N	7,065	1.70%	1,606	1.89%	1.25	0.213
O	12,046	2.90%	2,460	2.90%	-0.01	0.992
OTHER	270	.06%	53	.06%	-0.02	0.987
P	31,598	7.62%	6,362	7.51%	-0.74	0.456
Q	2,308	.55%	406	.47%	-0.50	0.617
R	14,245	3.43%	3,111	3.67%	1.55	0.121
S	36,682	8.85%	8,092	9.55%	4.75	0.000
T	21,710	5.23%	4,139	4.88%	-2.32	0.020
U	2,843	.68%	604	.71%	0.18	0.860
V	6,529	1.57%	1,473	1.73%	1.07	0.287
W	1,625	.39%	331	.39%	-0.01	0.994
X	453	.10%	113	.13%	0.16	0.876
Y	1,157	.27%	233	.27%	-0.03	0.979
Z	1,482	.35%	347	.40%	0.34	0.736
Total	414,460	100%	84,683	100%		

Source: Own calculations using STPS DNE and Inspections logs 2015.[40]
Notes: The test-stat column is the test statistic from a χ^2 test of proportions. Under random assignment, the share of establishments in the DNE with names starting with the corresponding row's letter equals the expected share of inspected establishments starting with that letter. The null hypothesis of random selection can be rejected for significance levels greater than the value indicated in the p-value column.
"Other" indicates non-alphabetic characters.

Table C.3: Haphazard Sampling Test 2015: Zip Code Order

First Zip Code Digits	No. of Establishments	% in DNE	No. of Inspections	% of Inspections	test-stat	p-value
10	2,175	0.52 %	385	0.45 %	-0.45	0.651
11	11,155	2.69 %	2,062	2.43 %	-1.67	0.094
12	943	0.22 %	195	0.23 %	0.02	0.986
13	1,592	0.38 %	265	0.31 %	-0.46	0.646
14	3,746	0.90 %	523	0.61 %	-1.85	0.064
15	5,010	1.20 %	779	0.91 %	-1.87	0.061
16	1,440	0.34 %	209	0.24 %	-0.65	0.516
17	423	0.10 %	95	0.11 %	0.07	0.948
18	22	0.00 %	3	0.00 %	-0.01	0.991
19	227	0.05 %	62	0.07 %	0.12	0.905
20	7,858	1.89 %	1,371	1.61 %	-1.80	0.072
21	6,046	1.45 %	977	1.15 %	-1.98	0.048
22	4,901	1.18 %	1,242	1.46 %	1.84	0.066
23	8,285	1.99 %	1,328	1.56 %	-2.80	0.005
24	6,754	1.62 %	901	1.06 %	-3.67	0.000
25	6,061	1.46 %	1,404	1.65 %	1.27	0.205
26	5,146	1.24 %	1,286	1.51 %	1.79	0.073
27	3,961	0.95 %	749	0.88 %	-0.46	0.645
28	5,284	1.27 %	1,024	1.20 %	-0.43	0.670
29	5,155	1.24 %	1,356	1.60 %	2.32	0.021
30	2,910	0.70 %	674	0.79 %	0.61	0.545
31	9,386	2.26 %	1,751	2.06 %	-1.28	0.200
32	3,909	0.94 %	608	0.71 %	-1.46	0.145
33	4,906	1.18 %	785	0.92 %	-1.66	0.096
34	6,383	1.54 %	1,715	2.02 %	3.15	0.002
35	2,503	0.60 %	348	0.41 %	-1.25	0.213
36	6,309	1.52 %	1,631	1.92 %	2.62	0.009
37	5,166	1.24 %	1,199	1.41 %	1.10	0.272
38	3,678	0.88 %	809	0.95 %	0.44	0.661
39	6,543	1.57 %	1,197	1.41 %	-1.07	0.284
40	2,154	0.51 %	386	0.45 %	-0.41	0.680
41	646	0.15 %	103	0.12 %	-0.22	0.825
42	6,221	1.50 %	1,217	1.43 %	-0.41	0.679
43	3,701	0.89 %	600	0.70 %	-1.19	0.233
44	12,023	2.90 %	3,636	4.29 %	9.10	0.000
45	6,592	1.59 %	1,640	1.93 %	2.25	0.025
46	453	0.10 %	82	0.09 %	-0.08	0.936
47	323	0.07 %	55	0.06 %	-0.08	0.933
48	2,632	0.63 %	819	0.96 %	2.14	0.032
49	469	0.11 %	77	0.09 %	-0.14	0.886
50	4,622	1.11 %	944	1.11 %	0.00	0.998
51	1,056	0.25 %	135	0.15 %	-0.61	0.539
52	5,196	1.25 %	1,106	1.30 %	0.34	0.734
53	5,026	1.21 %	846	0.99 %	-1.38	0.166
54	12,792	3.08 %	2,202	2.60 %	-3.18	0.001
55	3,956	0.95 %	761	0.89 %	-0.36	0.718
56	2,259	0.54 %	499	0.58 %	0.29	0.775

Continued on next page

Table C.3 – continued from previous page

First Zip Code Digits	No. of Establishments	% in DNE	No. of Inspections	% of Inspections	test-stat	p_value
57	708	0.17 %	203	0.23 %	0.44	0.657
58	7,866	1.89 %	1,918	2.26 %	2.39	0.017
59	700	0.16 %	143	0.16 %	0.00	1.000
60	5,099	1.23 %	1,078	1.27 %	0.28	0.782
61	1,228	0.29 %	266	0.31 %	0.11	0.909
62	7,435	1.79 %	1,329	1.56 %	-1.46	0.145
63	7,031	1.69 %	1,406	1.66 %	-0.23	0.815
64	7,049	1.70 %	1,389	1.64 %	-0.39	0.694
65	759	0.18 %	162	0.19 %	0.05	0.958
66	6,328	1.52 %	1,087	1.28 %	-1.58	0.115
67	6,098	1.47 %	769	0.90 %	-3.65	0.000
68	6,029	1.45 %	1,106	1.30 %	-0.96	0.335
69	222	0.05 %	36	0.04 %	-0.07	0.943
70	2,207	0.53 %	514	0.60 %	0.48	0.631
71	743	0.17 %	160	0.18 %	0.06	0.950
72	10,967	2.64 %	2,472	2.91 %	1.78	0.075
73	1,398	0.33 %	231	0.27 %	-0.42	0.677
74	1,432	0.34 %	211	0.24 %	-0.62	0.534
75	791	0.19 %	142	0.16 %	-0.15	0.881
76	11,428	2.75 %	2,642	3.11 %	2.37	0.018
77	7,556	1.82 %	1,756	2.07 %	1.63	0.104
78	12,452	3.00 %	2,003	2.36 %	-4.18	0.000
79	1,029	0.24 %	106	0.12 %	-0.79	0.427
80	7,879	1.90 %	2,211	2.61 %	4.61	0.000
81	2,733	0.65 %	406	0.47 %	-1.16	0.245
82	4,417	1.06 %	825	0.97 %	-0.59	0.554
83	7,811	1.88 %	2,028	2.39 %	3.32	0.001
84	3,521	0.84 %	433	0.51 %	-2.19	0.029
85	3,755	0.90 %	709	0.83 %	-0.44	0.657
86	7,398	1.78 %	1,589	1.87 %	0.59	0.553
87	3,001	0.72 %	475	0.56 %	-1.05	0.292
88	4,363	1.05 %	873	1.03 %	-0.14	0.888
89	3,553	0.85 %	867	1.02 %	1.08	0.282
90	10,368	2.50 %	1,830	2.16 %	-2.22	0.026
91	9,198	2.21 %	2,792	3.29 %	7.02	0.000
92	1,019	0.24 %	180	0.21 %	-0.21	0.830
93	3,665	0.88 %	786	0.92 %	0.28	0.777
94	4,052	0.97 %	896	1.05 %	0.52	0.603
95	1,325	0.31 %	226	0.26 %	-0.34	0.733
96	2,926	0.70 %	563	0.66 %	-0.27	0.790
97	8,331	2.01 %	1,471	1.73 %	-1.78	0.076
98	8,171	1.97 %	1,676	1.97 %	0.05	0.960
99	848	0.20 %	251	0.29 %	0.59	0.554
Missing	1,247	0.30 %	366	0.43 %	0.85	0.397
0	326	0.07 %	60	0.07 %	-0.05	0.960
Total	414,460	100 %	84,683	100 %		

Even if establishments are selected randomly, the order in which inspections are carried out need not be random. For example, inspectors may group geographically close establishments within the same month so as to minimize travel across locations. Such geographical clustering of the timing of random inspections, if anticipated by firms, could lead to spillovers to the control group. We evaluate whether the sequence of inspected establishments by location (zip code and geographical coordinates) is consistent with randomly allocated timing of inspections using the *k*-Category extension of the Wald-Wolfowitz Runs Test.[61]

Tables C.4 and C.5 show the results from separate runs tests for the sequence of inspected firms' zip code and coordinates for each year. In 2013, we reject the null hypothesis that the order of inspections is random across geographical coordinates at a 1% significance level. We also reject these null across zip codes between 2013 and 2015. After 2012, when the number of inspections per year more than doubled, there is evidence consistent with sequential inspections of geographically proximate establishments. It is important to note that the null hypothesis here is not whether inspections are assigned randomly across the universe of establishments in the DNE, but rather whether once selected for an inspection the order in which these occur is consistent with random allocation. Table C.3 indicates that inspections are proportionally assigned across zip codes. Clustering the dates of inspections by zip code does not pose a threat to our identification strategy unless establishments can accurately predict whether they are part of the randomly selected set for inspection within their zip code. Otherwise, grouping the dates for inspecting the randomly selected establishments for a location within a certain window of time could cause anticipatory behavior by all establishments (treated and in the control group) in the zip code. This would, if anything, bias our results against finding an effect from inspections.

[61] A runs test examines whether the sequence of repeated symbols (in this case, the first letter of each inspected firm's name or third number of a zipcode, ordered by inspection date) differs from the sequence that would arise if the symbols were selected from a population with an i.i.d. process. If selection is truly random, the probability of observing a very long/short or too few/many repeated sequence of symbols is low. See, for example, Agin and Godbole (1992) and Hall et al. (2012). We follow Sheskin (2004) to estimate the asymptotic distribution of the expected number of runs under random selection. We also check the estimated mean and variance of the expected number of runs using the procedure suggested by Smeeton and Cox (2003).

Table C.4: Random Inspection Timing: Runs Test of Geographical Coordinates

H0: Timing of Inspections is Random							
Year	No. of Inspections	No. of Runs	Expected Runs under H0 (μ_R)	σ_R^2 under H0	σ_R under H0	z-stat	p-value
2005	14,728	8,980	8,947	2,348.41	48.46	0.67	0.504
2006	13,233	7,983	8,015	2,000.83	44.73	-0.71	0.476
2007	13,541	8,202	8,169	2,079.24	45.60	0.72	0.474
2008	15,915	9,617	9,620	2,502.87	50.03	-0.05	0.963
2009	22,791	15,149	15,083	3,924.31	62.64	1.04	0.297
2010	26,293	17,285	17,264	4,639.77	68.12	0.29	0.768
2011	31,086	20,400	20,435	5,258.85	72.52	-0.47	0.639
2012	51,317	33,337	33,392	8,703.45	93.29	-0.58	0.560
2013	86,287	55,575	55,907	14,624.14	120.93	-2.74	0.006
2014	98,198	64,434	64,381	17,687.52	132.99	0.39	0.693
2015	84,683	55,800	56,079	15,146.68	123.07	-2.26	0.024
2016	29,046	19,034	19,051	5,009.50	70.78	-0.23	0.818

Source: Own calculations using STPS DNE and Inspections logs 2015.[40]

Notes: The z-stat column is the test statistic from a k-category Wald-Wolfowitz Runs Test for Randomness. Each category is the pair of the first digit of the establishment's geographical coordinates (latitude and longitude). Under random assignment, the expected number of runs follows a normal distribution with mean μ_R and variance σ_R^2. The null hypothesis of random selection can be rejected for significance levels greater than the value indicated in the p-value column.

Table C.5: Random Inspection Timing: Runs Test of Zip Codes

H0: Timing of Inspections is Random							
Year	No. of Inspections	No. of Runs	Expected Runs under H0 (μ_R)	σ_R^2 under H0	σ_R under H0	z-stat	p-value
2005	14,723	14,442	14,440	273.49	16.54	0.09	0.927
2006	13,230	12,930	12,963	256.52	16.02	-2.03	0.043
2007	13,538	13,225	13,270	256.51	16.02	-2.80	0.005
2008	15,909	15,588	15,609	288.14	16.97	-1.21	0.228
2009	22,767	22,346	22,366	390.30	19.76	-0.97	0.333
2010	26,264	25,788	25,798	452.77	21.28	-0.47	0.641
2011	31,030	30,509	30,507	510.63	22.60	0.08	0.939
2012	51,137	50,179	50,236	877.50	29.62	-1.90	0.058
2013	86,067	84,389	84,594	1,436.81	37.91	-5.38	0.000
2014	97,684	95,853	95,978	1,659.92	40.74	-3.06	0.002
2015	84,317	82,680	82,807	1,468.48	38.32	-3.29	0.001
2016	28,936	28,383	28,401	520.09	22.81	-0.77	0.442

Source: Own calculations using STPS DNE and Inspections logs 2015.[40]

Notes: The z-stat column is the test statistic from a k-category Wald-Wolfowitz Runs Test for Randomness. Each category is the first two digits of the establishment's zipcode. Under random assignment, the expected number of runs follows a normal distribution with mean μ_R and variance σ_R^2. The null hypothesis of random selection can be rejected for significance levels greater than the value indicated in the p-value column.

Table C.6: Testing for Random Probability of Inspections: Worker-level

	Sample: Informal Workers at Firms in the DNE					
	(1)		(2)		(3)	
	Pr(Inspected)		Pr(Formalized)		Pr(Separated to Unemp.)	
	$Z_{i,j,t}$		$IF_{i,j,t}$		$IU_{i,j,t}$	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Age	-0.001**	(0.001)	0.002	(0.002)	-.005**	(0.002)
Male	0.005	(0.013)	0.068*	(0.041)	-0.229***	(0.057)
Tenure (months)	0.001	(0.001)	0.008**	(0.003)	-0.014***	(0.002)
Education						
Less than HS	0.068	(0.066)	0.261*	(0.149)	0.287	(0.241)
Completed HS	0.043	(0.067)	0.522***	(0.154)	0.559**	(0.247)
Some College	0.075	(0.069)	0.619***	(0.167)	0.650**	(0.264)
College +	0.074	(0.068)	0.642***	(0.156)	0.554**	(0.250)
Children in daycare	0.000	(0.005)	0.050***	(0.015)	-0.784***	(0.270)
Household size	0.003	(0.006)	0.018	(0.017)	-0.008	(0.014)
Establishment size						
6-10	-0.046	(0.025)	0.372***	(0.015)	0.259***	(0.087)
11-15	0.014	(0.029)	0.516***	(0.073)	0.554***	(0.113)
16-50	0.002	(0.021)	0.727***	(0.055)	0.801***	(0.084)
51+	-0.003	(0.019)	0.892***	(0.052)	1.354***	(0.075)
LR	30.44		1,194.36		649.23	
p-value	0.595		0.000		0.000	
No. of Observations: 253,349						

Notes: This table presents evidence of random selection in STPS ordinary inspections. The baseline estimation sample is individuals who are informally employed at an establishment that is included in the DNE between 2005 to 2015. The dependent variables in columns 1, 2 and 3 are, respectively, $Z_{i,t}$ a dummy equal to 1 if individual i was informally employed at an establishment in period t, and 0 if the establishment is included in the DNE but was not subject to an inspection within this time frame; $IF_{i,j,t}$ a dummy equal to 1 if individual i was informally employed at establishment j in quarter t and was then formally employed at the same establishment in $t+1$ and 0 otherwise; and $IU_{i,j,t}$ a dummy equal to 1 if individual i was informally employed at establishment j in quarter t and became unemployed between quarters t and $t+1$ and 0 otherwise. Tenure is the numbers of months employed at the current establishment. Children in daycare are children within the ages of 0 to 4 in the household. All regressions include firm sector, worker occupation, and year fixed effects.