

## LABOR MARKET SEARCH, INFORMALITY, AND SCHOOLING INVESTMENTS\*

BY MATTEO BOBBA, LUCA FLABBI, AND SANTIAGO LEVY

*Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France; IZA, Bonn, Germany; Department of Economics, University of North Carolina – Chapel Hill, NC, U.S.A.; IZA, Bonn, Germany; Brookings Institution, Washington, DC, U.S.A.*

We develop a search and matching model where jobs can be formal or informal. Workers choose schooling and search for jobs. Firms post vacancies in each schooling market and decide the job's formality status. Upon meeting, workers and firms bargain over wages. The equilibrium size of the informal sector is an endogenous function of labor market fundamentals and institutions. We estimate the model using labor force survey data from Mexico and exploiting the exogenous variation induced by a noncontributory social program. Counterfactual experiments show that eliminating informal jobs increases schooling investments but decreases welfare for both workers and firms.

### 1. INTRODUCTION

High levels of informality characterize many labor markets, typically in medium and low-income countries.<sup>1</sup> Informality can be broadly defined as any deviation from labor regulations, such as avoiding payroll contributions and not conforming to labor law statutes. If these regulations are rigid and imperfectly enforced, noncompliance allows both firms and workers greater flexibility in their labor market decisions. The immediate costs of this flexibility are, on the firms' side, the possibility of being discovered and fined; on the workers' side, the loss of protections and benefits guaranteed by the law. The longer term costs include distortions in investments and allocations decisions. The existing literature has mainly considered the longer term costs of the demand side of the labor market, identifying distortions in firms' investment decisions as the main channel behind the correlation between productivity and informality (La Porta and Shleifer, 2008; de Paula and Scheinkman, 2010, 2011; Ulysea, 2018). We study instead the longer term costs of the supply side of the labor market, focusing on the main investment decision undertaken by workers before entering the labor market: schooling.<sup>2</sup>

\*Manuscript received July 2018; revised June 2021.

We thank for useful comments Giacomo De Giorgi, Chris Flinn, Marion Goussé, Gregory Jolivet (discussant), Lance Lochner, Thierry Magnac, Barbara Petrongolo, Imran Rasul, Stan Rabinovich, Jean-Marc Robin, Richard Rogerson, Todd Schoellman (discussant), as well as the anonymous referees. Marco Pariguana, Jose Mauricio Salazar, and especially Matias Morales provided excellent research assistance. Bobba acknowledges 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). Please address correspondence to: L. Flabbi, Department of Economics, University of North Carolina – Chapel Hill, NC, and IZA, Bonn, Germany. E-mail: [luca.flabbi@unc.edu](mailto:luca.flabbi@unc.edu).

<sup>1</sup> This issue is particularly relevant in Latin America where even large middle-income economies with well-developed labor market institutions feature more than half of the labor force in the informal sector (Perry et al., 2007; Levy and Schady, 2013). But the phenomenon is also common in other parts of the world (La Porta and Shleifer, 2014).

<sup>2</sup> Labor market frictions can potentially hinder this allocation by distorting the labor market returns to schooling (Flinn and Mullins, 2015). Yet, we lack systematic evidence on how labor market institutions responsible for the emergence of informality interact with labor market frictions in affecting the process of human capital accumulation. The macroliterature has either focused on the misallocation cost of labor market institutions (Hopenhayn and Rogerson, 1993; Ljungqvist and Sargent, 1998) or on the consequences of labor misallocation for aggregate productivity and growth (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013). At the microlevel, a number of studies have

Studying costs and benefits of informality—including consequences for productivity and welfare—requires an equilibrium model of the labor market that takes into account how workers and firms jointly sort between formal and informal jobs. It does also require to explain a set of stylized facts about informality that do not conform to either a segmented or a competitive view of the labor market but that could be explained by the labor market characterized by frictions that we propose. The evidence on labor market dynamics shows mobility of workers not only from formal to informal jobs but also from informal to formal jobs. The evidence on wage distributions shows that on average wages are higher in formal jobs but also that many informal jobs pay more than formal ones. Both pieces of evidence are in contrast with a segmented view where barriers restrict access to the formal sector. But they are also in contrast with a competitive view where the presence of the two types of jobs in equilibrium could only be justified by compensating differentials mapping into preferences and skills.<sup>3</sup>

In this article, we develop and estimate a search, matching, and bargaining model that takes into account sorting and endogenous decisions over formality regimes. The model is able to replicate the main empirical features of labor markets with high informality. Search and matching frictions support the presence of different types of contracts in equilibrium. Optimal decisions rules based on reservation values and the presence of termination shocks generate transitions between formal and informal jobs in both directions. Match-specific productivity and bargaining generate the overlapping wage distributions. An *ex ante* schooling decision allows for endogenous human capital accumulation. In this framework, we develop an identification strategy able to recover all the structural parameters of the model from standard labor market data from Mexico. We next perform policy experiments using the estimated model to study the equilibrium impact of informality and of the institutional parameters responsible for its emergence.

We are able to capture a wide range of costs and benefits related to informality. With respect to the immediate costs, we introduce parameters that capture the frequency at which firms are audited and, if hiring informally, fined. With respect to immediate benefits, we distinguish between the social security benefits received by formal workers and those received by the rest of the workers. We also identify and estimate preferences for each type of benefit. With respect to the impact on productivity and labor market performances, we allow all the institutional parameters responsible for the emergence of informality to have equilibrium effects on the steady-state distribution of productivity, wages, labor market states, and schooling. With respect to long-term costs, we can evaluate whether or not the same institutional parameters distort returns to schooling in a way that is detrimental to the accumulation of human capital.

Although previous studies on informality have employed similar models of the labor market (Albrecht et al., 2009; Bosch and Esteban-Pretel, 2012; Meghir et al., 2015), our approach incorporates additional features that take into account relevant but overlooked issues.<sup>4</sup> First, we allow individuals to choose their level of acquired schooling before entering the labor market. Holdup problems generate a dynamic inefficiency on human capital investments that is potentially exacerbated by the presence of the labor market frictions and institutions that give raise to informality.<sup>5</sup>

documented how plausibly exogenous institutional changes or labor demand shocks that alter returns to schooling in specific contexts affect human capital decisions (Munshi and Rosenzweig, 2006; Jensen, 2012; Abramitzky and Lavy, 2014; Heath and Mobarak, 2015; Atkin, 2016).

<sup>3</sup> Magnac (1991) presents these two competing views of the labor market. Maloney (1999) is a seminal contribution, showing that transitions between the formal and the informal sector in Mexico are equally probable in both directions. More recent evidence on Mexico is in Anton et al. (2012), whereas Perry et al. (2007) review evidence on a large number of Latin American countries. In Section 2, we show evidence that is consistent with these stylized facts using data for a sample of workers from the Mexican labor force survey (ENOE, 2005).

<sup>4</sup> Haanwinckel and Soares (2016) and Conti et al. (2017) are additional recent contributions that use empirical search models to study different aspects of labor market informality.

<sup>5</sup> Holdup problems arise since the investment leading to higher productivity (schooling) is made before the rewards are realized (output after matching with a firm). Acemoglu and Shimer (1999) examine the issue in frictional markets,

Second, the model parsimoniously accommodates some important institutional details. In response to the lack of health and pension coverage for informal workers, several countries are increasingly providing social security benefits that are not tied to payroll contributions. We allow for the fact that benefits financed by payroll contributions and those financed by resources collected outside the labor market may not be valued by individuals at full value. Recovering these preferences for the social security system turns out to be crucial to evaluate its impact on labor market outcomes.

Third, we propose a more nuanced definition of informality than previous literature by allowing workers to perform an informal job as either an employee or a self-employed. “Necessity” self-employment—that is, a form of self-employment requiring very limited skills and capital and imposing almost no barriers to entry—is so common and pervasive in labor markets across developing countries that is frequently used as a proxy for informality itself. However, if it is true that almost all these self-employed workers are informal, a prominent portion of informal jobs is organized as a genuine subordinate working relationship with a well-defined employer. Since self-employed and employees have markedly different labor market dynamics, these differences must be explicitly taken into account to provide a complete characterization of the informal sector.<sup>6</sup>

The model is estimated using a combination of individual-level data on workers from the Mexican labor force survey (ENOE) and demand-side information from aggregate data sources. The structure of the search equilibrium together with some distributional assumptions is enough to identify most of the model’s primitive parameters from the data at our disposal. The exception is the preference parameter for noncontributory benefits. To identify this parameter, we rely on an additional source of variation in the data stemming from the uneven geographical expansion of a social protection program, the *Seguro Popular*. The program increased access to health-care benefits for individuals not covered by the contributory system during the 2000s in Mexico. We use the same source of exogenous variation in a time period different from the one employed in estimation to validate the model.<sup>7</sup>

Estimation results show reasonable values of the model parameters, including those harder to identify like the firms’ costs of being discovered and punished for hiring informally and the workers’ preferences for the social security system. We uncover important differences by schooling in the labor market parameters leading to an overall return to schooling of about 30% in our sample of workers. This metric summarizes not only wages and self-employment incomes but also search and matching frictions, the probability of job termination, the selection over labor market states (including between formal and informal jobs), and the valuation of the nonwage benefits. The valuation of the noncontributory benefit is estimated to be very close to full monetary value. Instead, the valuation of the contributory benefit is found to be valued less than the monetary investment used to provide them, implying a net loss that reduces the incentive to work formally.

The counterfactual experiments show that the possibility to offer informal employee contracts is welfare increasing: removing this option would decrease steady-state welfare by

showing that higher frictions exacerbate the inefficiencies. More closely related to our paper, Flinn and Mullins (2015) develop and estimate a model extending the standard search and matching framework to allow for ex ante schooling decisions. In the context of the U.S. labor market, they find that the extent of the holdup inefficiency is sensitive to the workers’ bargaining power parameter.

<sup>6</sup> Fields (1975) is a seminal contribution pointing out the specific role played by self-employment in labor markets with high informality. Margolis (2014) and World Bank (2012) provide an overview of the evidence on many developing countries. Narita (2020) is one of the few contributions explicitly making the distinction between self-employed and employees in modeling informal workers.

<sup>7</sup> The econometric approach pursued in this article extends classic identification results from the empirical literature on search and matching models of the labor market (Flinn and Heckman, 1982; Eckstein and van den Berg, 2007; Keane et al., 2011) by adding exogenous sources of variation resulting from the roll-out of a policy intervention. As such, it fits within a recent strand of literature that combines field or natural experiments with econometric models for validation and/or estimation purposes (Todd and Wolpin, 2006; Attanasio et al., 2012; Galiani et al., 2015; Duflo et al., 2018; Gautier et al., 2018; Fu and Gregory, 2019).

about 6%. However, its benefits are very unequally distributed between firms and workers, with firms gaining almost six times more than workers on average. This asymmetry leads to substantial workers' gains in an economy able to reduce posting inefficiencies: under a Hosios-like condition, workers experience almost a 5% welfare gain when removing informal employee contracts. Informality is also found to distort the labor market returns to schooling and hence the underlying process of human capital accumulation. The proportion of agents acquiring additional schooling increases by 10% when informal employee contracts are eliminated. This relatively large effect on schooling suggests that the holdup problem is exacerbated by the possibility of offering informal job contracts.

We next consider changes in one institutional parameter that is at the center of the policy debate on informality—the social security contribution rate. Lowering the current contribution rate by about 10 percentage points would increase overall welfare by 3.5%. The overall optimal policy (setting the contribution rate to zero) would generate a 15.8% welfare increase compared to benchmark. These improvements are sensitive to the benefits' valuation. The closer the valuation is to full value, the larger is the improvement around the benchmark rate but the smaller is the improvement at the optimal value. The contribution rate significantly affects schooling decisions: the proportion of individuals who acquire the higher schooling level decreases monotonically with the payroll contribution rate. Interestingly, these changes in schooling are paired with an almost constant informality rate. The phenomenon is explained by major compositions effects resulting from the progressive features of the contributory social security benefits.

The article is organized as follows: In the next section, we describe the institutional context and the data we use in the empirical analysis. Section 3 presents the model and its empirical implications. Section 4 discusses the identification using the data at our disposal. Section 5 defines the estimation method, presents the estimation results, and reports the model validation. Section 6 contains the policy experiments and Section 7 concludes.

## 2. CONTEXT AND DATA

*2.1. Institutional Context.* We define informality with reference to compliance with salaried labor regulations. In Mexico, as in most countries, firms are required to enroll salaried workers in the social security registry (IMSS, for its Spanish acronym) and pay a contribution proportional to workers' wages. The revenues from the contributions are used to fund the social security benefits. The rate of the social security contribution is approximately 33% of the wage of salaried workers. Enrollment in the social security registry is considered the main indicator of compliance with labor regulations and, therefore, the main indicator of formality status (Levy, 2008; Kanbur, 2009).

The social security benefits in Mexico are *bundled*: firms and workers must pay for a fixed-proportions package that includes health benefits, housing benefits, some day care services, and death, disability, and retirement pensions. Some benefits are directly proportional to the worker's individual wage and contribution (pensions), whereas others are not (health benefits), implying redistribution within salaried workers. Other institutions related to labor regulations that may be relevant in other contexts—such as minimum wage, labor income tax, or unemployment insurance—are second order for the Mexican population at the center of our study.<sup>8</sup>

Labor regulations are only imperfectly enforced. Some firms thus avoid compliance or opt for partial compliance (i.e., by paying social security contributions for some but not all of their

<sup>8</sup> Mexico does not have significant unemployment insurance and thus no flow payments out of wages into an unemployment fund or individual accounts. Starting from the late eighties, the real value of the minimum wage in Mexico has been monotonically deteriorating and it is currently considered not binding (Bosch and Manacorda, 2010). Finally, labor income tax is mostly zero or very small over the wage support that we consider in our sample.

employees) to save on labor costs.<sup>9</sup> The IMSS attempts to enforce the payment of social security contribution by auditing the firms. Although the exact parameters that IMSS uses to determine which establishments to audit are confidential, anecdotal evidence suggests that firms' size, sector, the history of previous violations, and notifications made to IMSS by the Ministry of Labor may explain some of the variation in the occurrence of an audit (Samaniego de la Parra and Bujanda, 2020). When caught hiring illegally, employers have to pay monetary fines that range between 20 and 350 daily minimum wages for each nonregistered worker as well as back-due payroll contributions. The very wide support of monetary penalties mandated by the law leaves room for discretion in the amount of fines actually levied to noncomplying firms.

These rights and obligations do not generally apply to self-employed workers. The main self-employment activity of the individuals in our population can be described as a sort of "residual" labor market state where individuals not matched with firms set up their own microenterprises, whereas they keep searching for a job (Fields, 1975). Financial barriers to enter such self-employment state are minimal and do not constitute a significant obstacle (McKenzie and Woodruff, 2006; Bianchi and Bobba, 2013). This view of the self-employment state is consistent with the very low unemployment rate observed in these labor markets (Feng et al., 2018). It also shares some similarities with the "necessity" entrepreneurs literature in high-income countries (Ardagna and Lusardi, 2009; Fossen and Büttner, 2013).

In the early 2000s, Mexico's Federal Government designed a new noncontributory social security program, named *Seguro Popular*, aimed at expanding the scope of medical benefits for those not covered by the contributory social security system. The program started as a pilot during 2002 in five states and was then extended to municipalities across the country. In 2007, more than 21 million beneficiaries were reached by the program. During the same period, similar noncontributory programs were launched to expand the coverage of housing subsidies, retirement pensions, and day care facilities. There are no significant regional or quality differences between contributory and noncontributory pension, housing, and day care programs; with regard to health, differences have narrowed considerably as a result of a large expansion in the health infrastructure of state governments, which provide services to those not covered by the contributory program managed by IMSS (Levy, 2008).

Unlike social security benefits financed by payroll contributions, these noncontributory programs are financed by the federal government through a mix of general tax revenues and issuance of public debt. Spending in these programs doubled between 2002 and 2013, from 0.8% to 1.65% of GDP—a pattern that is common across many countries with a dual social security system (Frolich et al., 2014).<sup>10</sup> These increases correlate with favorable fiscal conditions in the region driven by the commodity price boom of the 2000s (Baffes and Haniotis, 2010).<sup>11</sup>

**2.2. Data.** The data we use in estimation are extracted from the 2005 Mexico's official labor force survey, the *Encuesta Nacional de Ocupación y Empleo* (ENOE). The survey covers a representative sample of the Mexican population aged 14 years and older and has a structure similar to the CPS: each household is interviewed every three months for a total of five interviews. As a result of the rotating scheme, a fifth of the sample is refreshed every quarter. We focus on 2005 since it is a period of relative stability in terms of informality rates and institutional parameters related to the social security system.

<sup>9</sup> For instance, Perry et al. (2007) show that in Mexico, 50%–70% of small-medium firms have used both formal and informal contracts simultaneously in a given point in time. Ulyssea (2018) documents that in small formal firms in Brazil, 40% of workers are informal. At the same time, 52% of all informal workers are employed in large firms that are unlikely to be fully informal.

<sup>10</sup> The corresponding figures for other Latin American countries document even steeper growth rates than Mexico over the same period. For instance, in Chile, spending in noncontributory social programs increased from 0.5% of GDP in 2002 to 1.5% of GDP in 2013. In Argentina, spending increased from 1% of GDP to 4% of GDP.

<sup>11</sup> For Mexico specifically, oil exports are an important revenue source: Mexico is currently the fifth largest exporter of crude oil in the world with a 5.6% share of the market.



To assure some homogeneity over observables in our estimation sample, we focus on non-agricultural, full-time, male, and private-sector workers between the ages of 25 and 55 who reside in urban areas (defined as localities with a population greater than 15,000 inhabitants). We focus our analysis on workers at the midrange of the skill distribution, that is, those with a secondary degree. Specifically, we only consider individuals who completed either lower secondary schooling (equivalent to middle school in the United States) or upper secondary schooling (equivalent to high school in the United States). These are the schooling groups for whom both education decisions and informality issues are very relevant (Bobba et al., 2012). We then split the resulting sample in two groups according to whether the worker has completed high school or not.<sup>12</sup>

We define a worker to be an *employee* if (i) he declares to be in a subordinate working relationship in his main occupation; and (ii) he receives a wage as a result of that working relationship (only 4% of individuals in our sample report having a secondary occupation). We identify the formal or informal status of the job depending on whether the employee reports having access to health benefits through his employer. This indicator is equivalent to the enrollment of the worker in the social security registry we discussed in Subsection 2.1.<sup>13</sup> We define a worker to be *self-employed* if he declares (i) not to be in a subordinate relationship in his main occupation and (ii) to have his own business. To obtain a population of self-employed individuals closer to the “residual” labor market state outlined above, we drop those business owners who report employing paid employees (roughly 30% of the self-employed sample) and who have access to contributory health benefits typically reserved for large business owners (1% of the self-employed sample). The entire subpopulation of self-employed workers that we consider is thus informal, as opposed to employee workers who can be formal or informal depending on employers’ decision to enroll some, none, or all of their employees in the social security registries.<sup>14</sup> We define an individual to be *unemployed* if he declares (i) not to be working during the last week and (ii) to be actively searching for a job.

The final estimation sample is composed of two data sets. The first data set is a *cross-sectional sample* extracted from the first quarter of the year 2005.<sup>15</sup> Given the structure of the survey, this cross-sectional sample includes the records of individuals who entered the window of observation in five different quarters assuring that: (i) each individual is counted only once, (ii) the sample remains representative, and (iii) we exploit all the available observations collected in the year. The sample size is 13,614 individual observations, with 8,570 observations belonging to the low schooling group (middle school completed) and 5,044 observations

<sup>12</sup> Appendix A.1 describes in more details the data used for the analysis. By discarding public sector workers, we do not consider the possible interactions between public, formal private, and informal private sectors. However, these workers constitute only 6% of the sample and have their own labor statute and social security system.

<sup>13</sup> We have cross checked this definition of informality with two auxiliary data sources. First, we use the nationally representative household survey (ENIGH) collected in the same period. This information allows us to construct the exact definition of informality that we have employed in the ENOE survey as well as an alternative definition based on more detailed information on respondents’ occupations and access to benefits through their job. The resulting discrepancies at the individual level in the categories of formal and informal employees are minimal. Second, using our definition, we use the survey weights in the ENOE to generate aggregate shares of formal workers at the national level. Those are by and large comparable with the share of formal workers resulting from aggregating the total number of individuals who are registered in the social security administration (IMSS) as a share of the total national workforce contained in the Mexican population census.

<sup>14</sup> To reduce the impact of outliers, individual earnings are winsorized at the 99th percentiles. The percentiles are computed conditioning on the two schooling levels and on the labor market status (formal employment, informal employment, and self-employment).

<sup>15</sup> As a result of the uneven geographical expansion of the *Seguro Popular*, roughly 68% of the individuals in our sample who are not formal employees belong to municipalities where the program was operating, whereas the remaining 32% belong to municipalities where the program was not yet operating. We use the differential exposure to the program in the year 2005 for identification of some of the model’s parameters (see Subsection 4.2) and the corresponding variation in program exposure in the year 2006 for out-of-sample validation of the model (see Subsection 5.3.2). The geographic roll-out of the program was completed in the year 2007, which is the reason why we cannot rely on more recent waves of labor market survey data for our analysis.

TABLE 1  
DESCRIPTIVE STATISTICS—CROSS SECTION

Labor Market State:	Unemployed	Formal Employee	Informal Employee	Self-Employed
<u>High Schooling (<math>N = 5,044</math>)</u>				
Proportion (percent)	5.06	58.90	19.77	16.28
Hourly Earnings: Mean		27.16	20.04	23.38
Hourly Earnings: SD		16.28	14.53	16.29
<u>Low Schooling (<math>N = 8,570</math>)</u>				
Proportion (percent)	4.36	56.16	23.40	16.08
Hourly Earnings: Mean		22.31	16.72	21.18
Hourly Earnings: SD		11.01	9.03	13.63

NOTES: Cross-sectional sample extracted from the Labor Market Survey (ENOE) in 2005:Q1 and composed of nonagricultural male employees and self-employed between the ages of 25 and 55 who reside in urban areas with completed secondary schooling (top panel) and uncompleted secondary schooling (bottom panel). Earning figures are reported in Mexican pesos (exchange rate: 10 Mex. pesos  $\approx$  1 U.S. dollars in 2005).

TABLE 2  
DESCRIPTIVE STATISTICS—YEARLY TRANSITION RATES

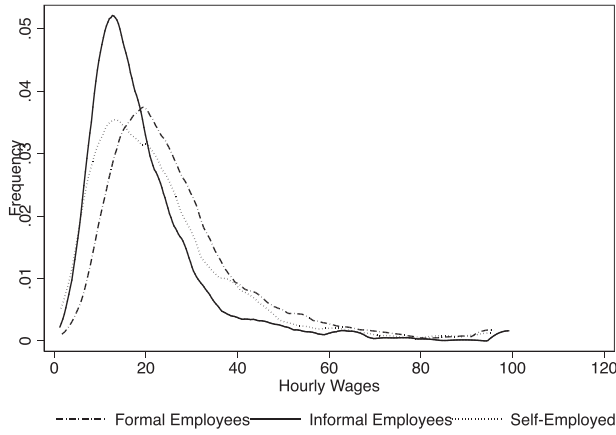
<u>High Schooling (<math>N = 1,330</math>)</u>				
<u>Labor Market State Q1:→</u>	Unempl.	Formal Empl.	Informal Empl.	Self-Empl.
<u>Labor Market State Q5:↓</u>				
Unemployed	31.58	52.63	15.79	0.00
Formal Employee	1.75	91.43	5.47	1.34
Informal Employee	3.49	21.49	68.86	6.14
Self-Employed	0.00	9.65	13.16	77.19
<u>Low Schooling (<math>N = 2,329</math>)</u>				
<u>Labor Market State Q1:→</u>	Unempl.	Formal Empl.	Informal Empl.	Self-Empl.
<u>Labor Market State Q5:↓</u>				
Unemployed	8.82	63.24	26.47	1.47
Formal Employee	2.10	87.80	8.66	1.44
Informal Employee	1.77	27.76	63.39	7.09
Self-Employed	0.44	10.09	16.23	73.25

NOTES: Balanced panel of four ENOE cohorts entering in 2005:Q1 through 2005:Q4 and observed over five quarters. Sample of nonagricultural male employees and self-employed between the ages of 25 and 55 who reside in urban areas with completed secondary schooling (top panel) and uncompleted secondary schooling (bottom panel).

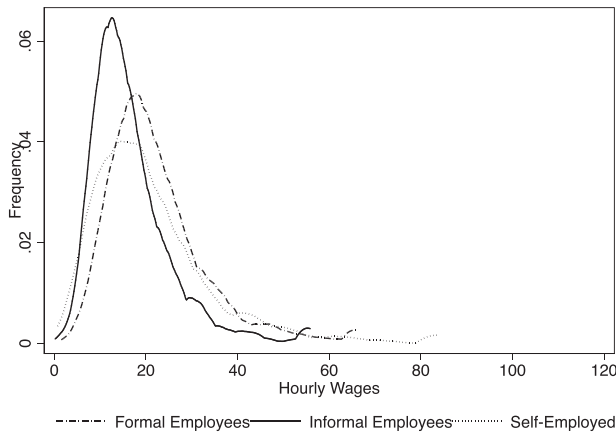
(37%) to the high schooling group (high school completed). We use this data set to construct moments that characterize the main steady-state patterns of the labor market under study, such as the relative shares of workers in each schooling group as well as the relative proportions and the earnings distributions in each labor market state. Although the cross-sectional data set contains information on ongoing spells in each labor market state, we have opted to use instead the transition rates across labor market states to better capture labor dynamics. To do so, we have constructed a *balanced panel* of workers who are followed over five consecutive quarters starting in one of the quarters of the year 2005. The sample size is 3,659 individuals observed in each quarter in municipalities in which the assignment of the *Seguro Popular* program did not change during the four quarters of 2005.<sup>16</sup> For a more detailed discussion on the construction of the panel data sets and on other specific features of the data, see Appendix A.1.

Tables 1 and 2 report descriptive statistics on the samples by schooling group. The patterns emerging from these data are consistent with the key stylized facts of labor markets characterized by high rates of informal employment. First, there is a significant mass of workers in each

<sup>16</sup> Transition rates may partly reflect changes over time in the noncontributory benefit  $B_0$  among the 753 individuals (17% of the Panel sample) who reside in municipalities that received the *Seguro Popular* program during any of the quarters of the year 2005.



(a) High Schooling



(b) Low Schooling

NOTE: The figure shows the empirical densities of the hourly earnings (in Mexican pesos) for nonagricultural male employees and self-employed between the ages of 25 and 55 who reside in urban areas with completed secondary schooling (Panel a) and uncompleted secondary schooling (Panel b). Earning figures are reported in Mexican pesos (exchange rate: 10 Mex. pesos  $\approx$  1 U.S. dollars in 2005). The formal/informal status of the job is defined according to whether or not workers report having access to health care through their employers.

FIGURE 1

WAGE DENSITY FUNCTIONS

employment labor market state: a bit more than half of the workers are in the formal sector, the other half is almost equally split between informal employment and self-employment. Unemployment rates are between 4% and 5% in the two schooling markets considered. There is also a fair amount of transitions of employees between the formal and informal status. Over a one-year period, about one quarter of the formal employees transit to informal employment and 5–8% of the informal employees transit to formal employment, possibly with a period of unemployment in between (see Appendix A.1). Second, there is a large overlap between the formal and the informal wage distributions (see also Figure 1), with the former first-order stochastically dominating the latter.<sup>17</sup> Self-employed earning distributions are approximately in between those of formal and informal employees but they display a larger standard

<sup>17</sup> The Komolgorov–Smirnov (KS) test statistic for the directional hypothesis that wages of formal employees FOSD wages of informal employees is equal to 0.306 for incomplete secondary and 0.294 for complete secondary.



deviation, especially in the high schooling group. Third, unemployed and self-employed workers have very different labor market dynamics. About 90% of the unemployed with incomplete secondary schooling accept an employee job over a one-year period compared with 27% of the self-employed. In addition, when they accept a job, they receive systematically different wages. The mean and standard deviations of accepted wages in the high schooling group are, respectively, 30% and 40% higher for individuals transiting to formal employment from self-employment than from unemployment.

We complement the microdata with two aggregate data sources. We obtain labor shares for Mexico in 2005 as measured by the total compensations per employee as percentage of GDP at market prices per person employed using data collected by AMECO (the *Annual Macroeconomic* database of the European Commission's Directorate General for Economic and Financial Affairs). In addition, we use a data set that comes from a free government-run employment service developed under the supervision of the Ministry of Labor. The service—called *Bolsa de Trabajo*—records monthly observations of number and characteristics of job vacancies posted in the urban areas of each of Mexico's 31 states and Federal District. This information allows us to analyze the job matching process in the segment of the Mexican labor market that is the focus of our analysis.<sup>18</sup> We compute schooling-specific vacancy rates for the year 2005 by aggregating over all the available job vacancies featuring the educational requirements that correspond to our two schooling groups and dividing the resulting statistics by the labor force in each category using the appropriate survey weights for the ENOE sample.

### 3. MODEL

3.1. *Environment.* The model assumes stationarity, continuous time and infinitely lived agents. All agents are subject to a common discount rate  $\rho$ . There are four possible labor market states: unemployment, self-employment, informal employment, and formal employment. The informal sector is composed by the self-employed and the informal employees.

Based on expectations in the labor market, agents make an irrevocable decision about which schooling level they want to acquire. For consistency with the empirical analysis, we assume only two schooling levels: high and low. They are denoted by  $h \in \{0, 1\}$ , with 1 indicating the higher level. Each agent incurs a direct, individual-specific cost  $\kappa \sim T(\kappa)$  when acquiring schooling level  $h = 1$ . They also sustain an opportunity cost equal to the value of participating in the market as an agent with  $h = 0$  over the period required to complete the additional schooling.

Agents enter the labor market and search for an employee job *only within* their own schooling submarket. They can search only as unemployed or as self-employed:<sup>19</sup> the searching state is denote by  $s \in \{0, 1\}$ , with 1 indicating searching as self-employed. The flow utilities of the searching states are:

$$(1) \quad (1 - s)\xi_h + sy + \beta_{0,h}B_0.$$

<sup>18</sup> See Arroyo Miranda et al. (2014) for more details on this data set. One concern with the job posting data is that they might be skewed toward the formal sector. It is unclear though how the employment service can guarantee that firms' postings comply with all labor regulations. Hence, we rather interpret the resulting vacancies as plausibly representative of both formal and informal jobs in urban Mexico.

<sup>19</sup> We rule out on-the-job search and therefore direct job-to-job transitions. Although this is a restriction that we impose for tractability, available data show that on-the-job search is relatively unimportant in Mexico. Our data source, ENOE 2005, does not collect information on continuous job spells, so we cannot use it to judge the importance of direct job-to-job transitions. However, an extra module that was added to ENOE in 2012 collects detailed retrospective records on labor market histories for a subset of the representative sample of workers covered in the survey. Out of the 872 observed employment spells, only 9% originated from a direct job-to-job transitions and only 2% involved a change in formality status.

If agents search as unemployed, they receive an instantaneous utility (or disutility) flow  $\xi_h$  that summarizes all costs and benefits of being an unemployed searcher.<sup>20</sup> If they search as self-employed, they enjoy an individual-specific self-employment income  $y \sim R(y|h)$ . Heterogeneity in  $y$  reflects differences in self-employment opportunities and skills. In both searching states, agents also receive a noncontributory benefit  $B_0$ . It is noncontributory because the agents receiving it do not provide any contributions to finance it. The benefit is fixed and distributed equally among all individuals.<sup>21</sup> The valuation that agents give to each pesos spent to provide the benefit is  $\beta_{0,h}$ , which represents the preferences for the nonmonetary components of the labor market state. Since we assume linear utility, it also has a direct interpretation as the marginal willingness to pay for the benefit.<sup>22</sup> The searching status also affects the rate at which workers meet firms, with unemployment being the state that is able to generate meetings with higher frequency. We describe in details this process in Subsection 3.4.1.

Just as agents search to find jobs, firms search to fill vacancies. They do so at a flow cost  $v_h$ . Firms can partially direct their search in each schooling submarket, within which we assume random matching with endogenous meetings governed by a matching function that we will define in Subsection 3.4.3. From the matching function, we can derive the Poisson rates at which workers meet firms and firms meet workers in each schooling submarket. A meeting between a potential employee and a firm produces a match-specific monetary value  $x \sim G(x|h)$ , which is time invariant and fully observed by both parties upon meeting.<sup>23</sup>

The labor relation when the match is formed may be formal or informal. We denote the formality status of the job with  $f \in \{0, 1\}$ , where  $f = 1$  indicates a formal job. The formality status of the job offer is decided by the firm optimally and it is taken as given by the worker. Assuming that the authority to set the formality status is in the hand of the firm is consistent with the institutional setting in Mexico and in other Latin American countries.<sup>24</sup>

Conditioning on  $x$  and  $f$ , workers and firms engage in bargaining to determine wages. At the end of the process, agents decide if accepting or rejecting the match. If the match is rejected, agents go back to search. If the match is accepted, agents stop searching and enjoy part of the output that is produced by the match. They will go back to search only if the match is terminated, an event that occurs at an exogenous Poisson rate  $\eta_h$ .

We will discuss the solution of the bargaining game in Subsection 3.4.3. Here, we simply anticipate that wages are an increasing function of the match-specific productivity  $x$ ; they depend on the worker's outside option governed by schooling level  $h$  and potential self-employment income  $y$ ; and they follow a different schedule depending on the formality status

<sup>20</sup> We let instantaneous utility  $\xi_h$  be schooling-specific to allow all the structural parameters that could be identified as schooling specific to be estimated as schooling specific. This empirical flexibility comes at a cost in terms of theoretical interpretation since a schooling choice translates in a preference change. We do not find completely impossible that an experience as formative as schooling could have some impact on preferences and search costs. In addition, these parameters may capture some residual selection on the portion of unobserved heterogeneity that we do not explicitly model. We adopt the same strategy for the preference over benefits  $\beta_{f,h}$ .

<sup>21</sup> This institution reflects the current social security system in Mexico and has become widespread in a large number of low-income and middle income countries (see Section 2 for details).

<sup>22</sup> A similar setting and interpretation is used by Dey and Flinn (2005) to evaluate health insurance and by Flabbi and Moro (2012) to evaluate job flexibility.

<sup>23</sup> This representation of firm-side and work-side heterogeneity is commonly used in search-matching-bargaining models of the labor market, including Eckstein and Wolpin (1995), Cahuc et al. (2006), and Flinn (2006). It is motivated by the theoretical work of Wolinsky (1987) and Jovanovic (1979). For a recent review, see Chapter 4.2 in Keane et al. (2011).

<sup>24</sup> As discussed on Subsection 2.1, it is the firms' responsibility to enroll their workers in the Social Security system and firms are the ones that have to pay fines and past contributions when found to hire illegally by the relevant authorities (IMSS). The same modeling choice is made by Bosch and Esteban-Pretel (2012) when describing the labor market in Brazil, another Latin American country sharing with Mexico a high informality rate and a wide range of institutional features.

$f$ . To summarize these relations, we denote the wage schedule with  $w_f(x; y, h)$ , leading to the following flow utilities of the employee states:

$$(2) \quad w_f(x; y, h) + \beta_{f,h}[fB_1(w_1(x; y, h)) + (1 - f)B_0],$$

where  $B_1(w_1(x; y, h))$  is the social security benefit received by formal employees and  $\beta_{1,h}$  is its valuation. Unlike the benefit  $B_0$  received by informal employees,  $B_1$  is contributory, that is, workers contribute to pay for it. The contribution is withdrawn at the source by firms and it is equal to a proportion  $t$  of the wage. A share  $\tau$  of the total contribution generates the proportional flow benefit  $\tau tw_1$ , which is meant to capture features of social security benefits such as defined contribution retirement plans. A share  $(1 - \tau)$  generates the constant flow benefit  $b_1$ , which is meant to cover health-related expenditures or similar benefits. As a result,  $b_1$  is an equilibrium object that depends on the proportion of formal employees in steady state and their (accepted) wage distributions.<sup>25</sup> This feature generates redistribution from high-wage earners to low-wage earners within the formal sector. Since  $b_1$  is not schooling-specific, it does also generate redistribution from workers with the high level of schooling to workers with the low level of schooling, introducing an equilibrium link between the two schooling submarkets.

Conditioning on  $x$  and  $f$ , firms' profit are equal to the match-specific productivity  $x$  net of labor costs. Labor costs include only wages when hiring informally but they also include social security contributions when hiring formally. Therefore, the firm's flow profits are:

$$(3) \quad x - w_f(x; y, h) - ftw_f(x; y, h),$$

where recall that  $t$  denotes the proportional contribution rates withdrawn at the source by the firms and split in proportions  $\tau$  and  $(1 - \tau)$  to finance, respectively, the proportional and the constant benefits defined above.

If hiring informally spare firms of paying the contribution  $t$ , it exposes them to be punished if they are audited by the authorities. Firms receive audits following an exogenous Poisson process with rate  $\chi_h$ . If they are audited while hiring formally, nothing happens. If they are audited while hiring informally, the informal match is terminated and the firm has to pay a one-shot monetary penalty  $c_h w_0(x; y, h)$ . This parameterization is meant to capture the partial enforcement of labor regulations discussed in Subsection 2.1. Firms are monitored with an intensity that allows for significant noncompliance; penalties are written in the law but levied with enough discretion that they can be approximated by an increasing function of observed wages. The simple linear form we use to approximate this function is driven by identification issues that we will discuss in Section 4.<sup>26</sup>

**3.2. Schooling Decision.** The first decision agents face is whether acquiring the high schooling level  $h = 1$  or remaining at the low schooling level  $h = 0$ . At this stage, agents are heterogeneous over two dimensions: the direct cost of schooling, denoted with  $\kappa \sim T(\kappa)$ , and the opportunity cost of schooling, which is heterogeneous since it depends on the self-employment income  $y \sim R(y|0)$ . Both  $\kappa$  and  $y$  are assigned by nature, they are time-invariant and fully observed.<sup>27</sup>

<sup>25</sup> Therefore,  $B_1(w_1(x; y, h)) \equiv \tau tw_1(x; y, h) + b_1$ . To ease notation, we do not make explicit the dependence of  $b_1$  on all the other parameters of the model. See Appendix B.4 for the formal derivation of  $b_1$  in equilibrium.

<sup>26</sup> This parameterization is similar to the crime model presented in Burdett et al. (2003) where employed workers who are caught while committing a crime have to leave their job and are forced into unemployment. In our model, the "crime" is working informally. The penalty is a function of  $w_0$  and not  $w_1$  since the inspector can only observe wages, not productivity. We rule out the possibility that after the inspection, the firm may hire the worker in a formal job since (i) the worker may not accept to keep working at the firm for a (lower) formal wage and benefit; (ii) it would essentially entail a new job because of the major restructuring happening at the firm; and (iii) firms may actually shut down as a result of the inspection.

<sup>27</sup> The cost of schooling parameter  $\kappa$  summarizes any monetary and utility costs associated with acquiring additional schooling. The cost of acquiring schooling level  $h = 0$  is normalized to zero.

Denote with  $Q(y, h)$  the present discounted value (PDV) of participating in the labor market with schooling level  $h$  and self-employment income  $y$ . Agents decide their schooling level based on:

$$(4) \quad \max \left\{ Q(y, 0); \int_0^{\bar{t}} -\kappa \exp(-\rho t) dt + \exp(-\rho \bar{t}) \int_y Q(y', 1) dR(y'|1) \right\},$$

where  $\bar{t}$  is the additional time required to complete the schooling level  $h = 1$ . To complete it, agents have to pay a direct cost  $\kappa$  while in school. In addition, they sustain an opportunity cost equal to the value of participating in the market with schooling  $h = 0$  over the period  $[0, \bar{t}]$ . At the time of the decision, the self-employment income with  $h = 0$ —denoted with  $y$  in the maximization problem (4)—is known, whereas the self-employment income with  $h = 1$ —denoted with  $y'$  in (4)—is not yet known. It will be assigned to the agent only when the additional schooling level is completed.

For given  $y$ , the cost of acquiring additional schooling is increasing in  $\kappa$ . Therefore, it exists a unique  $\kappa^*(y)$  that makes the agent indifferent between  $h = 1$  and  $h = 0$ . The optimal decision rule is then a reservation value rule where, for given  $y$ , only agents with cost of schooling low enough ( $\kappa < \kappa^*(y)$ ) acquire the additional schooling level  $h = 1$ .<sup>28</sup>

**3.3. Searching Status Decision.** Once schooling is completed, agents enter the labor market and decide if searching as self-employed or as unemployed. In deciding the searching state, agents trade off the flow values  $y$  and  $\xi_h$  with the different rates at which they meet employers. Since searching is the only activity of the unemployed, their search intensity is higher than the one the self-employed can provide. As a result, the unemployed meet employers at a higher rate than the self-employed: A searcher may then be willing to give up his self-employment income  $y$  to increase the probability to receive employee offers. We represent this dynamic in a reduced form way by setting:

$$(5) \quad \gamma_h = \psi_h \lambda_h, \psi_h \in (0, 1),$$

where  $\lambda_h$  is the unemployed's (Poisson) arrival rate of offers and  $\gamma_h$  the self-employed's one.<sup>29</sup>

We denote the value of searching while unemployed with  $U(h)$  and the value of searching while self-employed with  $S(y, h)$ .<sup>30</sup> The decision of the searching state is then equivalent to the following maximization:

$$(6) \quad Q(y, h) \equiv \max\{S(y, h); U(h)\}.$$

As shown in Lemma 1 in Appendix B.1, the dependence of  $S(y, h)$  on  $y$  is obvious but the sign of this dependence is not. Intuitively, even if the flow value of a self-employed searcher is clearly increasing in  $y$ , the option value may not be. This is the case because the option value includes the possibility to work as a formal or informal employee but the formality status is chosen by the firm, not the worker. Firms offer and commit to a formality status strategically, attempting to extract a higher share of the surplus—how effective they are in doing so depends on  $y$ , creating the nonmonotonicity stated above.

<sup>28</sup> For some  $y$ , the value of participating in the market with  $h = 0$  may be so high that the opportunity cost is enough to discourage acquiring additional schooling. Since we will assume that direct cost is extracted from a distribution with positive support, all the agents with that  $y$  will choose to stay at  $h = 0$ .

<sup>29</sup> If we assume search effort per unit of time to be the same for all searcher, a possible interpretation for  $\psi_h$  is the proportion of time the self-employed devote to search for each unit of time the unemployed devote to search. Suppose  $\psi_h = 0.1$ : if an unemployed individual searches for 40 hours a week,  $\psi_h = 0.1$  may be interpreted as the self-employed searching for four hours a week.

<sup>30</sup> Once the agents decide to search as unemployment, their value of  $y$  is irrelevant, which is why we do not index  $U(h)$  with  $y$ .

Since the sign of the dependence between  $S(y, h)$  and  $y$  is ambiguous, we cannot formally prove that the solution of the maximization problem (6) is characterized by a unique reservation value over  $y$ . This fact, together with our endogenous meeting rates formulation (see Subsection 3.4.3), may lead to equilibrium multiplicity—a major impediment when attempting to bring the model to the data. Faced with these difficulties, we choose to conjecture the sign of the dependency, and then, verify that it is satisfied at our parameter estimates through simulations. We therefore state the following:

Conjecture 1.  $S(y, h)$  is monotone increasing in  $y$ .<sup>31</sup>

Since  $U(h)$  is constant in  $y$ , Conjecture 1 is enough to conclude that exists a unique:

$$(7) \quad y^*(h) : S(y^*(h), h) = U(h).$$

Under (7), the solution of the maximization problem (6) has a simple reservation rule property: only agents with self-employment income high enough ( $y \geq y^*(h)$ ) search for an employee job while also working as self-employed.

At this stage, it is also useful to introduce notation to represent the workers' measures in the different labor market states. We denote with  $b(y|h)$ ,  $e(y|h)$ , and  $l(y|h)$  the steady-state measures for, respectively, searchers, informal employees, and formal employees. Each measure is a function of the potential self-employed income  $y$ , given schooling level  $h$ . The population is normalized by schooling so that  $b(y|h) + e(y|h) + l(y|h) = r(y|h)$  where  $r(y|h)$  denotes the PDF of the CDF  $R(y|h)$ . As a result:

$$(8) \quad \int_y \{b(y|h) + e(y|h) + l(y|h)\} dy = \int_y r(y|h) dy = 1.$$

To proceed, it is useful to write the following expression for the steady-state measure of searchers meeting firms:

$$(9) \quad \psi(h)b(h) \equiv \int_0^{y^*(h)} b(y|h, y < y^*(h)) dy + \psi_h \int_{y^*(h)}^{\infty} b(y|h, y \geq y^*(h)) dy,$$

where  $\psi(h)$  is a function equal to 1 if the searcher is unemployed and equal  $\psi_h$  if the searcher is self-employed. Recall from Equation (5) that  $\psi_h \in (0, 1)$  is the parameter that represents the lower search intensity of the self-employed with respect to the unemployed. The equilibrium version of expression (9) and its components are provided in Appendix B.3.

### 3.4. Labor Market Decisions.

3.4.1. *Firms.* When firms fill a vacancy, they receive the flow profits defined in Equation (3) but they are subject to the termination shock  $\eta_h$  and the auditing shock  $\chi_h$ , leading to the following value functions of filled jobs:

$$(10) \quad (\rho + \eta_h + \chi_h)F_0[x, y, h] = x - (1 + \chi_h c_h)w_0(x; y, h) + (\eta_h + \chi_h)V[h],$$

$$(11) \quad (\rho + \eta_h)F_1[x, y, h] = x - (1 + t)w_1(x; y, h) + \eta_h V[h],$$

<sup>31</sup> In Figure B.2 of Appendix B.3, we show that the conjecture holds at our parameter estimates.

where  $V[h]$  denotes the value function of a vacancy posted in schooling market  $h$  and it will be defined in Equation (14). When firms hire informally, they do not pay any social security contribution but if audited, they have to pay for the one-shot penalty  $c_h w_0(x; y, h)$  and terminate the match, as shown in the expression for the value function (10). When firms hire formally, they do pay for the social security contribution at the proportional rate  $t$ , whereas if audited, there are no consequences since they comply with the law, as shown in the expression for the value function (11).

When firms meet a worker, they observe his schooling level  $h$ , his outside option  $Q(y, h)$ , and the match-specific productivity  $x$ . They then decide the formality status, and commit to it, by solving:

$$(12) \quad \max\{F_0[x, y, h]; F_1[x, y, h]\}.$$

The set of match-specific values such that the firm offers a formal contract is denoted by:

$$(13) \quad \mathcal{A}(y, h) \equiv \{x : F_1[x, y, h] > F_0[x, y, h]\}.$$

We are now ready to propose the value functions of the posted vacancy in schooling market  $h$ :

$$(14) \quad \begin{aligned} (\rho + \zeta_h)V[h] &= v_h + \frac{\zeta_h}{\psi(h)b(h)} \int_0^{y^*(h)} b(y|h, y < y^*(h)) \\ &\quad \times \left\{ \int_x \max\{F_1[x, y, h], F_0[x, y, h], V[h]\} dG(x|h) \right\} dy \\ &\quad + \frac{\zeta_h}{\psi(h)b(h)} \int_{y^*(h)}^\infty \psi_h b(y|h, y \geq y^*(h)) \\ &\quad \times \left\{ \int_x \max\{F_1[x, y, h], F_0[x, y, h], V[h]\} dG(x|h) \right\} dy, \end{aligned}$$

where the flow cost of keeping a vacancy open is denoted by  $v_h$  and the optimal decisions about searching behavior (expressions (7) and (9)) are already taken into account. Employers meet potential employees at a rate  $\zeta_h$  but agents with different  $y$ s have a different probability to be in the searching state. This is captured by the integral of the steady-state measures over  $y$ . Once the employer meets a searcher, a match-specific productivity value is extracted. Based on  $x$  and the knowledge of the outside option of the potential employee (a function of  $y$  and  $h$ ), the employer optimally decides if offering a formal or informal job, as shown by the max operator over three possible options:  $F_0[x, y, h]$ ,  $F_1[x, y, h]$ , and the status quo option  $V[h]$ .

3.4.2. *Workers.* The value functions of the employee states are the sum of the flow utilities defined in Equation (2) and of the corresponding continuation values, leading to:

$$(15) \quad (\rho + \eta_h + \chi_h)E_0[w_0(x; y, h), y, h] = w_0(x; y, h) + \beta_{0,h}B_0 + (\eta_h + \chi_h)Q(y, h)$$

$$(16) \quad (\rho + \eta_h)E_1[w_1(x; y, h), y, h] = w_1(x; y, h) + \beta_{1,h}B_1[w_1(x; y, h)] + \eta_h Q(y, h).$$

The overall discount includes the intertemporal discount rate  $\rho$  and the Poisson rates of the two possible shocks that may hit the state: the termination shock at rate  $\eta_h$  and the auditing shock at rate  $\chi_h$ . Audit shocks occur both at formal and informal jobs but they have no con-



sequences when the job is formal, which is the reason why  $\chi_h$  does not appear among the discount factors in Equation (16).

The value functions of the searching states are defined as:

$$(17) \quad (\rho + \lambda_h)U(h) = \xi_h + \beta_{0,h}B_0 + \lambda_h \left\{ \int_{-\mathcal{A}(0,h)} \max\{E_0[w_0(x), 0, h], U(h)\}dG(x|h) \right. \\ \left. + \int_{\mathcal{A}(0,h)} \max\{E_1[w_1(x), 0, h], U(h)\}dG(x|h) \right\},$$

$$(18) \quad (\rho + \gamma_h)S(y, h) = y + \beta_{0,h}B_0 + \gamma_h \left\{ \int_{-\mathcal{A}(y,h)} \max\{E_0[w_0(x), y, h], S(y, h)\}dG(x|h) \right. \\ \left. + \int_{\mathcal{A}(y,h)} \max\{E_1[w_1(x), y, h], S(y, h)\}dG(x|h) \right\}.$$

The only possible shock is meeting a potential employers, an event occurring at the rate  $\lambda_h$  or  $\gamma_h$  depending on the searching state. The continuation values show that the firms are the ones deciding on the formality status. When the match-specific value belongs to the set  $\mathcal{A}(y, h)$  defined in Equation (13), the firm will offer a formal job; otherwise, it will offer an informal job. In the first case, the worker decides whether or not to accept the match by comparing the searching state with being employed in a formal job; in the latter case, he decides by comparing the searching state with being employed in an informal job.

This general behavior is the same if the workers search as an unemployed ( $U(h)$ ) or as a self-employed ( $S(y, h)$ ). The difference is the value of the outside option and the arrival rate of offers. The value of the outside option of the self-employment searchers is heterogeneous since it depends on the self-employment income  $y$ . The value of the outside option of the unemployed searchers is instead homogeneous since it is independent of the self-employment income. To emphasize this lack of dependence, we insert 0 in place of  $y$  in all the relevant functions and definitions. The second difference between the two searching state is the very reason why unemployed searchers may decide to give up on their self-employment income. They do so to receive offers at a higher rate than the self-employed ( $\lambda_h > \gamma_h$ ).

3.4.3. *Meetings and wages.* We assume random matching in the meetings between workers and firms but we allow the contact rates to be endogenous. Following previous literature,<sup>32</sup> we assume contact rates governed by a Cobb–Douglas matching function. By denoting with  $m(h)$  the number of matches per worker in schooling group  $h$ , we assume:

$$(19) \quad m(h) = [\psi(h)b(h)]^{\iota_h} [v(h)]^{1-\iota_h},$$

where  $\psi(h)b(h)$  is the measure of searchers in the economy introduced in Equation (9) and  $v(h)$  is the measure of vacancies. We can now make explicit that all the contact rates introduced in Subsections 3.4.2 and 3.4.1 are endogenous since they are all function of the tightness:

$$(20) \quad \omega(h) \equiv \frac{v(h)}{\psi(h)b(h)}.$$

Equations (9), (19), and (20) then imply:

$$\lambda_h = \frac{m(h)}{\int_0^{y^*(h)} b(y|h, y < y^*(h))dy} \frac{\int_0^{y^*(h)} b(y|h, y < y^*(h))dy}{\psi(h)b(h)} = \omega(h)^{1-\iota_h},$$

<sup>32</sup> See Petrongolo and Pissarides (2001) for a survey. See Meghir et al. (2015) and Bosch and Esteban-Pretel (2012) for applications to Latin American countries.

$$\gamma_h = \frac{m(h)}{\int_{y^*(h)}^{\infty} b(y|h, y \geq y^*(h))dy} \frac{\psi_h \int_{y^*(h)}^{\infty} b(y|h, y \geq y^*(h))dy}{\psi(h)b(h)} = \psi_h \omega(h)^{1-t_h},$$

$$\zeta_h = \frac{m(h)}{v(h)} = \omega(h)^{-t_h}.$$

We assume that wages are determined upon meeting by engaging in bargaining. Following the protocol typically used in search-matching-bargaining models estimated on micro data,<sup>33</sup> we assume the axiomatic Nash bargaining solution. Under our regularity conditions, the solution to the bargaining problem is equivalent to:

$$(21) \quad \max_{w|f} \{E_f[w, y, h] - Q(y, h)\}^{\alpha_h} \{F_f[x, y, h] - V[h]\}^{(1-\alpha_h)}, \alpha_h \in (0, 1).$$

3.5. *Equilibrium.* To define equilibrium conditions and optimal decision rules, it is useful to start from the firms' entry decision. We assume free entry of firms in both markets. Since firms enter only if the value of posting a vacancy is positive, in equilibrium, we obtain:

$$(22) \quad V[h] = 0.$$

Imposing condition (22), the solution to (21) generates the wage schedules:

$$(23) \quad w_0(x; y, h) = \frac{\alpha_h}{1 + \chi_h c_h} x + (1 - \alpha_h)[\rho Q(y, h) - \beta_{0,h} B_0],$$

$$(24) \quad w_1(x; y, h) = \frac{\alpha_h}{1 + t} x + \frac{(1 - \alpha_h)}{1 + \beta_{1,h} \tau t} [\rho Q(y, h) - \beta_{1,h} b_1].$$

The wage schedules have the usual structure generated by Nash bargaining in this context: they are a convex combination of the match-specific productivity values  $x$  and the values of the worker's outside option  $\rho Q(y, h)$ . The higher the working bargaining coefficient  $\alpha_h$ , the higher the weight on  $x$ . On top of this usual structure, the two wage schedules show the impact of the institutional parameters. Both the contribution rate  $t$  and the expected cost of hiring informally  $\chi_h c_h$  are partially transferred to the worker implying a negative relationships with wages at any  $x$ . The nonwage benefits of the employment relationship ( $\beta_{1,h} b_1$  and  $\beta_{0,h} B_0$ ) also decrease wages at any  $x$  since the benefits are valued by the worker.

Incorporating the wage schedules in the value functions leads to optimal decision rules based on reservation values defined over the match-specific productivity. The match-specific productivity value that makes the firm indifferent between offering a formal or an informal job is:<sup>34</sup>

$$\begin{aligned} \tilde{x}(y, h) : F_0[\tilde{x}(y, h), y, h] &= F_1[\tilde{x}(y, h), y, h] \\ &\Leftrightarrow \\ (25) \quad \tilde{x}(y, h) &= \frac{1}{\chi_h} \{ \Gamma_0 \beta_{0,h} B_0 - \Gamma_1 \beta_{1,h} b_1 + (\Gamma_1 - \Gamma_0) \rho Q(y, h) \}, \end{aligned}$$

where:

<sup>33</sup> See, for example, Eckstein and Wolpin (1995), Cahuc et al. (2006), Flinn (2006), Flabbi and Moro (2012), and Borowczyk-Martins et al. (2018).

<sup>34</sup> The reservation value  $\tilde{x}(y, h)$  only guarantees indifference on the firms' side but generally not on the workers' side. This is a direct implication of the firms' advantage in deciding first the formality status.

$$\begin{aligned}\Gamma_0 &= (\rho + \eta_h)(1 + \chi_h c_h), \\ \Gamma_1 &= (\rho + \eta_h + \chi_h)\phi_h, \\ \phi_h &\equiv \frac{1 + t}{1 + \beta_{1,h}\tau t}.\end{aligned}$$

The match-specific productivity value that makes both the firm and the worker indifferent between accepting the match as informal employee or continue searching is:<sup>35</sup>

$$\begin{aligned}x_0^*(y, h) : F_0[x_0^*(y, h), y, h] = 0 &\Leftrightarrow E_0[w_0[x_0^*(y, h)], y, h] = Q(y, h), \\ &\Leftrightarrow \\ (26) \quad x_0^*(y, h) &= (1 + \chi_h c_h)[\rho Q(y, h) - \beta_{0,h}B_0],\end{aligned}$$

while the value that makes them indifferent between accepting the match as formal employee or continue searching is:

$$\begin{aligned}x_1^*(y, h) : F_1[x_1^*(y, h), y, h] = 0 &\Leftrightarrow E_1[w_1[x_1^*(y, h)], y, h] = Q(y, h) \\ &\Leftrightarrow \\ (27) \quad x_1^*(y, h) &= \phi_h[\rho Q(y, h) - \beta_{1,h}b_1].\end{aligned}$$

Equations (26)–(27) state that the job formality status  $f \in \{0, 1\}$  has two opposite effects on the reservation productivity values at which the match is formed. It decreases the reservation value because employees receive additional benefits associated with the match ( $b_1$  or  $B_0$ ), but it also increases the reservation value because the firm faces some costs ( $t$  or  $\chi_h c_h$ ) to activate one or the other job contract. Because of these opposite effects, the equilibrium is characterized by potentially different optimal decision rules for different sets of parameters' values and different combinations of  $\{y, h\}$ . Still, all the decision rules retain the reservation value property.

To illustrate the equilibrium, we report an example in Figure 2. We condition on specific values of  $\{y, h\}$  and we use a specific set of parameters' values that describe well the patterns observed in the data since all labor market states survive in equilibrium: both searching states, informal employment, and formal employment. For low values of the match-specific productivity ( $x < x_0^*(y, h)$ ), firms prefer to keep the vacancy open and workers prefer to continue searching. For intermediate values ( $x_0^*(y, h) \leq x < \tilde{x}(y, h)$ ), firms offer informal jobs that are accepted by workers receiving a wage governed by (23). For larger values ( $\tilde{x}(y, h) \leq x$ ), firms offer formal jobs that are accepted by workers receiving a wage governed by (24).

The following proposition summarizes the equilibrium discussion and lists the potentially different optimal decision rules that arise for different sets of parameters and different combinations of  $\{y, h\}$ . The proof is provided in Appendix B.2.

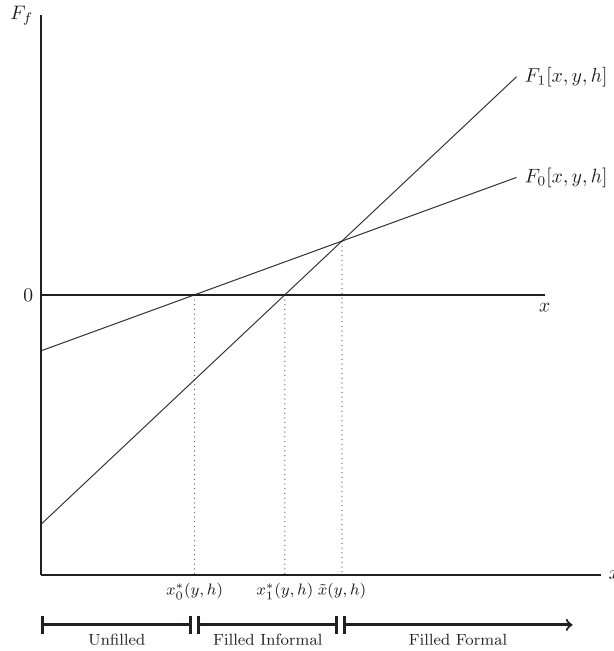
**PROPOSITION 1 (OPTIMAL DECISIONS RULES IN THE LABOR MARKET).** *For any  $\{y, h\}$ , the reservation values  $\{x_0^*(y, h), x_1^*(y, h), \tilde{x}(y, h)\}$  exist and are unique.*

*They characterize the optimal decision rules as follows:*

1. *If  $0 < \tilde{x}(y, h) < x_1^*(y, h) < x_0^*(y, h)$  or  $\tilde{x}(y, h) < 0 < x_1^*(y, h) < x_0^*(y, h)$ :*

$$\begin{array}{ll} \text{when } 0 &\leq x < x_1^*(y, h), & \text{the match is rejected} \\ \text{when } x_1^*(y, h) &\leq x, & \text{the match is accepted with formality status } f = 1 \end{array}$$

<sup>35</sup> The agreement result is assured by the axiomatic Nash bargaining solution.



NOTE: The figure shows the equilibrium when  $\bar{x}(y, h) > x_1^*(y, h)$  for given  $\{y, h\}$ . For the definitions of  $F_0$ ,  $F_1$ ,  $x_0^*$ ,  $x_1^*$ , and  $\bar{x}$ , see Equations (10), (11), (26), (27), and (25).

FIGURE 2

OPTIMAL DECISION RULES AND FIRMS' VALUE FUNCTIONS

2. If  $0 < x_0^*(y, h) < x_1^*(y, h) < \bar{x}(y, h)$ :

- when  $0 \leq x < x_0^*(y, h)$ , the match is rejected
- when  $x_0^*(y, h) \leq x < \bar{x}(y, h)$ , the match is accepted with formality status  $f = 0$
- when  $\bar{x}(y, h) \leq x$ , the match is accepted with formality status  $f = 1$

3. If  $x_0^*(y, h) < 0 < x_1^*(y, h) < \bar{x}(y, h)$  or  $x_0^*(y, h) < x_1^*(y, h) < 0 < \bar{x}(y, h)$ :

- when  $0 \leq x < \bar{x}(y, h)$ , the match is accepted with formality status  $f = 0$
- when  $\bar{x}(y, h) \leq x$ , the match is accepted with formality status  $f = 1$

4. Otherwise:

- when  $0 \leq x$ , the match is accepted with formality status  $f = 1$ .

The intuition for the existence and uniqueness of the reservation values  $\{x_0^*(y, h), x_1^*(y, h)\}$  is that the value functions for filled jobs are increasing in the match-specific value  $x$ , while the value of a vacancy is constant in  $x$ . Existence and uniqueness of  $\bar{x}(y, h)$  is instead a consequence of comparing the elasticity with respect to  $x$  of the value of jobs filled formally or informally: since the first is higher, for high enough values of match-specific productivity, firms will always prefer to hire formally; for low enough values, they will always prefer to hire informally or not at all. An implication of this result is that, for some parameters values, no firm offers informal jobs and therefore no worker is hired as an informal employee. This possibility is described in Cases 1 and 4 of Proposition 1. Instead, Case 2 describes a situation equivalent to the one discussed in Figure 2 where both formal and informal employees exist in equilibrium.

To close the definition of the equilibrium, the optimal decisions rules of Proposition 1 and the wage expressions (23)–(24) are incorporated in the value functions. The optimal decisions rules are also used to solve for the steady-state measures in each labor market state, conditioning on the schooling level. Finally, the steady measure of individuals with high schooling level is determined by the optimal decision rule described in expression (4). This process leads to equations that define a fixed point in  $\{U(h), S(y, h)\}$  as stated in the equilibrium definition 2 reported in Appendix B.3.

An important remark is that the schooling decision introduces the possibility of multiple equilibria. The possibility arises even under the Conjecture 1 that we proposed in Subsection 3.3. The intuition is as follows: Consider the process of firms entering the labor market for schooling level  $h = 1$ . The direct effect is a tighter market for firms but a better market for workers. As a result, more workers acquire schooling level  $h = 1$ , entering the market in larger numbers. But this additional entrance of workers makes the market more attractive to firms, potentially countering the direct effect. If these two opposing forces are enough to create multiple equilibria that depend on parameters but due to the complexity of the model, we are unable to characterize the parameter regions where multiplicity occurs. Since multiplicity greatly complicates the identification and estimation of the model with the data at our disposal, we follow the same strategy we proposed in Subsection 3.3: we make a conjecture that rules out equilibrium multiplicity and then verify that it is satisfied at our parameter estimates through simulations. We therefore state the following:

Conjecture 2.  $V[h]$  is monotone decreasing in the vacancy rate  $v(h)$ .<sup>36</sup>

Under Conjecture 2, the standard congestion externalities arise and this source of multiplicity is ruled out. For a more detailed discussion, a formal definition of the equilibrium and the simulation analysis generated by our estimated model, see Appendix B.3.

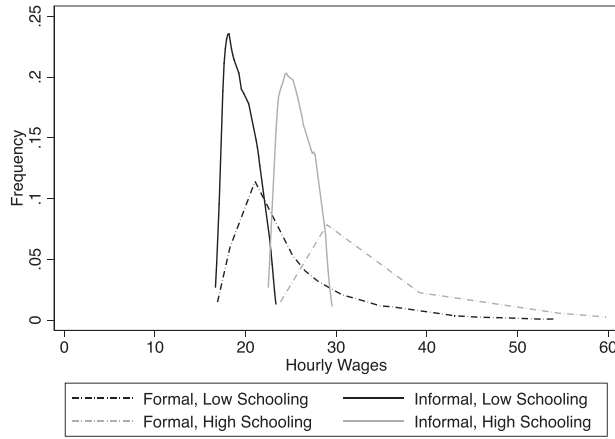
**3.6. Empirical Implications.** The model is able to replicate and explain the main empirical evidence that characterizes labor markets with high informality, as described in Section 2 for Mexico.

The first set of stylized facts is the significant mass of workers in each labor market state and the significant amount of transitions between the formal and informal status. In the model, individuals can accept jobs with different formality status as a result of different values of match-specific productivity. Since the same worker may receive different draws of match-specific productivity over his labor market careers, some draws may lead to formal jobs and other to informal jobs, generating the types of transitions and proportions we observe in the data.

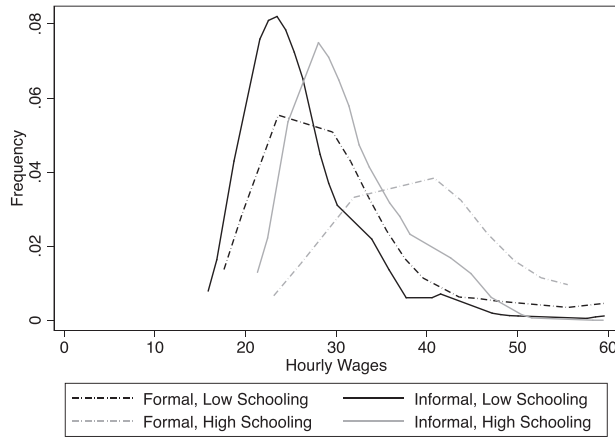
The second set of stylized facts refers to the wage distributions. Formal employees have on average higher wages than informal employees but the two wage distributions overlap over a large portion of their support. The evidence can be replicated thanks to two features of the equilibrium of the model. First, for given  $\{y, h\}$ , the reservation productivity to accept a formal job is higher than the one to accept an informal job. Since wages are increasing in productivity for given formality status, the ranking in reservation productivity maps into the ranking in reservation wages, generating a difference in accepted wages that is consistent with the data. Second, for given  $\{y, h\}$ , formal employees earn a lower net wage than informal employees with *same* productivity because they receive higher nonwage benefits—that is,  $\beta_{1,h}B_1[w_1(x; y, h)] > \beta_{0,h}B_0$ .<sup>37</sup> This channel generates an overlap in the neighborhood of  $\bar{x}(y, h)$ , the indifference point between formal and informal job, which is limited to a neighborhood of  $\bar{x}(y, h)$ . While necessary, this mechanism may not be sufficient to generate the

<sup>36</sup> In Appendix B.3, Figure B.3, we show that the conjecture holds at our parameter estimates.

<sup>37</sup> Almeida and Carneiro (2012) emphasize the same argument in an application focusing on enforcement of labor regulations in Brazil. Their empirical results based on regional variations in inspections are consistent with our empirical implications—that the jobs more susceptible to switch from formal to informal are those relatively lower paid.



(a) Employees With Outside Option of Unemployment



(b) Employees With Outside Option of Self-employment

NOTE: Simulated sample of 10,000 worker-level observations based on the estimates reported in Table 3. The figure shows the empirical densities of the accepted hourly wages (in Mexican pesos) for employees by formality status of the job, schooling level, and search status. The support of the distributions is restricted to observations below 60 Mexican pesos.

FIGURE 3

SIMULATED ACCEPTED WAGE DISTRIBUTIONS BY SEARCH STATUS

large overlap between the wage distributions of formal and informal jobs that we observe in the data. This result is delivered by the heterogeneity in the value of the outside options of the self-employed workers. As we mentioned, the results above are valid for given  $\{y, h\}$ . By changing  $y$  for given  $h$ , the reservation values change and the region of overlap changes. Enough heterogeneity in  $y$  generates a mixture distribution with a much larger region of overlap. Figure 3 shows these features using simulations generated by our estimated model. The top panel shows the overlap in the simulated accepted wages considering only workers exiting unemployment.<sup>38</sup> The overlap is present but is limited to a relative narrow portion of the support. The bottom panel considers only workers transiting from self-employment to formal and informal employment. The overlap is much larger, covering the entire support of the simulated wage distributions.

<sup>38</sup> This is equivalent to conditioning on the same  $y$  since unemployed workers share the same reservation values.



The third set of stylized facts refers to differences in the labor market dynamic of the unemployed and the self-employed. The two states exhibit systematic differences both in the hazard rates out of the state and in the wages accepted when leaving the state. The model delivers endogenously different reservation wages, generating the differences in accepted wages. It also allows for different arrival rates, adding to the ability of matching the different hazard rates.

The last set of stylized facts is the observation that not only workers but also firms are not neatly allocated to the formal and informal sector. Some firms hire formal or informal workers at different points in time just as workers transit over time between different formality status (Ulyseas, 2018). Our model captures this evidence well: ex ante identical firms may hire formally or informally depending on the match-specific productivity value. When the match is terminated, a new search cycle starts and may lead to a new match with a different formality status.

#### 4. IDENTIFICATION

The data sets that we use in the empirical analysis were presented in Subsection 2.2 and include both individual-level data and aggregate data. In turn, the microdata include two samples, the *cross-sectional* sample and the *panel* sample. From the cross-sectional sample, we extract all the earnings and income moments and the steady-state proportions in the four labor market states. From the panel sample, we extract all moments that describe the labor market dynamics: transitions between labor market states and hazard rates out of the searching states. The aggregate-level data include the economy-wide labor shares, the municipality-level roll out of the *Seguro Popular* program, and the schooling-specific vacancy rates.

All the identification discussions assume the knowledge of the institutional parameters  $\{B_0, \tau, t\}$ . We discuss their values in detail in Appendix A.2. They are assumed observed and set in accordance to the institutional setting of the Mexican labor market.

We organize the identification discussion into five parts to provide a better intuition of which data features have the greatest impact on particular parameters. However, there is no formal partition of the parameters that maps into a specific subset of moments. All moments effectively contribute to the identification of all the parameters. First, we discuss the usual search, matching and bargaining parameters. Second, we focus on the preferences for social security benefits and the costs to firms of hiring informally. Third, we consider the identification of the matching function and the other demand side parameters. Fourth, we discuss how we recover the parameters governing the distribution of the cost of schooling. Finally, we introduce and discuss the identification of additional unobserved workers heterogeneity.

*4.1. Search, Matching, and Bargaining Parameters.* We start this section focusing on the mobility parameters  $\{\lambda_h, \eta_h\}$ , on the match-specific productivity distribution  $G(x|h)$ , the flow unemployment utility  $\xi_h$ , and the discount rate  $\rho$ . Flinn and Heckman (1982) show that proportions in labor market states and duration information (or, as in our case, transitions between labor market states) identify hazard rates out of the searching state and termination rates out of employment. No additional progress in the identification of the model can be made without a parametric assumption on the exogenous match-specific productivity distribution. Under the assumption that the distribution is recoverable—that is, it can be identified by observing its truncation, the argument proceeds as follows: Observed wages in the data correspond to accepted wages in the model. Accepted wages in the model can be mapped into accepted match-specific productivity by inverting the wage schedules. The truncated productivity distribution then identifies the primitive  $G(x|h)$ .

The flow utility from unemployment  $\xi_h$  and the discount rate  $\rho$  can only be jointly identified, as shown in our case by the equilibrium equation (B.11). In conclusion, the implication of the Flinn and Heckman (1982)'s results for our model is that we have to assume a recoverable distribution and that we have to fix either  $\xi_h$  or  $\rho$ . We follow previous literature (Eckstein and van den Berg, 2007) by assuming that the productivity distribution belongs to a

two-parameter log-normal distribution and denote its schooling-specific parameters  $\{\mu_{x,h}, \sigma_{x,h}\}$ . The log-normal is recoverable and has shown to guarantee a good fit of wage data. We also follow previous literature by choosing to fix the discount rate at 5% a year.

Our identification departs from this standard case for three reasons. First, the mapping between wages and productivity contains both unknown parameters and individual unobserved heterogeneity. Second, the searching state is heterogeneous since it contains not only the set of individuals who are unemployed but also the self-employed workers who are heterogeneous in  $y$ . Third, the arrival rates of job offer from the searching states  $\lambda_h$  and  $\gamma_h$  are not primitive parameters but they are derived in equilibrium through the matching function (19).

The first departure from the standard case is shown in Equations (23)–(24). The mapping between wages and productivity contains the following unknown parameters: the preference parameters  $\{\beta_{1,h}, \beta_{0,h}\}$ , the cost of informality parameters  $\{\chi_h, c_h\}$ , and the Nash-bargaining coefficient  $\alpha_h$ . We will discuss the first two sets of parameters in the next section, clarifying that the identification of the parameters in this and the next section is closely connected. In other words, all the data features discussed in both sections jointly contribute to the identification of all the parameters discussed in both sections. With respect to the Nash-bargaining coefficient  $\alpha_h$ , previous literature has shown that the parameter is hard to identify without demand-side information. As a result, numerous contributions simply calibrate the parameter to a fixed value.<sup>39</sup> Instead, we have chosen to use a limited amount of demand-side information—aggregate labor shares—to identify  $\alpha_h$ .<sup>40</sup> Suppose that all the parameters of the wage schedules (24) and (23) are identified with the exception of  $\alpha_h$ . By rewriting the wage schedules as:

$$(28) \quad w_0(x; y, h) = \alpha_h \left\{ \frac{x}{1 + \chi_h c_h} - [\rho Q(y, h) - \beta_{0,h} B_0] \right\} + [\rho Q(y, h) - \beta_{0,h} B_0]$$

$$(29) \quad w_1(x; y, h) = \alpha_h \left\{ \frac{x}{1 + t} - \frac{[\rho Q(y, h) - \beta_{1,h} b_1]}{(1 + \beta_{1,h} \tau t)} \right\} + \frac{[\rho Q(y, h) - \beta_{1,h} b_1]}{(1 + \beta_{1,h} \tau t)},$$

it is immediate to see that the  $\alpha_h$  parameter is governing the portion of the surplus (net production minus the value of the outside option) appropriated by the worker through the wage. Since labor shares are the ratio between the aggregate value of worker's wages  $w_f(x; y, h)$  and the aggregate value of production  $x$ , their observation provides sufficient information to identify  $\alpha_h$ .<sup>41</sup> In our setting, it would be desirable to allow for schooling-specific  $\alpha_h$ . However, since we cannot observe schooling-specific labor shares, we have to impose:

$$(30) \quad \alpha_1 = \alpha_0 \equiv \alpha.$$

The other issue affecting the mapping between wages and productivity is the presence of individual unobserved heterogeneity. Consider the wage schedules (23)–(24): they both show that wages are a function of  $y$  through the value of the outside option  $Q(y, h)$ . However, when we observe an individual currently working as an employee, we observe his wage but we do

<sup>39</sup> See Cahuc et al. (2006) and Flinn (2006) for a formal discussion and Eckstein and Wolpin (1995) for a seminal contribution. A typical value at which  $\alpha_h$  is calibrated is 0.5, corresponding to the symmetric bargaining case (Flabbi and Moro, 2012). Other contributions set the value lower than 0.5 if they perceive, the workers have a weaker bargaining position (Borowczyk-Martins et al., 2018).

<sup>40</sup> See Flinn (2006) for a detailed and formal discussion. See Dey and Flinn (2005) and Flinn and Mullins (2015) for applications using the same type of information we use in our application.

<sup>41</sup> As shown in Flinn (2006), asymptotic consistency is attained either by observing a firm with a sufficiently large number of workers or by observing a sufficiently large number of firms. Since our labor shares come from aggregate data, we follow Flinn and Mullins (2015) by appealing to the second result.

not in general observe his previous searching state, and therefore, we do not observe his corresponding  $y$ . As a result,  $y$  works as the individual unobserved heterogeneity discussed in Subsection 2.5.2 of Flinn and Heckman (1982). As long as the unobserved heterogeneity belongs to an appropriate parametric distribution, it can be integrated out and its identification is helped by support conditions over the range of acceptable wages. This procedure can be applied to our case but with an important advantage: our setting delivers more information about the unobserved heterogeneity distribution than the standard setting of Flinn and Heckman (1982). In their case, the unobserved heterogeneity is driven by unobserved ability or by measurement errors about which nothing is known except the effect on accepted wages; in our case, the unobserved heterogeneity is driven by the self-employed income distribution about which we have additional information. The distribution of  $y$ ,  $R(y|h)$ , can be identified by observing the individuals currently searching as self-employed. The  $y$  distribution over the currently self-employed is a truncation of the primitive  $R(y|h)$  where the truncation is driven by the searching decision (see Subsection 3.3). We assume again log-normality denoting the location and scale parameters with  $\{\mu_{y,h}, \sigma_{y,h}\}$ . Identification is then attained by appealing to a recoverable distribution. As a result, the distribution over which we would need to integrate out is identified by additional information beyond the one delivered by accepted wages. Moreover, the support conditions in the standard Flinn and Heckman (1982) case only pertain to acceptable wages, while in our case pertain to acceptable wages in both formal and informal jobs. In other words, a given  $y$  not only determines the range of acceptable wages but also the range over which these wages belong to a formal or an informal job. Since we observe both wages and formality status, this is additional valuable information that can be used in identification.

The second departure from the standard case is the heterogeneity in the searching state. This setting creates two issues. First, the self-employment income distribution  $R(y|h)$  needs to be identified. As we described above, we assume that  $R(y|h)$  belongs to a recoverable parametric distribution and we recover it by observing its truncation. Its truncation is observable because we observe  $y$  for all the individuals currently self-employed. Second, an additional arrival rate parameters needs to be identified  $\gamma_h$ . Similar to  $\lambda_h$ , we identified it mainly through the transitions of the self-employed state to an employee job. As for  $\lambda_h$ , the contemporaneous identification of  $G(x|h)$  is essential to separate the hazard rate in the component due to the arrival rate  $\gamma_h$  and in the one due to the acceptance probability.

The third departure from the standard case is that the arrival rates  $\lambda_h$  and  $\gamma_h$  are not primitive parameters. However, the matching function has an impact on workers' decisions *only* through the arrival rates. We can therefore use the workers' side information discussed so far (accepted wages and transitions) to identify  $\lambda_h$  and  $\gamma_h$ . We can then use some firms' side information (the vacancy rate) to recover the primitive parameters of the matching function, given  $\lambda_h$  and  $\gamma_h$ . We discuss this procedure in more details in Subsection 4.3.

**4.2. Preferences and Informality Parameters.** The second set of parameters to be identified is specific to our labor market model with a dual social security system and imperfect enforcement of formality. This set includes the preference parameters  $\beta_{0,h}$  and  $\beta_{1,h}$ —representing the workers' valuation for each pesos spent to provide the social benefits—the cost parameter  $c_h$  and the audit shock parameter  $\chi_h$ —representing the expected costs of being caught hiring workers informally.

First, we discuss the identification of  $\chi_h$ . Recall from Subsection 3.1 that  $\chi_h$  is the Poisson rate at which firms are audited. If a firm is audited while hiring formally, nothing happens. If it is audited while hiring informally, the match is terminated and the firm has to pay the monetary penalty  $c_h w_0(x; y, h)$ . Therefore, the auditing shock creates a higher termination rate for informal jobs than for formal jobs: both types of jobs are subject to the exogenous termination rate  $\eta_h$  but only informal jobs terminate due to  $\chi_h$ . We can then identify  $\chi_h$  by looking at the different transition rates out of formal jobs and informal jobs (see Table 2).

Second, we discuss the identification of  $\beta_{1,h}$  and  $c_h$  assuming that  $\beta_{0,h}$  is known and recalling that  $\chi_h$  is identified by the termination rates information. We identify  $\beta_{1,h}$  and  $c_h$  by exploiting the location and extent of the overlap between the distribution of accepted wages for formal employees and the distribution of accepted wages for informal employees. This is a crucial feature observed in the data that our model is able to replicate. Recall from Subsection 3.6 that in the relevant range of the parameter space, we have:

$$(31) \quad w_0(\tilde{x}(y, h); y, h) - w_1(\tilde{x}(y, h); y, h) > 0,$$

that is at the reservation productivity value  $\tilde{x}(y, h)$ , the wage received working informally is higher than the one received working formally. This implies an overlap in the support of the formal and informal accepted wage distributions. The difference between the two wages represents the extent of the overlap, whereas the reservation value  $\tilde{x}(y, h)$  governs the location of the overlap.

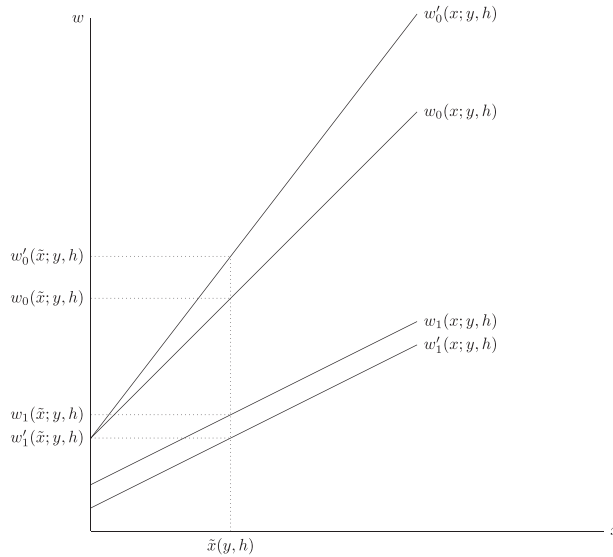
The parameters of interest  $\beta_{1,h}$  and  $c_h$  shape the extent and the location of the overlap in a very intuitive way. At any given value of the match productivity, formal employees receive lower net wages than informal employees because they receive higher nonwage benefits. The higher  $\beta_{1,h}$ , the more sensitive the worker to the added benefit, the larger the overlap. At the same time, informal employees receive higher net wages than formal employees with the same productivity because firms do not pay social security contributions. However, firms pay the penalty  $c_h w_0(x; y, h)$  if found to hire informally. The higher  $c_h$ , the less convenient to hire informally, the smaller the overlap. Finally, the location of the overlap is determined by  $\tilde{x}(y, h)$  that, in general, depends negatively on both  $\beta_{1,h}$  and  $c_h$ .<sup>42</sup>

If the previous discussion is informative on how both  $\beta_{1,h}$  and  $c_h$  have an impact on the location and on the extent of the overlap, it still does not indicate how the two parameters impact these data features differently. The intuition for the differential impact is illustrated in Figure 4. The figure reports the benchmark wage schedules—denoted by  $w_0(x; y, h)$  and  $w_1(x; y, h)$ —and the wage schedules resulting by changing  $\beta_{1,h}$  and  $c_h$ —denoted by  $w'_0(x; y, h)$  and  $w'_1(x; y, h)$ . To simplify the discussion, we focus only on the direct effects of the parameters, ignoring for the moment the equilibrium effects acting through the reservation value  $\tilde{x}(y, h)$ , the outside option  $Q(y, h)$ , and the redistribution component  $b_1$ .<sup>43</sup>

A decrease in  $c_h$  increases the sensitivity of informal wages to productivity  $x$  because it implies a lower cost of hiring informally. Graphically, it is equivalent to rotating the  $w_0$  wage schedule up. Ignoring equilibrium effects, a change in  $c_h$  has no direct impact on formal wages leaving the  $w_1$  wage schedule unaffected. As a result, the overlap increases because the upper bound moves up reaching  $w'_0(\tilde{x}(y, h); y, h)$ . The direct impact of an increase in  $\beta_{1,h}$  also leads to a larger overlap but by affecting a different margin. If  $\beta_{1,h}$  increases, formal wages decrease at each productivity value  $x$  because the nonmonetary benefits are now valued more. Graphically, it is equivalent to a parallel downward shift of the  $w_1$  wage schedule. A change in  $\beta_{1,h}$  has no direct impact on informal wages leaving the  $w_0$  wage schedule unaffected. As a result, the overlap increases because the *lower* bound moves down reaching  $w'_1(\tilde{x}(y, h); y, h)$ . In conclusion, if movement in  $\beta_{1,h}$  and  $c_h$  can achieve the same extent of the overlap, they do so by moving its location in different directions generating a different shape in the accepted wage distribution of formal and informal employees.

<sup>42</sup> This holds for most of the parameter space. Equilibrium effects work here through the outside option  $Q(y, h)$  and the endogenous redistribution component  $b_1$ . It is still possible that for a particular combination of the parameters' values and for specific values of  $y$ , the equilibrium effects are so large to change the sign of these predictions. Even when this is the case, the impact on the overall mixture distribution is limited because it involves only specific values of  $y$ .

<sup>43</sup> In the presence of these equilibrium effects, the differential impact may be stronger or weaker depending on the specific region of the parameter space and on the specific value of the outside option. However, the main identification argument for the differential impact of  $\beta_{1,h}$  and  $c_h$  does not change.



NOTE: Illustrative figure, not based on actual data. For the definitions of  $w_0(x; y, h)$ ,  $w_1(x; y, h)$ ,  $\bar{x}(y, h)$ , see Equations (23)–(25). The wage schedules resulting by changing  $\beta_{1,h}$  and  $c_h$  are denoted by  $w'_0(x; y, h)$  and  $w'_1(x; y, h)$ .

FIGURE 4

OVERLAP AND IDENTIFICATION OF  $\beta_{1,h}$  AND  $c_h$ 

Third, we use external sources of variation to separately identify  $\beta_{0,h}$ . It is not possible to make progress on the identification of  $\beta_{0,h}$  without additional sources of variation in the data since this preference parameter involves the same trade-offs in the model used to identify  $\beta_{1,h}$  and  $c_h$ . There is one relevant difference between  $\beta_{1,h}$  and  $\beta_{0,h}$ :  $\beta_{0,h}$  is the valuation of a fixed exogenous benefit ( $B_0$ ). If we were able to observe exogenous changes in the benefit, then we could exploit the additional information to identify  $\beta_{0,h}$ . The time-staggered entry across municipalities of the *Seguro Popular* program provides this additional source of variation.

As mentioned in Subsection 2.1, the *Seguro Popular* is a noncontributory social security program providing health services to everyone but formal workers. In terms of our model, it can be interpreted as an increase in the noncontributory benefit  $B_0$ . The magnitude of the increase corresponds to a change of  $B_0$  from 1.92 to 2.42 pesos per hour, or about a 25% increase in the per capita hourly extra-wage benefits for informal, self-employed, and unemployed workers.<sup>44</sup> The identification assumption that we use is that the structural parameters of the model do not differ between the set of municipalities that are exposed to the program during the year 2005 and the set of municipalities that are not yet exposed to the program as of 2005. We support this by first noticing that there were no other major policy changes simultaneous to the implementation of the *Seguro Popular* program.<sup>45</sup> We next provide supportive evidence for the identification assumption by looking at labor market outcomes before the *Seguro Popular* program was introduced. We can do this due to a previous round of the labor market survey conducted in 2001.<sup>46</sup> Since in 2001, no municipality received *Seguro Popular* benefits, we can use the 2001 data to check if the two groups are balanced in the pretreatment environment by running OLS regressions of the relevant labor market outcomes on an indicator variable for whether workers reside in municipalities that received the program in

<sup>44</sup> See Appendix A.2 for a detailed description of the computation of the two values of  $B_0$  with and without the *Seguro Popular* benefits.

<sup>45</sup> See, for example, Bosch and Campos-Vazquez (2014) and Azuara and Marinescu (2013) for a detailed discussion.

<sup>46</sup> The *Encuesta Nacional de Empleo* (ENE) was conducted from 1988 to 2004 and was replaced by the ENOE starting in 2005. The ENE contains less information than the ENOE but it is still a representative sample at the household level and has all the relevant information necessary to conduct the exercise.

TABLE 3  
PARAMETER ESTIMATES

	Low Schooling $h = 0$		High Schooling: $h = 1$	
	Coeff.	Std. Error	Coeff.	Std. Error
Search, Matching, and Bargaining				
$\lambda_h$	0.4679	0.0035	0.5167	0.0098
$\gamma_h$	0.0349	0.0042	0.0306	0.0014
$\eta_h$	0.0326	0.0007	0.0190	0.0052
$\mu_{x,h}$	2.7616	0.0367	2.6749	0.0382
$\sigma_{x,h}$	0.6243	0.0132	0.7970	0.0038
$\mu_{y,h}$	1.6718	0.0188	1.9497	0.0763
$\sigma_{y,h}$	0.7754	0.0028	0.8027	0.0258
$\xi_h$	-103.46	1.6661	-158.05	4.6038
$\alpha$	0.5630	0.0169	0.5630	0.0169
Preferences and Informality				
$\beta_{1,h}$	0.7949	0.0044	0.6091	0.0043
$\beta_{0,h}$	0.9862	0.0038	0.9807	0.0015
$\chi_h$	0.0079	0.0004	0.0113	0.0008
$c_h$	12.882	0.7045	16.574	1.3932
Matching Function and Demand Side				
$\psi_h$	0.0745	0.0088	0.0592	0.0034
$\iota_h$	0.7321	0.0253	0.7281	0.0184
$\zeta_h$	7.9718	1.6278	5.8569	0.8742
$\nu_h$	-496.01	288.80	-773.80	111.34
		Coeff.		Std. Error
Schooling and Ability				
$\delta$		0.0158		0.0010
$a_2^T$		0.6846		0.0059
$a_2^G$		1.1915		0.0037
$a_2^R$		1.0105		0.0004
$\pi_2$		0.5134		0.0003

NOTES: Estimates obtained via the method of simulated moments using the downhill simplex (Nelder–Mead) algorithm to minimize the quadratic form (36). Bootstrap standard errors based on 100 replications reported. For the definition of the parameters, see Subsection 3.1 and Section 4.

2005 or not.<sup>47</sup> Estimation results are reported in Table 5. The estimated coefficients are very small in magnitude and not statistically different from zero, suggesting balance between the two groups in the preprogram period for the same labor market variables that we use in estimation.<sup>48</sup> On the basis of this identification assumption, we will call “treated” those individuals who lived in municipalities with *Seguro Popular* and “control” those individuals who lived in municipalities without *Seguro Popular*.

**4.3. Matching Function and Demand Side Parameters.** Next, we focus on the parameters governing the matching function and characterizing labor demand. By Equation (19), the matching function parameters are the Cobb–Douglas coefficient  $\iota_h$  and the parameter denoting the lower search efficiency of the self-employed  $\psi_h$ . To identify them, we have to consider some firms’ side information. Specifically, we use the schooling specific vacancy rate  $v(h)$ , which, together with the unemployment and self-employment rates and the search efficiency  $\psi_h$ , generates  $\omega(h)$  (see Equation (20)). We can use the definition of the matching function

<sup>47</sup> To take into account the education levels we have in the model, we also add to the regression an indicator variable for secondary education completed or not.

<sup>48</sup> These findings are consistent with evidence reported in Bosch and Campos-Vazquez (2014) and Azuara and Marinescu (2013), which document that the roll-out of the *Seguro Popular* program was not correlated with labor market characteristics.



TABLE 4  
MEASURES OF RETURNS TO SCHOOLING

	Low $k = 1$	High $k = 2$
Ability:		
<u>PDV of Labor Market Search:</u>		
$\int_y Q(y, h) dR(y h)$	0.309	0.278
<u>PDV of Labor Income:</u>		
F: $\int_y \int_{\bar{x}(y, h)} E_1[w_1(x; y, h), y, h] dG(x h) dR(y h)$	0.291	0.303
I: $\int_y \int_{x_0^*(y, h)} E_0[w_0(x; y, h), y, h] dG(x h) dR(y h)$	0.267	0.266
SE: $\int_{y^*(h)} S(y, h) dR(y h)$	0.283	0.250
<u>Average Accepted Wages and Income:</u>		
F: $E_h[w_1   \bar{x}(y, h) \leq x]$	0.479	0.435
I: $E_h[w_0   x_0^*(y, h) \leq x < \bar{x}(y, h)]$	0.281	0.296
SE: $E_h[y   y^*(h) \leq y]$	0.167	0.167
<u>Average Offered Wages and Income:</u>		
F: $E_h[w_1   y < y^*(h)]$	0.213	0.210
F: $E_h[w_1   y \geq y^*(h)]$	0.213	0.204
I: $E_h[w_0   y < y^*(h)]$	0.133	0.134
I: $E_h[w_0   y \geq y^*(h)]$	0.142	0.136
SE: $E_h[y]$	0.349	0.349

NOTES: The table reports relative differences between the low and high schooling group. Denoting with  $X$  the variable of interest, differences are computed as:  $\frac{X_{h=1} - X_{h=0}}{X_{h=0}}$ . PDV denotes present discounted value. F, I, and SE denote, respectively, formal employee, informal employee, and self-employed. See main text for additional variable definitions.

TABLE 5  
ROLL-OUT OF THE *SEGURO POPULAR* (SP) PROGRAM AND PREDETERMINED LABOR MARKET CHARACTERISTICS

	Hourly Wages (log)			Labor Market Proportions			
	Formal	Informal	Self	Formal	Informal	Self	Unempl
SP in 2005 (1 = yes)	-0.041 (0.036)	0.048 (0.055)	-0.035 (0.062)	-0.034 (0.026)	0.035 (0.019)	-0.004 (0.014)	0.003 (0.006)
Complete Sec. (1 = yes)	0.218 (0.017)	0.288 (0.032)	0.092 (0.033)	0.061 (0.011)	-0.036 (0.008)	-0.029 (0.008)	0.003 (0.003)
Number of Obs.	7865	5474	2777	16458	16458	16458	16458

NOTES: OLS estimates. Standard errors clustered at the municipality level are reported in parenthesis. Data are drawn from the Mexican labor market survey (ENE, 2001) and matched at the municipality level with the roll-out of the *Seguro Popular* program. State dummies and municipality-level controls (log(population), log(population<sup>2</sup>), and poverty index) are included in all specifications.

and the equations defining the endogenous arrival rates to obtain:

$$(32) \quad \psi_h = \frac{\gamma_h}{\lambda_h},$$

$$(33) \quad \iota_h = \frac{\ln \omega_h - \ln \lambda_h}{\ln \omega_h}.$$

Since  $\gamma_h$  and  $\lambda_h$  are identified by workers' side information (see Subsection 4.1), Equations (32) and (33) are sufficient for the identification of the parameters of the matching function, which allows the computation of the arrival rate of workers to firms  $\zeta_h = \omega_h^{-\iota_h}$ . This information is enough to identify the demand side parameter: the flow value of keeping the vacancy open  $\nu_h$ . To see this, consider the equilibrium equation for the firms' value functions (Equation (B.10) in Appendix B.3) and notice that it can be solved for  $\nu_h$ .

An important restriction that we impose to obtain these identification results—on top of the usual constant returns to scale assumption—is the normalization of the matching efficiency coefficient to 1 in the definition of the matching function (19).<sup>49</sup> The normalization could in principle be avoided by using a more standard procedure to identify and estimate the matching function, which relies on a relatively long time series on vacancies, unemployment, and realized matches.<sup>50</sup> Although such data set is now available for Mexico, its quality is not very good. In addition, estimating a matching function under this approach would necessarily cover a period of large changes in the *Seguro Popular* roll-out. Given the importance of this institution in our application for both estimation and validation purposes (see Subsections 4.2 and 5.3.2), we prefer to use data on vacancy rates only for the year 2005 to keep the *Seguro Popular* coverage constant.

**4.4. Schooling Parameters.** The last element of the parameter set that needs to be identified is the distribution of the direct cost of acquiring the high schooling level,  $T(\kappa)$ . We do not have any direct information on schooling costs or on behavior at the time of the schooling decision. The only available information about this process is the schooling level completed for each individual in the sample. We can exploit this information using the threshold-crossing impact generated by the equilibrium of the model. Recall from Subsection 3.2 that agents with a cost of schooling higher than  $\kappa^*(y)$  do not acquire additional education, whereas those with a lower cost do. Using the sample analog, we can write the moment condition:

$$(34) \quad \frac{1}{n} \sum_{i=1}^n h_i = \int_y T(\kappa^*(y)) dR(y|0),$$

which allows for the identification of a one-parameter distribution. Based on the previous literature and computational convenience, we choose the negative exponential distribution and denote with  $\delta$  its parameter.<sup>51</sup>

**4.5. Workers' Unobserved Ability.** The model described so far contains a fair amount of both observed and unobserved heterogeneity. Still, two considerations led us to introduce additional workers' unobserved heterogeneity.

The first consideration is empirical. Hazard rates out of unemployment exhibit negative duration dependence conditioning on schooling level—they are about 12 percentage points lower at six months than at three months (see the bottom of Table C.5 in Appendix C). Although our model can replicate duration dependence for self-employed searchers, it cannot do that for unemployed searchers. Once agents have acquired their education level and they have decided to search as unemployed, they become *ex ante* identical. Therefore, all agents

<sup>49</sup> Common specifications of the Cobb–Douglas matching function include a scale factor that multiplies the measures of unemployment and vacancies to capture matching efficiency. This specification is less common in job search models estimated from microdata, which assume either exogenous contact rates or a very parsimonious specification of the matching function. Among the existing contributions allowing for endogenous contact rates, Flinn and Mullins (2015) employ the same matching function we use here, normalizing the Total Factor Productivity (TFP) parameter to one, whereas Flinn (2006) assumes a Constant Returns to Scale (CRS) matching function without unknown parameters.

<sup>50</sup> See Petrongolo and Pissarides (2001) for a review and Borowczyk-Martins et al. (2013) for a more recent contribution discussing relevant econometric issues. We perform some robustness exercises showing how the demand-side parameters estimates and the policy experiments results change for different values of the matching elasticity. Table C.1 in Appendix C shows the impact on parameter estimates and Figure D.1 in Appendix D shows the impact on a subset of the policy experiments.

<sup>51</sup> A similar strategy is used in other empirical search models to account for the extensive margin of the labor supply decision. The analogy is that even if no information about the value of nonparticipation is available, a threshold-crossing condition can be used to identify a one-parameter distribution from the proportion of labor market participants. For a recent example, see Flabbi and Mablí (2018).

with the same schooling level who are searching as unemployed share the same reservation value generating a constant hazard rate.

The second consideration relates to one of the main innovation we introduce in the model: the endogenous schooling decision. The institutional parameters responsible for the emergence of informality may also affect returns to schooling, distorting workers' investment decisions in human capital. Introducing an endogenous schooling decision allows us to evaluate the importance of this channel and its interaction with the institutional sources of informality. To do that, however, we need to recover appropriate returns to schooling. In a dynamic setting, the appropriate return should compare the value of participating in the market as a low schooling individual and the value of participating as a high schooling individual. The model described so far allows us to do that by comparing  $Q(1, 0)$  with  $Q(y, 0)$ . However, the comparison is subject to an important restriction: the flow cost of acquiring schooling is not correlated with future labor market performance (conditioning on schooling.) This is not an innocuous restriction. For example, higher "ability" individuals may be both more effective in acquiring schooling and more productive in the labor market. If this were the case, our estimated returns to schooling would be biased and our policy evaluations would miss composition effects in the two schooling markets. Removing this restriction may therefore correct for bias in the returns and lend more credibility to the policy experiments.

We follow the literature assuming that unobserved heterogeneity can be approximated by a finite number of discrete types.<sup>52</sup> The type is known to the agent but unobserved in the data. We denote each type with  $k$  and its proportion in the population with  $\pi_k$ . Types are time-invariant and affect individuals' productivity in the labor market and their direct cost of acquiring schooling. We introduce the following parsimonious specification to represent this dependence:

$$(35) \quad \begin{aligned} x|k &= a_k^G x, \\ y|k &= a_k^R y, \\ \kappa|k &= a_k^T \kappa, \end{aligned}$$

where  $a_k^J$  with  $J \in \{T, R, G\}$  and  $k \in \{1, 2, 3, \dots, K\}$  are positive scalars that represent scale factors similar to TFP parameters. To understand their role, assume that  $K = 2$ , type  $k = 1$  is normalized to  $a_1^T = a_1^R = a_1^G = 1$  and type  $k = 2$  exhibiting  $a_2^T < 1$ ;  $a_2^R > 1$ ;  $a_2^G > 1$ . In this economy, type  $k = 1$  is equal to the population in the model presented in Section 3; type  $k = 2$  is the "high-ability" type since individuals in this group have on average a lower cost of acquiring schooling and a higher productivity in the labor market. High-ability individuals have different reservations values than benchmark individuals in all the decisions they face. Everything else equal, they will acquire more schooling, they will search as unemployed in different proportions, and they will have different hazard rates out of the searching states.<sup>53</sup>

The last implication is crucial for identification. As discussed above, we observe some duration dependence in the data. The presence of unobserved types is the only model's feature that can generate duration dependence. Since different types have different hazard rates, individuals belonging to the type with the higher hazard rate will leave the state earlier. This

<sup>52</sup> An influential example in the labor literature is Keane and Wolpin (1997), for a more recent extension, see Keane and Wolpin (2010). Examples in the search literature include the seminal Eckstein and Wolpin (1995) and the more recent Flinn and Mullins (2019). Depending on data availability and model's specification, alternative forms of unobserved heterogeneity may be used in search models to capture features similar to ours, see, for example, Flinn and Mullins (2015) and Bagger et al. (2014).

<sup>53</sup> This heterogeneity creates a similar effect to the one found in Flinn and Mullins (2015). The conditional labor market outcomes of the two schooling groups are not only different because schooling gives access to a labor market with different structural parameters but also because the two schooling groups contain different proportions of high-ability individuals. This additional selection is relevant in the counterfactuals: any policy affecting the overall education level also affects the selection of agents that acquire additional education and, as a result, the returns of the policy.

dynamic changes the composition of workers in a given state as time passes. Since the hazard rate observed in the data is a mixture of the hazard rates of the different unobserved types, the observed duration will exhibit duration dependence. Due to data limitations (see Subsection 2.2 and Subsection A.1 in the Appendix), we are able to extract from the data only four credible moments that characterize duration dependence: the hazard rates out of unemployment at three and six months for both schooling levels. Given the limited number of moments, we have chosen to assume a limited number of types: we estimate the model with only two types, setting  $K = 2$ . In addition, we normalize the parameters of type  $k = 1$  to one to reduce the number of free parameters to be exactly equal to the number of moments that identify them. The four parameters are:  $\{a_2^T, a_2^G, a_2^R, \pi_2\}$ .

## 5. ESTIMATION

5.1. *Method.* We estimate the parameters of the model in two steps. In the first step, we jointly estimate the search, matching and bargaining parameters discussed in Subsection 4.1, the preferences and informality parameters discussed in Subsection 4.2, the schooling parameter discussed in Subsection 4.4, and the additional heterogeneity parameters discussed in Subsection 4.5. In the second step, we estimate the matching function and the demand-side parameters discussed in Subsection 4.3. The institutional parameters  $\{B_0, \tau, t\}$  are set to the values determined by the Mexican legislation (see Appendix A.2 for details) and the discount rate  $\rho$  is fixed to 5% a year.

The first step of the estimation procedure uses the method of simulated moments (MSMs).<sup>54</sup> Given the vector of parameters for each schooling group  $\Theta_h$ , the method defines a joint estimator  $\hat{\Theta} \equiv [\hat{\Theta}_0 | \hat{\Theta}_1]$  as:

$$(36) \quad \hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} [M_R(\Theta) - m_N]' W^{-1} [M_R(\Theta) - m_N],$$

where  $m_N$  is an appropriately chosen set of sample moments derived from our sample of size  $N$  and  $M_R(\Theta)$  is the set of the same moments derived from a simulated sample of size  $R$ , based on a steady-state equilibrium obtained at the parameter vector  $\Theta$ . Since we assume a continuous distribution function for the self-employed values  $y$ , the equilibrium generates infinitely many reservation thresholds. We therefore solve by approximation discretizing the  $R(y|h)$  distribution with 100 grid points. We set  $R$  at 20,000. We follow previous literature by defining the weighting matrix  $W$  as a diagonal matrix with diagonal elements equal to the inverses of the variance of each sample moment. While not achieving efficiency, the choice of this weighting matrix gives more weight to moments with lower sample variability and eases the computational burden by harmonizing the moments' scale.

The second step of the estimation procedure concerns the matching function parameters and the demand-side parameters. As mentioned in the identification Subsection 4.3, given a consistent estimator  $\hat{\Theta}$ , the parameters can be consistently estimated by solving Equations (32) and (33) and the equilibrium equation for the firms' value functions (Equation (B.10) in Appendix B.3).

We choose the moments to be used in the quadratic form (36) to capture the stylized facts described in Subsection 2.2 and the data features needed by for identification described in Section 4.

We use the cross-sectional sample described in Table 1 to capture the equilibrium proportions of workers in the four labor market states and the distributions of accepted wages and self-employment labor income. For wages and labor incomes, we extract mean and standard deviations. To capture the overlap between the distributions of formal and informal wages, we

<sup>54</sup> The method has become increasingly popular to estimate highly nonlinear models with value functions solved numerically such as ours. For the asymptotic properties of the MSM estimator defined in (36), see Pakes and Pollard (1989) and Newey and McFadden (1994).

compute quintiles over the distribution of accepted wages for formal workers.<sup>55</sup> For each interval, we compute: (i) the mean wage of informal employees, (ii) the mean wage of informal employees, and (iii) the proportion of employees in informal jobs earning a wage in that interval. The size of the cross-sectional sample allows us to extract all these moments conditioning on schooling and *Seguro Popular* treatment. In addition, we extract from this sample the proportion of workers in each schooling level by *Seguro* exposure.

We use the balanced panel sample described in Table 2 to capture the labor market dynamics. As discussed in Appendix A.1, it is more credible to capture the dynamics using transition rates instead of ongoing spells. We compute transition matrices over the four labor market states one year apart. As discussed in Subsection 4.5, we also use this sample to describe the duration dependence. We compute hazard rates out of unemployment at three months and at six months. Given the small sample size, we cannot compute these moments conditioning on both schooling and *Seguro Popular* exposure. Since most of the parameters we estimate are schooling-specific, this is the dimension most crucial for identification. We therefore compute these moments conditioning on schooling but aggregating over *Seguro* exposure.

Finally, we use two aggregate data sources to obtain labor shares and vacancy rates. As described in Subsection 2.2, labor shares are extracted from AMECO and computed as the total employee compensations as percentage of GDP at market prices. Vacancy rates are extracted from the Ministry of Labor's *Bolsa de Trabajo*. We extract schooling-specific vacancy rates by aggregating over all the available job vacancies featuring the educational requirements that correspond to the two schooling levels we use in the article. These national-level statistics are made comparable to our estimation sample using the survey weights provided in ENOE 2005.

We obtain a complete set of 137 sample moments targeted by our proposed estimator. They are reported, along with the simulated moments, in Appendix C, Tables C.3– C.6.

**5.2. Estimates.** The parameter estimates are reported in Table 3. The parameters governing the rates of job arrival and termination  $\{\lambda_h, \gamma_h, \eta_h\}$  are comparable to previous estimates for similar models on high-income countries. The main difference is in the relative higher termination rate.<sup>56</sup> There are differences between the two schooling groups, with lower arrival rates and higher termination rates for individuals who did not complete a high school degree. The differences in arrival rates between the unemployed and the self-employed are large, explaining in part the observed persistence in the self-employment state and the high turnover in the unemployment state. For example, an unemployed worker belonging to the low schooling group meets a firm on average every 2.1 months, while a self-employed worker on average every more than 2 years. Taking into account the endogenous acceptance probability, these rates translate in 75% of the low schooling unemployed accepting a wage offers within three months of searching. Unemployed workers belonging to the high schooling group receive more offers but they are pickier, leading to an acceptance probability at three months of 63%. These transitions result from movements to either a formal or an informal job.<sup>57</sup>

Important differences between the two schooling groups are also observed in the estimated values of the parameters of the match-specific productivity distribution  $\{\mu_{x,h}, \sigma_{x,h}\}$  and the self-employed earning distribution  $\{\mu_{y,h}, \sigma_{y,h}\}$ . As reported in Table C.2 in Appendix C, both average match-specific productivity and productivity in self-employment are higher in the

<sup>55</sup> This is the same procedure proposed by Flabbi and Moro (2012) to capture the overlap between the wage distributions of jobs offering flexibility and jobs that do not.

<sup>56</sup> See, for example, the review in Eckstein and van den Berg (2007) and specifically models of individual search without on-the-job search such as Flinn and Heckman (1982), Flinn (2006), and Flabbi and Moro (2012). A similar ranking in arrival rates by schooling levels is found in Flinn and Mullins (2015) (under the no renegotiation specification) and in Flabbi and Leonardi (2010), even if both papers use U.S. data and define schooling levels differently.

<sup>57</sup> Meghir et al. (2015) document that in Brazil, it takes on average three years to transit to a formal job from unemployment but only a few months to transit to an informal one. The main reason for the difference with our estimates—on top of some country-specific factors—lies in the fact that Meghir et al. (2015) do not differentiate between self-employed and informal employees in their definition of the informal sector.

high schooling group than in the low schooling group: the first is about 3.7% higher and the second is about 34% higher. These differences in productivity—along with the differences in the mobility parameters and in the valuation of the benefits—generate the returns to investing in additional schooling. We focus on this feature in Subsection 5.2.1.

The Nash bargaining coefficient  $\alpha$  is estimated at 0.56, which gives a slightly stronger bargaining position to the workers although it remains quite close to the value of 0.5 that defines symmetric bargaining. This value is higher but comparable to those estimated on different data sets using similar identification arguments.<sup>58</sup>

The estimated values of the preference parameters  $\{\beta_{1,h}, \beta_{0,h}\}$  show that both formal and informal extra-wage benefits are valued less than the monetary investment used to provide them. However, the valuation of the noncontributory benefit is very close to full monetary value (about 98 cents to the peso for both schooling groups), while the valuation of the contributory benefit is not (about 61 cents to the peso for the high schooling group, and about 79 for the low schooling group). Estimated values of  $\beta_{1,h}$  significantly lower than one imply that formal employees have a willingness to pay for the benefit that is substantially lower than the contribution paid to receive it, introducing a net loss that reduces the incentive to work formally. Based on parameters estimates and equilibrium matches, the worker at the average accepted formal wage in the low schooling group pays about 6.5 pesos per hour in payroll contributions while receiving a monetary benefit of about 7. But that amount of benefit is valued about 5.6 pesos by the worker, leading to a loss of 4% of the average wage. In the high schooling group, the average loss is much larger, reaching 14% of the average wage. This higher loss is due to the redistribution dynamics present in the system: only 55% of the contribution generates benefits that are proportional to the amount contributed, the rest is used to generate the fixed benefit  $b_1$ . Since individuals with higher schooling earn and contribute more but receive the same  $b_1$ , they redistribute resources toward individuals with lower schooling.<sup>59</sup> The estimated values have also a straightforward implication in terms of efficiency: since the valuation of the benefit is in linear utility terms and it is lower than one, a policy that eliminates the benefits in exchange for an equal monetary transfer would be Pareto-improving. However, a policy that eliminates benefits and taxes altogether would also eliminate the redistributing feature just discussed.

The estimates of the parameters describing the cost of informality  $c_h$  and  $\chi_h$  are better understood by recalling Equation (10). When firms hire informally, they do not pay any social security contribution but if audited they have to pay the one-shot penalty  $c_h w_0(x; y, h)$  and terminate the match. Audits come with a frequency governed by the Poisson process parameter  $\chi_h$ . The resulting cost of this process is equivalent to firms setting aside an amount  $\chi_h c_h w_0(x; y, h)$  in flow value terms. This is formally shown by the second term of the RHS of Equation (10). Under this interpretation, the estimates of  $c_h$  and  $\chi_h$  imply that the flow cost of hiring informally is about 10.1% of the wage in the low schooling group and about 18.7% in the high schooling group. In addition to the flow cost, firms face a cost in expectation since the match is terminated when the audit occurs. An informal employee match is then more likely to be terminated than a formal one: they are both subject to the exogenous termination

<sup>58</sup> Flinn (2006) estimates the Nash bargaining coefficient in the range of 0.39–0.43 in a sample of young low skilled U.S. workers. Flinn and Mullins (2015) estimate it at 0.25 on a sample of very young U.S. workers (25–34 years of age) but with a broader range of skills.

<sup>59</sup> We speculate about two possible explanations of why the valuation of the benefits is close to full value for informal workers and much lower for formal workers. The first relates to the contributory nature of the benefit. Formal employees contribute a proportion of their wages to obtain the benefit while informal employees do not. As a result, the attitude toward the service provided and its valuation may be different. The second possible interpretation relates to the composition of the benefit. The benefits in the formal sector,  $B_1(w_1(x; y, h))$ , bundle together two types of benefits: a retirement benefit and a health benefit—see Subsection 3.1. The health benefit ( $b_1$ ) is comparable to the noncontributory benefit ( $B_0$ ) received by informal and unemployed workers but the retirement benefit is qualitatively different. However, the valuation parameter for formal employees is estimated over the bundle of both types of benefits since this is the way the benefits are offered and the contributions are paid. A  $\beta_{1,h}$  lower than  $\beta_{0,h}$  may therefore reflect that retirement benefits are valued less than health benefits.



governed by  $\eta_h$  but only informal match are terminated due to the audit. This is formally shown in the third term of the RHS of Equation (10). The estimates of  $\chi_h$  imply that the average duration of low schooling informal jobs is about 20% shorter than the duration in low schooling formal jobs; the difference is 37% in high schooling jobs.

The estimates of the matching function and demand-side parameters report the following results. The matching function parameter  $\psi_h$  represents the lower search efficiency of the self-employed workers with respect to the unemployed workers. If we assume that the unemployed search full time, the estimated  $\psi_h$  implies that the average self-employed devotes between two and three hours a week to job search. The matching function parameter  $\iota_h$  represents the elasticity of the number of meetings with respect to the measure of searchers—see Equation (19). Our estimated values are very similar to average values estimated using macrodata on high-income countries (Petrongolo and Pissarides, 2001) and on Mexico (Arroyo Miranda et al., 2014).<sup>60</sup> As frequently found in empirical works, they are different from the estimates of the workers' bargaining weight  $\alpha$ , implying a violation of the Hosios condition (Hosios, 1990) and therefore the presence of inefficiencies. We will revisit the efficiency issue through counterfactual experiments in Section 6. The other demand-side parameter is the flow value  $v_h$  of keeping the vacancy open. We estimate a very high cost of keeping a vacancy open for both the low schooling and high schooling market. These high costs can be sustained thanks to the high arrival rate of workers to firm implied by  $\zeta_h$ : the estimates imply that the average duration for a firm to meet a worker is less than a week.<sup>61</sup>

The estimate of  $\delta$ , the parameter of the cost of schooling distribution, implies that the average direct flow cost of completing high school with respect to middle school is about 63.3 pesos per hour. This direct cost is approximately 2.2 times the average wage accepted by a high school graduate when working formally and approximately 3.1 times when working informally. Interestingly, these values are comparable to those found by Flinn and Mullins (2015) when estimating on U.S. data a similar model of endogenous schooling acquisition. As shown in problem (4), the direct cost of schooling should be integrated over the time required to complete the degree and summed up to the opportunity cost to obtain the total cost of acquiring additional schooling. At parameter estimates, the average direct cost is 11 pesos per hour, which is about half of the hourly wages of formal workers, for individuals who do not complete secondary schooling and it is much lower for individuals who complete secondary schooling (2 pesos per hour, or 7% of the hourly wages in the formal sector). Overall, direct costs are on average 63% of the overall costs of schooling, with the opportunity cost contributing for the rest.<sup>62</sup>

Finally, the bottom of Table 3 reports the ability parameters. In our specification with only two types, the reported  $\alpha_2^l$  refer to the parameters for type  $k = 2$ , while the parameters for type  $k = 1$  are all normalized to one. The interpretation is therefore very simple: type  $k = 2$  is the high-ability type since workers in this set are 19.1% more productive in an employee match with same productivity  $x$ , they receive about 1% more income in self-employment at same  $y$ , and they have on average a 32% lower flow cost of acquiring additional schooling. A little more than half of the population belongs to this type, as indicated by the estimate for  $\pi_2$ .

*5.2.1. Returns to schooling.* Table 4 reports the relative difference in labor market outcomes between the low and high schooling group. The differences are computed from a simulated sample generated by the parameter estimates reported in Table 3. The first two measures are the PDVs of lifetime utility and lifetime income. These are the relevant dimensions

<sup>60</sup> Arroyo Miranda et al. (2014) estimate a value of 0.68 (see columns (2) in Tables 4 and 5) using Mexican data from 1993 to 2013. The broad review in Petrongolo and Pissarides (2001) reports estimates ranging from 0.35 to 0.81, with many estimates concentrated in the upper third of the range.

<sup>61</sup> Sensitivity of these parameters estimates with respect to the match elasticity  $\iota_h$  (see Table C.1 in Appendix C) shows that an estimate of  $\iota_h$  closer to average values found in the literature would generate lower  $v_h$ .

<sup>62</sup> Opportunity costs are a function of two sources of individual-level heterogeneity: self-employed income with  $h = 0$  and ability  $k$ . The reported average is obtained by integrating over both sources of heterogeneity.

of returns to schooling since they summarize both the cross-sectional and the dynamic aspects affecting labor market outcomes. The return in the first measure is the closest to the expression in Equation (4) in Subsection 3.2: it represents the difference between the value of participating in the market as a low schooling individual,  $Q(y, 0)$ , and as a high schooling individual,  $Q(y, 1)$ . Since both values depend on  $y$ , we integrate over the two  $R(y|h)$  distributions to obtain average values and returns. Our predicted return is about 30% for both ability types.<sup>63</sup> The labor market returns according to lifetime income are further decomposed into the contribution of labor income in formal employment, informal employment, and self-employment. The returns in formal employment are the highest, more so for the high-ability type.

To provide a comparison with previous literature, the last two statistics at the bottom of Table 4 report more traditional measures of returns to schooling. The labor market returns calculated on average accepted wages and incomes are comparable to those obtained by estimating wage equations or by comparing conditional means on observational data. Completing high school with respect to completing at most middle school, increases average accepted wages by more than 40% when working as formal employee and by less than 30% when working as informal employee. These returns are higher than those found by previous works on Mexico.<sup>64</sup> Relative differences on accepted wages generate a distorted measure of wage opportunities by schooling levels in a market characterized by search frictions and labor market dynamics. First, accepted wages in a given point in time are just one episode in the individuals' labor market careers. Second, accepted wages are selected since agents can reject job offers (Belzil, 2007). By reporting average *offered* wages and incomes, we follow Eckstein and Wolpin (1995) in proposing a statistic that at least addresses the selection issue.<sup>65</sup> Except for self-employment, the implied returns to schooling are lower than those generated by accepted wages, pointing out a possible large bias in linear regression models.

### 5.3. Model Fit.

5.3.1. *Within-sample fit.* Tables C.3 and C.4 report the moments describing the cross-sectional features of the data, which are computed conditioning on both the schooling level and the exposure to the *Seguro Popular* program.<sup>66</sup> Table C.3 reports results for the sample receiving the program (treated) and Table C.4 reports results for the sample not receiving the program (control). The estimated model fits well the distributions over the four possible labor market states and over wages and self-employment incomes. The only mismatch is over the unemployment rate, which the model predicts to be higher than in the data. The reason for the mismatch is the relatively high persistency in unemployment of the high-ability type. The overlap between the formal and informal wage distributions is also qualitatively replicated but the mismatch is more significant. The main issue is that the support of the overlap,

<sup>63</sup> The lack of significant differences by ability is in part driven by our functional form assumptions. Since we do not allow for complementarities between ability and skills, higher ability individuals are more productive across the board, including when they are in the low schooling group.

<sup>64</sup> Levy and Lopez-Calva (2018) use ENOE 2006 to estimate returns on our same schooling levels by linear regressions. They find returns of about 23% over formal wages and of about 14% over informal wages.

<sup>65</sup> In our model, as in Eckstein and Wolpin (1995), wages are bilaterally bargained, so they are not properly "offered" by the firm but they are outcomes that depend on the entire labor market dynamics. Still, for given match-value  $x$ , wages corresponding to a given firm-worker meeting can be recovered using Equations (23) and (24). Unlike Eckstein and Wolpin (1995), our model also includes a formality status. In computing offered wages, we can take into account the productivity ranges over which wages at different formality status would be offered or not. To favor the comparison to wages that are not selected, we present values that do not take into account this additional truncation. In short, we obtain mean offered wage by formality status by integrating the wage schedules (23) and (24) over the primitive productivity distribution  $G(x|h)$ . Like Eckstein and Wolpin (1995), we also find large differences between returns estimated on accepted wages and those estimated on offered wages. In their sample of low-skilled black workers, they also find the same sign of our estimates: returns based on accepted wages are higher than those on offered wages.

<sup>66</sup> They are also computed unconditionally on the labor market state to guarantee a smoother and well-defined quadratic form during the optimization procedure.

in particular for the unemployed, is shifted up with respect to the data. Therefore, even if the shape and overall size of the overlap is comparable in the model and in the data, the wages and the proportions over some quintiles are quite different.<sup>67</sup> The bottom of the two tables reports the schooling shares, which are matched very well by the model on both the control and treated sample.

Table C.5 reports the moments capturing the longitudinal features of the data. As mentioned in Subsection 2.2 and Appendix A.1, the small size of the panel sample forces us to aggregate these moments over *Seguro Popular* exposure. Since it is necessary for identification, we still compute them conditioning on the schooling level. We fit well the transitions involving most of the transitions over labor market states.<sup>68</sup> The first three transitions listed in the table cover 78% of the movements in the low schooling sample and 61% in the high schooling sample. The model is able to replicate this feature, predicting 74% in the low schooling sample and 71% in the high schooling. We fit less well transitions that occur at lower frequency: for example, less than 1% of the overall transitions in the high schooling group are from informal employment to unemployment but we predict that about 2% are. A data feature we wanted to capture is the negative duration dependence implied by the different hazard rates at three and six months. We are able to replicate the negative duration dependence but we predict a more significant gradient than in the data, in particular on the high schooling group.

Finally, Table C.6 reports the moments extracted from aggregate data. Vacancy rates are matched exactly by construction (see Subsection 4.3). The labor shares are matched as part of the MSM procedure and return a reasonable fit.

**5.3.2. Out-of-sample validation.** In Subsection 4.2, we describe how we use the roll-out of the *Seguro Popular* program in 2005 for identification purposes. The roll-out of the program continued in 2006, creating an opportunity for model validation. We can look at municipalities that started receiving the program in 2006 and collect data on their labor markets. We can then compare the 2006 data with predictions from our model estimated on data collected in 2005.

Specifically, we extract from the ENOE survey a balanced panel of individuals observed in every quarter of the year 2006 and we impose the same sample restrictions we used for the 2005 estimation sample (see Subsection 2.2). The resulting sample contains 1,360 individual observations, of which 87% reside in municipalities that are exposed to the *Seguro Popular* program.<sup>69</sup> We estimate the effects of this new policy on labor market outcomes  $y_{i,q}$  for individual  $i$  observed in quarter  $q$  through the following linear specification:

$$(37) \quad y_{i,q} = \theta d_{m(i),q} + \vartheta h_i + \varrho_{m(i)} + \varphi_q + \epsilon_{i,q},$$

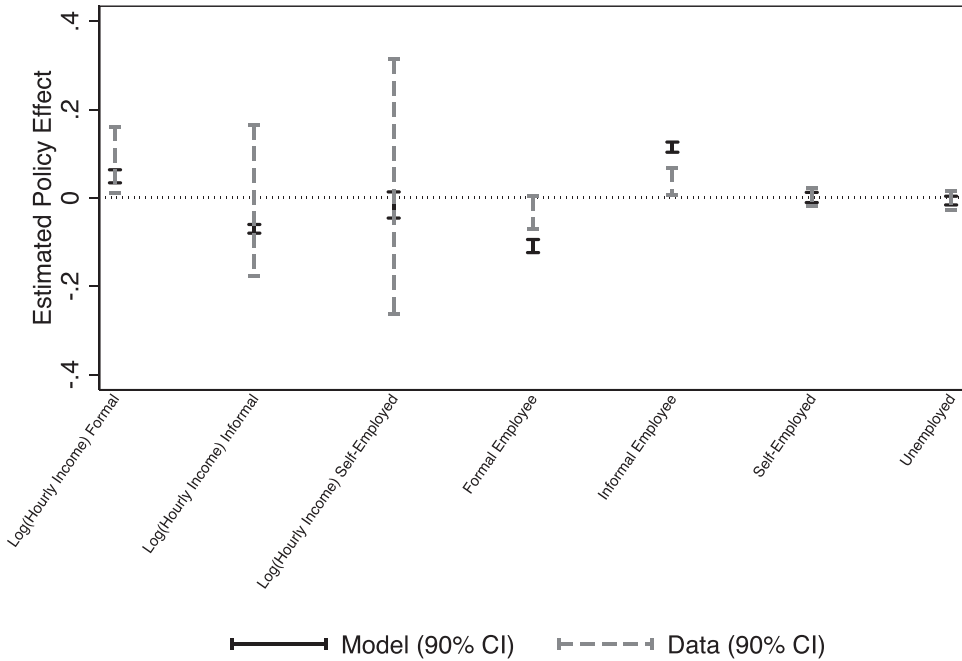
where  $d_{m(i),q}$  is an indicator variable for whether or not the individual  $i$  residing in municipality  $m$  receives the *Seguro Popular* program in quarter  $q$ ,  $h_i$  is an indicator for the level of education attained by individual  $i$  (completed secondary schooling or not),  $\varrho_{m(i)}$  is a municipality fixed effect that controls for (time invariant) unobserved heterogeneity at the local labor market level, and  $\varphi_q$  captures time-varying aggregate shocks during the year 2006. The error term  $\epsilon_{i,q}$  is clustered at the municipality level. The  $\theta$  parameter identifies the effect of the program for those individuals who were not covered in 2005 but started to be covered at some point during the year 2006.

We next predict the estimated change in labor market outcomes from the policy implemented in 2006 using our model with the parameter estimates obtained with information up

<sup>67</sup> The problem is particularly acute for the control sample with low schooling. As reported in Table C.2, the reason is the low  $\bar{x}(0, h)$  for the low-ability type. This equilibrium result is generated in our simulation by the relatively generous redistribution toward the low-ability, low-productivity matches operated by the formal social-security system.

<sup>68</sup> The table orders the transitions by the frequency observed in the data. We sort them based on the low schooling group.

<sup>69</sup> In the 2005 panel sample, about 77% individual observations reside in municipalities that receive the program.



NOTE: The figure shows OLS estimates with robust standard errors of the  $\theta$  coefficients displayed in Equation (37), which capture the effect of receiving the *Seguro Popular* program in 2006 on labor market outcomes both in the data sample and in the simulated sample. The “Data” sample is drawn from the Mexican labor market survey (ENOE 2006) and matched at the municipality level with the roll-out of the *Seguro Popular* program in the year 2006. The “Model” sample is based on simulated data generated by the estimated model parameters reported in Table 3. The full set of two-sided and one-sided T-tests for the estimated  $\beta$  coefficients in the simulated data and in the survey data are reported in Table C.7 in the Appendix.

FIGURE 5

OUT-OF-SAMPLE MODEL VALIDATION

to 2005. Specifically, we simulate the model at estimated values and generate a counterfactual labor market using the increased share of individuals who are exposed to the program and the increase in the level of  $B_0$  from the year 2005 to the year 2006.<sup>70</sup> We estimate linear specifications similar to the one depicted in (37) on the simulated data to recover the estimated average differences of labor market outcomes between the two  $B_0$  regimes under the policy in 2006.<sup>71</sup> To the extent that (i) our model captures the main equilibrium effects due to the policy changes and (ii) the linear model depicted in (37) is able to isolate the average impact of *Seguro Popular* in the data sample, the estimates for the data sample and the simulated data should be comparable.<sup>72</sup>

Figure 5 displays the OLS estimates of the  $\theta$  coefficients in Equation (37) along with the 90% confidence intervals of the estimated coefficients from the data sample and the simulated sample. The figure reports results from seven regressions. Each regression shares the same specification but includes a different dependent variable. The first three dependent variables are the log of the hourly labor income; the last four are indicator variables that are equal to

<sup>70</sup> The specific increase is from 2.42 pesos per hour to 2.84 pesos per hour in the treatment group and from 1.92 pesos per hour to 2.13 in the control group. In the simulation of the model, we keep the schooling decisions fixed as in the benchmark model estimated on data for the year 2005 given the short time span of the policy change.

<sup>71</sup> We do not condition on municipality fixed effects or quarter fixed effects when estimating the linear specification (37) on simulated data. These variables are meant to capture sources of unobserved heterogeneity in the survey data that are not present in the data generated from the model.

<sup>72</sup> See also Table 5 on the absence of systematic differences between these two sets of municipalities before the roll-out of the program.

one when the individual is in the reported labor market state. The model predicts correctly all the directions of the policy changes in 2006. It does also deliver point estimates that are very close to the data.<sup>73</sup> The only mismatch in the point estimates is found in the relative proportions of formal and informal employees: in these cases, the model predicts slightly larger impacts of the policy. These smaller elasticities in the data with respect to the model may be due to the transition dynamics. The data are obtained from an environment where the changes in  $B_0$  are relatively recent, for some municipalities significantly less than a year. As a result, the labor market data may not yet be in an equilibrium well approximated by our model's steady state.

## 6. COUNTERFACTUAL EXPERIMENTS

In this section, we use our estimates to examine the labor market impact of changes in some primitives of the model environment. We focus on counterfactual experiments closely related to the emergence of informality. We start with a scenario that prevents firms from offering jobs with informal contracts. Next, we study a more realistic policy where we change the payroll contribution rate for formal employees. These experiments shed some lights on the nature of the interaction between the institutional parameters and three sources of inefficiency that are present in our setting.<sup>74</sup>

The first is the holdup problem affecting the education decision (Acemoglu and Shimer, 1999; Flinn and Mullins, 2015). Acquiring additional schooling is an investment decision for the worker but a decision that must be taken before meeting an employer and observing the realized output of the match. Firms absorb part of the future benefits of the workers' investment due to the presence of wage bargaining. The result is a wedge between the private return to education and the aggregate return, leading to inefficiencies that are a function of the amount of frictions in the economy and of the way workers and firms split the surplus.

The second source of inefficiency relates to the job posting behavior. Again, through bargaining, firms receive only a portion of the surplus and therefore may not post enough vacancies. In addition, firms ignore their impact on the overall meeting rates when deciding to post, which may lead them to post too many vacancies. In a simpler model than ours, Hosios (1990) shows that the two effects can balance each other leading to efficient outcomes. The balance occurs when the workers' Nash-bargaining coefficient  $\alpha$  is set equal to the elasticity of the number of meetings with respect to the measure of searchers  $\iota$ .

Finally, the third source of inefficiency relates to the externalities created by worker sorting by unobserved ability (Charlot and Decreuse, 2010; Flinn and Mullins, 2015). In a model with workers' ability, any policy change that affects schooling levels will also affect the proportions of high- and low-ability types that acquire education. The externality takes place because firms tend to post more or less vacancies in each schooling submarket based in part on their expectations about the ability composition of the workers.

*6.1. The Impact of Informality.* As mentioned in Section 1 and Subsection 2.1, informality is both an optimal reaction to a given institutional environment and a feature of that same environment that may magnify its distortions. In our model, informality arises in equilibrium given regulations and partial enforcement. To give a quantitative assessment of the impact of informality, we perform a simple counterfactual experiment. In the model, firms have the option of offering both a formal and an informal contract to the workers they meet. In the counterfactual, we simply impose that firms do not have this option: they can only offer a formal

<sup>73</sup> We report formal tests on the difference in the signs and the magnitudes of the estimated coefficients in the simulated data when compared to those estimated in the data sample in Table C.7 of Appendix C.

<sup>74</sup> All the labor market outcomes in this section are obtained by simulation. For each new value of the policy parameter, we find and compute the new equilibrium holding fixed the other institutional and estimated parameters. Then, we simulate the labor market careers of 20,000 individuals in these counterfactual labor markets. Finally, we compute the relevant statistics on the simulated data.

contract. We do this without taking a stand of what enforcement policy may be able to achieve this result. Therefore, we simply provide a decomposition exercise highlighting how important is the option of hiring informally on labor market outcomes, welfare, and schooling investments.

Table 6 reports changes in summary statistics of the resulting counterfactual labor market without informal employment with respect to the benchmark model discussed in Subsection 5.2 with estimated parameters reported in Table 3. Since we will also be interested in studying some redistributive effects of the system, we focus on an ex post welfare measure: the overall flow value in steady state. Given our linear utility assumption, the welfare for workers in the searching and the employee states is fully described by the sum between income (or, if unemployed, search cost) and the valuation of the social security benefit. Analogously, the welfare for firms is described by their flow payoff. When the vacancy is filled, the payoff is the level of flow profit; when the vacancy is unfilled, the (negative) payoff is the flow cost  $v$ . We weight each flow value by its equilibrium measure in steady state and we integrate over the sources of workers' heterogeneity  $\{k, \kappa, y\}$ —since the ability type ( $k$ ), the direct cost of schooling ( $\kappa$ ), and the self-employment income ( $y$ ) are assigned by nature and are not affected by the decisions of the agents.

The welfare outcomes are reported in the top panel of Table 6. The second panel reports the steady-state proportions in each labor market state, showing how the optimal decisions rules in the labor market adjust to the new environment. Finally, the last panel of Table 6 shows the new equilibrium channel we introduce in the model: the proportion of individuals who choose to acquire additional schooling, together with a measure of their selection with respect to unobserved ability.

The first column reports the main experiment: firms cannot offer informal contracts but everything else in the model remains the same. Imposing this constraint leads to a significant loss in overall welfare, about 6%. This result was expected: agents are prevented from using an option they had optimally chosen; therefore, their welfare decreases. More interesting is the asymmetry of the welfare loss: firms loose more, about 28%, showing that the demand side of the labor market reaps most of the benefits from the option of offering informal contracts. This is driven by the firms' "first-mover" advantage in proposing the formality status of the job. Workers also loose but by a much smaller 4.6%. In addition, the heterogeneity over ability is substantial. Low-ability workers experience a 7.2% loss compared to a 2.3 loss for high-ability workers. The loss in welfare is accompanied by a large increase in the formality rate: the proportion of formal employees increases by 33.5%. The result has an important corollary for the policy debate as it clearly shows that policy interventions aimed at decreasing informality have likely negative welfare consequences. Therefore, looking at the overall informality rate as the outcome variable of interest is not very informative as it may actually mask relevant labor market dynamics.

Equally relevant for the policy debate is the impact on schooling: the proportion of agents acquiring additional schooling increases by a nonnegligible 10.3%. The result confirms that informality actually distorts returns to schooling and human capital accumulation. The distortion is not limited to the overall schooling level but extends to selection into schooling. The increase in the proportion of workers completing high school is almost eight times larger for the high-ability group than for the low-ability group. Without the option of working as informal employees, not only the stock of human capital in the economy improves but also its quality. These findings suggest that both the holdup problem and the negative externalities created by selection over ability are exacerbated by the possibility of offering informal job contracts.

The second column of Table 6 shows the relative importance of endogenous schooling decisions in generating the results presented in the first column of the table. We keep the schooling decisions fixed as in the benchmark model and in such environment we impose that firms can only offer a formal contract. Results show that the loss in welfare is more pronounced when compared to the baseline specification with endogenous schooling decisions. On the workers' side, all the effect is driven by the low-ability group. Ignoring endogenous schooling



TABLE 6  
THE EQUILIBRIUM EFFECTS OF INFORMAL EMPLOYMENT

Model: Specifications:	Firms Can Only Offer a Formal Contract			
	Baseline Model	Exogenous Schooling	Exogenous Contact Rates	Hosios-Like Condition ( $\alpha = \iota$ )
<u>Flow Welfare:</u>				
Total (Firms + Workers)	-0.0596	-0.0750	-0.0020	0.0478
Firms	-0.2821	-0.3219	-0.3055	-0.1589
Workers	-0.0460	-0.0599	0.0166	0.0570
Workers—Low Ability	-0.0725	-0.1141	-0.0355	0.0713
Workers -High Ability	-0.0231	-0.0130	0.0617	0.0320
<u>Labor Market Proportions:</u>				
Unemployed	0.0213	0.0636	0.0019	-0.0459
Self-employed	0.3353	0.3526	0.3625	0.2329
Formal Employees	0.0275	-0.0146	-0.0376	0.0076
<u>Schooling Proportions:</u>				
HS Completed	0.1029	-	0.0781	0.1501
HS Completed—Low Ability	0.0244	-	-0.0030	0.0528
HS Completed—High Ability	0.1623	-	0.1394	0.2202
High Ability among HS	0.0538	-	0.0569	0.0628

NOTES: The table reports relative changes with respect to the benchmark model, that is, the model estimated in Subsection 5.2 with estimated parameters reported in Table 3. For a more detailed description of the counterfactuals, see Subsection 6.1. All the labor market outcomes are obtained by simulating the labor market careers for 20,000 individuals. HS denotes the high Schooling level.

decision would lead to overestimate the negative impact on welfare of removing informality by about one-fourth. The third column of Table 6 studies the importance of endogenous contact rates. Similarly to the experiment reported in the second column of the table, we keep the contact rates fixed at estimated values and we impose that firms can only offer a formal contract. Preventing firms from adjusting their posting behavior leads to a 31% welfare loss for them and to a welfare gain for (high-ability) workers. The overall loss of welfare is greatly reduced when compared to the main experiment. Once again, the result shows that the option of offering informal job contracts is of greater advantage to firms as it magnifies the benefits of posting vacancy in each schooling submarket.

Given the trade-off implied by the job posting behavior, firms receive only a portion of the surplus and at the same time ignore their impact on the overall meeting rates. We investigate this issue further by generating a last counterfactual that impose the Hosios condition (Hosios, 1990), which balances out these two effects by equating the workers Nash-bargaining coefficient,  $\alpha$ , with the elasticity of the number of meetings with respect to the measure of searchers,  $\iota$ . It is important to notice that, in our environment, the condition provides just a useful reference point and not a parameterization that fully restores efficiency.<sup>75</sup> In column four of Table 6, we impose this condition both in the counterfactual (where informal employees are not allowed), and in the benchmark model, we compare it with (where informal employees are allowed). Results show that a Hosios-like condition in our environment would indeed eliminate the welfare loss of removing informality. Overall welfare would actually increase by almost 5%, suggesting that increasing the workers' bargaining position moves the economy toward higher efficiency. The distribution of the welfare benefits remains very asymmetric between firms and workers: the entire improvement is captured by workers, while firms experience a smaller but still significant loss. This dynamics improves human capital accumulation: both the increase in schooling and the improvement over ability are higher than in the baseline experiment.<sup>76</sup> The welfare gains are higher for the low-ability workers, while the increase in schooling is higher for high-ability workers.

*6.2. Changing the Contribution Rate.* In this section, we study more realistic policy interventions focusing on the institutional features responsible for the emergence of informality. Among them, we have chosen to concentrate on the payroll contribution rate  $t$ . This parameter is at the center of both the recent and the historical policy debate on informality, with countries with relatively high rates considering a reduction and countries with relatively small contributory benefits considering an increase (Antón, 2014; Bernal et al., 2017).

The policy experiments varying  $t$  can be performed under a balanced budget constraint in our setting. All the contributions collected through  $t$  are redistributed to the workers: A proportion  $(1 - \tau)$  finances the common health benefit  $b_1$ , while the remaining is paid as a retirement benefit increasing in the worker's contribution. One complication is that the noncontributory benefit  $B_0$  is instead financed by resources collected outside the labor market (see Subsection 2.1). Since the relative proportion of formal and informal workers is endogenous and as such it reacts to changes in  $t$ , the resources necessary to finance  $B_0$  are also endogenous. As a result, a change in  $t$  that leads to an endogenous increase in informality will require more outside resources to finance the same per capita noncontributory benefits. To satisfy the balanced budget constraint, we have decided to conduct the policy experiments keeping the *overall* amount of resources devoted to  $B_0$  unchanged from baseline.

Changes in the contribution rate generate interesting redistribution and selection effects. Since part of the benefit is fixed (the  $b_1$  component), the policy introduces redistribution from

<sup>75</sup> Two reasons prevents the model from delivering such clean theoretical result. First, the distortions introduced by the institutional parameters interact with the main sources of inefficiency and externality discussed above. Second, the expected match output in each schooling market is endogenous because of the composition effects over ability. Mangin and Julien (2021) show that efficiency cannot be restored through a specific choice of the Nash-bargaining coefficient when this source of endogeneity is present.

<sup>76</sup> In Appendix D, we study in more general terms the impact of  $\alpha$  and  $\iota$  on welfare and schooling outcomes.

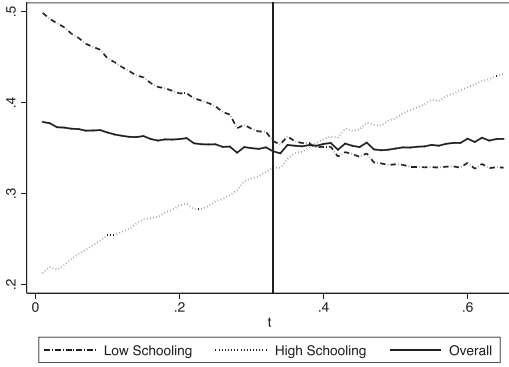
formal workers with relatively high wages to formal workers with relatively low wages. In addition, since high-ability/high schooling individuals are more likely to receive high formal wages, the policy also introduces redistribution from high schooling to low schooling workers and from high-ability to low-ability individuals. Ultimately, these effects impact the overall proportion and selection of agents who acquire additional schooling. The impact of the policy experiments on selected outcome variables of interest is reported in Figures 6 and 7.

Figure 6 reports the baseline experiment. Panels (a) and (b) show that the impact on the overall informality rate (informal employee and self-employed) is different from what usually assumed in the policy debate. The informality rate remains remarkably stable over a large interval around the baseline value. The main reason for this lack of elasticity is the presence of significant composition effects over both schooling and ability. Moving from low to high contribution rates, the informality rate among the high schooling individuals steadily increases, while it decreases and later stabilizes among the low schooling individuals. The informality rate by ability shows an even starker difference. At the same time, the overall proportion of high schooling individuals in the population decreases as shown in Panel (e), leading to the roughly constant informality rate we observe. The composition effects are driven by the redistribution implied by the endogenous benefit  $b_1$ : as the contribution rate increases, a proportionally larger benefit is available to lower earnings individuals.

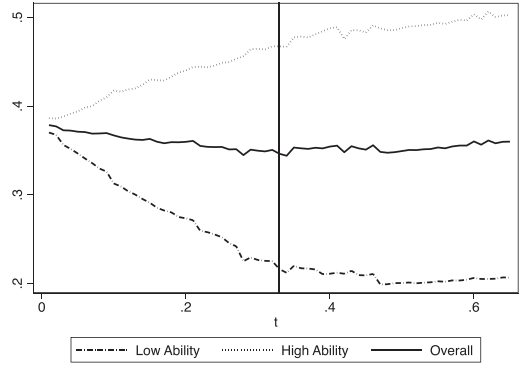
The middle panels of Figure 6 report the impact on welfare. The experiments are all taking into account the balance budget constraint: all the resources collected through the contribution rate  $t$  are fully redistributed to the workers through the  $b_1$  and  $\tau tw_1(x; y, h, k)$  benefits. However, the valuation of the benefits by the workers is not at full value: for each pesos contributed, the workers value the benefit between 21% and 39% less than full value (see the estimates for  $\beta_{1,h}$  in Table 3). We then expect the value of the contribution rate that maximizes welfare to be as low as possible. The experiments confirm this claim: welfare is maximized at a contribution rate around zero when it reaches a value 15.8% higher than at the benchmark contribution rate. This global optimum is accompanied by some local nonlinearity. For example, looking at a more policy-feasible neighboring interval around the benchmark value equal to 0.2–0.4, it is possible to pick a locally optimal contribution rate of 0.30 that generates a 3.5% increase in welfare with respect to baseline. While this increase is relatively small, it does show that politically feasible policy changes could lead to welfare improvements. Composition effects over schooling play an important role in generating the result, while composition effects over ability are less relevant. The welfare of high schooling individuals is much more sensible to the contribution rate than the one of low schooling individuals. A contribution rate twice as large as the baseline barely changes the low schooling welfare but decreases high schooling welfare by about 40%. These results suggest that welfare functions weighting individuals differently based on individuals characteristics such as schooling or income could be maximized for positive values of  $t$ .

The last two panels of Figure 6 report the impact on schooling. Since the returns to schooling are positive, redistribution through  $b_1$  transfers resources from the higher educated to the lower educated. As a result, the overall schooling level decreases monotonically in  $t$ , as shown in Panel (e). The negative impact is present in both ability levels: the slightly higher elasticity for high-ability individuals is not enough to significantly change the positive selection in high schooling. As for the case of welfare, schooling investments are maximized at a contribution rate around zero—with a 10% increase in the proportion of individuals who complete secondary education with respect to the baseline level at  $t = 0$ .<sup>77</sup>

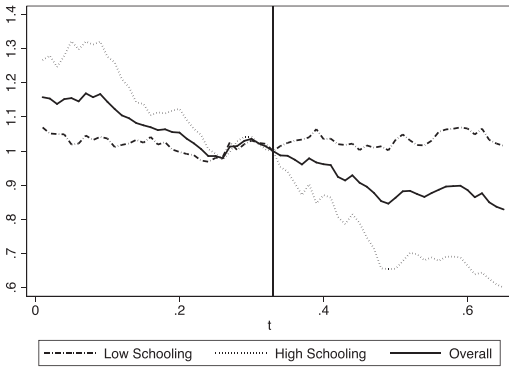
<sup>77</sup> To add some perspective on the magnitude of the schooling elasticity that we find in this counterfactual, consider the following benchmark. The *Progres-Oportunidades* program is a large-scale welfare program covering millions of families in Mexico, which aims to increase school enrollment and attainments by means of cash transfers that are conditional on specific household investments in health and education. Behrman et al. (2012) find an increase of 0.13 years of schooling for boys aged 15–18 using data for an urban sample of households that received cash transfers from *Oportunidades* for one year during the same period of observation of our study. In our sample, the average years of



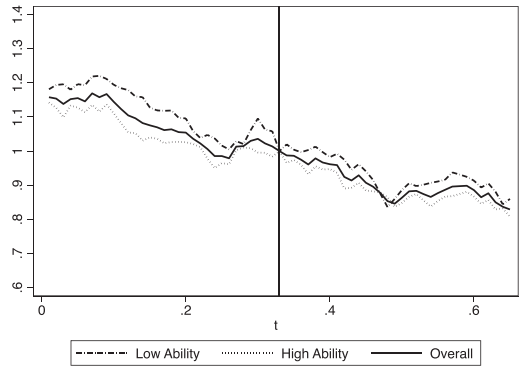
(a) Informality by Schooling



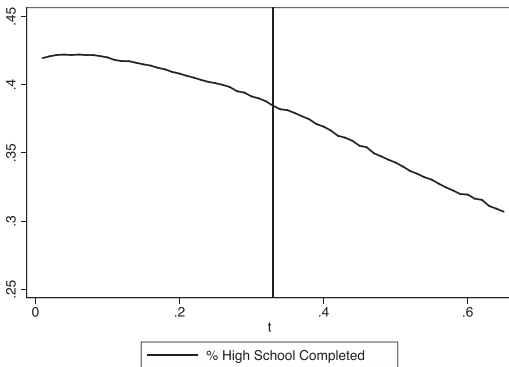
(b) Informality by Ability



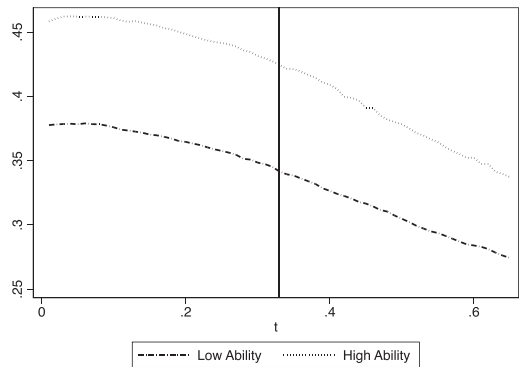
(c) Welfare by Schooling



(d) Welfare by Ability



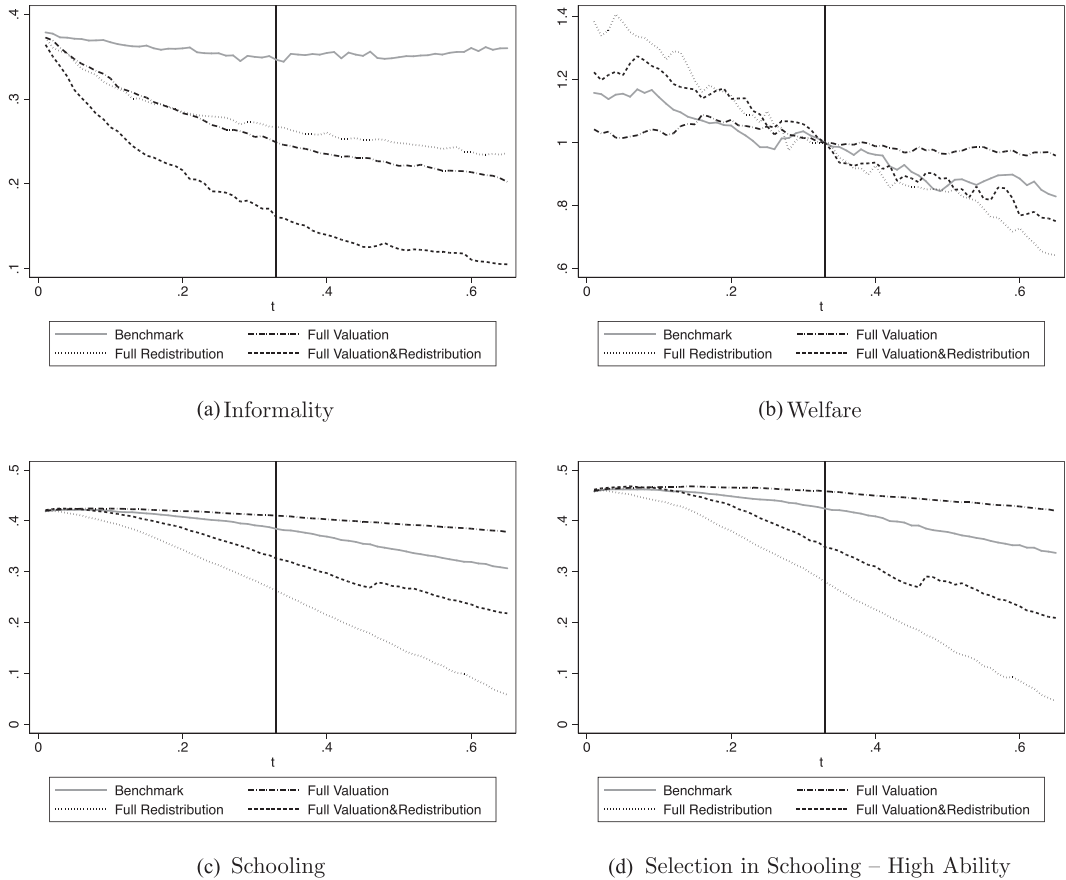
(e) Schooling



(f) Schooling by Ability

NOTE: The figures report outcomes from policy experiments that change the social security contribution rate  $t$  from 0 to 0.66 (a 0.33 change below and above the contribution rate at baseline, represented by the vertical line). *Informality* is the proportion of informal employee and self-employed in the population. *Inequality* is the ratio between the top 90% and the bottom 10% in the wages and self-employment income distribution. *Welfare* is the overall flow welfare in steady state (see Section 6). *Production* is the overall production ( $x$  and  $y$ ) in steady state. *Schooling* is the proportion of High Schooling in the population. *Selection in Schooling* is the same proportion by ability type.

FIGURE 6



NOTE: The figures report outcomes from the same policy experiments reported in Figure 6 with the addition of three changes in the environment: *Full Redistribution* sets  $\tau$  to zero; *Full Valuation* sets  $\beta_{1,h}$  to one; *Full Valuation&Redistribution* combines the two experiments. *Informality* is the proportion of informal employee and self-employed in the population. *Inequality* is the ratio between the top 90% and the bottom 10% in the labor income distribution. *Welfare* is the overall flow welfare in steady state (see Section 6). *Schooling* is the proportion of individuals with high schooling in the population. *Selection in Schooling* is the same proportion by ability type.

FIGURE 7

POLICY EXPERIMENTS, REDISTRIBUTION, AND BENEFITS' VALUATION

Figure 7 reports additional simulation results that clarify which channels matter the most in generating the observed effects documented above. We move two levers: the intensity of the redistribution and the valuation of the benefits. In the counterfactual, we label *Full Redistribution*, we set  $\tau$ —the portion of the contribution that is devoted to proportional benefits—to zero so that the entire payroll contribution is used to finance the fixed benefit  $b_1$ . This experiment provides the maximum amount of redistribution possible within this institutional framework. The baseline value for  $\tau$  is 0.55. In the counterfactual we label *Full Valuation*, we set

schooling completed for workers in the low schooling group is 9 years. Hence, the observed increase in the proportion of individuals who complete secondary (12 years of schooling) from the baseline value of 38.5% would translate in our sample into a 0.12 [= 0.1 \* 0.385 \* (12 - 9)] average increase in schooling attainment. The underlying costs of achieving a very similar policy objective are remarkably different. The experiment that we perform alters the labor market returns to schooling through the infrastructure of the social security system and it is essentially balanced-budget. Instead, the estimated cost for the year 2005 of the overall monetary grants and school supplies related to the education component of the *Progres-Oportunidades* program is 1.4 billion USD (Levy, 2006), or 0.1% of Mexico's GDP in 2005. For a systematic review of the effects of the *Progres-Oportunidades* program in both rural and urban areas of Mexico, see Parker et al. (2008).

the benefit valuation parameters  $\beta_{1,h}$  to one, imposing that agents value the benefits at full value. The baseline-estimated values for  $\beta_{1,0}$  and  $\beta_{1,1}$  are, respectively, 0.79 and 0.61. This experiment clarifies how important the valuation of the preferences is in generating the loss of welfare as  $t$  increases. Finally, in the counterfactual we label *Full Valuation&Redistribution*, we combine the two experiments setting both  $\tau$  to zero and  $\beta_{1,h}$  to one. For each of the three parameter combinations, we repeat the policy experiments we presented in Figure 6, moving the contribution rate  $t$  over the same interval.

As reported in Panel (a), both full redistribution and full valuation decrease informality at same  $t$ . The result was expected because both margins make formality more attractive. More surprising is the fact that the magnitudes of the two impacts are so strong that informality decreases as the payroll contribution rate increases. The reason is that they both make informality more attractive exactly for the matches at the margin between formality and informality—that is, those with productivity close enough to  $\bar{x}(y, h, k)$ . Panel (b) reports the welfare impacts, showing that redistribution was the main source of the local nonlinearities observed in Figure 6. With full redistribution, local nonlinearities remain and become more frequent while with full valuation they are smoothed out. Full redistribution also significantly increases the elasticity with respect to  $t$ : for values around zero, welfare would increase by almost 40% compared to baseline. Full valuation of benefits has exactly the opposite effect. The overall elasticity becomes almost zero if the benefits are considered at full value, with the negative impact of redistribution almost exactly compensating the positive impact of valuing the benefits more. Such offsetting effect does not hold anymore for higher levels of benefits' redistribution, as shown by the case where both full valuation and full redistribution are present. Finally, Panels (c)–(d) report the impact on schooling. As expected, full redistribution reduces the proportion of individuals completing high schooling. The reduction is larger for high-ability individuals, leading to a worse selection into high schooling as  $t$  increases. Full valuation of the benefits has exactly the opposite effects, reducing the negative impact both on overall schooling and on selection into schooling.

The main conclusion we draw from the experiments is that there is room for improving welfare when using the contribution rate as policy lever. For example, lowering the current contribution rate by about 10 percentage points would increase overall welfare by 3.5%. The overall optimal policy (setting the contribution rate to zero) would generate a 15.8% welfare increase compared to benchmark. These improvements are sensitive to the benefits' valuation. The closer the valuation is to full value, the larger is the improvement around the benchmark rate but the smaller is the improvement at the optimal  $t$  value. As in the previous subsection, we find that the informality rate is not a good guide for policy: all the welfare-improving policies we just mentioned leave the informality rate almost unaffected.

## 7. CONCLUSION

Informality is a defining feature of many labor markets. In Latin America, over 50% of the labor force is employed informally. Studying costs and benefits of informality requires an equilibrium model of the labor market that takes into account how workers and firms endogenously sort between formal and informal jobs. If the model wants to generate credible estimates and relevant counterfactual policy scenarios, it also needs to replicate the empirical regularities and the salient institutional features observed in these markets.

We attempt to accomplish both objectives by developing and estimating a search and matching model of the labor market where firms and workers endogenously decide to form matches (jobs) that can be formal or informal. The model replicates the main features of labor market dynamics in Mexico by allowing endogenous formality posting and endogenous wage determination through bargaining. The sources of heterogeneity determining wages and formality status are the match-specific productivity, the income of self-employed workers, and unobserved workers' ability. Meeting rates are also endogenous, as they are governed by the equilibrium proportions of searchers and vacancies. In this environment, we introduce three



relevant features ignored by previous literature. First, recognizing that the same labor market institutions that generate informality may affect not only short-run labor market outcomes but also long-run investment decisions, we allow for endogenous schooling decisions. Second, observing the recent policy changes in Latin America and similar regions with high informality, we model an important institutional innovation: the introduction of a dual social security system where noncontributory benefits targeting informal workers coexist with standard contributory benefits reserved for formal workers. Finally, on the basis of significant differences in labor market dynamics, we differentiate between informal workers hired as employees and informal workers working as self-employed.

We identify and estimate the model parameters using a combination of individual-level data on labor market dynamics, exogenous variation induced by institutional changes, and aggregate data on vacancies and labor shares from Mexico. The exogenous variation (the roll-out of the *Seguro Popular program*) is crucial to recover hard-to-identify preferences for the social security benefits. The estimates of the model parameters generate a good fit of the data, both in-sample and out-of-sample. The out-of-sample model validation is performed using the same source of exogenous variation used in estimation but over a different time period.

Estimation results deliver reasonable and precise point estimates. The parameters governing the rates of job arrival and termination and the match-specific productivity are comparable to previous estimates for similar models on high-income countries. The estimate of the Nash-bargaining coefficient implies a slightly stronger bargaining position to the workers although it remains quite close to the value of 0.5 that defines symmetric bargaining. The estimated values of the preference parameters show that the valuation of the noncontributory benefit is very close to full monetary value, while the valuation of the contributory benefits is about 61 cents to the peso for the high schooling group and about 79 for the low schooling group. In general, there are important differences between the primitive parameters estimated for the two schooling groups. These differences generate a return to schooling on the PDVs of participating in the market of about 30%.<sup>78</sup> Unobserved ability plays a role in the schooling choice. High-ability workers are about 19% more productive when matching with firms leading about 43% of them to complete high school compared with about 35% of low-ability workers.

We use the estimated model to examine the labor market impact of changes in the parameters closely related to the emergence of informality. We first implement a radical policy that prevents firms from offering jobs with informal contracts. We then study a more realistic policy where we change the payroll contribution rate for formal employees. From the first experiment, we find that offering informal jobs is welfare increasing but at the same time decreases schooling investments. Completely removing informal employee contracts in the current Mexican labor market would decrease steady-state welfare by about 6%. Implementing the same experiment in a different labor market that reduces posting inefficiencies would instead *increase* the steady-state welfare of workers by about 5%. From the second experiment, we find that the overall informality rate remains remarkably stable over a wide range of values of the contribution rate. The result is driven by significant composition effects induced by the redistributing nature of the social security system. While the contribution rate affects only marginally the level of informality, it has a more pronounced impact on welfare and schooling. Lowering the current contribution rate by about 10 percentage points would increase overall welfare by 3.5%. The optimal policy (setting the contribution rate to zero) would generate a 16% increase in welfare and a 10% increase in the proportion of individuals who complete secondary education. From both experiments, we conclude that using the overall informality rate as the variable of interest to guide policy may be misleading.

We see two main limitations in our work. The first concerns the demand side. If we allow workers to decide on long-term investments in schooling, we do not let firms make investment

<sup>78</sup> The return refers to completing secondary schooling with respect to completing at most lower secondary schooling.

decisions. In our model, firms can enter the market and post vacancies for different schooling levels but they cannot change their capital/labor mix and they are restricted to constant returns to scale technologies. A unified model able to merge the demand and supply side's investment decisions in labor markets with high informality is still missing in the literature.<sup>79</sup> The second limitation concerns the supply side of the labor market. Workers are allowed to decide on human capital investments before entering the labor market but not after they do. Just as our results show that informality affects the first decision, so it is likely to matter for the second. Evidence on the higher instability of informal jobs and on the reluctance of firms to invest in specific human capital when hiring informally indicates that the associated consequences in terms of human capital decisions and earning dynamics over the life cycle can be large.<sup>80</sup> Given our results, allowing for both firms' investment decisions and human capital accumulation on the job should enrich our understanding of the benefits and costs of informality in labor markets where agents have shown to respond to the incentives and limitations created by the institutional system.

### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure B.1:** Optimal Decision Rules

**Figure B.2:** Workers' Values of Searching:  $S(y, h)$  and  $U(h)$

**Figure B.3:** Firms' Values of Posting:  $V[h]$

**Table C.1:** Demand-side Parameters for Different Values of  $\iota$

**Table C.2:** Predicted Values

**Table C.3:** Cross-Sectional Moments and Model Fit—Treated Sample

**Table C.4:** Cross-Sectional Moments and Model Fit—Control Sample

**Table C.5:** Longitudinal Moments and Model Fit

**Table C.6:** Aggregate Moments and Model Fit

**Table C.7:** Out-of-Sample Model Validation

**Figure D.1:** Robustness for Different Values of the Bargaining Parameter ( $\alpha$ ) and the Matching function Parameter ( $\iota$ )

suppmat

### REFERENCES

- ABRAMITZKY, R., AND V. LAVY, "How Responsive Is Investment in Schooling to Changes in Redistributive Policies and in Returns?" *Econometrica* 82 (2014), 1241–72.
- ACEMOGLU, D., AND R. SHIMER, "Holdups and Efficiency with Search Frictions," *International Economic Review* 40 (November 1999), 827–49.
- ALBRECHT, J., L. NAVARRO, AND S. VROMAN, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal* 119 (2009), 1105–29.
- ALMEIDA, R., AND P. CARNEIRO, "Enforcement of Labor Regulation and Informality," *American Economic Journal: Applied Economics* 4 (2012), 64–89.
- ANTÓN, A., "The Effect of Payroll Taxes on Employment and Wages under high Labor Informality," *IZA Journal of Labor & Development* 3 (2014), 20.
- ANTON, A., F. HERNANDEZ, AND S. LEVY, *The End of Informality in Mexico?* (Washington, D.C.: The Inter-American Development Bank, 2012).

<sup>79</sup> Meghir et al. (2015) develop and estimate a model able to explain labor market dynamics and firms' size but they do not allow for physical capital investments and productivity shocks. Haanwinkel and Soares (2016) propose a search and matching model of an informal labor market that allows for imperfect substitutability across types of workers and decreasing returns to scale technology.

<sup>80</sup> See Lagakos et al. (2018) for recent cross-country evidence. In a companion paper (Bobba et al., 2021), we have started to study the issue, finding that the rate of human capital accumulation on the job is higher while working formally than informally.

- ARDAGNA, S., AND A. LUSARDI, "Where Does Regulation Hurt? Evidence from New Businesses Across Countries," Technical Report, National Bureau of Economic Research, 2009.
- ARROYO MIRANDA, J., R. GÓMEZ CRAM, AND C. LEVER GUZMÁN, "Can Matching Frictions Explain the Increase in Mexican Unemployment after 2008?" Technical Report, Working Papers, Banco de México, 2014.
- ATKIN, D., "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review* 106 (August 2016), 2046–85.
- ATTANASIO, O. P., C. MEGHIR, AND A. SANTIAGO, "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA," *Review of Economic Studies* 79 (2012), 37–66.
- AZUARA, O., AND I. MARINESCU, "Informality and the Expansion of Social Protection Programs: Evidence from Mexico," *Journal of Health Economics* 32 (2013), 938–50.
- BAFFES, J., AND T. HANIOTIS, "Placing the 2006/08 Commodity Price Boom into Perspective," World Bank Policy Research Working Paper, 2010.
- BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J.-M. ROBIN, "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *American Economic Review* 104 (June 2014), 1551–96.
- BEHRMAN, J. R., J. GALLARDO-GARCIA, S. W. PARKER, P. E. TODD, AND V. VELEZ-GRAJALES, "Are Conditional Cash Transfers Effective in Urban Areas? Evidence from Mexico," *Education Economics* 20 (February 2012), 233–59.
- BELZIL, C., "The Return to Schooling in Structural Dynamic Models: A Survey," *European Economic Review* 51 (2007), 1059–105.
- BERNAL, R., M. ESLAVA, M. MELÉNDEZ, AND A. PINZÓN, "Switching from Payroll Taxes to Corporate Income Taxes: Firms' Employment and Wages after the 2012 Colombian Tax Reform," *Economía* 18 (2017), 41–74.
- BIANCHI, M., AND M. BOBBA, "Liquidity, Risk, and Occupational Choices," *Review of Economic Studies* 80 (2013), 491–511.
- BOBBA, M., L. FLABBI, AND S. LEVY, "Labor Market Policies and Schooling Investments," Technical Report, Draft prepared for the Workshop Labor Market Opportunities and Schooling Investments, IDB, 2012.
- , ———, ———, AND M. TEJADA, "Labor Market Search, Informality, and on-the-job Human Capital Accumulation," *Journal of Econometrics* 223 (2021), 433–453.
- BOROWCZYK-MARTINS, D., J. BRADLEY, AND L. TARASONIS, "Racial Discrimination in the US Labor Market: Employment and Wage Differentials by Skill," *Labour Economics* 50 (2018), 45–66.
- , G. JOLIVET, AND F. POSTEL-VINAY, "Accounting for Endogeneity in Matching Function Estimation," *Review of Economic Dynamics* 16 (2013), 440–51.
- BOSCH, M., AND R. M. CAMPOS-VAZQUEZ, "The Trade-Offs of Welfare Policies in Labor Markets with Informal Jobs: The Case of the "Seguro Popular" Program in Mexico," *American Economic Journal: Economic Policy* 6 (2014), 71–99.
- , AND J. ESTEBAN-PRETEL, "Job Creation and Job Destruction in the Presence of Informal Markets," *Journal of Development Economics* 98 (2012), 270–86.
- , AND M. MANACORDA, "Minimum Wages and Earnings Inequality in Urban Mexico," *American Economic Journal: Applied Economics* 2 (October 2010), 128–49.
- BURDETT, K., R. LAGOS, AND R. WRIGHT, "Crime, Inequality, and Unemployment," *American Economic Review* 93 (2003), 1764–77.
- CAHUC, P., F. POSTEL-VINAY, AND J.-M. ROBIN, "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica* 74 (2006), 323–64.
- CHARLOT, O., AND B. DECREUSE, "Over-Education for the Rich, Under-Education for the Poor: A Search-Theoretic Microfoundation," *Labour Economics* 17 (2010), 886–96.
- CONTI, G., R. GINJA, AND R. NARITA, "Non-Contributory Health Insurance and Household Labor Supply: Evidence from Mexico," Technical Report, Mimeo, 2017.
- DE PAULA, A., AND J. A. SCHEINKMAN, "Value-Added Taxes, Chain Effects, and Informality," *American Economic Journal: Macroeconomics* 2 (October 2010), 195–221.
- DE PAULA, A., "The Informal Sector: An Equilibrium Model and Some Empirical Evidence from Brazil," *Review of Income and Wealth* 57 (2011), S8–S26.
- DEY, M. S., AND C. J. FLINN, "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica* 73 (2005), 571–627.
- DUFLO, E., M. GREENSTONE, R. PANDE, AND N. RYAN, "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India," *Econometrica* 86 (2018), 2123–60.
- ECKSTEIN, Z., AND G. J. VAN DEN BERG, "Empirical Labor Search: A Survey," *Journal of Econometrics* 136 (February 2007), 531–64.
- , AND K. I. WOLPIN, "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model," *Review of Economic Studies* 62 (April 1995), 263–86.

- FENG, Y., D. LAGAKOS, AND J. E. RAUCH, "Unemployment and Development," NBER Working Papers 25171, National Bureau of Economic Research, Inc., October 2018.
- FIELDS, G. S., "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs," *Journal of Development Economics* 2 (June 1975), 165–87.
- FLABBI, L., AND M. LEONARDI, "Sources of Earnings Inequality: Estimates from an on-the-job Search Model of the US Labor Market," *European Economic Review* 54 (2010), 832–54.
- , AND J. MABLI, "Household Search or Individual Search: Does It Matter?" *Journal of Labor Economics* 36 (2018), 1–46.
- , AND A. MORO, "The Effect of Job Flexibility on Female Labor Market Outcomes: Estimates from a Search and Bargaining Model," *Journal of Econometrics* 168 (2012), 81–95.
- FLINN, C., "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining and Endogenous Contact Rates," *Econometrica* 73 (2006), 1013–62.
- , AND J. HECKMAN, "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics* 18 (1982), 115–68.
- , AND J. MULLINS, "Labor Market Search and Schooling Investment," *International Economic Review* 56 (2015), 359–98.
- FLINN, C. J., AND ———, "Firms' Choices of Wage-Setting Protocols in the Presence of Minimum Wages," Technical Report, New York University and University of Minnesota, 2019.
- FOSSEN, F. M., AND T. J. BÜTTNER, "The Returns to Education for Opportunity Entrepreneurs, Necessity Entrepreneurs, and Paid Employees," *Economics of Education Review* 37 (2013), 66–84.
- FROLICH, M., D. KAPLAN, C. PAGES, J. RIGOLINI, AND D. ROBALINO, eds., *Social Insurance, Informality, and Labor Markets: How to Protect Workers While Creating Good Jobs*, number 9780199685233 in OUP Catalogue (Oxford: Oxford University Press, 2014).
- FU, C., AND J. GREGORY, "Estimation of an Equilibrium Model with Externalities: Post-Disaster Neighborhood Rebuilding," *Econometrica* 87 (2019), 387–421.
- GALIANI, S., A. MURPHY, AND J. PANTANO, "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment," *American Economic Review* 105 (November 2015), 3385–415.
- GAUTIER, P., P. MULLER, B. VAN DER KLAUW, M. ROSHOLM, AND M. SVARER, "Estimating Equilibrium Effects of Job Search Assistance," *Journal of Labor Economics* 36 (2018), 1073–125.
- HAANWINCKEL, D., AND R. R. SOARES, "Workforce Composition, Productivity, and Labor Regulations in a Compensating Differentials Theory of Informality," IZA Discussion Papers 9951, Institute for the Study of Labor (IZA), May 2016.
- HEATH, R., AND M. A. MOBARAK, "Manufacturing Growth and the Lives of Bangladeshi Women," *Journal of Development Economics* 115 (2015), 1–15.
- HOPENHAYN, H., AND R. ROGERSON, "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy* 101 (October 1993), 915–38.
- HOSIOS, A. J., "On the Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies* 57 (1990), 279–98.
- HSIEH, C.-T., AND P. J. KLENOW, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics* 124 (2009), 1403–48.
- JENSEN, R., "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *The Quarterly Journal of Economics* 127 (2012), 753–92.
- JOVANOVIC, B., "Job Matching and the Theory of Turnover," *Journal of political economy* 87 (1979), 972–90.
- KANBUR, R., "Conceptualizing Informality: Regulation and Enforcement," *Indian Journal of Labour Economics* 52 (2009), 33–42.
- KEANE, M. P., P. E. TODD, AND K. I. WOLPIN, "The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications," *Handbook of Labor Economics* 4 (2011), 331–461.
- , AND K. I. WOLPIN, "The Career Decisions of Young Men," *Journal of Political Economy* 105 (June 1997), 473–522.
- , AND ———, "The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women," *International Economic Review* 51 (2010), 851–92.
- LA PORTA, R., AND A. SHLEIFER, "The Unofficial Economy and Economic Development," *Brookings Papers on Economic Activity* 2008 (2008), 275–352.
- LA PORTA, R., AND ———, "Informality and Development," *Journal of Economic Perspectives* 28 (September 2014), 109–26.
- LAKAKOS, D., B. MOLL, T. PORZIO, N. QIAN, AND T. SCHOELLMAN, "Life Cycle Wage Growth across Countries," *Journal of Political Economy* 126 (2018), 797–849.
- LEVY, S., *Progress against Poverty: Sustaining Mexico's Progresa-Oportunidades Program* (Washington, DC: The Brookings Institution, 2006).

- , *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico* (Washington, DC: The Brookings Institution Press, 2008).
- , AND L. F. LOPEZ-CALVA, “Persistent Misallocation and the Returns to Education in Mexico,” *The World Bank Economic Review* 34 (2018), 284–311.
- , AND N. SCHADY, “Latin America’s Social Policy Challenge: Education, Social Insurance, Redistribution,” *Journal of Economic Perspectives* 27 (Spring 2013), 193–218.
- LJUNGQVIST, L., AND T. J. SARGENT, “The European Unemployment Dilemma,” *Journal of Political Economy* 106 (June 1998), 514–50.
- MAGNAC, T., “Segmented or Competitive Labor Markets,” *Econometrica* 59 (1991), 165–87.
- MALONEY, W. F., “Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico,” *World Bank Economic Review* 13 (May 1999), 275–302.
- MANGIN, S., AND B. JULIEN, “Efficiency in Search and Matching Models: A Generalized Hosios Condition,” *Journal of Economic Theory* 193 (2021), 105208.
- MARGOLIS, D. N., “By Choice and by Necessity: Entrepreneurship and self-employment in the Developing World,” *The European Journal of Development Research* 26 (2014), 419–36.
- MCKENZIE, D. J., AND C. WOODRUFF, “Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises,” *Economic Development and Cultural Change* 55 (October 2006), 3–42.
- MEGHIR, C., R. NARITA, AND J.-M. ROBIN, “Wages and Informality in Developing Countries,” *The American Economic Review* 105 (2015), 1509–46.
- MUNSHI, K., AND M. ROSENZWEIG, “Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy,” *American Economic Review* 96 (September 2006), 1225–52.
- NARITA, R., “Self-employment in Developing Countries: A Search-Equilibrium Approach,” *Review of Economic Dynamics* 35 (2020), 1–34.
- NEWBY, W. K., AND D. MCFADDEN, “Large Sample Estimation and Hypothesis Testing,” in S. Durlauf, L. P. Hansen, J. J. Heckman, and R. L. Matzkin, eds., *Handbook of Econometrics*, Volume 4 (San Diego: Elsevier, 1994), 2111–245.
- PAKES, A., AND D. POLLARD, “Simulation and the Asymptotics of Optimization Estimators,” *Econometrica* 57 (September 1989), 1027–57.
- PARKER, S. W., L. RUBALCAVA, AND G. TERUEL, “Evaluating Conditional Schooling and Health Programs,” in T. P. Schultz and J. A. Strauss, eds., *Handbook of Development Economics*, Volume 4, Chapter 62 (Amsterdam: Elsevier, 2008), 3963–4035.
- PERRY, G. E., W. F. MALONEY, O. S. ARIAS, P. FAJNZYLBER, A. D. MASON, AND J. SAAVEDRA-CHANDUVI, *Informality: Exit and Exclusion*, Number 6730 in World Bank Publications (Washington, DC: The World Bank, 2007).
- PETRONGOLO, B., AND C. A. PISSARIDES, “Looking into the Black Box: A Survey of the Matching Function,” *Journal of Economic literature* 39 (2001), 390–431.
- RESTUCCIA, D., AND R. ROGERSON, “Misallocation and Productivity,” *Review of Economic Dynamics* 16 (2013), 1–10.
- SAMANIEGO DE LA PARRA, B., AND L. F. BUJANDA, “Increasing the Cost of Informal Workers: Evidence from Mexico,” Technical Report, Mimeo, 2020.
- TODD, P. E., AND K. I. WOLPIN, “Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility,” *American Economic Review* 96 (December 2006), 1384–417.
- ULYSSEA, G., “Firms, Informality, and Development: Theory and Evidence from Brazil?” *American Economic Review* 108 (August 2018), 2015–47.
- WOLINSKY, A., “Matching, Search, and Bargaining,” *Journal of Economic Theory* 42 (1987), 311–33.
- Bank World, *World Development Report 2013: Jobs* (Washington, DC: The World Bank, 2012).