Labor Market Search, Informality, and On-The-Job Human Capital Accumulation

Matteo Bobba* Luca Flabbi† Santiago Levy‡ Mauricio Tejada§

March 16, 2018

Abstract

We develop a search and matching model where firms and workers produce output that depends both on match-specific productivity and on worker-specific human capital. The human capital is accumulated while working but depreciates while searching for a job. Firms optimally post the formality status; wages are determined by bargaining. The equilibrium is characterized by an endogenous steady state distribution of human capital and by an endogenous formality rate. We estimate the model on labor market data from Mexico. We find that human capital accumulation on-the-job is responsible for about one third of the overall value of production. We also find that human capital upgrading is slower while working informally than formally. This magnifies the negative impact on productivity of the labor market institutions that give raise to informality.

Keywords: Labor market frictions, Search and matching, Nash bargaining, Informality, On-the-Job human capital accumulation.

JEL Codes: J24, J3, J64, O17

*Toulouse School of Economics, University of Toulouse Capitole, Manufacture des Tabacs, 21 Allée de Brienne 31015 Toulouse France. E-mail: matteo.bobba@tse-fr.eu.

†Department of Economics, University of North Carolina – Chapel Hill, 107 Gardner Hall, CB 3305 Chapel Hill, NC 27599-3305 USA. E-mail: luca.flabbi@unc.edu.

‡Vice-Presidency for Sectors and Knowledge, Inter-American Development Bank, 1300 New York Avenue 20577 Washington DC USA. E-mail: slevy@iadb.org.

§Department of Economics (ILADES), Universidad Alberto Hurtado, Erasmo Escala 1835, Santiago, Chile. E-mail: matejada@uahurtado.cl

1 Introduction

Most labor markets in medium- and low-income countries are characterized by high levels of informality.[1] Informality refers to the non-compliance with labor market regulations, including the failure to contribute to the social security system. The result is a lower contribution base and the loss of benefits for a large portion of the labor force. The advantage is the reduction of the negative employment effects induced by some labor market institutions.

If the presence of informality may be seen as an optimal reaction to a given institutional context, it is also correlated to other labor market features that may impact overall productivity. A growing literature is focusing on the firm side, showing strong correlations between firm's productivity and formality status and identifying an important channel of the relation in the distortions of firms' investment decisions.[2] The literature focusing on productivity and the worker side is smaller and rarely takes into account workers's investment decisions – such as human capital accumulation – in presence of high informality.[3] In a companion paper [Bobba et al., 2017], we study the issue focusing on human capital accumulation decisions *before* entering the labor market. In this paper, we move our attention to the dynamic of human capital that takes place *after* entering the labor market. In particular, we look at human capital accumulation on-the-job, its possible depreciation while searching for a new job, and we study whether and how the formality status of the job significantly affects this dynamic. The evidence on the higher instability of informal jobs and on the reluctance of firms to invest in specific human capital when hiring informally, suggests that this relation is potentially very relevant.[4] We develop and estimate a model where formality status and job search decisions are updated optimally every time the human capital levels change either as a result of upgrading on the job or downgrading while searching for a new job. As in our previous work, we model in detail the structure of social security, allowing for the presence of a "dual" system where formal jobs enjoy benefits financed by payroll contributions while informal workers and labor market searchers receive benefits financed by resources collected outside the labor market.

Specifically, we consider an environment where workers and employers search for potential partners to enter a job relation. When they meet, they observe a match-specific productivity

[1] See for example La Porta and Shleifer [2014]; Perry et al. [2007]; Levy and Schady [2013].

[2] For evidence on informality and firm productivity in Mexico, see Busso et al. [2012]. Studies of firms' investment decisions in the presence of informality include de Paula and Scheinkman [2011]; Ulyssea [2015].

[3] Meghir et al. [2015] is an important contribution taking into account worker's behavior and firms' productivity but it does not allow for workers's investment decision.

[4] See Lagakos et al. [forthcoming] for recent cross-country evidence.

that contributes to the overall output of the match together with the workers' human capital. Firms optimally post the formality status for each specific match and wages are determined by bargaining. At a given point in time, each worker can be in one of four possible labor market states: formal employee, informal employee, self-employed, unemployed. The human capital evolution while participating in the labor market captures the additional productivity that may be acquired on the job (human capital upgrading). This additional productivity may depreciate if not working (human capital downgrading). While working, human capital upgrading may occur at different rates depending on the formality status of the job and on the current human capital level. The process of human capital upgrading and downgrading generates endogenous changes in wages and labor market states. For example, an informal employee who upgrades his human capital may endogenously negotiate a higher wage, a different formality status or even quit. The possibility of a change in formality status at the *same* job is a non-negligible event in the data that few previous labor market models are able to replicate.[5]

We estimate the model on labor market data from Mexico.[6] We find that human capital accumulation on-the-job is important: in steady state, it is responsible for about one-third of the overall value of production. As expected, we find that human capital upgrading is slower while working informally than formally: for first entrants in the labor market, it takes on average 2 years to start upgrading their human capital if they work formally and about 6 years to do so if they work informally. We also estimate that the upgrading is harder the higher the level of human capital already acquired on the job. Still, at any human capital level, the probability of upgrading remains higher if working formally. There are two main sources of identification in the data for the parameters governing the human capital dynamic: transitions between jobs and labor market states; wages growth within and between jobs, conditioning on formality status.

Using the point estimates, we generate by simulation counterfactual labor markets characterized by different policy parameters. We focus on two policy parameters: the contribution rate paid by formal employees and the level of the non-contributory social security benefits received by informal workers and unemployed searchers. These two parameters are considered crucial in generating the high level of informality observed in Mexico and other countries in Latin America.[7] The first experiment shows that formality and productivity

[5]For example, neither Bosch and Esteban-Pretel [2012], nor Meghir et al. [2015], nor Bobba et al. [2017] can account for this type of labor market dynamics.

[6]We extract the data from the 2013 and 2014 Mexico's official labor force survey: the *Encuesta Nacional de Ocupación y Empleo* (ENOE). See Section 2.2 for more details on the data.

[7]For a focus on the payroll contribution, see Albrecht et al. [2009] and Rocha et al. [2017]. For a review

decrease as the contribution rate increases. The elasticity of the negative impact on overall output is magnified by the negative impact on human capital accumulation on the job. The second experiment shows that an increase of the non-contributory benefit from the current level would increase informality, decrease human capital accumulation and, ultimately, decrease output. However, the experiment also shows that a positive non-contributory benefit of about half the current level would be able to increase overall output by generating more productive matches.

The paper is organized as follows. Section 2 describes the data and the institutional context. Section 3 develops and discusses the model. Section 4 describes the identification of the model's parameters with the data at our disposal. Section 5 defines the estimation method and presents the estimation results. Section 6 reports the policy experiments. Section 7 concludes.

2 Context and Data

2.1 Institutional Setting

We define informality with reference to compliance with regulations on salaried labor.[8] In Mexico, as in most countries, firms are obligated to enroll salaried workers in the social security registry (IMSS, for its Spanish acronym) and pay a contribution proportional to workers' wages whose revenue is used to fund social security benefits. Unlike in the United States, those benefits are bundled in the sense that workers receive a package that includes health benefits, housing benefits, some day care services, and pensions. Some benefits are directly proportional to the worker's individual wage and contribution (pensions) while others are not (health benefits). Since the contribution are proportional to wages, this implies redistribution within salaried formal workers. There is no unemployment insurance and thus no flow payments out of wages into an unemployment fund or individual accounts. In Mexico, the rate of the social security contribution is approximately 33 percent of the wage of salaried workers. Since these regulations are imperfectly enforced, non-compliance occurs as a device for firms to save on labor costs. When caught hiring illegally, firms have to pay monetary

on Mexico, see Levy [2008].

[8]See Kanbur [2009] and Levy [2008]. This definition abstracts from other dimensions of the formal-informal divide that are possibly relevant for worker behavior, such as the compliance with income taxes. While conceptually relevant, in our context this is of second-order importance as labor income tax is small over the wage support that we consider in our sample.

fines that range between 20-350 daily minimum wages for each non-registered worker.[9] Many firms operate in both the formal and informal sector because they hire workers both legally and illegally.[10]

To the extent that there is no firm-worker relationship, these rights and obligations do not apply to self-employed workers. For most of the individuals engaged in those activities, the notion of self-employment differs quite fundamentally from its counterpart in high-income countries. It can be mostly ascribed as a "necessity" labor market state whereby individuals who are not matched with firms engage in self-employment activities while also searching for a job [Fields, 1975]. A typical example of such activity is working as a street vendor. Financial barriers to enter into self-employment do not appear as an important obstacle [Bianchi and Bobba, 2013], which is consistent with the fact that unemployment is in general very limited in those labor markets.

In response to the lack of social security coverage for informal workers, starting from the early 2000s non-contributory programs were launched to expand the coverage of housing subsidies, retirement pensions and day care facilities. Spending in those programs doubled between 2002 and 2013, from 0.8 to 1.65 percent of GDP – a pattern that is in common across many countries with a dual social security system [Frolich et al., 2014].[11] The voluntary, unbundled, and practically free nature of non-contributory programs implies that valuation issues are substantially less complex than in the case of contributory programs. There are no significant regional or quality differences between contributory and non-contributory pension, housing and day care programs; with regards to health, differences have narrowed considerably as a result of a large expansion in the health infrastructure of state governments, which provide services to those not covered by IMSS [Levy, 2008].

[9]The exact parameters that IMSS uses to determine which establishments to inspect are confidential. However, according to IMSS officers in charge of inspections, when deciding which firms to inspect they take into account firm size, industry, history of previous violations and notifications made to IMSS by the Ministry of Labor.

[10]For instance, Perry et al. [2007] show that in Mexico 50% to 70% of small-medium firms have used both formal and informal contracts simultaneously in a given point in time. Ulyssea [2015] documents that in small formal firms in Brazil 40% percent of workers are informal. At the same time, 52% of all informal workers are employed in large firms that are unlikely to be fully informal.

[11]The corresponding figures for other Latin American countries document even steeper growth rates than Mexico over the same period. For instance, in Chile spending in non-contributory social programs increased from 0.5 percent of GDP in 2002 to 1.5 percent of GDP in 2013. In Argentina, spending increased from 1 percent of GDP to 4 percent of GDP.

2.2 Data

The data is extracted from Mexico's official labor force survey, the *Encuesta Nacional de Ocupación y Empleo* (ENOE). Similar to the US Current Population Survey, the dataset has a panel component – households stay in the sample for five consecutive quarters. In the first quarter of each year, employed individuals are inquired about the date in which they started working with their current employer. We stack together two cohorts of individuals entering in the first quarter of the year 2013 and in the first quarter of 2014, respectively. This information combined with quarterly panel data on wages and labor market status allows us to fully characterize the labor market trajectories (job-to-job transitions,) for the individuals in our sample. The two years under consideration are the most recent available in ENOE at the start of this project. We are forced to merge the two years together to gain sample size. While the overall sample size is not small, some labor market transitions important for the identification of the model are relatively rare: combining two years allows us to compute more credible moments. The Mexican economy was quite stable over the period, so assuming that agents were participating in the same labor market over the two years is not unreasonable.[12]

We restrict the sample to nonagricultural, full-time, male, private-sector workers between the ages of 25 and 55. We focus our analysis on workers at the mid-range of the skill distribution – i.e. those with a secondary schooling degree, which comprise more than half of the labor force. We thus drop from the sample those who did not complete junior secondary schooling (i.e. below 9th grade) and those who completed college or a higher educational degree. We define a worker to be an *employee* if he declares (i) being in a subordinate working relationship in their main occupation; and (ii) receiving a wage as a result of that working relationship. We identify the formal or informal status of the job depending on whether the employee reports having access to health benefits through their employers, which is common practice in the literature.[13] We define the *self-employed* workers as those who declare (i) not being in a subordinate relationship in their main occupation and (ii) having a business by their own. In order to obtain a more homogenous population of self-employed individuals and to be consistent with the "necessity" self-employment we are interested in, we drop those who report having paid employees and those who report having access to contributory

[12]For example, unemployment rate was 4.9% in 2013 and 4.8% in 2014; real GDP growth was, respectively, 1.4% and 2.6%. Source: World Development Indicators (WDI), *The World Bank*.

[13]In the literature on Latin America the informality status of an employee is typically defined in reference to firms' compliance with the social security regulation. See Bobba et al. [2017] for more details on this measure of labor informality in Mexico.

health benefits. The entire sub-population of self-employed workers that we consider is thus informal, as opposed to employee workers who can be formal or informal depending on employers' decision to enroll some, none or all of their employees in the social security registries. We define the *unemployed* as those who declare (i) not to be working during the last week; and (ii) being actively searching for a job. Earning distributions are trimmed at the top and bottom 1% in each labor market state (formal employees, informal employees and self-employed).

The final sample that we use in our empirical analysis is a balanced panel dataset comprised of 3,860 individuals observed every quarter for five quarters, either starting in the first quarter of 2013 or in the first quarter of 2014. Table 1 and Figure 1 depict the main cross-sectional facts; Table 2 reports statistics on labor market dynamics. The observed patterns are broadly consistent with previous evidence for Mexico and with aggregate evidence from Latin America. First, there is a significant mass of workers in each labor market state: about 60% of workers are employed formally and 35% informally. Among the informals, 2/3 are employees and 1/3 self-employed. Unemployment rates are less than 5%. Second, there is a large overlap between the wage distributions of formal employees and informal employees. Self-employed earnings distributions are approximately in between those of formal and informal employees, with a larger standard deviation. Third, there is a significant amount of transitions between labor market states and formality regimes. Looking at the second row of Table 2, we observe that more than 30% of the informal employees change labor market status after a year. In the case of the most persistent state – formal employee – about 12% change labor market status after a year. Changes of formality status are also significant, with more than 18% of informal employees becoming formal after a year. Fourth, and frequently neglected by the literature, changes in formality status may frequently occurs at the *same* job. Out of all the informal employees becoming formal within a year, more than 60% of them do so at the same job. Interestingly, the opposite is also taking place: out of the 8% of the formal employees becoming informal, 53% of them do so at the same job. Fifth, transition rates out of unemployment are on average more frequent than those out of self-employment, suggesting different dynamics in the two labor market states. Roughly 82% of the unemployed find a job over a period of one year compared with 28% of the self-employed. Also, while the majority of the unemployed transit toward a formal job most of the self-employed transit toward an informal one.

3 Model

3.1 Environment

The model assumes stationarity and continuous time. All agents are subject to a common discount rate ρ and to a common probability of death, modeled as a Poisson process with parameter δ. When an agent dies, a new agent is born as a draw from the initial population of agents.

The labor market is characterized by search frictions: workers and employers search for potential partners to enter a job relation but meetings do not happen instantenously and they require time. When they meet, they decide if entering the job relationship or continue searching for a new partner. Crucial in the decision is the productivity generated by the specific match of a given worker with a given employer. We model the match-specific productivity as a draw $x \sim G(x)$.[14]

Workers can be in four labor market states: unemployment, self-employment, informal employment and formal employment. The informal sector is composed by the self-employed and by the informal employees. Agents only receive job offers as employees while unemployed or self-employed. Formality status as an employee is denoted by $f \in \{0, 1\}$, with 1 indicating a formal labor contract. Searching status as an agent receiving employee offers is denoted by $s \in \{0, 1\}$, with 1 indicating self-employment.

We focus on the human capital that accumulates and depreciates *while* participating in the labor market. We condition on the human capital accumulated *before* entering the labor market.[15] The human capital evolution while participating in the labor market captures the additional productivity that may be acquired on the job (human capital upgrading). This additional productivity may depreciate if not working (human capital downgrading). While working, human capital upgrading may occur at different rates depending on the formality status of the job and on the current human capital level. Our prior is that human capital upgrading is slower while working informally than formally but the actual rates will be estimated in the empirical section of the model.[16] We also expect the upgrading to be

[14]This is the most commonly used productivity representation in search-matching-bargaining models of the labor market, including our previous Bobba et al. [2017] and Eckstein and Wolpin [1995], Cahuc et al. [2006] and Flinn [2006]. For theoretical foundations, see Wolinsky [1987] and Jovanovic [1979]. For a recent review, see Chapter 4.2 in Keane et al. [2011].

[15]In the empirical analysis, the pre-labor market human capital will be fully described by education. As discussed in Section 2, we focus on individuals with Secondary School education level.

[16]The prior is based evidence suggesting that informal jobs are characterized by higher instability and lower firms' investmente in human capital. See for example Lagakos et al. [forthcoming] for cross-country evidence.

harder the higher the level of human capital already acquired on the job, consistently with decreasing returns in human capital investments on the job. Again we want to allow for this flexibility in the theoretical model. In the empirical section of the paper we will show if the estimation results confirm this prior. We represent the evolution of human capital in the labor market by assuming a discrete distribution of human–capital–upgrading values $1 = a_1 < ... < a_K < \infty$. The total productivity of the match of a worker with labor market human capital a_k meeting a firm in a match generating productivity x is:

$$y(x, k) = a_k x \tag{1}$$

A worker in such relationship receives a human capital upgrading shocks following a Poisson process with rate $\tau_{f,k}$. When an upgrading shock arrives, the labor market human capital of the workers 'upgrades' to the next, higher level $a_{k+1} > a_k$. A searcher in a labor market state s with labor market human capital a_k receives human capital downgrading shocks following a Poisson process with rate $\gamma_{s,k}$. When a downgrading shock arrives, it decreases the labor market human capital to the next, lower level $a_{k-1} < a_k$. Notice the limiting cases: $\tau_{f,K} = 0$ and $\gamma_{s,1} = 0$.

On top of the human capital process, the usual labor market dynamic is taking place. While searching, agents meet employers at the Poisson rate λ_s. While working as employee, matches are terminated at the Poisson rate η_f. Agents have the faculty to accept or reject job offers but they cannot reject a termination: when the shock hits, they have to revert to their optimal searching state. Termination may also occur endogenously, as a result of human capital upgrading.

Formality and searching status are endogenous. The formality status while working as employee (f) is posted by the firms optimally, based on the observed labor market human capital a_k, and the match-specific productivity x. Assuming that the authority to post the formality status is in the hand of the firm is consistent with the institutional setting in Mexico and most Latin American countries. Conditioning on x, f and k, workers and firms engage in bargaining to determine wages. The searching status (s) is decided by the workers optimally, based on their labor income generated as self-employed: $q \sim R(q)$. q is heterogenous in the population but time-invariant within individuals. The flow utility while searching as unemployed is homogenous and denoted by ξ.

We follow previous literature by assuming linear utility.[17] We follow our previous work

[17]Search models of the labor market typically assume linear utility. The exception are household (or dual) search model of the labor market, such as Dey and Flinn [2008], Guler et al. [2012] and Flabbi and Mabli

on Mexico (Bobba et al. [2017]) in defining flow utility as composed by labor income and by a social security benefit component. The social security benefit component depends on the formality status and includes both the preferences for the the benefit and the monetary input used to provide the benefit. This setting leads to the following four flow utility definitions:

$$\xi + \beta_0 B_0 \tag{2}$$

$$q + \beta_0 B_0 \tag{3}$$

$$w_0(x;k,q) + \beta_0 B_0 \tag{4}$$

$$w_1(x;k,q) + \beta_1 B_1[w_1(x;k,q)] \tag{5}$$

The first flow utility refers to the unemployed: they receive the (dis)utility of being unemployed and searching ξ and the non-contributory benefit B_0, which they value β_0 to the peso. Exactly the same benefit is received in all the other labor market states with the exception of formal employment. Formal employees receive a contributory benefit B_1, which they value β_1 to the peso. We discuss the exact form of this benefit in the next paragraph. On top of the benefits, agents in self-employment receive labor income q and agents working as employees receive the wage $w_f(x;k,q)$.

The benefit B_1 is received only by formal employees and it is a contributory benefit, i.e. the firm contributes to the benefit of each employee by withdrawing at the source a rate t of the employee's wage. This contribution provides two benefits: a proportional benefit, which represents institutions such as a defined contribution retirement plan; and a fixed benefit, which represents institutions such as health benefits. The contributions to the first benefit is a proportion ϕ of the total contribution. Formally, the benefit B_1 is defined as:

$$B_1[w_1(x;q,h)] \equiv \phi t w_1(x;k,q) + b_1 \tag{6}$$

where b_1 is the notation we use for the fixed benefit. As discussed in more detail in our previous work on Mexico (Bobba et al. [2017]), the system has important distributional effects. Since the collection of contributions is proportional to wages and b_1 is equal for all formal employees, the system implies redistribution from high-wage earners to low-wage earners within the formal sector.

The employers side of the model is very stylized. Employers post vacancy at no cost and earn revenues equal to the match-specific productivity, scaled by the worker's human capital.

[2018].

The labor costs include wages and social security contribution when hiring formally. They include wages and the probability and penalty of being caught when hiring informally. This setting leads to the following two flow profit definitions:

$$\pi_1(x;k,q) = y(x,k) - (1+t)w_1(x;k,q) \tag{7}$$

$$\pi_0(x;k,q) = y(x,k) - w_0(x;k,q) - cy(x,k) \tag{8}$$

Notice that in order to fit some data regularity we impose that the penalty for being caught hiring informally is an increasing function of productivity. We choose a very simple linear form for this function: cy.

3.2 Value Functions

Assume an individual searching in the labor market with human capital a_k, and potential self-employment income q. This agent will receive two possible shocks: meeting an employer and incurring human capital downgrading. The value of this state can be written in recursive form as follows, where we denote the searching states with V_s and the employee states with E_f:

$$\begin{aligned}(\tilde{\rho} + \lambda_s + \gamma_{s,k})V_s(k,q) &= (1-s)\xi + sq + \beta_0 B_0 \\ &+ \lambda_s \int_x \max\{(1-f)E_0(x;k,q) + fE_1(x;k,q), V_s(k,q)\}dG(x) \\ &+ \gamma_{s,k}\max\{V_0(k-1,q), V_1(k-1,q)\}\end{aligned} \tag{9}$$

where to simplify the notation we define $\tilde{\rho} \equiv \rho + \delta$. The first row represents the flow value, which is a function of the searching state (either unemployment or self-employment). When the worker meet an employer, a match-specific productivity x is drawn and the workers receive either a formal or informal employee offer. The worker decides if accepting the offer or not by maximizing over the two possible value function. When the worker receives a human capital downgrading shock, he will move to the lower level a_{k-1} and decide if continue searching in the current state – being that unemployment or self-employment – or switch to the other state. Note that the formality status f is endogenous and posted by the firm, as we show in Section 3.3.

When an agent is working as an employee, two shocks are possible: termination and

human capital upgrading. The value of the employee state in recursive form is therefore:

$$\begin{aligned}(\tilde{\rho}+\eta_f+\tau_{f,k})E_f(x;k,q) &= w_f(x;k,q)+(1-f)\beta_0 B_0+fB_1[w_1(x;k,q)] \\ &+ \tau_{f,k}\max\{(1-f)E_0(x;k+1,q)+fE_1(x;k+1,q),V_s(k+1,q)\} \\ &+ \eta_f V_s(k,q)\end{aligned} \tag{10}$$

The first row represents the flow value, which is a function of the wage and the formality–status–specific benefit (either B_0 or B_1). The second row shows that when the worker upgrades the labor market human capital, the formality status and the searching state are both updated optimally. This generates an interesting dynamic usually ignored in the literature: formality status may change *within* the same employer and job termination may occur *endogenously*.[18] Finally, the third row shows that when the match is exogenously terminated, the agent has to go back to the searching state.

The value functions for the demand side of the market are as follows. Employers post vacancies and search for workers to fill them. We assume there is no cost of posting and keeping the vacancy open.[19] The value of a filled job is consistent with the worker's side. It is denoted by F_f and defined as:

$$\begin{aligned}(\tilde{\rho}+\eta_f+\tau_{f,k})F_f(x;k,q) &= (1-f)\pi_0(x;k,q)-f\pi_1(x;k,q) \\ &+ \tau_{f,k}\max\{F_0(x;k+1,q),F_1(x;k+1,q),0\}\end{aligned} \tag{11}$$

The flow value is defined by the firm's profit, defined in equation (7) and (8). A filled job is subject to the same shocks we discussed for the worker's side: a termination shock η_f, which sends the firm back to a value of zero, and a human capital upgrading shock $\tau_{f,k}$. When the human capital upgrading shock hits, the employer enters a new negotiation with the worker and decides optimally the formality regime and if keeping the worker.

3.3 Wages and Formality Status

The formality status decision is taken by the firm upon observing the labor market human capital a_k, the outside option $V_s(k,q)$, the match-specific productivity x and with the knowledge that wages will be set by bargaining. The decision involves comparing the value

[18] For example, neither Bosch and Esteban-Pretel [2012], nor Meghir et al. [2015], nor Bobba et al. [2017] can account for this type of labor market dynamics.

[19] A foundation for this result may be given by assuming free-entry of firms together with congestion effects, as in Bobba et al. [2017] and Flinn and Mullins [2015]. For a more complete discussion, see Pissarides [2000].

of filling the vacancy hiring formally or informally. The endogenous formality status f is therefore determined as:

$$f \equiv f(x;k,q) = \begin{cases} 1 & \text{if } F_1(x;k,q) \geq F_0(x;k,q) \\ 0 & \text{otherwise} \end{cases}$$

Note that throughout the paper we simplify notation by dropping the dependence of f on $(x;k,q)$.

Wages are set by bargaining upon observing the labor market human capital a_k, the outside option $V_s(k,q)$, the match-specific productivity x and the formality status posted by the firm f. We assume the axiomatic Nash-bargaining solution leading to:

$$w_f(x;k,q) \quad = \quad \arg\max_w \left[E_f(x;k,q) - V_s(k,q)\right]^{\alpha} \left[F_f(x;k,q)\right]^{(1-\alpha)} \tag{12}$$

The solution is a quite involved analytical expression that we report in the Appendix (equations (A.1) and (A.1)). But the interpretation is the usual one: wages are a linear combination of productivity y and the outside option $V_s(k,q)$. The higher the worker's bargaining coefficient α, the more weight is given to productivity in determing wages.

3.4 Equilibrium

First entrants in the labor market start at the lowest level of human capital a_1. This level is a lower bound and it does not depreciate. Based on q, they decide if start searching for an employee job as unemployed ($s=0$) or self-employed ($s=1$). They decide based on the following maximization:

$$\max_s \{V_0(1,q), V_1(1,q)\}$$

where $V_s(1,q)$ is the value of searching for an employee job (equation 9). Since $V_1(1,q)$ is increasing in q faster than $V_0(1,q)$, there exists a unique:

$$q^*(1) : V_0(1,q^*(1)) = V_1(1,q^*(1)) \tag{13}$$

Only agents with $q < q^*$ search as unemployed, whereas agents with $q \geq q^*$ search at lower intensity while working as self-employed.

After accepting employee offers, workers start to accumulate human capital, upgrading from a_1 to a_2 to potentially any a_k up to a_K. Once they go back to a searching state with a generic a_k, that value may depreciate and may affect the searching status decision.

The searching status decision is updated using the same reservation value rule based on the generic $q^*(k)$.

A worker with searching status s, labor market human capital a_k, and potential self-employment income q observes a match-specific productivity value x when meeting an employer. The employer observes the same information and knows the wage determination process. Based on this information, posts a formality status f. The worker observes the formality status, bargains with the firm leading to the wage schedule defined in (12), and decides if accepting the match or not. The firm also is deciding if completing the match or not. Both firm and worker will arrive to the same optimal decision thanks to the no disagreement result implied by Nash bargaining. Since the outside option for both agents are constant in x while the value of the match is increasing in x, the optimal decision rule will again be a reservation decision rule. The reservation value is defined by:

$$x_f^*(k,q) : F_f(x_f^*;k,q) = 0 \quad \Longleftrightarrow \quad E_f(x_f^*;k,q) = V_s(k,q) \tag{14}$$

For any $x \geq x_f^*$, the match is realized.

The formality status is posted by the firm following the optimal decision rule described in Section 3.3. As shown in Bobba et al. [2017], this decision is also characterized by a reservation value property based on x. The indifference point is determined as:

$$\tilde{x}(k,q) : F_1(\tilde{x};k,q) = F_0(\tilde{x};k,q) \tag{15}$$

For any $x \geq \tilde{x}(k,q)$, the firm is posting a formal job ($f = 1$); for any $x < \tilde{x}(k,q)$, the firm is posting an informal job ($f = 0$).

With the optimal decision rules in place, the equilibrium is defined by the set of value functions that satisfies equations (9)–(11), once the optimal decision rules – including the optimal determination of wages and formality status – are taken into account. The equilibrium also determines steady state values for the measures of workers in each labor market state and for the distribution of human capital. We solve the model numerically by value function iteration. Appendix A.2 provides a detailed description of our procedure.

3.5 Empirical Implications

The equilibrium just described is able to capture the main characteristics of a labor market with high informality, such as many markets in Latin America, in general and Mexico, in particular (see Section 2).

First, the model equilibrium can generate a positive mass of workers in each labor market state and produce the significant amount of transitions between formality and informality. Transitions between formality and informality can take place not only when agents change job but also within the same job. The human capital upgrading process is the reason why a worker may change formality status within the same job. For example, worker i with human capital a_k may have accepted a job working informally as an employee because the match-specific productivity x_i was:

$$x_0^*(k,q) \leq x_i < \tilde{x}(k,q)$$

While in the informal job, he may receive a human capital upgrading shock, moving him from a_k to a_{k+1}. The upgrading may be such that the new reservation value to work formally is now lower than the match-specific productivity x_i:

$$\tilde{x}(k+1,q,h) \leq x_i$$

since it is possible that $\tilde{x}(k+1,q,h) < \tilde{x}(k,q)$. As a result, the worker will remain in the same job but at the same time will change his formality status from informal to formal.

Second, the model is able to generate wage distributions in line with the data. The data show two main features: a ranking between average wages conditioning on formality status and a lot of wage dispersion within formality status. If formal employees have on average higher wages than informal employees, the heterogeneity within informality status generates a large overlap between the distributions. We discuss in detail the importance of this stylized fact in our previous contribution, Bobba et al. [2017]. On top of the arguments made there, the dynamic of human capital upgrading and downgrading may generate larger or smaller overlaps depending on the different rates of human capital accumulation in the two formality statuses.

Third, the model is able to generate wage growth not only across jobs – as common in related literature – but also within jobs.[20] The reason is the renegotiation process taking place when the human capital upgrading occurs. Assume a worker i with match-specific productivity x_i and human capital level a_k upgrades his human capital while working as a formal employee. Further assume that x_i is such that $\tilde{x}(k+1,q_i) < x_i$. Then the worker

[20]Most estimated search model of the labor market impose constant wages at the same job. The main exceptions include models allowing for on-the-job search and wage renegotiation, such as Cahuc et al. [2006]. Very few introduce human capital accumulation on-the-job: notable examples are Bagger et al. [2014] and Flinn et al. [2017].

will remain matched with the same employer but his wage will increase from $w_1(x_i; k, q_i)$ to $w_1(x_i; k+1, q_i)$. These wage changes are observed in our data and are essential in the identification of the human capital upgrading shocks.

4 Identification

The model is characterized by the following parameters set:

$$\{\rho, \delta, \tau_{f,k}, \gamma_{s,k}, \lambda_s, \eta_f, \xi, \alpha\} \tag{16}$$

and by the following distributions:

$$\{G(x), R(q)\} \tag{17}$$

In addition, we have to define the support of the human capital dynamics: $\{a_k\}_{k=1}^K$ and the set of parameters that characterizes the institutional setting: $\{\beta_0, B_0, \phi, t, \beta_1, b_1, c\}$.

Starting with the institutional parameters, we set $\{\phi, t\}$ at the values present in Mexico during the surveying period of our sample. The parameters are stable over the entire decade that include our two years and they are respectively equal to 0.55 and 0.33.[21] The non-contributory benefit B_0 is calibrated from aggregate data following the same procedure described in Bobba et al. [2017] and it is equal to 4.27 pesos per hour for the year 2013. The portion of the contributory benefit that is distributed equally across all the formal employee after collecting their individual contributions (b_1) is estimated from the data by assuming that the formal system runs a balanced budget. Denoting with i a generic observation in our sample, the estimator is:

$$\hat{b}_1 = t(1-\phi) \sum_{i \in N_{E_1}} \frac{w_1(i)}{N_{E_1}} \tag{18}$$

where N_{E_1} denotes the set of formal employees.

With the institutional parameters in place, Bobba et al. [2017] proposes an identification strategy for β_0, β_1 and c. It builds upon observing the large overlap in accepted wages between formal and informal employees and providing an explanation for such overlap based on the model. The intuition is that at the reservation value $\tilde{x}$ – and in a small enough neighborhood around it – workers accept lower wages to work formally than informally because they receive higher non-monetary benefits. The amount of this overlap is driven by the preference and quantity of the benefits and by the cost of informality c. Adding this observation to the

[21] See Appendix C in Bobba et al. [2017] for more details on the institutional sources of these values.

quasi-random roll-out of a non-contributory social program (the *Seguro Popular* program) concludes the identification strategy we proposed there.[22] In the current setting, we cannot rely on the differential roll-out of the *Seguro Popular* program because at the time of our surveying period virtually everybody was covered by that program. Moreover, adding the human capital dynamic on the job weakens the separate identification of the preference parameters, β_0 and β_1, from the cost parameter of offering an informal job, c. Under the assumption that preferences for social security benefits are stable over the nine years that separate the data of the two papers, we have chosen to calibrate the preferences with the point estimates obtained in Bobba et al. [2017]. With the preference in place, we can use the overlap of the accepted wage distributions for formal and informal employees to identify c.

We exploit classic results from Flinn and Heckman [1982] to identify the labor market parameters $\{\rho, \lambda_s, \eta_f, \xi, \alpha\}$ and the match-specific distribution $G(x)$. They show that by assuming a recoverable distribution for $G(x)$, the entire set of parameters – up to two restrictions – is identified from observing accepted wages and transitions between labor market states.[23] The recoverable distribution we assume for $G(x)$ is a lognormal with parameters that we denote $\{\mu_x, \sigma_x\}$.[24] The two restrictions refer to the parameters $\{\rho, \xi\}$ and α. Flinn and Heckman [1982] show that the first two parameters are only jointly identified. We follow previous literature by setting ρ to 5% a year and recovering ξ by exploiting the equilibrium equation (9). Flinn and Heckman [1982] do not provide an identification strategy for α because they impose a sharing rule that splits productivity equally between worker and employer. We lack the demand side information necessary to identify α and we therefore choose to assume symmetric Nash bargaining which leads to a value of α equal to 0.5.[25] In addition to the inter-temporal discount, we also have to identify δ, the Poisson parameter describing the death shock. Since the risk of death is constant in the model, we can identify it by the average duration of the (labor market) lives of our sample.

The same recoverability condition necessary and sufficient to identify $G(x)$ from accepted wage distributions can be applied to identify $R(q)$ from observed self-employed labor income.

[22] For a more detailed and more formal discussion of this identification strategy, see Section 4 of Bobba et al. [2017].

[23] In their original contribution, Flinn and Heckman [1982] use durations to describe labor market dynamic. Using transitions across labor market states (both across and within jobs), as we do in our estimation procedure, does not change the source of identification; it just describes in a different way the dynamic over labor market states.

[24] Virtually all the search-matching-bargaining literature assume log-normality, from the seminal Eckstein and Wolpin [1995] to our recent Bobba et al. [2017].

[25] Recent works setting α using a similar strategy include Flabbi and Moro [2012] and Borowczyk-Martins et al. [2018].

The observed distribution of q is a truncation of the primitive $R(q)$ at the reservation value $q^*(1)$. If we assume a recoverable distribution, the primitive can be identified from its truncation. We assume a lognormal distribution with parameters that we denote $\{\mu_q, \sigma_q\}$.

We finally consider the parameters describing the novel feature of our model: the human capital dynamic while working in the labor market. The dynamic is characterized by human capital upgrading on the job and by human capital downgrading while searching. Both processes are characterized by shocks moving agents over the support $\{a_k\}_{k=1}^K$. We do not have direct information about events that may change human capital on the job, such as training, specific knowledge acquisition, or testing of skills. To identify the process, we can only rely on standard labor market dynamics, i.e. wages and transitions. Given this data limitation, we do not attempt to estimate the support $\{a_k\}_{k=1}^K$. Instead, we follow Flinn et al. [2017] by imposing an upper and lower bound for the support of the a_k distribution and we discretize the resulting range in equal intervals. The breakpoints generated by the intervals define the different a_k. After some robustness checks, we have set the upper bound at $a_K = 3$ and we have divided the support in 10 discrete intervals. The lower bound has a natural normalization at $a_1 = 1$. This means that the productivity on the job of first entrants is equal to the actual match-specific productivity x. The productivity of agents with level of human capital equal to the midpoint of the support is double their match-specific productivity draw. The maximum productivity boost is equal to three times the match-specific productivity.

Given the support, we can propose an identification strategy for the parameters characterizing the shocks: the Poisson rates $\tau_{f,k}$ and $\gamma_{s,k}$. The human capital upgrading shock is governed by $\tau_{f,k}$ and has three important consequences for the labor market dynamic on the job. First, it induces wage renegotiation: as a result of human capital upgrading, the surplus increases and the wage of a worker at the same job increases. Second, the formality status may change as a result of the renegotiation since – as detailed in Section 3.5 – the reservation values are all dependent on the human capital level of the worker. A formality status change on the job is therefore additional valuable information to identify the occurrence of an upgrading shock. Third, the renegotiation may lead neither to a wage change nor to a formality regime change but to a labor market status change, i.e. firm and worker may agree to dissolve the match and go back to search. Transitions between labor market states is the final piece of information useful to identify the shock, albeit it is less valuable than the other two because transitions out of the employee state may also be induced by an exogenous termination shock.

In the data at our disposal, we can observe the rich labor market dynamic just described. We can see how wages evolve within the same job, we can observe transitions between labor market states and we can observe changes in formality status within the same job. We can therefore directly use this information to identify the frequency of the human capital upgrading shock. Even if this information is relatively rich, it remains limited: we observe individuals only at quarterly intervals and only for five quarters. We have therefore decided to impose a functional form that is very parsimonious in terms of parameters but still maintains enough flexibility to describe the process. We propose the following specification for the human capital upgrading shock:[26]

$$\tau_{f,k} \equiv \begin{cases} \tau_{f,1} a_k^{\tau_{f,2}} & \text{if } 1 \leq k < K \\ 0 & \text{if } k = K \end{cases}$$

The functional form reduces the number of parameters from (K-1) to 2 for each formality status f. The unknown parameters in the most flexible specification are (K-1) because – as shown in the second row – $\tau_{f,K}$ is zero by definition (no additional upgrading can take place when reaching the upper bound of the human capital distribution). Even if clearly restrictive, the proposed specification allows to capture that a more modest human capital upgrading may be easier to achieve than a really significant one. A positive $\tau_{f,1}$ combined with a negative $\tau_{f,2}$ implies that the probability to upgrade at a low a_k is higher than at a high a_k.[27]

In the case of the human capital downgrading shock, the amount of information that we have from the data is more limited because – as common in most standard labor market data – we do not observe much about the searching process. Specifically, we only observe durations and transitions over the searching states but we do not observe offers actually received. The impact of the downgrading shock during search is to make workers less picky in accepting jobs. We should then observe impacts on durations and transition rates. However, this is the same information that is identifying arrival rates of offers so we need additional information to separately identify the downgrading shock. The additional information we use is comparing if the optimal decision rules change between different search episodes for the *same* individual. As mentioned, our panel is short – 5 quarters – so we only rarely see two or more search episodes for the same individual. However, we see a significant number of

[26] A similar specification is adopted by Flinn et al. [2017].

[27] This is consistent with decreasing returns in human capital investment and it is actually what we find in estimation *without* imposing any sign constraints.

individuals quitting their job, searching, and finding another job all within our observation window. Comparing wages accepted in a previous job with wages accepted in a job following a search period is informative about the shocks received during the search episode in between. For example, if wages accepted in the following period are systematically lower than those in the previous period, it is very likely that a downgrading shock has occurred. If, for same length of search, this is more likely the case while searching as unemployed than as self-employed, then the depreciation shock should be more frequent in the first searching state than in the second. Finally, a downgrading shock may also induce a change in searching state. While this is relatively rare in the data, it is very valuable in terms of identification because it signals a depreciation shock has taken place with probability one (of course conditioning on the model). In conclusion, comparisons of wages before and after a search period and changes of searching states for the same individual is the information we use to identify the downgrading shock. However, only the change in searching state unequivocally identifies a depreciation shock. The wage information is also driven by different draws of the match-specific productivity x. As a result, we propose the most parsimonious specification possible by assuming that the downgrading shocks only depend on the searching state s:

$$\gamma_{s,k} \equiv \begin{cases} \gamma_s & \text{if } 1 < k \leq K \\ 0 & \text{if } k = 1 \end{cases}$$

The second row simply states that no additional depreciation can take place at the lower bound of the human capital distribution.

5 Estimation

5.1 Method

We estimate the parameters of the model using the Method of Simulated Moments (MSM).[28] To define the estimator, we introduce the following notation: Θ is the parameter vector; m_N is an appropriately chosen set of sample moments derived from our sample of size N; $M_R(\Theta)$ is the set of the same moments derived from a simulated sample of size R, extracted from the steady state equilibrium realized at the parameter vector Θ. We set R at 5,000, which is

[28]The method is commonly used to estimate highly nonlinear models with value functions solved numerically such as ours. For the asymptotic properties of the MSM estimator defined in (19), see Pakes and Pollard [1989] and Newey and McFadden [1994]. For applications similar to ours, see Bobba et al. [2017], Flinn et al. [2017], Flabbi and Moro [2012], and Dey and Flinn [2008].

slightly larger than the sample size N, in order to gain more precision in capturing those labor market transitions that are relatively rare. W is a symmetric, positive-definite weighting matrix that we introduce to harmonize the different scales of the moments and to weight them according to their sampling variability. We thus build W by replacing the diagonal of an identity matrix with the bootstrapped sample variances of the sample moments. We are now ready to define the estimator as:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \left[M_R(\Theta) - m_N\right]' W^{-1} \left[M_R(\Theta) - m_N\right], \tag{19}$$

We choose the moments to be used in the quadratic form (19) in order to capture the data features described in Section 4. We match the proportion of workers in each labor market state in a given point in time (we choose the first quarter) and the transitions rates obtained by observing agents one year apart in order to describe the distribution over labor market states and the dynamics between them (see Tables 1 and 2). From working as either a formal or informal employee in the first quarter, the worker may end up in the fifth quarter in one of the following six possible states: working formally at the same job, working formally at a different job, working informally at the same job, working informally at a different job, working as self-employed, and being unemployed. From the search states of self-employment or unemployment, the agent may end up in the fifth quarter in one of the four labor market states: formal employee, informal employee, unemployed and self-employed.

We use mean and standard deviation of employees' wages and self-employment labor income in a given point in time (we choose the first quarter) in order to describe the wage information. We add wage growth on the job and across jobs one year apart, taking into account if there is an episode of search longer or shorter than a quarter when changing job. To describe the overlap between the accepted wages in formal and informal jobs, we follow the procedure proposed in Flabbi and Moro [2012] and Bobba et al. [2017]. We build quintiles over the distribution of accepted wages for formal workers and for each interval, we compute mean wages of formal and informal employees; and the proportion of employees in informal jobs earning a wage in that interval. The complete set of 62 sample moments and the corresponding simulated moments and weights used in the quadratic form (19) are reported in Appendix B.

5.2 Results

The estimated parameter values are reported in Table 3. The values of parameters governing the rates of job arrival and termination $\{\lambda_0, \lambda_1, \eta_0, \eta_1\}$ are comparable to previous estimates for similar models.[29] The differences in arrival rates between the unemployed and the self-employed are very large, explaining in part the observed persistency in the self-employment state and the high-turnover in the unemployment state. Taking into account the endogenous acceptance probability, these rates translate in unemployed workers accepting a job after on average 9.3 months while self-employed workers do so in 9.8 years.

The estimated values of the parameters of the match-specific productivity distribution $\{\mu_x, \sigma_x\}$ and the self-employed earning distribution $\{\mu_q, \sigma_q\}$ are smaller in magnitudes when compared with the previous estimates on Mexico obtained by Bobba et al. [2017]. Beyond some differences in the data on employees' wages and self employed incomes across the time periods considered in the analysis, it is the model's specification that explains the difference. Here, the total productivity of the match between a worker and a firm is augmented by the role of the worker's human capital (see equation (1)) and hence it is not surprising that the primitive match-specific productivity is estimated to be smaller. Also, the model allows for transitions between the two searching states, unemployment and self-employment, which affect the value of self-employment in equilibrium. As reported in the bottom panel, average productivity is about twice as large when working as employee when compared to working as self-employed.

The estimate of the cost of hiring informally – the parameter c – is roughly 10% of job productivity, which is similar to the value estimated in Bobba et al. [2017]. The parameter captures all the costs associated with hiring informally, including the probability and penalty of getting caught. While the estimated cost is economically important – at the mean productivity of the realized informal matches is approximately 1.79 pesos per hour – it is still lower than the cost of hiring formally – as measured by the payroll tax rate times the wage schedule at the same productivity level, 3.71 pesos per hour, justifying the significant but not dominant presence of informal employees observed in our labor market.[30]

Finally, the flow value of being an unemployed searcher ξ is estimated to be negative. A negative value was expected in order to generate enough wage dispersion in accepted wages.[31]

[29] See for example the review in Eckstein and van den Berg [2007] and specifically models of individual search without on-the-job search such as Flinn and Heckman [1982], Flinn [2006] and Flabbi and Moro [2012].

[30] Notice that these are only direct costs, i.e. they do not take into account that through bargaining firms are able to partially transfer them to the workers, as seen in the equilibrium wage schedules (12).

[31] See Hornstein et al. [2011] for an extensive treatment of the issue.

The estimates for the arrival rates of the human capital downgrading shocks $\{\gamma_0, \gamma_1\}$ imply that on average individuals who are unemployed depreciate their stock of human capital every 5 years. The depreciation rate is much lower during spells of self-employment: approximately 14 years. The estimated values for the arrival rates of the human capital upgrading shocks $\{\tau_{f,1}, \tau_{f,2}\}$ reveal concave patterns in the expected time of arrival of these shocks, which are depicted in Figure 2. In this case too there are systematic differences between the two labor market states. We find that the rate of human capital upgrading is slower while working informally than formally at any level of human capital. For first entrants in the labor market it takes about 2 years to start upgrading their human capital if they work formally and 6 years if they work informally. At the average value of the distribution of human capital ($\bar{a} = 1.41$, see Figure 3), it takes about 4 years to upgrade while working formally and 15 years to upgrade while working informally. Table 4 reports some statistics conditional on employment status and level of human capital implied by these estimates. The value of productivity per worker increases steeply with the level of human capital. The relative contribution on overall productivity is about one third in the overall sample, it is slightly higher while working formally than while working informally, and it monotonically increases with the level of human capital acquired on the job.

5.3 Model Fit

Tables B.2-B.3 in the Appendix report the complete set of moments targeted by the MSM estimator. When compared to the data, the distribution over the four labor market states in the simulated data tend to understate the share of formal employees and to overstate slightly the share of self-employed and unemployed workers. The match is quite good on means and standard deviations of the accepted wages and of the self-employment incomes – with differences in the 10% range – with the exception of the standard deviation of self-employed incomes.[32]

A peculiar and relevant feature of Mexico's and other labor markets with high informality is the overlapping of the formal and informal accepted wages distributions. We are able to replicate the overlap quite well, both in terms of the proportions of informal employees in each quantile and in the mean accepted wages by quantiles. This result is obtained through two channels. First, the endogenous mapping between match-specific productivity and wages

[32]These set of moments is computed unconditionally on the labor market state in order to guarantee a smoother and well-defined quadratic form during the optimization procedure. As expected, the match between the corresponding conditional sample moments and the conditional simulated moments is slightly worse, notably for both the mean and the standard deviation of the self-employed incomes.

implied by bargaining. Second, the flexibility introduced by allowing the self-employment state to be a searching state, with heterogenous productivity levels that are pinned down by the (observed) income generated while in self-employment.

There are no major mismatches in the moments describing the yearly transitions between labor market states both within and across jobs targeted by the procedure but some data features are captured better than others. The model is able to generate wage growth rates within jobs that are lower but not too far away from the growth rates observed in the data. This is quite remarkable since the process of human capital accumulation is the only possible source of variation in wages within the same job that is embedded in the model. Finally, while wage changes across jobs are well matched for formal employees, they are somewhat off for informal employees. This is arguably due to the limited longitudinal dimension of the available data, which allows us to observe transitions only on a relative short time window.

6 Policy Experiments

Our model incorporates the structure of the social security system implemented by several countries in response to the lack of coverage for informal workers. The resulting 'dual system' is characterized by contributory benefits – governed by a payroll contribution rate, a benefit level increasing in wages and a redistributive component – and non-contributory benefits. To evaluate the impact of this complex system of incentives and disincentives, we use the estimated model to generate counterfactual labor markets where the crucial policy parameters assume a wide range of different values. The counterfactual labor markets are characterized by new steady state equilibria where labor market outcomes – including informality levels and value of production – and human capital levels are endogenously determined. We focus on changes in two policy parameters: the payroll contribution rate in formal jobs t and the per-capita level of the non-contributory social benefits B_0. These two parameters are considered crucial in generating the high level of informality observed in Mexico and other LAC countries since they directly affect the differential benefits and cost of working formally.[33]

The policy experiments procedure works as follows. For each value of the policy parameter, we find and compute the new equilibrium holding fix the other institutional parameters and setting the structural parameters at the point estimates reported in Table 3. Then,

[33]For a focus on the payroll contribution, see Albrecht et al. [2009] and Rocha et al. [2017]. For a review on Mexico, see Levy [2008]. for similar experiments in an environment with endogenous schooling choice, see our companion paper Bobba et al. [2017].

we simulate labor market careers for 5,000 individuals and compute the relevant statistics reported in Figures 4 and 5 from the resulting simulated data. We vary the policy parameters over a wide range: the contribution rate (currently at 33%) from 10% to 90%; the non-contributory benefit (currently at about 4.5 pesos per hour) from 0 to 8.

6.1 Policy Experiment 1: Contribution Rate

In the first experiment, we vary the contribution rate t (currently at 33%) over a wide range: from 10% to 90%. Figure 4 reports simulation results on labor market outcomes computed at the various contribution rates over the range. We denote the benchmark value with a vertical dashed line. Panel (a) describes the size of informality in the market. We report two informality rates: one computed by including only informal employees; the other by including both informal employees and the self-employed. The rate is monotonically increasing in t: as the realtive flow cost to work in a formal job increases, the level of formality decreases. Recall that while the contribution generates a benefit for formal workers, the evaluation of the benfit for a specific worker is in general different from the contribution. For the average worker, it is lower because the valuation parameter β_1 is smaller than one.[34] For a worker at a wage lower than the mean, it may be higher because of the redistribution occouring through b_1.[35] For a worker at a wage higher than the mean, it is always lower than the contribution paid. As a result for most workers an increase in t is a net loss that will induce them to work more informally. This decision has an impact on human capital accumulation on the job since the rate of human capital upgrading is lower when working as an informal employee than as a formal employee.

We see the impact in Panel (c): if the current contribution rate were to increase by 10 percentage points, the aggregate human capital would decrease by 2 percentage points.[36] Both the fact that the agents work more informally and that they accumulate less human capital have an impact on output: for the same 10 points increase in the contribution rate, the overall value of production would decrease by about 5 percentage points (see Panel (b)).[37] The final panel (Panel (d)) reports another dimension of the impact of human

[34]Specifically, as reported in Table B.1, is equal to 0.6705 meaning that one peso of contribution generates a benefit for which the workers would be willing to pay about 67 cents.

[35]Recall that the contribution rate is proportional to the wage but only a portion of the benefit is. The other portion of the benefit is fixed and therefore people with higher wages that contribute realtively more are redistributing benefit to people earnings relatively less.

[36]The aggregate human capital in the economy is calcualated using the equilibrium distribution over the discrete values of a_k.

[37]The value of production is the sum over the productivity of all the realized matches in equilibrium (the x's above $x^\star(k,q)$) and of all the self-employed active in equilibrium (the q's above $q^\star$).

capital on productivity: the contribution of human capital on overall production. To use the notation we introduced in the model (equation (1)), we are reporting the contribution of the a_k component to the overall value of production y. The contribution depends both on the equilibrium distribution of human capital in the economy and on the equilibrium distribution of human capital over the set of realized matches. Since an increase in t has a negative impact on both distribution, the contribution that human capital can give in increasing the output of match-specific productivities is decresing in t, as shown by the monotone decreasing curve reported in the panel.

In conclusion, the experiment confirms the expected result: formality and productivity decrease as the contributon rate increase. The elasticity of the negative impact on overall output is magnified by the negative impact on human capital accumulation on the job.

6.2 Policy Experiment 2: Non-contributory Benefit

In the second experiment, we vary the non-contributory benefit B_0 (currently at about 4.5 pesos per hour) from 0 to 8. Figure 5 reports simulation results on labor market outcomes computed at the different benefit levels over the range.

The most interesting result from the experiment is that the impact on the overall value of production is non-monotone. Over a range between 0 and 2 pesos per hour, a positive benefit has a positive impact on productivity, increasing the overall value of production by up to 4%. After the 2 pesos per hour level, an increase in the benefit starts to be detrimental to the economy, reducing productivity and human capital accumulation, increasing informality and human capital contribution to production. The non-monotone result is the equilibrium outcome of different channels. A positive non-contributory benefit favors informal employment with respect to formal employment. However, low levels of the benefit are not enough to induce agent to work as informal employees: in panel (a), the informal employees rate is zero up to $B_0 = 2$. As a result the negative impact on the human capital level is negligible – see panel (c) – because all employees work in jobs with the higher rate of human capital accumulation. At the same time, a low but positive level of the benefit allows for the creation of better matches between employers and employees and for the generation of more self-employment income. Both effects are due to the fact that the benefit effectively acts as a subsidy to search. The final result is the increase in overall production that we observe at the beginning of the range.

However, after a threshold of about 2 pesos per hour, this favorable channel cannot balance the negative channel implied by higher informality. When the benefit is high enough

to induce workers to accept jobs as informal employees, their human capital accumulation is hindered and productivity decreases, in some ranges quite sharply. One of the ranges where the decrease is sharpest is the range that includes the current benchmark value. If the current benefit were to be increased by an additional peso per hour from 2013 Mexican levels, the overall value of production would decrease by almost 4 percentage points (panel (b)). Partially responsible for this drop would be a 2 percentage points decrease in the overall level of human capital (panel (c)).

In conclusion, the experiment confirms that an increase of the non-contributory benefit from the current level would increase informality, decrease human capital accumulation and, ultimately, decrease output. However, the experiment also shows that exists a range – about up to half the current level – where a positive benefit may be beneficial by subsidizing search and allowing the creation of more productive matches.

7 Conclusions

We study how the different rate of human capital accumulation in formal and informal jobs may impact labor market outcomes and output values. Recognizing that formality status and labor market states are endogenous choices interacting with the human capital dynamic on the job, we develop a search and matching model where firms and workers produce output that depends both on match-specific productivity and on worker-specific human capital. Human capital accumulate on-the-job, potentially at a different rate in a formal or informal job. It may depreciate while searching for a job. This setting is able to generate a very rich dynamic: not only it generates a mixture of formal and informal jobs with overlapping wage distributions but also allows for changes in formality status at the same job; not only it generates wage growth as a result of a job change but also produces wage growth on-the-job as a result of human capital accumulation.

We propose and implement an identification strategy for the structural parameters of the model using standard and representative labor market data for Mexico, an economy that share a significant informality rate with many other middle income economies. The parameters of the human capital accumulation and depreciation process are identified exploiting the (short) panel dimension available in the data set where the same individual is interviewed every quarter for five consecutive quarters. We condition on the human capital accumulated before entering the labor market by considering only the individuals with Secondary School completed and no more. The crucial information we use are wage changes within

and between jobs and transitions between labor market states. The estimation results show that human capital on the job accumulates at a lower rate when working informally. The difference is substantial: for first entrants, it will take on average three times longer to obtain the first human capital upgrade if working informally instead of formally. We also estimate that the contribution of the human capital accumulated after entering the labor market is important: about a third of overall production is due to this component of human capital. These findings speak to a recent and growing literature that seeks to explain differences in the wage profiles over the life cycle across countries [Lagakos et al., forthcoming]

We use the estimated model to perform policy experiments where we change two crucial parameters over a wide range of values: the contribution rate while working formally and the level of non-contributory benefit received while working informally and while searching. The results on the contribution rate confirm previous literature [Levy, 2008; Albrecht et al., 2009; Rocha et al., 2017; Bobba et al., 2017] but at the same time provide a novel channel by showing that an increase in the contribution rate not only reduces productivity by imposing a tax on more productive matches but also by reducing human capital upgrading on the job. The results of the policy experiments on the non-contributory benefit partially contradicts previous literature. As expected, a relatively high level of contribution incentivizes informality and reduces overall production by reducing human capital accumulation and generating lower productivity matches. However, this is not true for positive but lower levels of the benefit: a benefit at about half the current Mexican level, would not be enough to sway employees toward informality but would be enough to subsidize their search and generate better matches.

In conclusion, this paper focusing on human capital accumulation after entering the labor market reinforces the results we found in a companion paper focusing on human capital accumulation before entering the labor market [Bobba et al., 2017]. Labor market informality results from optimal reactions to distortionary labor market institutions. But the presence of such labor market state may magnify the negative impact of the distortions on labor market outcomes. A crucial channel of this magnifying effect is the endogenous accumulation of human capital.

References

Albrecht, J., Navarro, L. and Vroman, S. [2009], 'The Effects of Labour Market Policies in an Economy with an Informal Sector', *Economic Journal* **119**(539), 1105–1129.

Anton, A., Hernandez, F. and Levy, S. [2012], *The End of Informality in Mexico?*, The Inter-American Development Bank.

Bagger, J., Fontaine, F., Postel-Vinay, F. and Robin, J.-M. [2014], 'Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics', *American Economic Review* **104**(6), 1551–96.

Bianchi, M. and Bobba, M. [2013], 'Liquidity, Risk, and Occupational Choices', *Review of Economic Studies* **80**(2), 491–511.

Binmore, K., Rubinstein, A. and Wolinsky, A. [2006], 'The nash bargaining solution in economic modelling', *The RAND Journal Of Economics* **17**(1), 176–188.

Bobba, M., Flabbi, L. and Levy Algazi, S. [2017], Labor market search, informality and schooling investments, IZA Discussion Papers 11170, Institute for the Study of Labor (IZA).

Borowczyk-Martins, D., Bradley, J. and Tarasonis, L. [2018], 'Racial discrimination in the us labor market: Employment and wage differentials by skill', *Labour Economics* .

Bosch, M. and Esteban-Pretel, J. [2012], 'Job creation and job destruction in the presence of informal markets', *Journal of Development Economics* **98**(2), 270–286.

Busso, M., Fazio, M. and Levy, S. [2012], (in) formal and (un) productive: the productivity costs of excessive informality in mexico, IDB Working Paper Series IDB-WP-341, The Inter-American Development Bank.

Cahuc, P., Postel-Vinay, F. and Robin, J.-M. [2006], 'Wage Bargaining with On-the-Job Search: Theory and Evidence', *Econometrica* **74**(2), 323–364.

de Paula, Á. and Scheinkman, J. A. [2011], 'The informal sector: An equilibrium model and some empirical evidence from brazil', *Review of Income and Wealth* **57**, S8–S26.

Dey, M. and Flinn, C. [2008], 'Household search and health insurance coverage', *Journal of Econometrics* **145**(12), 43 – 63.

Eckstein, Z. and van den Berg, G. J. [2007], 'Empirical labor search: A survey', *Journal of Econometrics* **136**(2), 531–564.

Eckstein, Z. and Wolpin, K. I. [1995], 'Duration to first job and the return to schooling: Estimates from a search-matching model', *Review of Economic Studies* **62**(2), 263–86.

Fields, G. S. [1975], 'Rural-urban migration, urban unemployment and underemployment,

and job-search activity in LDCs', *Journal of Development Economics* **2**(2), 165–187.

Flabbi, L. and Mabli, J. [2018], 'Household search or individual search: Does it matter?', *Journal of Labor Economics* **35**(1).

Flabbi, L. and Moro, A. [2012], 'The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model', *Journal of Econometrics* **168**(1), 81–95.

Flinn, C. [2006], 'Minimum wage effects on labor market outcomes under search, bargaining and endogenous contact rates', *Econometrica* **73**, 1013–1062.

Flinn, C., Gemici, A. and Laufer, S. [2017], 'Search, matching and training', *Review of Economic Dynamics* **25**, 260–297.

Flinn, C. and Heckman, J. [1982], 'New methods for analyzing structural models of labor force dynamics', *Journal of Econometrics* **18**(1), 115 – 168.

Flinn, C. and Mullins, J. [2015], 'Labor Market Search and Schooling Investment', *International Economic Review* **56**, 359–398.

Frolich, M., Kaplan, D., Pages, C., Rigolini, J. and Robalino, D., eds [2014], *Social Insurance, Informality, and Labor Markets: How to Protect Workers While Creating Good Jobs*, number 9780199685233 *in* 'OUP Catalogue', Oxford University Press.

Guler, B., Guvenen, F. and Violante, G. L. [2012], 'Joint-search theory: New opportunities and new frictions', *Journal of Monetary Economics* **59**(4), 352 – 369.

Hornstein, A., Krusell, P. and Violante, G. L. [2011], 'Frictional wage dispersion in search models: A quantitative assessment', *The American Economic Review* **101**(7), 2873–2898.

Jovanovic, B. [1979], 'Job matching and the theory of turnover', *Journal of political economy* **87**(5, Part 1), 972–990.

Kanbur, R. [2009], 'Conceptualizing informality: Regulation and enforcement.', *Indian Journal of Labour Economics* **52**(1), 33–42.

Keane, M. P., Todd, P. E. and Wolpin, K. I. [2011], 'The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications', *Handbook of Labor Economics* **4**, 331–461.

La Porta, R. and Shleifer, A. [2014], 'Informality and development', *Journal of Economic Perspectives* **28**(3), 109–26.

Lagakos, D., Moll, B., Porzio, T., Qian, N. and Schoellman, T. [forthcoming], 'Life-cycle wage growth across countries', *Journal of Political Economy* .

Levy, S. [2008], *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic*

Growth in Mexico, Washington DC, The Brooking Institution Press.

Levy, S. and Schady, N. [2013], 'Latin America's Social Policy Challenge: Education, Social Insurance, Redistribution', *Journal of Economic Perspectives* **27**(2), 193–218.

Meghir, C., Narita, R. and Robin, J.-M. [2015], 'Wages and informality in developing countries', *The American Economic Review* **105**(4), 1509–1546.

Newey, W. K. and McFadden, D. [1994], Large sample estimation and hypothesis testing, Vol. 4 of *Handbook of Econometrics*, Elsevier, pp. 2111 – 2245.

Pakes, A. and Pollard, D. [1989], 'Simulation and the Asymptotics of Optimization Estimators', *Econometrica* **57**(5), 1027–1057.

Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D. and Saavedra-Chanduvi, J. [2007], *Informality : Exit and Exclusion*, number 6730 *in* 'World Bank Publications', The World Bank.

Pissarides, C. A. [2000], *Equilibrium Unemployment Theory*, Cambridge, MA, MIT Press.

Rocha, R., Ulyssea, G. and Rachter, L. [2017], Do lower taxes reduce informality? evidence from brazil, mimeo, PUC-Rio.

Ulyssea, G. [2015], Firms, informality and development: Theory and evidence from brazil, Technical report, Mimeo.

Wolinsky, A. [1987], 'Matching, search, and bargaining', *Journal of Economic Theory* **42**(2), 311–333.

Tables and Figures

Table 1: Descriptive Statistics, Cross-Section

Labor Market State	Proportions	Mean Hourly Wages	SD Hourly Wages
Formal Employees	0.613	24.178	10.714
Informal Employees	0.223	18.051	7.101
Self-employed	0.123	22.394	15.449
Unemployed	0.041	.	.

NOTE: Data extracted from the first quarters of 2013 and 2014 of the Mexican labor force survey (N=3,860). Wages for employees and incomes for self-employed individuals are reported in Mexican pesos (exchange rate: 1 US dollars $\approx$ 13.5 Mex. pesos in 2014). The Formal status of the job is defined according to whether or not workers report having access to health care through their employers.

Table 2: Yearly Transition Rates

LMK State Q5:	Formal Employees		Informal Employees		Self-empl.	Unempl.
Job change:	(No	Yes)	(No	Yes)		
LMK State Q1:						
Formal Employee	88.34		7.73		1.27	2.66
	(78.92	9.42)	(4.10	3.63)		
Informal Employee	18.63		68.92		8.73	3.73
	(11.76	6.87)	(49.01	19.91)		
Self-employed	5.26		20.42		72.00	2.32
Unemployed	47.80		23.27		10.69	18.24

NOTE: Stacked panel of individuals who were followed for five quarters starting in the first quarters of 2013 and 2014 of the Mexican labor force survey (N=19,300). The Formal status of the job is defined according to whether or not workers report having access to health care through their employers..

Table 3: Estimates of the Model Parameters

	Coefficient	Standard Error
Estimated Parameters		
$\lambda_{\{s=0\}}$	0.2995	0.01490
$\lambda_{\{s=1\}}$	0.0435	0.00098
$\eta_{\{f=0\}}$	0.0152	0.00017
$\eta_{\{f=1\}}$	0.0132	0.00006
μ_x	2.1217	0.00250
σ_x	1.1182	0.00076
μ_q	1.8950	0.00816
σ_q	0.6320	0.00167
$\gamma_{\{s=0\}}$	0.2027	0.00311
$\gamma_{\{s=1\}}$	0.0735	0.00294
$\tau_{\{f=0\},1}$	0.0160	0.00070
$\tau_{\{f=0\},2}$	-2.6241	0.02095
$\tau_{\{f=1\},1}$	0.0424	0.00229
$\tau_{\{f=1\},2}$	-2.7243	0.02391
c	0.1013	0.00113
ξ	-7.2292	0.10100
Predicted Values		
$E(x)$	15.59	
$SD(x)$	24.62	
$E(q)$	8.12	
$SD(q)$	5.69	
Number of Individuals	3860	

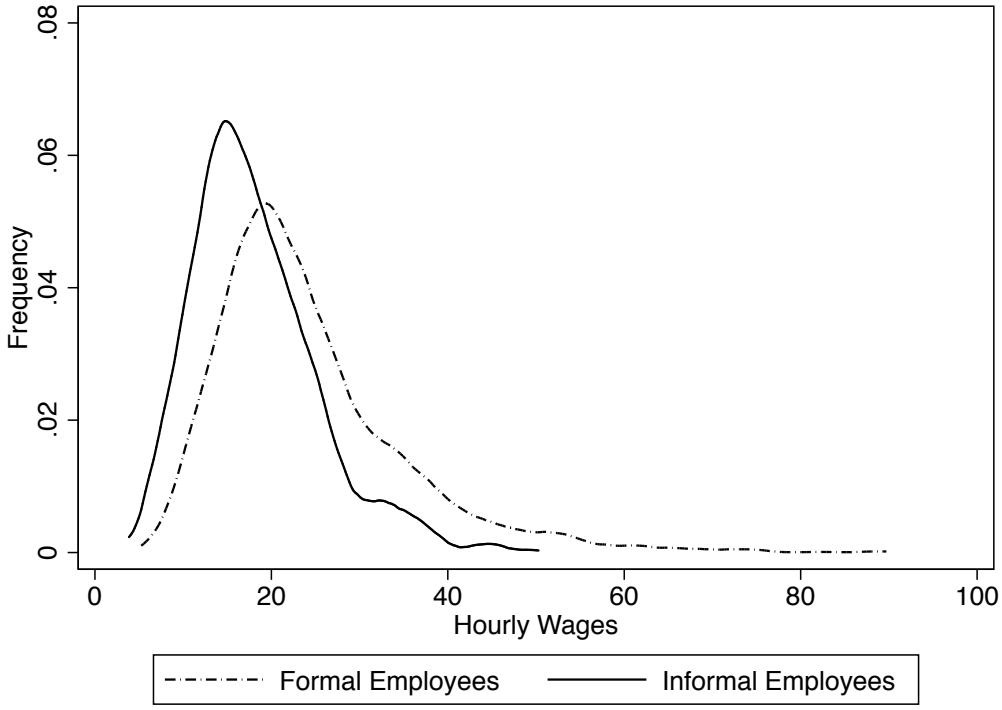
NOTE: Bootstrap standard errors reported. For the definition of the parameters, see Section 3.1 and Section 4.

Table 4: Output per Workers by Employment Status

	Proportion Over All Employees	Average Value of Production	Contribution of Human Capital
All Employees	1.0000	42.2084	0.3393
By Formality Status			
Formal Employees	0.7099	52.1555	0.3429
Informal Employees	0.2901	17.8710	0.3137
By Human Capital Level			
a_1	0.1391	24.6769	0.0000
a_2	0.2237	33.5628	0.1818
a_3	0.2359	39.5116	0.3077
a_4	0.1966	47.9012	0.4000
a_5	0.1204	56.6714	0.4706
a_6	0.0573	65.6566	0.5263
a_7	0.0217	70.5683	0.5714
a_8	0.0041	78.0250	0.6087
a_9	0.0010	51.1451	0.6400
a_{10}	0.0003	171.7176	0.6667

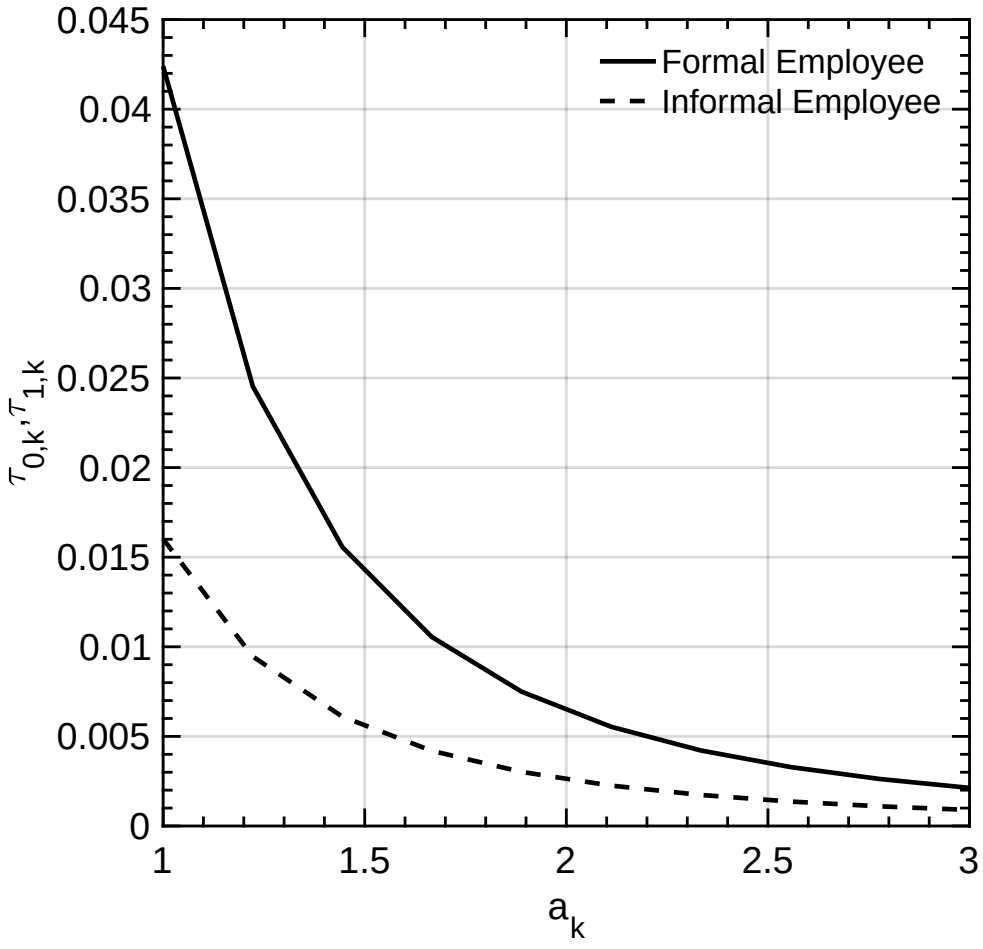
NOTE: Simulated samples of 5,000 worker-level observations for each quarter based on the estimates reported in Table 3. Let e_i be an indicator variable denoting the employee status (both formal and informal) for individual i in the simulated data, then the Average Value of Production can be expressed as $S(y) = \frac{1}{\sum_i e_i} \sum_i y_i$, while the average value of the match-specific productivity is given by $S(x) = \frac{1}{\sum_i e_i} \sum_i x_i$, and the Contribution of Human Capital can be expressed as $1 - \frac{S(x)}{S(y)}$.

Figure 1: Observed Wages, Density Functions



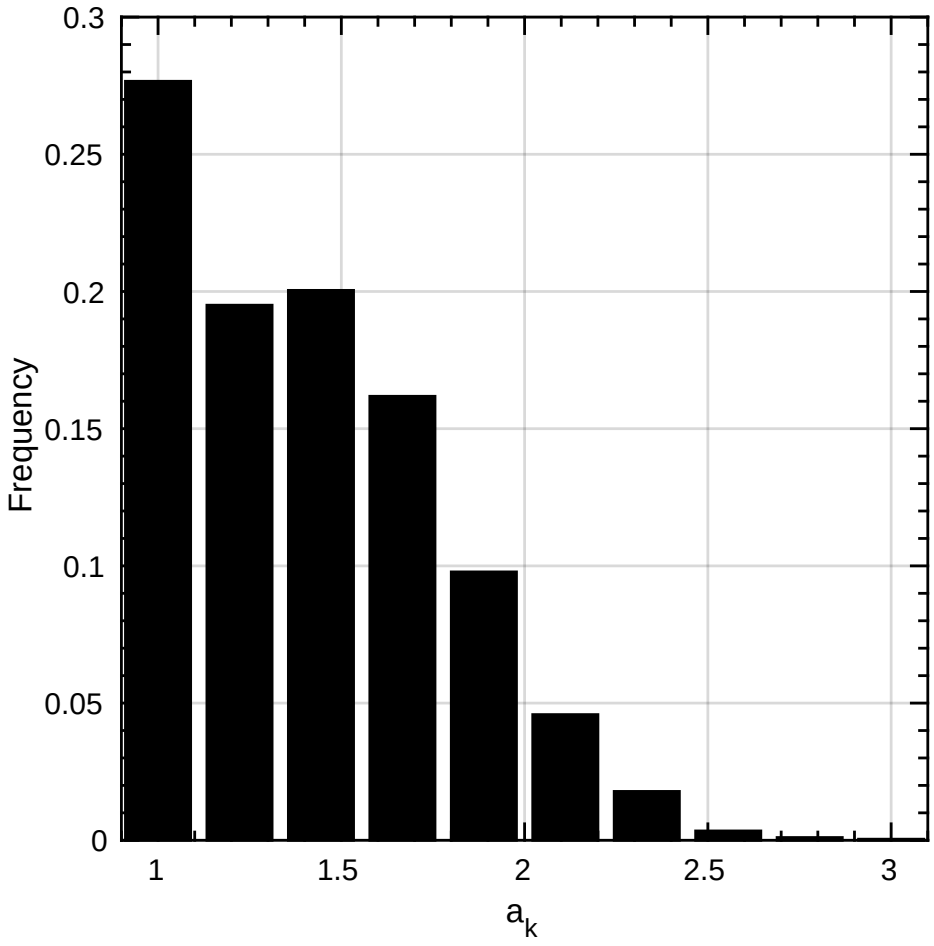
NOTE: Data extracted from the first quarters of 2013 and 2014 of the Mexican labor force survey (N=3,860). Wages for employees are reported in Mexican pesos (exchange rate: 1 US dollars $\approx$ 13.5 Mex. pesos in 2014). The Formal status of the job is defined according to whether or not workers report having access to health care through their employers.

Figure 2: Distribution of Arrival Rates of Human Capital Upgrading Shocks



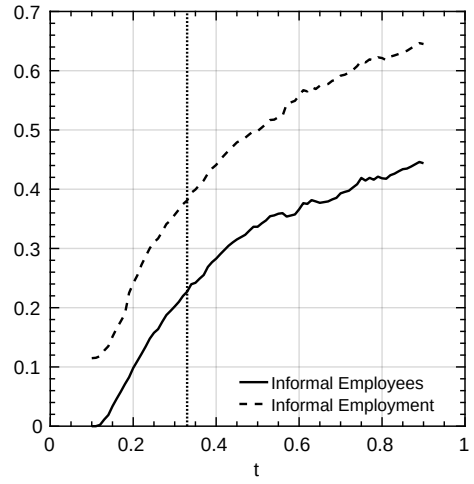
NOTE: Figure based on the estimates reported in Table 3.

Figure 3: Distribution of On-The-Job Human Capital

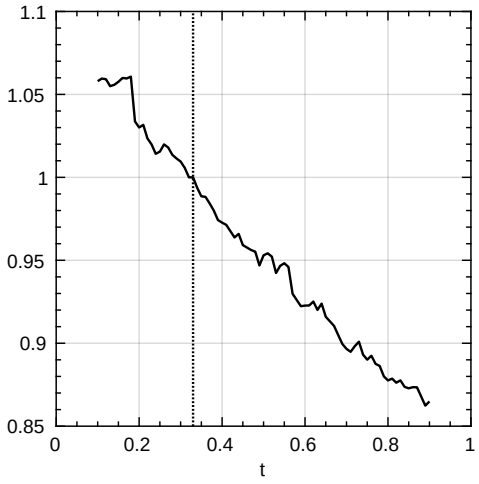


NOTE: Figures based on the estimates reported in Table 3.

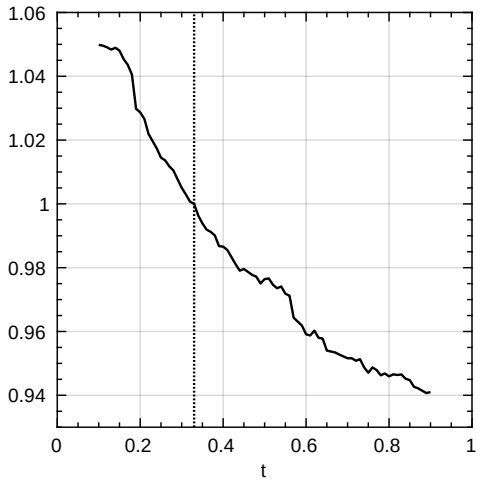
Figure 4: Impacts of Policy 1: Changes in the Contribution Rate t



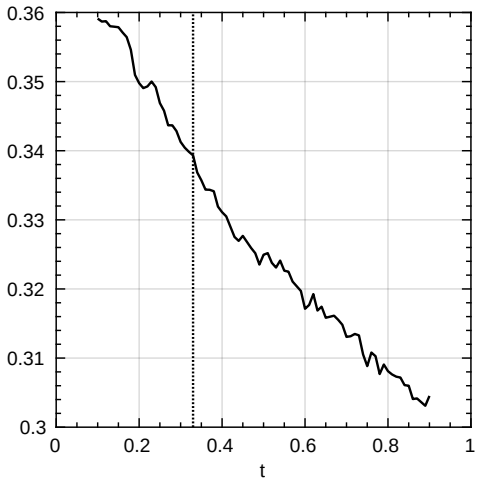
(a) Informality Rates



(b) Value of Production (Benchmark = 1)



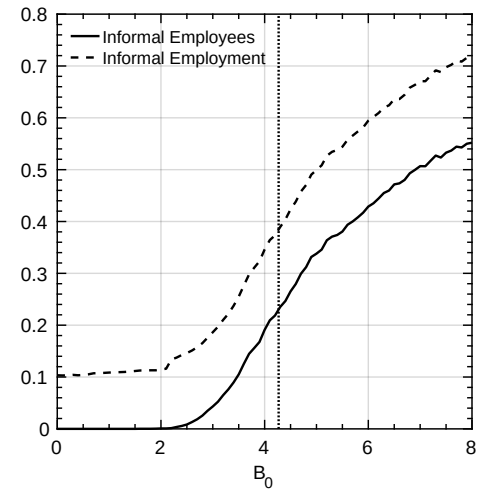
(c) Aggregate Human Capital (Benchmark = 1)



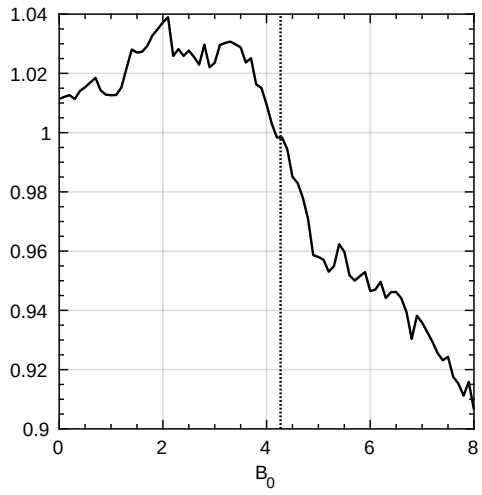
(d) Contribution of HC to Production

NOTE: Informal employment is the sum of informal employees and the self-employed. Simulated samples of 5,000 worker-level observations for each quarter based on the estimates reported in Table 3. The vertical lines are set at the institutional values for the Mexican labor market in 2013-2014. See Table B.1 for details.

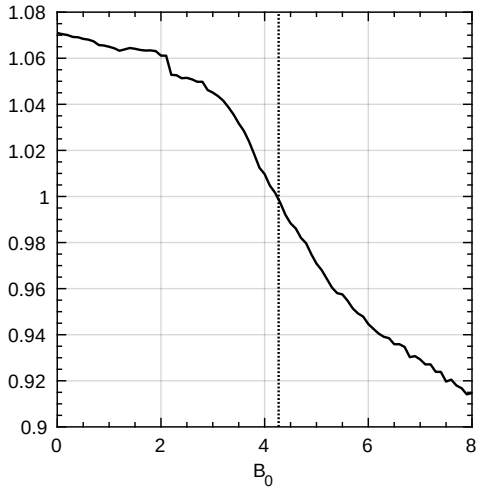
Figure 5: Impacts of Policy 2: Changes in the Non-Contributory Benefit B_0



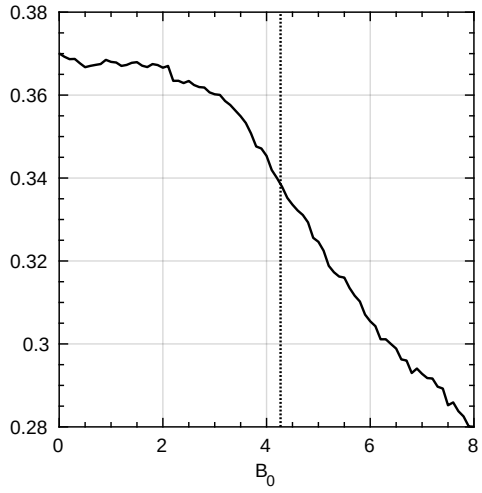
(a) Informality Rates



(b) Value of Production (Benchmark = 1)



(c) Aggregate Human Capital (Benchmark = 1)



(d) Contribution of HC to Production

NOTE: Informal employment is the sum of informal employees and the self-employed. Simulated samples of 5,000 worker-level observations for each quarter based on the estimates reported in Table 3. The vertical lines are set at the institutional values for the Mexican labor market in 2013-2014. See Table B.1 for details.

Appendix

A Additional Material on the Model

A.1 Wages

Wages are set by bargaining upon observing the the labor market human capital a_k, the outside option $V_s(k,q)$, the match-specific productivity x and the formality status posted by the firm f. We assume the axiomatic Nash-bargaining solution reported in equation (12). The resulting analytical expression for wages of employees hired formally and informally are:

$$\begin{aligned} w_1(x;k,q) &= \alpha(1+t)^{-1}\left[y(x,k)+\tau_{1,k}\max\{F_0(x;k+1,q),F_1(x;k+1,q),0\}\right] \\ &+(1-\alpha)(1+\beta_1\tau t)^{-1}\left[(\tilde{\rho}+\tau_{1,k})V_s(k,q)-\beta_1 b_1\right. \\ &\left.-\tau_{1,k}\max\{(1-f)E_0(x;k+1,q)+fE_1(x;k+1,q),V_s(k+1,q,h)\}\right] \end{aligned} \tag{A.1}$$

$$\begin{aligned} w_0(x;k,q) &= \alpha\left[(1-c)y(x,k)+\tau_{0,k}\max\{F_0(x;k+1,q),F_1(x;k+1,q),0\}\right] \\ &+(1-\alpha)\left[(\tilde{\rho}+\tau_{0,k})V_s(k,q)-\beta_0 B_0\right. \\ &\left.-\tau_{0,k}\max\{(1-f)E_0(x;k+1,q)+fE_1(x;k+1,q),V_s(k+1,q,h)\}\right] \end{aligned} \tag{A.2}$$

A.2 Numerical Solution and Simulation

We solve the model using value function iteration. We discretize the state space by using a grid of 100 equally spaced points in the interval $[0,150]$ for both x and y. The human capital distribution is already assumed discrete over 10 equally spaced points in the interval $[1,3]$. Since a match can change the formality status or a worker can decide to search for a new job when receiving an upgrading shock, all the value functions are dependent on each other and therefore the value function iteration is performed as a block. Specifically, we guess $V_s(k,q)$, $E_0(x;k,q)$, $E_1(x;k,q)$, $F_0(x;k,q)$ and $F_1(x;k,q)$ over the grid points in the state space and then we jointly iterate the Bellman's equations (9) to (11) (using the definitions of wages and profits) until convergence is achieved on these value functions. To approximate the integral in equation (9), we discretize the distribution $G(x)$ over the grid points of x (using the midpoint intervals between the grid points as support) and we compute the expected value as in a discrete probability distribution.

We simulate the model constructing labor market careers. To characterize all the optimal decisions involved in the dynamics of each career, we use direct comparisons between the

solved value functions. Since we discretize the state space, we use linear interpolation to approximate the value functions and wages off the grid. We simulate 5,000 individual careers for 540 months. Each individual is assigned a potential self-employment income q drawn from $R(q)$ and starts his career searching for a job with an initial human capital level equal to a_1. The lifetime duration is drawn from a negative exponential distribution with rate δ. The optimal decision in the search state with respect to being unemployed or self-employed is made comparing $V_0(k, q)$ and $V_1(k, q)$ given k and q. In the search state, individuals meet firms and receive downgrading shocks. The durations of these events are draws from negative exponential distributions with rates λ_s and $\gamma_{s,k}$, respectively. If the meeting with a firm occurs first, a productivity x is drawn and firms and individuals decide whether to complete the match and at what wage and formality status. If the match is realized, the individual leaves the searching state with a human capital level of a_k and if not, the searching process continues. If a downgrading shock hits, a new search process starts for the same individual but with human capital a_{k-1}. While working as employees, individuals receive termination and upgrading shocks. As before, we simulate a competing risk model where the durations of these events are draws from negative exponential distributions. In this case, the rates are η_f and $\tau_{f,k}$, respectively. If the termination shock arrives first, the individual starts a new search process with human capital a_k. On the contrary, if the upgrading shock arrives first, then the individual is upgraded to a_{k+1} and an optimal decision is made regarding whether to remain in the match and at what wage and formality status. This process continues until the arrival of the termination shock, which sends the agent back to search. Once the lifetime is complete, the individual dies and he is replaced by a new individual that starts his career with q potential self-employment income and with a_1 human capital level.

As time passes in the simulation, the distributions of the labor market states and the human capital levels stabilize, which means that the model has reached the steady state invariant distributions. For estimation, we use a panel of five quarters extracted from a time window in which these distributions are in steady state.

B Complete Estimation Results

Table B.1: Fixed Parameters

Parameter	Value	Source
α	0.5000	Symmetric Bargaining case [Binmore et al., 2006]
β_0	0.9082	Bobba et al. [2017]
β_1	0.6705	Bobba et al. [2017]
B_0	4.2700	Updated from Bobba et al. [2017]
ϕ	0.5500	Levy [2008]
t	0.3300	Anton et al. [2012]
b_1	4.5470	Based on average observed wages (see equation 18)
ρ	0.0500	Previous literature [Flinn and Heckman, 1982; Eckstein and van den Berg, 2007]
δ	0.0013	Based on average life of 65 years

Table B.2: Matched Moments: Cross-Section

Moments	Simulated	Data	Weight
Proportions:			
Formal Employee	0.554	0.613	126.91
Informal Employee	0.228	0.223	151.43
Self-Employed	0.154	0.123	199.80
Unemployed	0.064	0.041	309.00
Wages and Income:			
Formal Employee: Mean	14.107	14.826	4.20
Formal Employee: SD	16.753	14.459	5.25
Informal Employee: Mean	3.220	4.017	7.49
Informal Employee: SD	6.205	8.222	6.35
Self-Employed: Mean	2.162	2.756	6.96
Self-Employed: SD	5.692	9.135	2.02
Share Informal Employee:			
Quintile 1	0.679	0.451	44.14
Quintile 2	0.230	0.210	55.32
Quintile 3	0.071	0.169	59.37
Quintile 4	0.016	0.113	80.70
Quintile 5	0.004	0.058	108.60
Mean Wages by Quintiles:			
Formal Employee - Q1	12.400	12.944	5.84
Formal Employee - Q2	16.989	17.970	6.00
Formal Employee - Q3	21.192	21.834	4.99
Formal Employee - Q4	27.912	27.039	3.32
Formal Employee - Q5	48.819	41.109	1.63
Informal Employee - Q1	12.088	12.235	5.51
Informal Employee - Q2	16.804	17.883	4.30
Informal Employee - Q3	20.843	21.795	3.95
Informal Employee - Q4	27.026	26.617	2.62
Informal Employee - Q5	37.613	36.192	1.36

Table B.3: Matched Moments: Dynamics

Moments	Model	Data	Weight
Transition Probabilities (Yearly):			
Formal → Formal (Same Job)	0.834	0.789	122.08
Formal → Formal (New Job)	0.052	0.094	166.89
Formal → Informal (Same Job)	0.010	0.041	248.73
Formal → Informal (New Job)	0.027	0.036	260.01
Formal → Self-Employment	0.019	0.013	431.94
Formal → Unemployment	0.058	0.027	304.32
Informal → Formal (Same Job)	0.000	0.118	93.35
Informal → Formal (New Job)	0.045	0.069	120.48
Informal → Informal (Same Job)	0.833	0.490	57.79
Informal → Informal (New Job)	0.029	0.199	75.75
Informal → Self-Employment	0.025	0.087	106.78
Informal → Unemployment	0.067	0.037	158.61
Self-Employment → Formal	0.089	0.053	96.44
Self-Employment → Informal	0.030	0.204	54.43
Self-Employment → Self-Employment	0.880	0.720	49.05
Self-Employment → Unemployment	0.000	0.023	150.53
Unemployment → Formal	0.562	0.478	25.50
Unemployment → Informal	0.188	0.233	30.45
Unemployment → Self-Employment	0.022	0.107	40.69
Unemployment → Unemployment	0.227	0.182	32.32
Wage Growth Rates Within Jobs (Yearly):			
Formal Employee: Mean	0.031	0.045	144.02
Informal Employee: Mean	0.012	0.033	67.42
Wage Growth Rates Within Jobs by Quintiles (Yearly):			
Formal Employee: Mean - Q1	0.044	0.202	58.12
Formal Employee: Mean - Q2	0.037	0.090	61.37
Formal Employee: Mean - Q3	0.027	0.035	66.99
Formal Employee: Mean - Q4	0.023	0.004	71.53
Informal Employee: Mean - Q1	0.020	0.185	27.38
Informal Employee: Mean - Q2	0.013	0.090	32.51
Informal Employee: Mean - Q3	0.016	0.016	33.87
Informal Employee: Mean - Q4	0.005	-0.012	31.63
Wage Growth Rates Between Jobs:			
Less than a Quarter:			
Formal Employee: Mean	0.192	0.194	23.98
Informal Employee: Mean	-0.002	0.044	21.93
More than a Quarter:			
Formal Employee (with Unemployment in Between): Mean	0.035	0.048	67.15
Informal Employee (with Unemployment in Between): Mean	0.014	0.035	77.03
Formal Employee (with Self-Employment in Between): Mean	0.005	0.004	105.25
Informal Employee (with Self-Employment in Between): Mean	0.000	0.057	51.36