ARTICLE IN PRESS

Journal of Econometrics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Labor market search, informality, and on-the-job human capital accumulation

Matteo Bobba [a,*], Luca Flabbi [b], Santiago Levy [c], Mauricio Tejada [d]

[a] Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France
[b] Department of Economics, University of North Carolina – Chapel Hill, Chapel Hill, USA
[c] Brookings Institution, Washington, DC, USA
[d] Department of Economics (ILADES), Universidad Alberto Hurtado, Santiago, Chile

ARTICLE INFO

Article history:
Received 17 March 2018
Received in revised form 14 January 2019
Accepted 22 May 2019
Available online xxxx

JEL classification:
J24
J3
J64
O17

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

ABSTRACT

We develop a search and matching model where firms and workers produce output that depends both on match-specific productivity and worker-specific human capital. The human capital is accumulated while working but depreciates while searching for a job. Jobs can be formal or informal. The model is estimated on labor market data for Mexico. Human capital accumulation is responsible for more than half of the overall value of production, and upgrades more quickly while working formally than informally. Policy experiments reveal that human capital accumulation magnifies the negative impact on productivity of the labor market institutions that give rise to informality.

1. Introduction

Most labor markets in medium- and low-income countries are characterized by high levels of informality (see, e.g., La Porta and Shleifer (2014) and Levy and Schady (2013)). Informality refers to 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 health and retirement benefits for a large portion of the labor force. The advantage is the reduction in the negative employment effects caused by costly contractual arrangements between firms and workers.

If the presence of informality may be seen as an optimal response 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, identifying distortions in investment decisions as the main channel behind the correlation between productivity and formality status (de Paula and Scheinkman, 2011; Ulyssea, 2018). The literature focusing on the worker side is smaller and rarely takes into account human capital accumulation in the presence of high informality. 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

* Corresponding author.
E-mail addresses: matteo.bobba@tse-fr.eu (M. Bobba), luca.flabbi@unc.edu (L. Flabbi), santiagolevy4@gmail.com (S. Levy), matejada@uahurtado.cl (M. Tejada).

https://doi.org/10.1016/j.jeconom.2019.05.026
0304-4076/

particular, we look at human capital accumulation on-the-job, its possible depreciation while searching for a new job, and whether and how the formality status of the job significantly affects this dynamic. Recent evidence on lower wage profiles over the life-cycle in countries with a large informal sector (Lagakos et al., 2018) suggests that this relationship is potentially very relevant. Mexico is a good example of a middle-income country with a large informal sector. In our representative dataset, one third of all workers are informal and frequently transit in and out of formal jobs. About 20 percent of informal employees transit to a formal job within a year and about 10 percent of formal employees transit to an informal job within a year.

We develop and estimate a search and matching model where formal and informal jobs arise endogenously and where transitions between formality regimes occur both between and within jobs. Firms face monetary penalties for hiring informally. Workers and employers search for potential partners to enter a job relationship. When they meet, they observe the value of the match-specific productivity 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, and unemployed. The human capital evolution while participating in the labor market captures the additional productivity that may be acquired on the job. We allow this dynamic to depend on the formality status of a job. While off-the-job and searching (either as unemployed or as self-employed), individuals may even lose previously accumulated knowledge, leading to a depreciation of human capital.

We estimate the model using longitudinal data from Mexico's official labor force survey, the *Encuesta Nacional de Ocupación y Empleo* (ENOE). There are two main sources of variations in the data for the identification of the parameters governing the human capital dynamics: (i) transitions between jobs and labor market states; and (ii) wage growth within and between jobs, conditional on formality status. In the dataset available to us, we can observe both for a balanced panel of individuals for up to five quarters. We find that human capital accumulation on-the-job is important: in steady state, it is responsible for more than half of the overall value of production. Human capital upgrading is slower while working informally than formally: for first entrants in the labor market, it takes on average 1.4 years to start upgrading their human capital if they work formally and about two 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.

We use the estimated model to perform policy experiments focusing on the two parameters considered crucial in generating the high level of informality observed in Mexico and other countries in Latin America: the contribution rate paid by formal employees and the level of non-contributory social security benefits received by any non-formal employee.[1] Increasing the contribution rate leads to an increase in informality and a decrease in the stock of human capital. However, the negative impact on aggregate human capital is almost neutralized when the contribution rate increase is paired with a proportional increase in the benefit. Increasing the non-contributory benefit also leads to an increase in informality and a decrease in the stock of human capital but has a different, and major, impact on the selection of workers into formal jobs.

We contribute to the growing empirical literature on equilibrium search models by introducing and quantifying an additional mechanism behind wage growth over the life cycle: on-the-job human capital accumulation. Most estimated search models of the labor market impose constant wages for the same job. The main exceptions include models allowing for on-the-job search and wage renegotiation, such as Cahuc et al. (2006) and Dey and Flinn (2005). Very few introduce on-the-job human capital accumulation: notable exceptions are Yamaguchi (2010), Bowlus and Liu (2013) and Bagger et al. (2014).[2] Flinn et al. (2017) is the only paper that identifies accumulation of both general and specific human capital. The authors are able to achieve that by using information about training on-the-job, which is not available in our dataset.

We also contribute to the small literature estimating search models of the labor market in which informality arises endogenously. Meghir et al. (2015) is the only published work to have accomplished this result but it does not allow for human capital accumulation. Within the general literature of search and informality, we are unique in providing a theoretical foundation and an empirical implementation for the possibility of a change in formality status at the *same* job. This is a significant empirical regularity that is often neglected by the literature focusing on the supply side of the labor market. For example, Bosch and Esteban-Pretel (2012), Garcia (2015), Meghir et al. (2015) and Bobba et al. (2017) do not account for this type of labor market transition.

A key result of our policy experiments is that a higher contribution rate not only increases informality but also decreases human capital. This is a direct consequence of taking into account the dynamic of the human capital accumulation: more informality in steady state means a lower rate of human capital upgrading over time. This result creates a link with the literature on life-cycle labor supply with learning-by-doing. In policy experiments based on estimated models, Imai and Keane (2004) and Keane (2015) show that higher – respectively, temporary, and, as in our case, permanent – tax rates reduce human capital accumulation by inducing a dynamic feedback loop with labor supply.

The paper is organized as follows: Section 2 describes the institutional context and the data. Section 3 develops and discusses the model. Section 4 describes the identification of the model's parameters with the data available to us. Section 5 defines the estimation method and presents the estimation results. Section 6 reports the policy experiments, and Section 7 concludes. Additional material is available in a Web Appendix.

1 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).

2 Search-theoretic models of the labor market with human capital accumulation on the job include Rubinstein and Weiss (2006), Michelacci and Pijoan-Mas (2012), Menzio et al. (2016) and Herkenhoff et al. (2018).

2. Context and data

2.1. Institutional setting

We sketch a few prominent features of the Mexican labor market that are relevant for the analysis. For a more detailed description of the institutional setting, see Bobba et al. (2017). Informality is defined with reference to compliance with salaried labor regulations. In Mexico, as in most countries, firms are obligated to enroll salaried workers in the social security registry and pay a contribution proportional to workers' wages. 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 fines for each non-registered worker. Many firms operate in both the formal and informal sectors because they hire workers both legally and illegally.[3] This fact, together with the frequent and significant flow of workers who transit back and forth from formal to informal jobs (Maloney, 1999, 2004; Meghir et al., 2015), is in contrast with a segmented view of the labor market, where barriers restrict access to the formal sector. Yet, there is evidence to suggest that the returns to formal and informal jobs are potentially different.[4]

To the extent that there is no firm–worker relationship, labor market regulations do not apply to self-employed workers. For most of the individuals engaged in those activities, the notion of self-employment differs 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 an activity is working as a street vendor. Financial barriers to entering 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 (Feng et al., 2018).

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. 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 for the individuals in our sample (job-to-job transitions, wage growth within and between job spells, as well as changes in the formality status within the same job). Section A of the Web Appendix provides further details about the construction of the longitudinal sample employed in the analysis.[5]

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. We define the *self-employed* workers as those who declare: (i) not being in a subordinate relationship in their main occupation; and (ii) having their own business. In order to obtain a more homogeneous 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 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' decisions 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) actively searching for a job. Earning distributions are trimmed at the top and bottom 1 percent in each labor market state (formal employees, informal employees and self-employed).

The final sample used in our empirical analysis is a balanced panel dataset comprised of 4936 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 Fig. 1 depict the main cross-sectional facts; Table 2 reports statistics on labor market dynamics. The observed patterns

[3] Perry et al. (2007) show that in Mexico 50–70 percent of small–medium firms have used both formal and informal contracts simultaneously in a given point in time. Samaniego de la Parra (2016) finds on Mexican data between 2005 and 2016 that 25 percent of all employees at formal firms are actually hired informally. Ulyssea (2018) documents that in small formal firms in Brazil, 40 percent of workers are informal. At the same time, 52 percent of all informal workers are employed in large firms that are unlikely to be fully informal.

[4] World-Bank (2019) reports evidence from a variety of countries that the returns to experience for a worker are higher in the formal sector than in the informal sector. In emerging economies, on average, the earnings increase for an additional year of work for informal wage workers is 1.4 percent, whereas it is 1.8 percent for formal wage workers. Also, job training for active workers takes place largely in formal firms. Alaimo et al. (2015) document striking differences between the two sectors in the proportion of workers that receive on-the-job training during their work life.

[5] We restrict the sample to nonagricultural, male, private-sector workers between the ages of 20 and 55. We focus our analysis on workers at the mid-range of the skill distribution that is, those with at most secondary schooling completed. We thus drop from the sample those who did not complete middle school (i.e., below 9th grade) and those with at least some tertiary education (i.e., some College or more). We consider individuals with at most secondary schooling completed instead of primary or tertiary for three main reasons: (i) they are the most numerous, comprising more than half of the labor force in most Latin American countries, including Mexico (Bobba et al., 2012); (ii) they are significantly affected by informality (a statement which is true for primary but much less so for tertiary); (iii) they are at a skill level where human capital accumulation on the job is relevant (a statement which is clearly true for tertiary but more questionable for primary).

Table 1
Descriptive statistics, cross-section.

Labor market state	Proportions	Mean hourly wages	SD hourly wages
Formal employees	0.597	24.525	12.406
Informal employees	0.262	18.857	9.975
Self-employed	0.090	22.521	16.650
Unemployed	0.051	.	.

Note: Data extracted from the first quarters of 2013 and 2014 of the Mexican labor force survey (N = 4936). Wages for employees and incomes for self-employed individuals are reported in Mexican pesos (exchange rate: 1 US dollar ≈ 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		86.43			9.16		1.12	3.29
	(57.50		28.93)	(2.78		6.38)		
Informal employee		19.66			68.96		6.50	4.88
	(7.51		12.15)	(37.54		31.42)		
Self-employed		6.55			26.19		64.79	2.48
Unemployed		43.48			29.25		8.70	18.58

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 = 24,680). The Formal status of the job is defined according to whether or not workers report having access to health care through their employers.

Fig. 1. Observed wages, density functions. Note: Data extracted from the first quarters of 2013 and 2014 of the Mexican labor force survey (N = 4936). Wages for employees are reported in Mexican pesos (exchange rate: 1 US dollar ≈ 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.

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 percent of workers are employed formally and 35 percent informally. Among informal workers, two thirds are employees and one third self-employed. The unemployment rate is around 5 percent. 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 number of transitions between labor market states and formality regimes. Looking at the second row of Table 2, we observe that more than 30 percent of the informal employees change labor market status after a year. In the case of the most persistent state – formal employee – about 14 percent change labor market status after a year. Changes in formality status are also significant, with about 20 percent of informal employees becoming formal after a year. Fourth, and frequently neglected by the literature, changes in formality status may frequently occur at the *same* job.[6] Out of all the informal employees becoming formal within a year, almost 40 percent of them do so at the same job. Interestingly, the opposite is also taking place: out of the 9 percent of the formal employees becoming informal, 30 percent 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.

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 require time. When potential partners meet, they decide whether to enter 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 $x \sim G(x)$.[7] Since the productivity is match-specific, it is realized only upon meeting the employer. Therefore individual workers ex-ante identical may end up either in a formal or informal job.

Workers can be in one of four different labor market states: unemployed self-employed, informal employee and formal employee. 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.[8] 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.[9] 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). Notice that neither process results from explicit investment decisions but is a result of the worker's labor market state. In other words, choosing the labor market status means also choosing the human capital accumulation process. This approach is consistent with a learning-by-doing view of human capital accumulation.[10] While working on the job, the worker has the possibility to practice his skills and to learn, potentially leading to higher productivity. While off the job, the worker has less opportunities to practice his skills and may even lose previously accumulated knowledge, potentially leading to a depreciation of his human capital. We let the rate of human capital upgrading depend on the formality status of the job. This flexibility allow us to empirically study if human capital accumulation on the job is a channel through which the presence of informality imposes costs on the system. We also allow for a flexible specification of rates at which the shocks arrive in order to take into account that human capital upgrading may be harder the higher the level of human capital already acquired on the job.

We represent the evolution of human capital as movements along a discrete distribution of human capital values sorted in increasing order: $1 = a_1 < \cdots < a_K < \infty$. The total productivity of the match of a worker with labor market human capital a_k meeting a firm in a match generating x is:

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

[6] The main reason the previous literature has neglected this empirical regularity is the lack of direct observation. This paper and the work in progress (Samaniego de la Parra, 2016) are the first to directly document this fact using worker-level data. Other contributions have resorted to indirect evidence derived from firm-level data, such as in Levy (2018). Samaniego de la Parra (2016) uses a different dataset – matched employer–employee data – and looks at the different time period – from 2005 to 2016 – but finds transitions rates comparable to ours. As it is the case in our estimation sample, her computed quarterly transitions rates report higher transition from informal to formal than from formal to informal but her estimated differential is higher than ours.

[7] 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). Notice that this assumption imposes a restriction because the primitive match-specific productivity is the same for both formal and informal jobs. However, in estimation we allow all the other relevant labor market parameters (mobility parameters and human capital upgrading parameters) to be formality-specific.

[8] We rule out the possibility of receiving employee offers while working as an employee, i.e. there is no on-the-job search while working as a formal or informal employee. The main reason for this modeling assumption is data limitation. A good identification of on-the-job search parameters together with the related renegotiation mechanism requires a longer panel than the one available to us and typically needs information from matched employer–employee data as done, for example, in Cahuc et al. (2006) and Bagger et al. (2014). On top of the lack of a longer panel, we also lack relevant information on the relative short panel we actually observe. Specifically, we observe labor market status only quarterly so we cannot recover precisely the job-to-job transitions timing and the associated wages.

[9] 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.

[10] Seminal contributions are Arrow (1962) and Lucas Jr. (1993); for a recent review, see Thompson (2010).

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 worker 'upgrades' to an higher level, from the starting a_k to a new $a_{k'}$ with $k' > k$. 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 a lower level of a, denoted as $a_{k''}$ with $k'' < k$. Notice the limiting cases: $\tau_{f,K} = 0$ and $\gamma_{s,1} = 0$. Notice also that this human capital is only valuable while working as an employee but has no impact on self-employment income. This assumption is driven by the type of self-employment we are observing on our sample of medium- to low-educated individuals.[11]

In addition to 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 an employee (f) is posted by the firm 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 in most Latin American countries. Specifically, as we mention in Section 2.1, the legislation mandates the firm to be responsible to enroll the worker in the social security registry. It is also the firm that is legally bound to pay fines if this registration does not occur and if the correct amount of contributions is not collected. 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 heterogeneous in the population but time-invariant within individuals. The flow utility while searching as unemployed is homogeneous and denoted by ξ.

We follow previous literature by assuming linear utility.[12] We follow our previous work 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 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 is a contributory benefit; that is, 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 are a proportion ϕ of the total contribution. Formally, the benefit B_1 is defined as:

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

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.

On the demand side, employers post vacancy at no cost and earn revenues equal to $y(x, k)$ as defined in Eq. (1). They are only faced with labor costs, which include wages and social security contributions, when hiring formally. When hiring informally, they include wages and the probability, as well as the monetary penalties, of being caught. This setting leads to the following two flow profit definitions:

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

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

11 As we discuss in Section 2.1, self-employment in this education range mainly consists of very low-skill activities such as reselling modest quantity of food, drinks or clothing in public spaces. They are activities requiring some talent that may be heterogeneous in the population but they should be relatively unaffected by the human capital accumulated while working as an employee.

12 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 (2018).

The linear specification of the cost function for hiring informally is meant to capture the notion that imperfect enforcement creates a size-dependent distortion in the economy: larger firms face a significantly higher probability of being audited (see for example de Paula and Scheinkman (2011) and Ulyssea (2018)). The same empirical literature also points out that the larger firms are also the more productive firms. Since our model cannot incorporate firm size, we can only match the evidence by imposing a positive correlation between productivity and the cost of informality.

3.2. Wages and formality status

Before discussing the determination of wages and formality status, we introduce the notation for the value functions. The value functions definition is provided in Section B.2 of the Web Appendix. On the workers' side, we denote the searching states with V_s and the employee states with E_f; on the firms' side, we denote the value of a filled vacancy with F_f. Since we assume there is no cost of posting and keeping the vacancy open, we do not introduce notation for the value of an unfilled vacancy.[13]

The formality status decision is taken by the firm upon observing the labor market human capital a_k, the outside option $V_s(k, q)$, and the match-specific productivity x. The decision involves comparing the value 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:

$$w_f(x; k, q) = \arg\max_w \left[E_f(x; k, q) - V_s(k, q)\right]^{\alpha} \left[F_f(x; k, q)\right]^{(1-\alpha)}. \tag{9}$$

The solution is a quite involved analytical expression that we report in the Web Appendix (equations B.1 and B.2). However, 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 determining wages.

3.3. Equilibrium

3.3.1. Definition

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 whether to 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 B.3 in the Web Appendix). Since $V_1(1, q)$ is increasing in q faster than $V_0(1, q)$, there exists a unique:

$$q^*(1) : V_0(1, q^*(1)) = V_1(1, q^*(1)). \tag{10}$$

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 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 described in (10) but for the current a_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 as the worker and posts a formality status f. Given the formality status, both worker and employer engage in bargaining over wages – leading to the wage schedule defined in (9) – and decide whether to accept the match or not. Both firm and worker arrive at the same optimal decision thanks to the no disagreement result implied by Nash bargaining. Since the outside option for both agents is 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:

$$x_f^*(k, q) : F_f(x_f^*; k, q) = 0 \iff E_f(x_f^*; k, q) = V_s(k, q) \tag{11}$$

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

[13] 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).

The formality status is posted by the firm following the optimal decision rule described in Section 3.2. 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:

$$\tilde{x}(k,q) : F_1(\tilde{x}; k, q) = F_0(\tilde{x}; k, q). \tag{12}$$

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 ($f = 0$). Notice that $\tilde{x}(k,q)$ is determined by equating the firm's value functions because the formality status is posted by the firm.[14]

With the optimal decision rules in place, the equilibrium is defined by the set of value functions that satisfies equations (B.3)–(B.5) in the Web Appendix, 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 measure of workers in each labor market state and for the distribution of human capital. We solve the model numerically by value function iteration. Section B.3 of the Web Appendix provides a detailed description of our procedure.

3.3.2. Discussion

We highlight some features of the equilibrium that are useful to understand both the empirical implications of the model and the identification strategy with the data at our disposal.

A first crucial decision concerns the formality status. When firms post a formal job instead of an informal job, they trade-off the cost $tw_1(x; k, q)$ of contributing to maintain formal status with the cost $cy(x,k)$ of covering for the risk of being discovered hiring informally and having to pay a fine. The workers also face a trade-off when accepting to work formally: they are willing to give up some monetary wage in exchange for better benefits ($B_1 > B_0$). The combination of these two mechanisms, together with the determination of wages through bargaining, implies that the value function of a filled formal job is more sensitive to x than the value function on an unfilled formal job, generating the unique (for each k, q) reservation value $\tilde{x}(k,q)$ defined in Eq. (12). Exactly at $\tilde{x}(k,q)$, the formal wage is lower than the informal wage because the benefit is higher in the former than in the latter. This is true also in a neighborhood around $\tilde{x}(k,q)$ which is more or less large depending on the willingness to pay for the additional benefits, on the difference between contributory and non-contributory benefits and on the cost of informality c. For example, the higher the valuation of the formal benefit (the β_1), the larger the portion of wage the worker is willing to give up to work formally.[15]

Another important dynamic concerning the formality status is how it is affected by the human capital process. In other words, is the reservation value $\tilde{x}(k,q)$ increasing or decreasing in k? The answer is that the impact of the human capital upgrading process on $\tilde{x}(k,q)$ is ambiguous. The source of the ambiguity is that human capital upgrading is valuable under both formality statuses. The impact on both value functions $F_1(x; k, q)$ and $F_0(x; k, q)$ is therefore positive but in a non-linear way since wages are determined by bargaining, and human capital upgrading also changes the outside options. Looking at Eq. (12) – which defines $\tilde{x}(k,q)$ – this means that both the left-hand side and the right-hand side will increase as a result of human capital upgrading. At different points of the support of the match-specific distribution $G(x)$, one or the other will increase more, leading to an increase or a decrease of $\tilde{x}(k,q)$.

The second crucial decision is about accepting an employee job or not. The relevant reservation value is now $x_f^*(k,q)$, defined in Eq. (11). It is again unique (for each k, q) but it differs by formality status. In the most typical configuration, the reservation value to accept an informal job is lower than the one to accept a formal job; that is, $x_0^*(k,q) < x_1^*(k,q)$. When this is the case, the support of the match-specific productivity x is divided in three regions: the first for $x \in [0, x_0^*(k,q))$, the second for $x \in [x_0^*(k,q), \tilde{x}(k,q))$, and the third for $x \in [\tilde{x}(k,q), +\infty)$. This creates the following behavior: for low enough productivity (first region) the agent continues searching, for intermediate values of productivity (second region) the agent accepts to work informally, for high enough productivity the agent accepts to work formally.[16] Unlike the previous case, the impact of the human capital updating process on $x_f^*(k,q)$ is now unambiguous: since both value functions for a filled job increase in k while the value function for an unfilled vacancy does not depend on k, the $x_f^*(k,q)$ are decreasing in k. The economics intuition is straightforward, for given x, a higher k means a higher multiplicative factor in generating the overall productivity y (Eq. (1)) and therefore firms and workers will be more willing to accept the match.

The final crucial decision is about looking for an employee job as an unemployed or as a self-employed. The relevant reservation value is now defined over the labor income generated as self-employed, q. The reservation value is denoted by $q^*(k)$ and it is defined by equating the value of the two searching states (see Eq. (10) for an example with $k = 1$). The trade-off in this case is between receiving more employee offers by searching full-time (unemployment) and receiving less offers but earning income while searching (self-employment). Neither search efficacy nor self-employment income is

14 If formality posting were done by the worker or if formality status were part of a joint bargaining game with the wage, the threshold values would be different. Intuitively, firms and workers value formality at the margin differently because firms pay full value for the benefit but workers value the benefit at less (or more) than full value due to the preference parameters β_0 and β_1.

15 We discuss in detail this feature of the equilibrium in our previous contribution (Bobba et al., 2017) where we also provide a graphical interpretation of the result.

16 There is only another possible configuration: $x_0^*(k,q) > x_1^*(k,q)$. In this case, the agent will only accept formal jobs. See Proposition 1 in Bobba et al. (2017).

affected by human capital and therefore the impact of k on $q^*(k)$ works only through the continuation (or option) values. The impact is positive on the continuation values of both searching states leading to a situation similar to the one observed about the formality status: the impact is ambiguous, and transitions between unemployment and self-employment are possible in both directions.

3.3.3. 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, including Mexico (see Section 2).

First, the model's equilibrium can generate a positive mass of workers in each labor market state and produce the significant number 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. 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 $a_{k'}$, with $k' > k$. The upgrading may be such that the new reservation value to work formally is now lower than the match-specific productivity x_i:

$$\tilde{x}(k', q) \leq x_i,$$

since it is possible that $\tilde{x}(k', q) < \tilde{x}(k, q)$. 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. Notice that, as discussed in Section 3.3.2, the impact of the human capital upgrading process on $\tilde{x}(k, q)$ is ambiguous and therefore both transitions from informal to formal and from formal to informal may take place at the same job.

Second, the model is able to generate wage distributions in line with the data. The data show two main features: the average wages of formal employees are on average higher than the average wages of informal employees, and the two distributions significantly overlap. As we discussed in Section 3.3.2, both results are a direct consequence of the property of the equilibrium.

Third, the model is able to generate wage growth not only across jobs – as common in related literature – but also within jobs. The reason for this 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 to $a_{k'}$, with $k' > k$. Further assume that x_i is such that $\tilde{x}(k', q_i) < x_i$. Then the worker will remain matched with the same employer and with the same formality status but his wage will increase from $w_1(x_i; k, q_i)$ to $w_1(x_i; k', q_i)$.

4. Identification

The model is characterized by the following parameters:

$$\left\{\rho, \delta, \omega_{f,1}, \omega_{f,2}, \gamma_s, \lambda_s, \eta_f, \xi, \alpha, c\right\}_{f\in\{0,1\}, s\in\{0,1\}}, \tag{13}$$

and by the following distributions:

$$\left\{G(x), R(q), \{a_k\}_{k=1}^K\right\}. \tag{14}$$

In addition, we have to set the parameters that characterize the institutional setting:

$$\{\beta_0, B_0, \phi, t, \beta_1, b_1\}. \tag{15}$$

We split the identification discussion in two parts. We first focus on the preferences for social security benefits, the cost to firms of hiring informally and the usual search, matching and bargaining parameters. We then consider the identification of the parameters describing the novel feature of our model: the human capital dynamic while working in the labor market.

4.1. Labor market parameters

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.[17] The non-contributory benefit B_0 is calibrated from aggregate data following the same procedure described in Bobba et al. (2017), and 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 employees after collecting their individual contributions (b_1) is estimated from

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

the data by assuming that the formal system runs a balanced budget. Denoting with i a generic observation in our sample, the estimator is:

$$\hat{b}_1 = t(1-\phi)\sum_{i\in N_{E_1}} \frac{w_1(i)}{N_{E_1}}, \tag{16}$$

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

With the institutional parameters in place, Bobba et al. (2017) propose 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 an 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 quasi-random roll-out of a non-contributory social program (the *Seguro Popular* program) concludes the identification strategy we proposed there.[18] 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 $\left\{\rho, \lambda_s, \eta_f, \xi, \alpha\right\}$ 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.[19] The recoverable distribution we assume for $G(x)$ is a lognormal with parameters that we denote $\{\mu_x, \sigma_x\}$.[20] 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 percent a year and recovering ξ by exploiting the equilibrium equation (B.3). 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.[21] In addition to the inter-temporal discount rate, 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. 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 $\{\mu_q, \sigma_q\}$.

4.2. Human capital parameters

The human capital dynamic while participating in the labor market is characterized by upgrading on the job and 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; that is, 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 a 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 $a_K = 5.5$ 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 equal to three times their match-specific productivity draw. The maximum productivity boost is equal to 5.5 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 $\tau_{f,k}$: when the shock hits, it induces renegotiation of the job relationship. The renegotiation has potential impacts on three observables: wages, that may change as a result of

---

18 For a more detailed and more formal discussion of this identification strategy, see Section 4 of Bobba et al. (2017).

19 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 it in a different way.

20 Virtually all the search–matching–bargaining literature assume log-normality, from the seminal Eckstein and Wolpin (1995) to our recent Bobba et al. (2017).

21 Recent works setting α using a similar strategy include Flabbi and Moro (2012) and Borowczyk-Martins et al. (2018).

bargaining; formality status, that may change as a result of a different productivity level; and labor market status, that may change because firms and workers may decide to dissolve the match. Wage growth on-the-job, formality status changes at the same job and transitions between labor market states are therefore the information providing the identification.

In the data at our disposal, we can observe all of the elements of the labor market dynamic just described but 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 allow for a dependency between the arrival rate and the human capital level:

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

The functional form reduces the number of parameters from $(K-1)$ to 2 for each formality status f but still captures that human capital upgrading may be a function of current levels of human capital.[22] For example, a positive $\omega_{f,1}$ combined with a negative $\omega_{f,2}$ implies that the probability to upgrade at a low a_k is higher than at a high a_k. This is consistent with decreasing returns in human capital accumulation and it is actually what we find in estimation *without* imposing any sign constraints.

Another important dimension of the human capital upgrading shock is the extent of the upgrading. In terms of our parametrization, it is equivalent to ask how much higher is k' with respect to the starting k. Wage growth and formality status changes within the same job are valuable information to identify this dynamic: longer 'jumps' imply more wage growth and a higher probability to change formality status. As we will show in Section 5.3, the second event is not infrequent in our data, in particular when looking at switches from informal to formal within the same job. The additional flexibility allowed by different jump's lengths for given upgrading shocks is crucial in matching this dimension of the data.[23] Another source of the separate identification of the parameters governing the arrival rate of human capital upgrades and the parameters governing the size of those upgrades is the distribution of the wage growth. For example, frequent and small jumps will lead to a wage growth distribution with a lot of probability mass right above zero and very rare changes in formality status, the opposite would be true for infrequent but large jumps. Still, both sets of parameters are jointly contributing to this dynamic therefore we keep a tight parametrization in estimating the distribution of the jumps' length: let m be the size of the jump in the human capital grid, we assume that $m \sim Q_f(m; \nu_f)$ with $Q_f(\cdot)$ being a negative exponential distribution with parameter ν_f. In Section B.3 of the Web Appendix, we provide additional details on how we implement the truncation and discretization of this distribution in the simulations used in estimation. The discretization of the continuous exponential draws is needed because we have assumed a discrete grid for the support $\{a_k\}_{k=1}^{K}$.

In the case of the human capital downgrading shock, we dispose of more limited information since – as common in most standard labor market datasets – we do not observe much about the searching process. Specifically, we observe durations and transitions over the searching states but we do not observe either the number or the amount of offers actually received. The downgrading shock during the search makes the worker more willing to accept jobs, a fact that should be reflected in durations and transition rates. But durations and transition rates are also crucial to identify the arrival rate of offers: we need additional information to separately identify the downgrading shock from the job offer 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 – five quarters – but we see a significant number of individuals quitting their job, searching, and finding another job all within our observation window. Comparing wages accepted before and after a search episode is informative about the shocks received during the episode. 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. Moreover, a downgrading shock may also induce a change in searching state. While this is a relatively rare event in the data, it is very valuable in terms of identification because it signals that a depreciation shock has taken place with probability one.

If this information is enough to identify the frequency of the shock while searching, it is not rich enough to identify different shock frequencies for different human capital levels. As a result, we assume that the downgrading shock only depends on the searching state s:

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

[22] 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).

[23] Even if our proposed model environment can account for the whole labor market dynamic just described, there could be other explanations for it. A prominent one is the presence of on-the-job search together with a renegotiation mechanism determining wages when offers on the job are received, a mechanism we rule out due to data limitation (see footnote 8). Dey and Flinn (2005), Cahuc et al. (2006) and a number of follow-up papers propose a renegotiation mechanism where a new firm and the incumbent firm engage in Bertrand competition for the services of the worker. In so doing, firms may transfer some surplus to the worker increasing her wage even within the same job. However, an on-the-job search model with renegotiation will have difficulty in generating changes in formality status. Without human capital accumulation, the reservation value for switching between formal and informal is fixed. Formally, in our model the reservation value $\tilde{x}$ is a function of both k and q (Eq. (12)), without human capital accumulation it is only a function of q. As a result, either the incumbent firm and worker match is within the match range for informal job or it is within the range for formal job: the presence of a new firm bidding for the services of the worker does not change this result.

We also fix the downgrading length at one step: every time that a downgrading shock is received, human capital moves from the current a_k to a_{k-1}.

5. Estimation

5.1. Method

We estimate the parameters of the model, denoted by the parameter vector Θ, using the Method of Simulated Moments (MSM)[24]:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \; [M_R(\Theta) - m_N]' \, W^{-1} \, [M_R(\Theta) - m_N] \,, \tag{17}$$

where: m_N is an appropriately chosen set of statistics derived from our data sample of size N; $M_R(\Theta)$ is the corresponding set of simulated statistics extracted from a sample of size R obtained from the steady state equilibrium implied by Θ[25]; and 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.[26]

The different sets of moments $M_R(\Theta) - m_N$ that enter in the quadratic form (17) are chosen in order to capture the data features described in Section 4. We match the proportion of workers in each labor market state in 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). We use mean and standard deviation of employees' wages and self-employment labor income in the first quarter in order to describe the wage information. To describe the overlap between the accepted wages in formal and informal jobs, we build quintiles over the distribution of accepted wages for formal workers. For each quintile, we compute the mean wages of formal and informal employees and calculate the proportion of employees in informal jobs. To describe the wage dynamic, we consider the 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. We compute average wage growth both on the overall sample and by quintiles. The final estimators include 62 moments and 18 parameters.

To assess the reliability of our estimator, we perform a Monte Carlo procedure in which the point estimates obtained by applying our estimation procedure on the original data are compared to the point estimates obtained by applying the same estimation procedure on synthetic data generated by known parameters. We find them close enough to lend credibility to our estimation method. Details and results on the Monte Carlo procedure are reported in Section C.1 of the Web Appendix.

5.2. Results

The estimated parameter values are reported in Table 3, together with the bootstrap standard errors.[27] The parameters calibrated to values derived from the Mexican institutional setting are reported in Section C.2 of the Web Appendix. There is a large difference in arrival rates between the unemployed and the self-employed, explaining in part the observed persistency in the self-employment state and the high-turnover in the unemployment state. The estimated parameters of the job destruction rates imply average durations of 17 months in informal jobs and of 31 months in formal jobs.[28] Out of those total destruction rates, 95.7 percent are due to exogenous separations between firms and workers and 4.3 percent are due to endogenous quits following human capital updating. The majority of workers do not quit as a result of human capital updating because their employee status is renegotiated: both the formality status and the wage may change to reflect the higher productivity generated by the human capital increase.

The match-specific productivity distribution is estimated to generate a lower average value than the one we obtained for Mexico in Bobba et al. (2017). The difference is explained both by data, more recent in the current paper, and by modeling differences. The main modeling feature responsible for the difference is that the total productivity of the match between a worker and a firm is augmented by the worker's human capital (Eq. (1)) in the current paper while it had to be fully explained by the primitive match-specific productivity in Bobba et al. (2017).

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 by 10 percent their stock of human capital every half a year. The depreciation rate is much

24 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 (17), 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).

25 We set R at 5000, which is slightly larger than the sample size N.

26 We build W by replacing the diagonal of an identity matrix with the bootstrapped sample variances of the sample moments.

27 Bootstrap standard errors are based on 35 replications. Each replication is started from the point estimates.

28 Although we do not directly observe employment durations in our sample, these numbers are broadly comparable with existing estimates from Mexico and other Latin American countries (see, e.g., Alaimo et al. (2015)). They confirm the presence of high turnover and churning in labor markets characterized by high informality rates, causing workers to change jobs frequently (and to frequently transit between formal and informal employment).

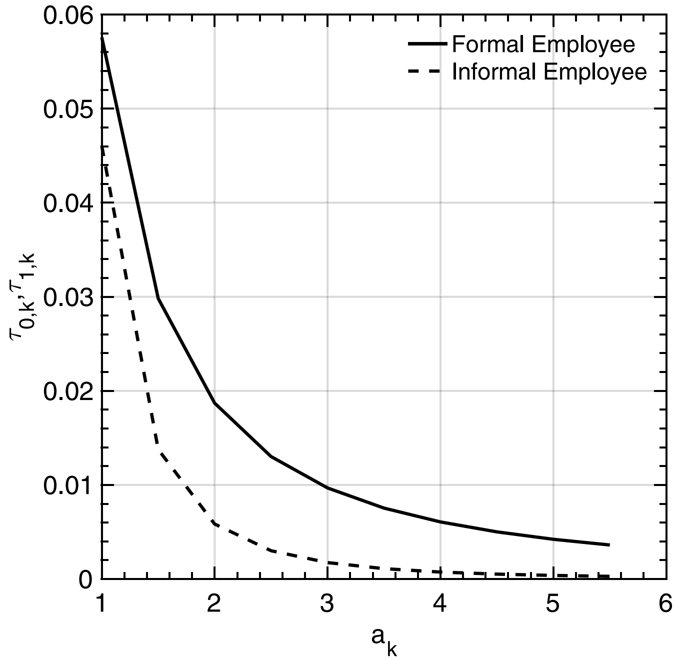


Fig. 2. Distribution of arrival rates of human capital upgrading shocks. NOTE: Figure based on the estimates reported in Table 3.

lower during spells of self-employment: as it takes approximately 1.8 years to decrease human capital by the equivalent amount of 10 percent. The estimated values for the arrival rates of the human capital upgrading shocks $\{\omega_{f,1}, \omega_{f,2}\}$ reveal concave patterns in the expected time of arrival of these shocks, which are depicted in Fig. 2. The rate of human capital upgrading is estimated to be slower while working informally than formally, at any level of human capital. For individuals at the lower bound of the human capital support, it takes about 1.4 years to start upgrading their human capital if they work formally and about two years if they work informally. At the average value of the distribution of human capital, it takes about 5.2 years to upgrade while working formally and 20 years to upgrade while working informally.

The observed advantage in the human capital accumulation process while working formally is partially offset by the size of the upgrade when the shock hits. As shown by the estimated values for the parameters ν_0 and ν_1 in Table 3, the average size of the jump is larger when working informally, as it amounts to two discrete intervals of the distribution of human capital or 20 percent of the support when compared to 0.7 of an interval or 3.5 percent of the support when working formally. The final result of the upgrading and downgrading human capital process in equilibrium is the steady state distribution reported in Fig. 3. The majority of workers possess some positive human capital: for example, about 15 percent will double the productivity of the match-specific productivity when working as employees ($a_k = 2$).

Table 4 reports some statistics on productivity and human capital implied by these estimates. The average worker's productivity (second column) increases steeply with the level of human capital. This is partially due to the selection over the match-specific productivity and to the effect of a higher level of human capital. The relative contribution of human capital on overall productivity is presented in the third column and it is estimated at about 60 percent in the overall sample, with a value of about 62 percent when working formally and a value of about 50 percent when working informally. The relative contribution of human capital is monotonically increasing in its level, reaching more than 80 percent at the upper bound of the human capital values.

The estimate of the cost of hiring informally (the parameter c) is roughly 5 percent of job productivity. This 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 it is approximately 0.85 pesos per hour – it is still lower than the cost of hiring formally. As a comparison, the payroll tax rate applied to the formal wage that corresponds to the same productivity level would be 3.26 pesos per hour.[29]

5.3. Model fit

Tables 5 and 6 report the complete set of moments targeted by the MSM estimator. A relevant feature of labor markets with high informality is the substantial overlap in the formal and the informal accepted wage distributions. As shown in

[29] Notice that these are only direct costs, i.e. they do not take into account that through bargaining firms are able to partially transfer costs to the workers, as seen in the equilibrium wage schedules (9).

Table 3
Estimates of the model parameters.

	Coefficient	Standard error
Estimated parameters		
$\lambda_{\{s=0\}}$	0.5051	0.00101
$\lambda_{\{s=1\}}$	0.0782	0.00054
$\eta_{\{f=0\}}$	0.0573	0.00008
$\eta_{\{f=1\}}$	0.0317	0.00003
μ_x	1.6835	0.00424
σ_x	1.0099	0.00003
μ_q	1.5733	0.00668
σ_q	0.9464	0.00193
$\gamma_{\{s=0\}}$	0.1617	0.00026
$\gamma_{\{s=1\}}$	0.0472	0.00030
$\omega_{\{f=0\},1}$	0.0460	0.00008
$\omega_{\{f=0\},2}$	−2.9775	0.01376
$\omega_{\{f=1\},1}$	0.0576	0.00014
$\omega_{\{f=1\},2}$	−1.6230	0.00344
c	0.0514	0.00039
$\nu_{\{f=0\}}$	0.4958	0.00107
$\nu_{\{f=1\}}$	1.3988	0.00126
ξ	−8.9533	0.02147
Predicted values		
$E(x)$	8.9664	0.00148
$SD(x)$	11.9388	0.00265
$E(q)$	7.5473	0.00344
$SD(q)$	9.0853	0.00850
Value of the loss function	3665	
Number of individuals	4936	
Number of observations	24 680	

Note: Estimates obtained via the Method of Simulated Moments using the downhill simplex (Nelder–Mead) algorithm to minimize the quadratic form (17). Standard errors calculated with 35 bootstrap replications are reported in parenthesis. The complete set of 62 sample moments and the corresponding simulated moments and weights used in the quadratic form (17) are reported in Table C.2 of the Web Appendix. For the definition of the parameters, see Sections 3.1 and 4.

Table 4
Output and contribution of human capital.

	Proportion over all employees	Average value of production	Contribution of human capital
All employees	1.0000	42.8322	0.6052
By formality status			
Formal employees	0.6965	54.0746	0.6171
Informal employees	0.3035	17.0267	0.5183
By human capital level			
a_1	0.1172	16.6469	0.0000
a_2	0.1350	26.9492	0.3333
a_3	0.1668	33.3405	0.5000
a_4	0.1882	41.7311	0.6000
a_5	0.1558	50.2172	0.6667
a_6	0.1131	59.6796	0.7143
a_7	0.0589	63.2288	0.7500
a_8	0.0351	77.0307	0.7778
a_9	0.0200	85.4184	0.8000
a_{10}	0.0100	112.4372	0.8182

Note: Simulated samples of 5000 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 $1 - \frac{S(x)}{S(y)}$.

the bottom panels of Table 5, we are able to replicate the overlap in the data quite well, both in terms of the proportions of informal employees in each quintile and in terms of the mean accepted wages by quintiles.

However, matching the overlap comes at a cost. The parameter estimates able to fit the overlap imply exit rates from the searching states that are lower than in the sample, producing a poor fit on the proportions of workers in the searching

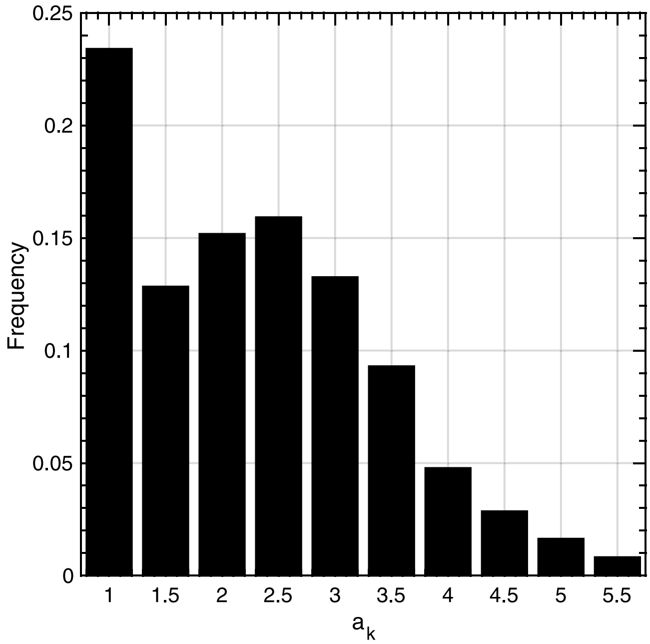


Fig. 3. Distribution of human capital. Note: Figure based on the estimates reported in Table 3.

Table 5
Matched moments: Cross-section.

Moments	Simulated	Data	Weight
Proportions:			
Formal employee	0.518	0.597	144.99
Informal employee	0.224	0.262	161.05
Self-Employed	0.163	0.090	248.09
Unemployed	0.095	0.051	324.51
Wages and income:			
Formal employee: Mean	27.018	24.525	4.78
Formal employee: SD	18.042	12.406	4.26
Informal employee: Mean	14.316	18.857	7.33
Informal employee: SD	8.139	9.975	3.89
Self-Employed: Mean	18.155	22.521	8.61
Self-Employed: SD	13.217	16.650	1.77
Quintiles — Proportions:			
Informal employee — Q1	0.592	0.401	55.67
Informal employee — Q2	0.174	0.247	60.04
Informal employee — Q3	0.127	0.159	69.81
Informal employee — Q4	0.091	0.118	88.63
Quintiles — Mean wages:			
Formal employee — Q1	12.000	12.517	7.03
Formal employee — Q2	16.990	17.560	6.29
Formal employee — Q3	22.433	21.708	5.67
Formal employee — Q4	29.950	27.158	3.39
Informal employee — Q1	8.860	11.620	6.49
Informal employee — Q2	16.831	17.386	5.33
Informal employee — Q3	22.219	21.444	4.29
Informal employee — Q4	29.400	27.096	2.47

states. As reported in the first panel of Table 5, we overestimate the proportion of workers in self-employment and unemployment; as reported in the bottom of the first panel of Table 6, we also overestimate the persistency in each of the two states. To better match the proportions and transitions in the searching states, we would need to estimate a model with lower reservation wages but such lower reservation wages would generate a worse fit on the overlap and on the accepted wages distributions. Since self-employment and unemployment are the two least important labor market

Table 6
Matched moments: Dynamics.

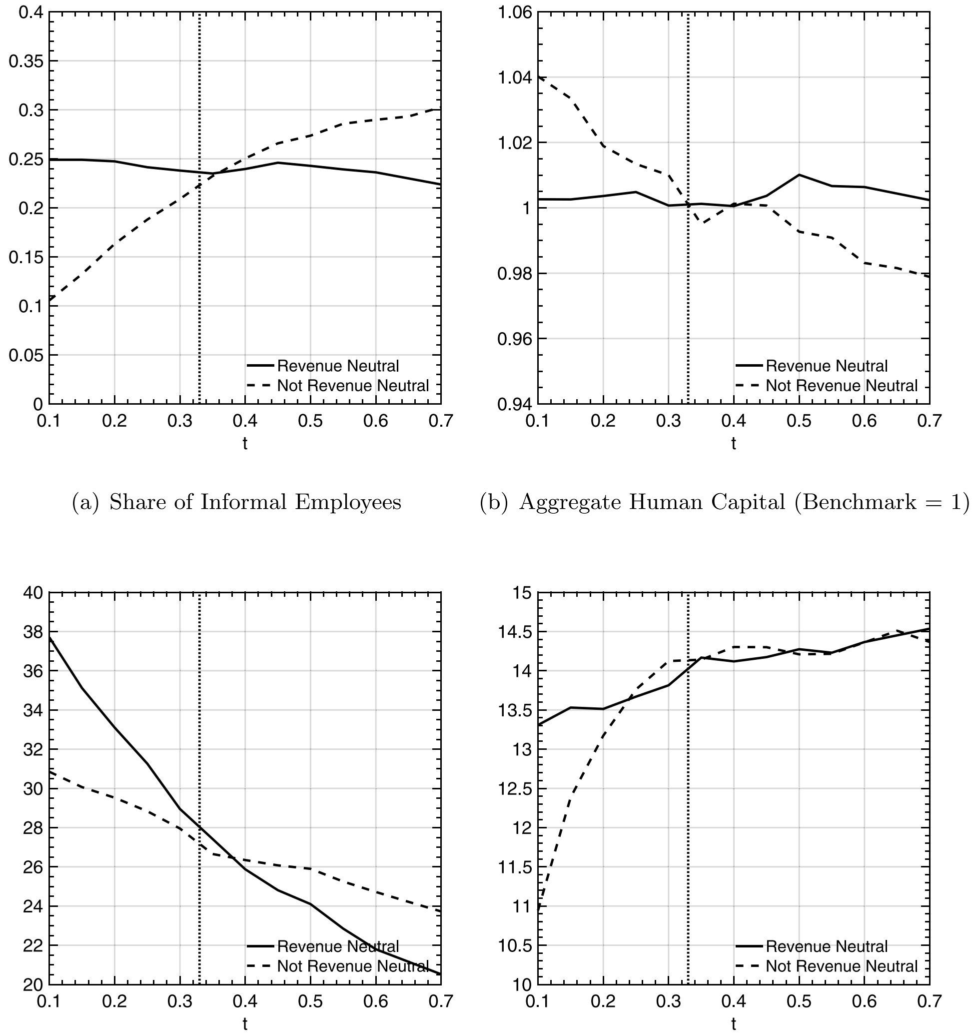
Moments	Model	Data	Weight
Transition probabilities (Yearly):			
Formal → Formal (Same Job)	0.667	0.575	111.02
Formal → Formal (New Job)	0.138	0.289	120.20
Formal → Informal (Same Job)	0.003	0.028	328.69
Formal → Informal (New Job)	0.069	0.064	220.70
Formal → Self-Employment	0.025	0.011	527.15
Formal → Unemployment	0.098	0.033	300.31
Informal → Formal (Same Job)	0.032	0.075	142.92
Informal → Formal (New Job)	0.180	0.122	109.33
Informal → Informal (Same Job)	0.471	0.375	75.15
Informal → Informal (New Job)	0.129	0.314	78.87
Informal → Self-Employment	0.066	0.065	142.15
Informal → Unemployment	0.122	0.049	168.98
Self-Employment → Formal	0.067	0.065	83.04
Self-Employment → Informal	0.098	0.262	47.43
Self-Employment → Self-Employment	0.833	0.648	43.84
Self-Employment → Unemployment	0.001	0.025	133.74
Unemployment → Formal	0.451	0.435	31.99
Unemployment → Informal	0.334	0.292	36.91
Unemployment → Self-Employment	0.015	0.087	57.04
Unemployment → Unemployment	0.200	0.186	40.36
Wage growth rates within jobs (Yearly):			
Formal employee: Mean	0.058	0.083	82.88
Informal employee: Mean	0.091	0.071	44.64
Wage growth rates within jobs by quintiles (Yearly):			
Formal employee: Mean – Q1	0.105	0.440	29.86
Formal employee: Mean – Q2	0.067	0.176	41.39
Formal employee: Mean – Q3	0.042	0.070	42.31
Formal employee: Mean – Q4	0.040	−0.032	48.73
Informal employee: Mean – Q1	0.335	0.441	18.25
Informal employee: Mean – Q2	0.090	0.108	21.89
Informal employee: Mean – Q3	0.033	0.009	21.69
Informal employee: Mean – Q4	0.021	−0.021	21.18
Wage growth rates between jobs:			
Less than a quarter:			
Formal employee: Mean	0.107	0.115	55.59
Informal employee: Mean	0.112	0.129	35.92
More than a quarter:			
Formal employee (with Unemployment in Between): Mean	0.087	0.025	74.97
Informal employee (with Unemployment in Between): Mean	0.033	0.023	73.74
Formal employee (with Self-Employment in Between): Mean	0.001	0.003	136.42
Informal employee (with Self-Employment in Between): Mean	0.006	0.029	54.99

states in the sample, we think that our estimation algorithm is striking a good balance in trading-off some matching over searching states' proportions in exchange for a better match of the accepted wages distributions.

Table 6 reports the fit of the model in terms of labor market dynamics. The moments describing the yearly transitions between labor market states both within and across jobs are all qualitatively replicated; most of them are also quantitatively matched with reasonable precision. An original feature of our theoretical model is the ability to generate transitions in formality status within the same job. Our estimated model generates a significant number of these transitions but still underestimates the proportions observed in the data, in particular the transitions from formal to informal at the same job. This mismatch results again from a tension in fitting competing features of the data with a relatively parsimonious set of parameters. More transitions could be generated but at the cost of a worse fit on the distributions of accepted wages.

In spite of the limited longitudinal dimension of the data, the estimated model matches the growth rates of wages between jobs well. The fit on the wage growth within jobs is slightly worse. The estimated model generates growth rates that are lower for formal employees when compared to informal employees whereas the corresponding statistics from the sample imply the opposite pattern.[30] This last mismatch feature arises due to the tension between the faster arrival rates of human capital shocks and the smaller jumps of the upgrading in the formal sector when compared to the informal sector. Again, improving these moments using different combinations of the parameters governing the human capital upgrading and the size of the jump would result in a worse fit for the transition rates within jobs targeted by the MSM estimator.

30 The distributions of wage growth rates within jobs in the sample have been further trimmed for the presence of outliers.



(a) Share of Informal Employees (b) Aggregate Human Capital (Benchmark = 1)

(c) Mean Wages Formal (d) Mean Wages Informal

Fig. 4. Impacts of Policy 1 — Changes in the contribution rate t. NOTE: Simulated samples of 5000 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 C.2 for details. Wages are hourly and reported in Mexican pesos (exchange rate: 1 US dollar ≈ 13.5 Mex. pesos in 2014).

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 and non-contributory benefits. We focus on changes in the two main policy parameters that characterize these benefits: 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 between benefits and costs of working formally.[31]

The policy experiments procedure works as follows: For each value of the policy parameter, we find and compute the new equilibrium holding fixed the other institutional parameters and setting the structural parameters at the point estimates reported in Table 3. We simulate labor market careers for 5000 individuals. Figs. 4 and 5 display relevant statistics from the resulting simulated data.

[31] 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).

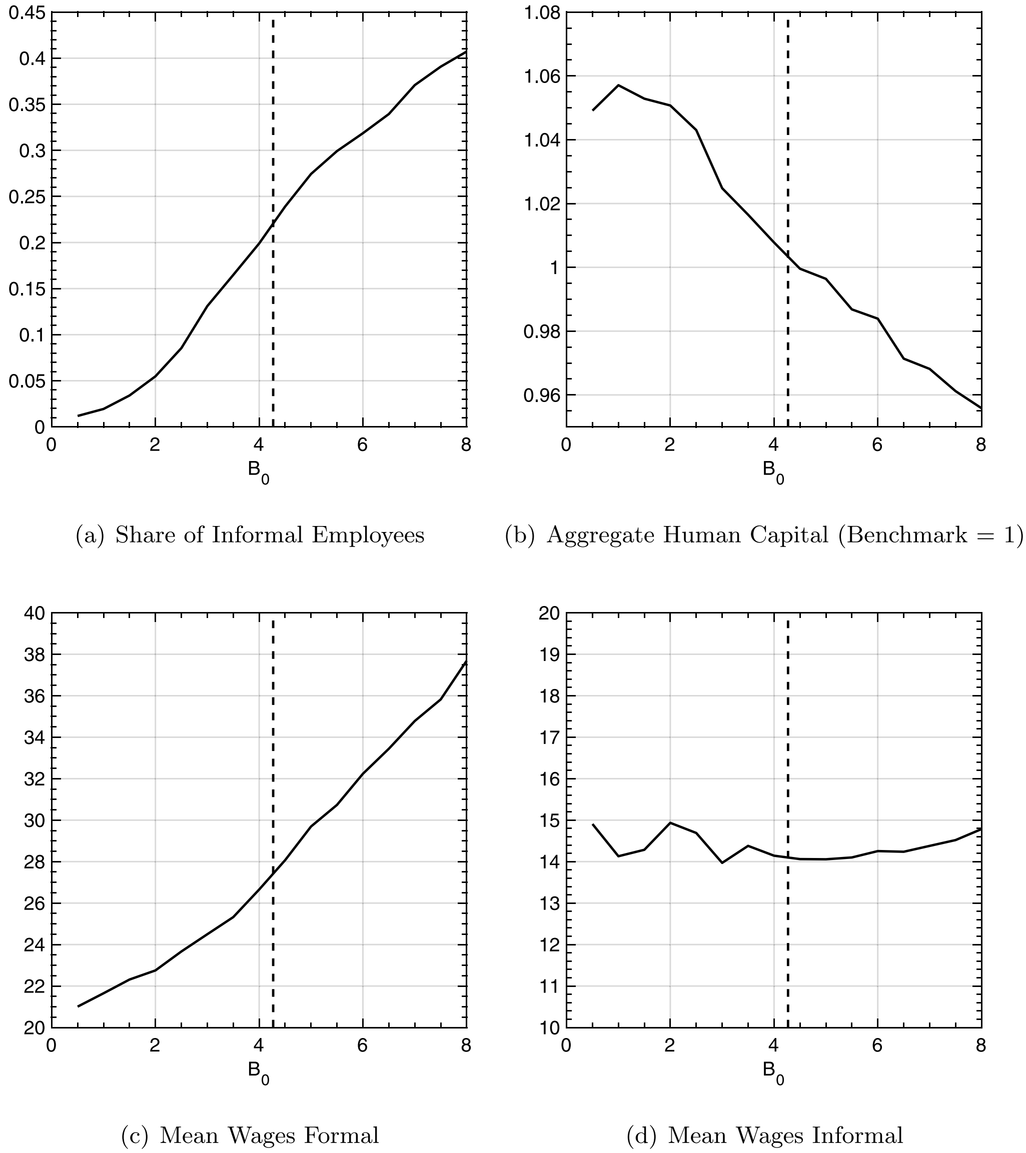


Fig. 5. Impacts of Policy 2 — Changes in the non-contributory benefit B_0. NOTE: Simulated samples of 5000 worker-level observations for each quarter based on the estimates reported in Table 3. The vertical lines are set at the institutional values for the Mexican labor market in 2013–2014. See Table C.2 for details. Wages are hourly and reported in Mexican pesos (exchange rate: 1 US dollar ≈ 13.5 Mex. pesos in 2014).

6.1. Policy Experiment 1: Contribution rate

We vary the contribution rate t in a wide neighborhood around the benchmark level and we run the experiments under two different scenarios. In the first scenario, labeled "Not Revenue Neutral", we keep constant the redistributive component of the social security benefits for formal employees, b_1 (see Eq. (6)). In this way, we can isolate the impact of the policy lever we are changing (t) but we lose the link between contributions and benefits. For example, if the change in t decreases the proportions of formal workers, the revenue generated by their lower amount of contribution may or may not be enough to cover for the expenses of the benchmark benefit b_1. In the second scenario, labeled "Revenue Neutral", we adjust the value of b_1 according to the endogenous wages and proportion of formal workers in the counterfactual economy so that the equilibrium contributions fully pay for the benefit (see Eq. (16)).

Fig. 4 reports the simulation results on labor market outcomes computed at the various contribution rates over the range. We denote the benchmark value of t with a vertical dashed line in all Panels. Panel (a) shows that as the contribution rate in a formal job increases, the share of informal employees in the labor force increases substantially in the "Not Revenue Neutral" scenario. The mechanism is straightforward: formality becomes more costly and only a portion of the benefit (retirement) increases while the other portion (health) remains constant. As a result, workers and firms prefer to realize informal matches. In the "Revenue Neutral" case, instead, the proportion remains roughly constant because both the retirement and health component of the formal benefit increases with the contribution rate.

This dynamic has a direct link with the amount of human capital accumulated in the economy, as shown in Panel (b). Since we estimate that the probability of receiving a human capital upgrading shock is higher for formal employees, the aggregate human capital roughly follows the behavior of the proportion of formal employees: decreasing in the "Not Revenue Neutral" case; more or less constant in the "Revenue Neutral" case. The first scenario is consistent with recent results from the literature on life-cycle labor supply with learning-by-doing. This literature shows that not only higher tax rates decrease labor supply today but also induce a dynamic feedback loop by affecting human capital accumulation. Even a transitory tax rate increase may therefore have permanent impacts: a higher tax rate today decreases labor supply today, but a lower labor supply today reduces human capital tomorrow, affecting labor supply tomorrow even if the tax rate goes back to its original value (Imai and Keane, 2004). Closer to our experiments are the permanent tax changes impacts estimated by Keane (2015), who shows that permanent tax changes can have larger current effects on labor supply than transitory tax changes. In our model, the impact is not so much on the overall reduction of labor supply but on the overall reduction of formal work since the alternative to (formal) labor is not leisure but either informal work (either as employee or as self-employed) or search. Conditioning on these differences (and of course the lack of a life-cycle component) both the overall and feedback effects found in this literature are reflected in our results. Looking at Panels (b) and (c), we observe that both aggregate human capital and formal employees' average wages significantly decrease as the contribution rate increases. However, this happens only in the first scenario because, as mentioned above, in the second "Revenue Neutral" scenario the worker who is marginal between formality and informality is compensated by the added benefit.

When looking at wages in Panels (c) and (d), it is useful to recall that in our Nash-bargaining context the impact of the contribution rate is highly non-linear, and it affects the wage schedule through four distinct channels: a direct channel, an equilibrium channel, a policy channel, and a selection channel. The direct channel, as discussed in Section 3.2 and shown in equation (B.1) in the Web Appendix, refers to the fact that the contribution rate enters directly into the wage equation because bargaining implies that firms can partially pass-through the cost of the tax to the worker. The equilibrium channel refers to the fact that the contribution rate affects the value of participating in the market which itself enters the wage schedule (B.1) as the value of the worker's outside option. In the revenue neutral case, the policy channel is at work because higher contribution rates mean a higher benefit, that in turn allows workers to give up more wages when working formally. Finally, the contribution rate impacts the average wages of the formal and informal employees because it affects who becomes a formal or an informal employee; that is, the selection over the match-specific productivity value x. Since wages are proportional to x, a higher threshold value between formality and informality (the $\tilde{x}(k, q)$ defined in Eq. (12)) means that formal employees are more positively selected in terms of productivity and therefore everything else being equal they earn higher wages.

All of these components generate the highly heterogeneous impact we see on average wages in the formal and informal sectors when comparing the two scenarios. If in both cases a higher contribution rate decreases average formal wages and increases average informal wages, the two elasticities are quite different. Formal wages are more sensitive to contribution rate changes in the "Revenue Neutral" scenario while the opposite is true for informal wages. The main reason for this difference is that in the "Not Revenue Neutral" scenario the selection into formal jobs is more positive and partially offsets the wage drop on formal wages induced by the direct and equilibrium effects mentioned above.

6.2. Policy Experiment 2: Non-contributory benefit

Fig. 5 reports simulation results on labor market outcomes computed at different values of the non-contributory benefit B_0. We denote the benchmark value of B_0 with a vertical dashed line in all Panels. An increase in B_0 can predict the impact of current policy proposals in Mexico and other Latin American countries that are focusing on broadening the coverage and amount of the non-contributory benefits (Levy, 2018).

Panels (a) and (b) report expected results: as more resources are given to informal employees at no cost, their share in the labor market increases. Since human capital upgrading shocks arrive faster while working formally, the human capital accumulation slows down and the aggregate level of human capital in the steady state decreases. The impact can be quite substantial. For example, the increase in benefit from the inception of the *Seguro Popular* program in 2002 to our period of observation in 2013 has been of about two and half pesos per hour (from 1.82 to 4.27). Such an increase would be associated in our experiments with a drop in aggregate human capital of about 5 percentage points. In addition, an increase in the non-contributory benefit is by definition not budget neutral because the benefit is not paid by any contribution. The fiscal cost is therefore increasing in B_0 and, relative to GDP, is increasing at a higher rate because the overall productivity of the labor market decreases due to the lower human capital.[32]

Panels (c) and (d) generate results that are less obvious but that become clear when recalling the discussion of Section 6.1. A higher non-contributory benefit has a very different impact on the average wage for the formal and informal

[32] In Figure D.3 of the Web Appendix, we report the cost of providing B_0 relative to the overall value of production: for example, if the benefit were to double from current values, the relative cost would increase almost three times.

employees: the first increases, while the second remains quite stable. As was the case for the contribution rate t, the non-contributory benefit B_0 impacts average wages through multiple channels: the direct impact on the wage schedule, the equilibrium impact through the value of the outside option, and the selection impact through the reservation value $\tilde{x}(k, q)$. The net results are the ones shown in the two Panels. Panel (c) shows an increase in the average wage for formal employee because they are increasingly more positively selected over productivity.[33] Panel (d) shows a stable wage for informal employees because the positive impact of better selection over productivity is compensated by the workers' willingness to accept lower wages in exchange for higher benefits.

7. Conclusions

We study how the different rate of human capital accumulation in the formal and informal jobs impact labor market outcomes. Recognizing that formality status and labor market states are endogenous outcomes interacting with the dynamics of human capital 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. Worker-specific human capital accumulates on-the-job in a learning-by-doing fashion and depreciates while searching. This setting is able to generate the specific features of labor market with high informality: a mixture of the formal and informal jobs with overlapping wage distributions and changes in formality status both between and within jobs.

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 sharing a significant informality rate with many other middle income countries. The human capital accumulation and the depreciation process are identified by exploiting the panel dimension of the data; specifically, wage changes within and between jobs, transitions between labor market states, and changes in formality status both between and within jobs.

The estimation results show that the probability of human capital upgrading is lower when working informally than formally. The relative contribution of human capital to overall productivity is estimated to be substantial: it is about 60 percent in the overall sample, reaching more than 80 percent for workers with the highest level of human capital.

We use the estimated model to perform policy experiments where we change the two crucial benefit parameters responsible for the high level of informality observed in Mexico: the payroll contribution rate in formal jobs and the per-capita level of the non-contributory social benefits. The first experiment generates some expected results: informality increases and human capital decreases when the contribution rate increases. However, it also shows that these expected results are very sensitive to the design of the policy: if the contribution rate increase is paired with a proportional increase in the benefit (a "revenue neutral" setting), the negative impact on aggregate human capital is almost neutralized while the increase in informality is limited to the increase in self-employment. The second experiment confirms that an increase of the non-contributory benefit would increase informality and decrease human capital accumulation. The decrease could be substantial: an increase in benefit equal to the one implemented under the *Seguro Popular* program in the last decade is associated in our experiments with a drop in aggregate human capital of about 5 percentage points.

In conclusion, this paper focusing on human capital accumulation after entering the labor market reinforces the results we have 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 specific features of the labor market. However, the presence of an informal labor market state may magnify the negative impact that such features have on labor market outcomes.

Acknowledgments

We thank the Editor and the anonymous referees for very useful remarks that substantially improved the paper, as well as the audiences at Barcelona GSE Summer Forum on Structural Microeconometrics, PUC Chile, SOLE, the UCL-Cemmap Workshop on Dynamic Structural Models, University of Chile, and the Cowles Summer Conference on Structural Microeconomics. Sara Compton Meade provided assistance proof-reading the text. Bobba thanks financial support from the AFD, the H2020-MSCA-RISE project GEMCLIME-2020 GA No 681228, and the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2019.05.026.

33 As reported in Figure D.2 of the Web Appendix, the proportion of formal employees decreases since the reservation value $\tilde{x}(k, q)$ increases.

References

Alaimo, V., Bosch, M., Kaplan, D.S., Ripani, L., et al., 2015. Jobs for Growth. Inter-American Development Bank.
Albrecht, J., Navarro, L., Vroman, S., 2009. The effects of labour market policies in an economy with an informal sector. Econom. J. 119 (539), 1105–1129.
Arrow, K.J., 1962. The economic implications of learning by doing. Rev. Econ. Stud. 29 (3), 155–173.
Bagger, J., Fontaine, F., Postel-Vinay, F., Robin, J.-M., 2014. Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. Amer. Econ. Rev. 104 (6), 1551–1596.
Bianchi, M., Bobba, M., 2013. Liquidity, risk, and occupational choices. Rev. Econom. Stud. 80 (2), 491–511.
Bobba, M., Flabbi, L., Levy, S., 2012. Labor Market Policies and Schooling Investments. Technical Report, IDB.
Bobba, M., Flabbi, L., Levy Algazi, S., 2017. Labor Market Search, Informality and Schooling Investments. Discussion Papers 11170, Institute for the Study of Labor (IZA).
Borowczyk-Martins, D., Bradley, J., Tarasonis, L., 2018. Racial discrimination in the US labor market: Employment and wage differentials by skill. Labour Econ..
Bosch, M., Esteban-Pretel, J., 2012. Job creation and job destruction in the presence of informal markets. J. Dev. Econ. 98 (2), 270–286.
Bowlus, A.J., Liu, H., 2013. The contributions of search and human capital to earnings growth over the life cycle. Eur. Econ. Rev. 64, 305–331.
Cahuc, P., Postel-Vinay, F., Robin, J.-M., 2006. Wage bargaining with on-the-job search: Theory and evidence. Econometrica 74 (2), 323–364.
Dey, M.S., Flinn, C.J., 2005. An equilibrium model of health insurance provision and wage determination. Econometrica 73 (2), 571–627.
Dey, M., Flinn, C., 2008. Household search and health insurance coverage. J. Econometrics 145 (12), 43–63.
Eckstein, Z., Wolpin, K.I., 1995. Duration to first job and the return to schooling: Estimates from a search-matching model. Rev. Econom. Stud. 62 (2), 263–286.
Feng, Y., Lagakos, D., Rauch, J.E., 2018. Unemployment and Development. Working Paper 25171, National Bureau of Economic Research.
Fields, G.S., 1975. Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. J. Dev. Econ. 2 (2), 165–187.
Flabbi, L., Mabli, J., 2018. Household search or individual search: Does it matter?. J. Labor Econ. 35 (1).
Flabbi, L., Moro, A., 2012. The effect of job flexibility on female labor market outcomes: estimates from a search and bargaining model. J. 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., Laufer, S., 2017. Search, matching and training. Rev. Econ. Dyn. 25, 260–297.
Flinn, C., Heckman, J., 1982. New methods for analyzing structural models of labor force dynamics. J. Econometrics 18 (1), 115–168.
Flinn, C., Mullins, J., 2015. Labor market search and schooling investment. Internat. Econom. Rev. 56, 359–398.
Garcia, I.L., 2015. Human Capital and Labor Informality in Chile A Life-Cycle Approach. Working Papers WR-1087, RAND Corporation.
Guler, B., Guvenen, F., Violante, G.L., 2012. Joint-search theory: New opportunities and new frictions. J. Monetary Econ. 59 (4), 352–369.
Herkenhoff, K., Lise, J., Menzio, G., Phillips, G.M., 2018. Production and Learning in Teams. NBER Working Papers 25179, National Bureau of Economic Research, Inc.
Imai, S., Keane, M.P., 2004. Intertemporal labor supply and human capital accumulation. Internat. Econom. Rev. 45 (2), 601–641.
Jovanovic, B., 1979. Job matching and the theory of turnover. J. Polit. Econ. 87 (5, Part 1), 972–990.
Keane, M.P., 2015. Effects of permanent and Transitory tax changes in a life-cycle labor supply model with human capital. Internat. Econom. Rev. 56 (2), 485–503.
Keane, M.P., Todd, P.E., Wolpin, K.I., 2011. The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. Handb. Labor Econ. 4, 331–461.
La Porta, R., Shleifer, A., 2014. Informality and development. J. Econ. Perspect. 28 (3), 109–126.
Lagakos, D., Moll, B., Porzio, T., Qian, N., Schoellman, T., 2018. Life cycle wage growth across countries. J. Polit. Econ. 126 (2), 797–849.
Levy, S., 2008. Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico. The Brooking Institution Press, Washington DC.
Levy, S., 2018. Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico. Inter-American Development Bank.
Levy, S., Schady, N., 2013. Latin america's social policy challenge: Education, social insurance, redistribution. J. Econ. Perspect. 27 (2), 193–218.
Lucas Jr., R.E., 1993. Making a miracle. Econometrica 251–272.
Maloney, W.F., 1999. Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. World Bank Econ. Rev. 13 (2), 275–302.
Maloney, W.F., 2004. Informality revisited. World Dev. 32 (7), 1159–1178.
Meghir, C., Narita, R., Robin, J.-M., 2015. Wages and informality in developing countries. Amer. Econ. Rev. 105 (4), 1509–1546.
Menzio, G., Telyukova, I., Visschers, L., 2016. Directed search over the life cycle. Rev. Econ. Dyn. 19, 38–62.
Michelacci, C., Pijoan-Mas, J., 2012. Intertemporal labour supply with search frictions. Rev. Econom. Stud. 79 (3), 899–931.
Newey, W.K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Handbook of Econometrics, vol. 4, Elsevier, pp. 2111–2245.
Pakes, A., Pollard, D., 1989. Simulation and the asymptotics of optimization estimators. Econometrica 57 (5), 1027–1057.
de Paula, A., Scheinkman, J.A., 2011. The informal sector: an equilibrium model and some empirical evidence from Brazil. Rev. Income Wealth 57, S8–S26.
Perry, G.E., Maloney, W.F., Arias, O.S., Fajnzylber, P., Mason, A.D., Saavedra-Chanduvi, J., 2007. Informality : Exit and Exclusion. In: World Bank Publications, vol. 6730, The World Bank.
Pissarides, C.A., 2000. Equilibrium Unemployment Theory. MIT Press, Cambridge, MA.
Rocha, R., Ulyssea, G., Rachter, L., 2017. Do Lower Taxes Reduce Informality? Evidence from Brazil. mimeo, PUC-Rio.
Rubinstein, Y., Weiss, Y., 2006. Post schooling wage growth: Investment, search and learning. Handb. Econ. Educ. 1, 1–67.
Samaniego de la Parra, B., 2016. Formal Firms, Informal Workers and Household Labor Supply in Mexico. Technical Report, Mimeo, University of Chicago.
Thompson, P., 2010. Chapter 10 - learning by doing. In: Hall, B.H., Rosenberg, N. (Eds.), Handbook of the Economics of Innovation, Vol. 1. In: Handbook of the Economics of Innovation, vol. 1, North-Holland, pp. 429–476.
Ulyssea, G., 2018. Firms, informality, and development: Theory and evidence from Brazil. Amer. Econ. Rev. 108 (8), 2015–2047.
Wolinsky, A., 1987. Matching, search, and bargaining. J. Econom. Theory 42 (2), 311–333.
World-Bank, 2019. World Development Report 2019, Vol. 30435.
Yamaguchi, S., 2010. Job search, bargaining, and wage dynamics. J. Labor Econ. 28 (3), 595–631.