Sources and Implications of Labor Market Informality: The Case of Minimum Wage Policies

Mariano Bosch Mossi* Luca Flabbi† William F. Maloney‡
Mauricio Tejada§

December 10, 2025

Abstract

The paper develops and estimates a search and matching model of the labor market that can generate all the empirical features we observe when a binding mandatory minimum wage is introduced in markets with high informality. The estimated model delivers measures of overall welfare that are more appropriate than wage distributions or employment rates to evaluate the impact of minimum wage policies on the entire labor market dynamics. In addition, the estimated model allows to perform counterfactual policy experiments taking into accounts equilibrium effects. Preliminary results on Colombia show that reducing by a small amount the minimum wage from current levels increases workers' welfare, but reducing it by a large amount decreases it. Interestingly, significant increases of the minimum wage from current levels lead to a substitution of formal employee jobs with informal self-employment jobs and not with informal employee jobs.

Keywords: Informality, Minimum Wage, Latin America, Search and Matching, Bargaining.

JEL Codes: J24, J3, J64, O17

*Inter-American Development Bank, Washington, DC. MBOSCH@iadb.org
†University of North Carolina - Chapel Hill, NC. luca.flabbi@unc.edu.
‡World Bank, Washington, DC. Wmaloney@worldbank.org
§Universidad Diego Portales, Santiago, Chile. mauricio.tejada@udp.cl.

1 Introduction

A comprehensive theoretical framework for analyzing labor market informality under a mandatory and binding minimum wage policy must satisfy several key stylized facts. It should generate an equilibrium featuring a substantial share of jobs with varying formality statuses and allow for transitions in both directions between informal and formal employment. Formal sector wages should, on average, exceed informal wages while still exhibiting considerable overlap in their distributions. Additionally, the formal wage distribution must reflect a significant concentration of workers earning exactly the minimum wage, and the informal wage distribution should respond to changes in the minimum wage, a phenomenon often described as the lighthouse effect.

In this paper, we develop a search and matching model of the labor market that meets these key stylized facts. Beyond replicating observed wage and employment patterns, the model provides welfare measures that more accurately capture the overall impact of minimum wage policies on labor market dynamics than traditional metrics, such as wage distributions or employment rates.

Furthermore, we propose an identification and estimation strategy tailored to standard labor market data commonly available for several Latin American and Caribbean (LAC) countries. The model is then estimated using data from Colombia and Mexico, two countries with high levels of informality and different levels of how binding the minimum wage is, enabling the simulation of policy experiments grounded in these empirical estimates.

The paper's modeling approach and identification strategy build on previous empirical studies in the search and matching literature. Two main strands of literature are particularly relevant. The first relates directly to labor market informality and has rapidly developed in recent years (see for example Albrecht et al., 2009; Bosch and Esteban-Pretel, 2012; Meghir et al., 2015; Bobba et al., 2021, 2022). While this literature studies and estimates labor market informality within the search and matching framework, it has yet to provide a convincing treatment and estimation of binding and mandatory minimum wages in markets with high informality. The second strand focuses on minimum wage policies but is limited to high-income countries only (see in particular Flinn, 2006; Eckstein et al., 2011; Flinn and Mullins, 2021). A first theoretical attempt to integrate minimum wage and informality in a search and matching model is Flabbi (2021).

2 Model

2.1 Environment

The model's environment is characterized by frictions in the labor market, by workers and firms that make decisions based on expectations about the future, by match-specific productivity, and by bargaining as a wage-determination mechanism. There are four labor market states: unemployment, self-employment, informal salaried employment, and formal salaried employment. Agents are infinitely lived and discount the future at a rate ρ.

While unemployed, agents receive a flow utility β_u that captures unemployment benefits net of any cost of searching for a job as they search for both salaried employment and self-employment opportunities. Search frictions in the market for salaried jobs are characterized by a constant returns to scale matching function $Q(u, v)$, where u represents the unemployment rate, and $v = \sum_{f=0,1} v_f$ denotes the total vacancy rate. We denote with $f = 0, 1$ respectively an informal or a formal job. The aggregate market tightness is defined as $\theta = v/u$, and the probability of a job seeker encountering an employer of type f is given by $\pi_f = v_f/(v_1 + v_0)$. With this setup, the Poisson arrival rate for an unemployed meeting a prospective employer of type f is $\lambda_f = \pi_f \lambda(\theta) = \pi_f q(\theta)$ for $f = 1, 0$. When a meeting occurs, both the worker and the prospective observe the potential productivity of the match x and the formality regime of the job f. We assume that $x \sim G_f(x)$, where G_f is an exogenous, formality-specific cumulative distribution function (cdf). Conditioning on the match-specific productivity and the offered formality status, workers and firms engage in bargaining to determine the wage w and decide if accepting the match or not. If the match is accepted by both sides, production will take place; if not, no production will take place and the agent will go back to unemployment and search. An accepted match will last with productivity x until it is exogenously terminated by a shock with Poisson rate η_f, a shock summarizing any firing, termination or quitting taking place at a job.

For a firm with an unfilled vacancy, the cost of posting one is κ. We assume that posting a vacancy in the formal sector requires technological specificity and that firms vary in their ability to meet these requirements. We define $\nu \sim F(\nu)$ as the additional cost of posting a formal vacancy. Only firms for which the expected value of a formal job, once filled, is sufficient to cover this additional cost will post a formal vacancy (i.e., those with a low enough ν) and pay $\kappa + \nu$; otherwise, they will post an informal vacancy and pay only κ. As in Acemoglu (2001), all firms, regardless of whether they post a formal or informal vacancy, compete to meet workers in the same market. Consequently, the Poisson arrival rate for them to meet with a prospective employee from the pool of unemployed is $\varphi(\theta) = q(\theta)/\theta$.

Unemployed searchers can also find a self-employment opportunity, and they do so at exogenous Poisson rate γ. We denote the monetary (net) value of such self-employment opportunity with y and assume $y \sim R(y)$, where $R(y)$ is an exogenous cdf.[1] As in the previous case, if the opportunity is taken, production will take place; if not, no production will take place and the agent will go back to unemployment and search. An accepted self-employment opportunity will last with output y until it is exogenously terminated by a shock with Poisson rate δ, a shock summarizing any failure, termination or quitting.

On top of the structural differences described above, the formality status interacts with the institutional context. Being in a formal salaried job guarantees compliance with labor regulations, delivers the associated benefits but requires to pay the mandatory costs. Being in an informal salaried job excludes workers from the benefits provided by a formal job but allows them to avoid the associated costs.

[1] By assuming y represents the full net productivity, we are abstracting from any setup costs or flow costs.

Self-employed jobs share some aspects of informal jobs and some aspects of formal jobs but they are distinctive enough to warrant the introduction of a separate labor market state able to describe their specific features. The debate around the nature of self-employed jobs remains active but mounting evidence suggests that they should be seen as a possible alternative to salaried work since they offer a distinct set of characteristics that workers value. Recently, Donovan et al. (2023) argue that the relative patterns of entry from self-employment into formal salaried work suggest a more traditional view (Fields, 1975) of the sector as primarily a residual or disguised unemployed sector. However, the same data confirms earlier findings (Maloney, 1999; Bosch and Maloney, 2010) that duration of self-employment jobs is long and more comparable to that of formal salaried work than to unemployment, and underlying transition patterns are more consistent with self-employment being a closer substitute for salaried work than for unemployment. Motivational responses suggest that roughly 2/3s of workers enter informal self- employment voluntarily from formal salaried work (Maloney, 1999) in Mexico and a similar share report choosing not to leave the sector in Brazil for reasons of greater flexibility, being one's own boss and higher earnings. Willingness to pay responses in Mexico suggest that the value of these amenities may reach up to 100% of the wage. Similarly, measures of self-reported wellbeing in Ghana of informal micro firm owners and formal salaried are comparable (Falco et al., 2015). Finally, gross worker flows between formal salaried and self-employment behave pro-cyclically and symmetrically, and sector size pro-cyclically in some countries where nominal rigidities are low such as Mexico (Bosch and Maloney, 2010) but may become a-cyclical or counter-cyclical in more rigid contexts like Colombia (Fiess et al., 2010) suggesting the mix between those selecting into self-employment and those preferring other sectors varies with how binding rigidities are. In addition, starting this type of activities does not require significant set-up costs (McKenzie and Woodruff, 2006; Bianchi and Bobba, 2013) and it is therefore comparable to searching for a salaried job. We have therefore taken the view that the unemployed are searching for both employee jobs and self-employed jobs but possibly at different rates and facing different income distribution offers.

From the point of view of regulations and institutions, self-employment activities are typically required to pay no or very little contributions. In Mexico, self-employed workers are not legally required to make any social security contribution and, in practice, avoid them completely. Similar institutional setting are present in other LAC countries, such as Peru, Honduras, Dominican Republic, and, partially, Colombia. For this large share of the self-employed labor force, it is therefore reasonable to assume that agents do not pay social security contributions, similar to informal salaried workers. At the same time, it is also reasonable to assume that they do not face the same auditing and fines that agents in informal salaried jobs do.[2]

[2] Another portion of the self-employed labor force is significantly different. It is constituted by skilled agents working in their own professional activities, such as consultancies, medical professions and similar. It is also constituted by actual entrepreneurs for whom the enterprise is not "micro" but actually of small-medium size, involving a number of salaried workers working for them. This type of self-employment activities involves a relative small number of agents. It is also concentrated on the most skilled strata of the labor force. Since this is not the labor force most likely to be affected by minimum wage regulations, the model will ignore this form of self-employment

We formalize the benefits and costs of the institutional context affecting these three different labor market states in the following way.

We assume that salaried formal jobs are in full compliance with the labor legislation. We describe the cost side of the labor legislation with two parameters: a proportional payroll contribution t and a mandatory minimum wage w_M. We assume that the payroll contribution is withdrawn at the source by the firm. We formalize compliance with the minimum wage by introducing a wage floor constraint when workers and firms bargain for wages. We describe the benefit side of the labor legislation with one parameter: a flow benefit β_1 that workers receive while they are employed in a formal salaried job. It captures the entire bundle of benefits that workers receive from social protection and social security institutions, as well as preferences for being protected by labor legislation. These costs and benefits are in addition to the labor market structural differences between formality regimes described above and affecting mobility rates and productivity distributions.

Salaried informal jobs do *not* comply with labor legislation, potentially incurring the cost of being discovered and punished. We parameterize it as a flow cost c, without attempting to separate the probability of being discovered from the extent and amount of the punishment. Consistently with most legislation in the LAC region, we assume that the cost is paid by the firm. The benefit of non-compliance is not paying the payroll contribution t and acquiring the freedom to set wages without the added constraint w_M.

Self-employed workers-like informal salaried workers-avoid any payroll contribution payments and receive a flow utility β_s that captures preferences for flexibility. Unlike informal salaried workers, they do not face any auditing or punishment for their behavior because labor market legislation imposes far fewer requirements on them. In addition, they are not subject to any mandatory minimum wage requirement.

The last ingredient to characterize the labor market environment is the specification of workers' flow utility function and firms' flow profit. As common in most of the search literature, we assume linear utility.[3] As common in most of the search literature with informality, we assume that the job's monetary benefit is linearly separable from the additional contributory or non-contributory benefits.[4] Assuming linear utility is of course a major restriction but makes the model more tractable and improves the chance to identify the model on multiple LAC countries because it reduces the necessary data requirements. In addition, its loss of generality is less costly in a model like ours that does not focus on saving and borrowing behavior. The firms' flow profits are simply the difference between the value of production x and the labor costs. Labor costs include not only the wage cost $w_f(x)$ but also the institutional costs imposed on the firm: the proportional payroll tax t if the job is formal or the cost c of being found and punished if the job is informal.

and it will concentrate on the "residual" type of self-employment described above.

[3]For example, all the papers cited in two of the most influential reviews of the literature assume linear utility (Eckstein and van den Berg, 2007; Rogerson et al., 2005). For exceptions, see Dey and Flinn (2008); Flabbi and Mabli (2018).

[4]See for example, Conti et al. (2018); Bobba et al. (2021, 2022).

To summarize the assumptions of the model:

1. The technology determines the distribution of match-specific productivity $G_f(x)$, the distribution of vacancies π_f and the distribution of self-employment opportunities $R(y)$.

2. The labor market structure determines the extent of the frictions in the meeting of labor supply and labor demand, affecting the arrival rates λ, γ, the job meeting rate φ, and the termination rates η_f, δ.

3. The institutional context sets the policy parameters: the features of the social security system t, β_f; the mandatory minimum wage w_M; and the cost c of being found in an informal job and punished.

4. Workers and firms endogenously determine the accepted wage distributions, the employment level, the informality level, and the distribution over the different labor market states.

2.2 Value Functions

Thanks to the characteristics of the environment described in Section 2.1, it is possible to write the present discounted value of being in each labor market state in recursive form. The values are reported below and each of them has the usual interpretation as the sum of the flow value and the continuation value. We denote the value of unemployment U, the value of self-employment in an activity producing y with $S[y]$, the value of employment in a salaried job producing x with formality status f with $E_f[w_f(x)]$

The value of an unemployed worker is determined as follows:

$$\begin{aligned}(\rho+\lambda+\gamma)U &= \beta_u + \lambda \left\{ \begin{array}{l} \pi_1 \int \max\left[E_1[w_1(x)], U\right] dG_1(x) \\ +\pi_0 \int \max\left[E_0[w_0(x)], U\right] dG_0(x) \end{array} \right\} \\ &+ \gamma \left\{ \int \max\left[S(y), U\right] dR(y) \right\} \ , \pi_1 + \pi_0 = 1 \end{aligned} \tag{1}$$

where it should be clear that the contact rate λ depend on the aggregate market tightness θ and π_f depend on the number of vacancies created.

The value of a self-employed in a job with productivity y is determined by the following expression:

$$(\rho+\delta)S[y] = y + \beta_s + \delta U \tag{2}$$

The value of a salaried employee producing x with formality status f is determined by the following expression:

$$(\rho+\eta_f)E_f[w_f(x)] = w_f(x) + f \times \beta_1 + \eta_f U \tag{3}$$

On the firm side, the value of a filled job $J_f(x)$ producing x with formality status f is simply the discounted value of flow profits minus the value of an unfilled vacancy V_j whenever a termination shock arrives:

$$(\rho+\eta_f)J_f[x] \;=\; x-w_f(x)-ftw_f(x)-(1-f)c+\eta_f V_f \tag{4}$$

The value of an unfilled vacancy V_f for a type ν firm, in turn, is the discounted value of the flow cost of posting a vacancy plus the expected value in case of a filled job:

$$(\rho+\varphi)V_f(\nu) = -(\kappa+\nu\times f)+\varphi\int \max\{J_f[x],V_f\}\,dG_f(x) \tag{5}$$

where it should be clear that the contact rate φ depend on the aggregate market tightness θ.

2.3 Wage Determination, Free Entry Condition and Optimal Decision Rules

As mentioned, the wages are determined by bargaining. A common and tractable way to solve such bargaining problem is to assume the axiomatic Nash bargaining solution.[5] This procedure assigns the wage at a value that maximizes the sum of the surplus of both agents involved in the negotiations: the worker and the firm. The solution is efficient in the sense that a match producing positive surplus will always be realized. It does also avoid disagreement results, i.e. the firm and the worker always agree on the matches that should be realized and on how to share the surplus that results from them.

The workers surplus is the difference between the value if the match is realized $E_f[w_f(x)]$ and the status quo value the agent brings to the negotiation U:

$$E_f[w_f(x)]-U=\frac{w_f(x)+f\times\beta_1-\rho U}{\rho+\eta_f} \tag{6}$$

where the final expression is obtained by substituting expression (3) in equation (6).

The firm surplus is the difference between the value if the match is realized $J_f[x]$ and the value of an unfilled vacancy V_j (the outside option of the firm):

$$J_f[x]-V_j=\frac{x-w_f(x)-ftw_f(x)-(1-f)c-\rho V_f}{\rho+\eta_f} \tag{7}$$

where the final expression is obtained by substituting expression (4) in equation (7).

The free entry of firms posting both formal and informal jobs drives the value of both types of vacancies to zero ($V_f(\nu)=0$), which induces the expected value of a salaried job of type f to be equal to the cost of posting a vacancy, $\kappa+f\times\nu$, multiplied by the expected duration of the

[5] In the empirical search literature see, for example, Eckstein and Wolpin (1995); Flinn (2006); Bobba et al. (2022).

vacancy posting, $1/\varphi$. Thus, we write the free-entry condition as

$$\frac{\kappa + f \times \nu}{\varphi} = \int \max\left\{\frac{x - w_f(x) - ftw_f(x) - (1-f)c}{\rho + \eta_f}, 0\right\} dG_f(x); \quad i = 1, 0 \tag{8}$$

The decision to post formal or informal vacancies is determined as follows: Each firm entering the market receives a value ν, drawn from distribution $F(\nu)$. There exists a cost threshold above which a firm will never post a formal vacancy and below which it will always post a formal vacancy. This threshold is a *reservation cost* ν^*, representing the point of indifference in deciding between posting formal and informal vacancies. This is determined by the free-entry condition for a formal vacancy:

$$\nu^* = \varphi \int \max\left\{\frac{x - (1+t)w_1(x)}{\rho + \eta_1}, 0\right\} dG_1(x) - \kappa \tag{9}$$

Intuitively, a firm with a ν that is at least covered by the expected profit of a filled formal job–calculated as the arrival of a worker times the expected profit per worker in a formal job relative to the base cost of posting the vacancy–will post a formal vacancy. For any firm with $\nu > \nu^*$, the cost of posting a formal vacancy exceeds the expected profit of a filled formal job; thus, it will opt to post an informal vacancy and incur the cost κ, which is exactly offset by the expected profit of a filled informal job, as shown in Equation (8) with $i = 0$. Consequently, the probability of a posted vacancy being formal in the labor market is defined as

$$\pi_1 = \Pr[\nu \leq \nu^*] = F(\nu^*) \tag{10}$$

Using the worker and firm surplus (plus free entry) we obtain two bargaining problems, one for workers and employers in a formal job and one for workers and employers in an informal one:[6]

$$\begin{aligned} w_1(x) &= \text{argmax}[w_1(x) + \beta_1 - \rho U]^{\alpha}[x - (1+t)w_1(x)]^{(1-\alpha)} \qquad (11) \\ &\quad \text{subject to: } w_1(x) \geq w_M \\ w_0(x) &= \text{argmax}[w_0(x) - \rho U]^{\alpha}[x - w_0(x) - c]^{(1-\alpha)} \qquad (12) \end{aligned}$$

where α is a measure of the bargaining power of the worker with respect to the firm. If $\alpha = 0.5$, there is 'symmetric' bargaining; if $\alpha < 0.5$ the firm is in a relative stronger bargaining position; if $\alpha > 0.5$ the worker is in a relative stronger bargaining position. A stronger bargaining position simply means that the agent is able to obtain a larger share of the surplus.

When the job is informal, the solution to this surplus-maximizing problem generates the fol-

[6] In the maximization we ignore the denominators in expressions (6)-(7) since the result is invariant to their inclusion.

lowing convenient one-to-one mapping between wages and match-specific productivity:

$$w_0(x) \quad = \quad \alpha(x-c) + (1-\alpha)\rho U \tag{13}$$

The wage is a combination of the match-specific productivity x and the worker's outside option U. The first element introduce a positive relationship between productivity and wages. The second element introduce a direct link between the wage and the worker's search behavior and outside option: the higher the worker's outside option, the larger the share of the surplus going to the worker and the higher the wage. Match-specific productivity x and the worker's outside option U are weighted in the wage schedule by the parameter α: the higher α, the higher the relative bargaining position of the worker, the more weight is given to productivity in determining the wage. Finally, the wage equation includes the relevant institutional parameters: the non-contributory benefit β_0 and the cost of being found and punished c. Since the first is a benefit to the worker, the higher its value the lower the wage the worker will accept at same productivity x. The second is a cost to the firm but it is also reducing the overall surplus, therefore the higher the cost of being found and punished the lower the wage at same productivity x. The expression denotes a crucial feature of bargaining: even if the cost is paid by the firm, some of it can be transferred to the worker as lower wage, the more so the larger the portion of the surplus going to the worker.

Now that the wage schedule is well-defined, it is straightforward to obtain the optimal decision rules to the dynamic programming problem stated in Section 2.2. The wage schedule (13) shows that wages are monotone increasing in the productivity x. The value function (3) shows that, for the worker, the value of accepting a job is increasing in wages. The value function (1) shows, instead, that the value of unemployment is constant in wages. As a result, there exists a wage level above which the worker will always accept the job and below which he will always reject the job and continue searching. This wage level defines a *reservation wage* w_0^* – the point of indifference in the decision – that uniquely determine a *reservation productivity* x_0^*. Formally, they are defined as:

$$E_0[w_0(x_0^*)] = U \Leftrightarrow w_0^* = \rho U \Leftrightarrow x_0^* = \rho U + c \tag{14}$$

where the analytical expression are obtained by using the definition of the value function (3) and the definition of the wage schedule (13). In conclusion, the optimal decision rules state that when a worker and a firm meet to decide if accepting or rejecting an informal job, they will always accept if the match productivity value x is larger than $x_0^* = \rho U + c$. The economics of this expression is intuitive: the larger the outside option, the larger the productivity required to induce the worker to accept the match (the worker is "pickier"), and the larger the cost of being found and punished, the pickier the worker.

When the job is formal, the solution to the surplus-maximizing problem is more complex but the logic is similar.[7] The added complexity is due to the presence of the mandatory minimum wage w_M, which creates a price floor constraint. Consider again the maximization problem stated in (11). The interior solution delivers a wage schedule very similar to the one we have found for the informal wage:

$$w_1(x) \quad = \quad \alpha \frac{x}{1+t} + (1-\alpha)[\rho U - \beta_1] \tag{15}$$

It is the same combination of match-specific productivity x and outside option U but the institutional parameters. First, the wage is affected by β_1. Second, and more relevant, the payroll contribution plays a role, lowering the wage for given x value. Again the bargaining structure shows that, even if the payroll tax is withdrawn at the source and paid by the firm, the employer can transfer some of the cost to the worker: The higher the payroll contribution, the lower the wage for given match-specific productivity value x. Using the same logic reported in the equivalences (14), there will exist a reservation wages and a reservation productivity value such that the worker and the firm are indifferent between accepting or rejecting the match:

$$E_1[w_1(x_1^*)] = U \Leftrightarrow w_1^* = \rho U - \beta_1 \Leftrightarrow x_1^* = (1+t)[\rho U - \beta_1] \tag{16}$$

Again the reservation values have an intuitive interpretation: as in the informality case, they are increasing in the outside option and decreasing in the benefit. Differently from the informality case, they are positively affected by the payroll contribution: the higher the contribution, the pickier the worker will be to accept a match.

This interior solution is the only solution if the mandatory minimum wage is *non-binding*, i.e. it does not affect agents' decisions in equilibrium. Within our model, it is possible to formally define what a binding or non-binding mandatory minimum wage is:

Definition 1 (Binding Minimum Wage) *A mandatory minimum wage w_M is binding if and only if $w_M > w_1^*$.*

If the minimum wage is binding, a corner solution is present in addition to the interior solution just described. The corner solution arises for those productivity values that are still able to generate a surplus but that map into a bargaining wage lower than the mandatory minimum wage. In this range of values, agents have an incentive to deviate from the wage schedule (15) and conclude a match that pays *exactly* the minimum wage. To see this result formally, recall that (15) implies that formal wages are monotone increasing in productivity. Since the minimum wage is binding, it does exist a productivity value $\overline{x}_M$ that the bargaining wage schedule predicts

[7] See Flinn (2006, 2011) for seminal contributions introducing mandatory minimum wages in labor market characterized by search frictions.

should pay exactly the minimum wage:

$$\begin{aligned} \overline{x}_M &: \quad w_1(\overline{x}_M) = w_M \\ &\Leftrightarrow \\ \overline{x}_M &= \frac{(1+t)}{\alpha}\{w_M - (1-\alpha)[\rho U - \beta_1]\} \end{aligned} \tag{17}$$

If $x = \overline{x}_M$, the match is realized and the wage obeys the usual wage schedule which *happens* to dictate a value exactly equal to the minimum wage. Now consider an x' just slightly lower than this value: $x' = \overline{x}_M - \epsilon$ for an arbitrary small and positive ϵ. We know the firm would like to pay a slightly lower wage and that the worker would accept it. We also know that a match at x' would generate a positive surplus to share between the parts. But the agents cannot agree on a lower wage because it would be lower than the mandated w_M. However, they can agree on deviating from the wage schedule and pay the mandatory minimum wage imposed by the labor legislation. The worker will extract a larger share of the surplus from the match but the firm would still receive a positive surplus which is better than leaving the vacancy unfilled. This process will be optimal for all the range of productivities that generate positive surplus for both agents. The lower bound of such a range is the value $\underline{x}_M$ that guarantees exactly zero flow profits to the firm:

$$\begin{aligned} \underline{x}_M &: \quad \underline{x}_M - (1+t)w_M = 0 \\ &\Leftrightarrow \\ \underline{x}_M &= (1+t)w_M \end{aligned} \tag{18}$$

We can now fully characterize the optimal decision rules for formal jobs in the presence of a binding mandatory minimum wage. When $x \geq \overline{x}_M$, the match is realized and the wage obeys the bargained wage schedule (15). When $\underline{x}_M \leq x < \overline{x}_M$, the match is realized and the wage is exactly equal to the minimum wage w_M. When $x < \underline{x}_M$, the match is rejected and the agents continue to search. The full wage schedule for formal salaried employment is therefore:

$$w_1(x) = \begin{cases} w_M & \text{if } \underline{x}_M \leq x < \overline{x}_M \\ \alpha\frac{x}{1+t} + (1-\alpha)[\rho U - \beta_1] & \text{otherwise} \end{cases} \tag{19}$$

The last optimal decision rule we have to discuss concerns the self-employment decision. The reasoning is analogous but simpler. No bargaining or minimum wage is involved because self-employed jobs directly provide income y and is not subject to any wage floor. Equation (2) shows that the value of a self-employed job is increasing in monetary net value that the job provides y while the value of unemployment is constant in y. Therefore, it does exist a *reservation income* y^* that guarantees indifference between the two states:

$$S[y^*] = U \Leftrightarrow y^* = \rho U - \beta_s \tag{20}$$

where the analytical expression is obtained by using the definition of the value function (2). The optimal decision rules state that when a worker meets a self-employment opportunity that pays more than y^*, he will take it; he will reject otherwise.

2.4 Equilibrium

To complete the definition of equilibrium, notice that the optimal decision rules solve the maximization problems embedded in the dynamic programming problem, as specified by the values (1)–(5), and that the optimal decision rules are entirely characterized by the reservation values, which are functions of known parameters and U. The value of unemployment U can be determined by solving the fixed point defined by (1), given the market tightness θ and the proportion of formal vacancies π_1. Additionally, market tightness must satisfy the free entry condition (8) for informal vacancies, while π_1 should satisfy (10) given the reservation cost in (9). Ultimately, both must be consistent with the fixed point of U. Therefore, we can conclude the following:

Definition 2 (Equilibrium) *Given:*

1. *The primitive parameters* $\{\rho, \eta_f, \gamma, \delta, \alpha, \kappa\}, f = 0, 1$ *and the matching function* $Q(u, v)$*;*
2. *The distributions of wage offers* $G_f(x), f = 0, 1,$ *of self-employment income opportunities* $G(y)$*, and of the additional cost of posting formal vacancies* $F(\nu)$*;*
3. *The institutional parameters* $\{t, w_M, c\}$ *and utility parameters* $\{\beta_u, \beta_1, \beta_s\}$*;*

The steady state equilibrium is defined as a value U *that satisfies the Bellman equation* (1)*, alongside the market tightness* θ *that satisfies the free entry condition* (8) *for informal vacancies and the proportion of formal vacancies* π_1*, which is determined by the free entry condition* (8) *for formal vacancies. All of the above, together with the invariant distribution of individuals across various labor market states, are considered.*

2.5 Empirical Implications

As discussed in Section 1, an important task of the model is the ability to generate five crucial empirical facts.

First, the equilibrium should allow for a significant proportion of ex-ante identical workers being employed in jobs with different formality regimes. The model allows for it because unemployed workers may find it optimal to choose any of the three employment states (formal salaried employee, formal salaried employee, self-employed), depending on the value of x or the value of y they meet when matching with, respectively, employers or self-employment opportunities.

Second, the model should generate not only transitions from informal to formal jobs but also from formal to informal jobs. The labor market history of any given workers in the model can exhibit both types of transitions because every time a job ends, the worker goes back to unemployment and search. At that stage, as stated before, any job transition is possible, depending on which offers arrive first and what productivity values they carry.

Third, the accepted wage distributions for formal and informal employees should report higher average wages for formal workers than informal workers but also allow for a large overlap between the two wage distributions. Both predictions are possible in the model for appropriate values of the $G_f(x)$ distributions. For example, a $G_1(x)$ distribution that first-order stochastically dominates the $G_0(x)$ distribution will deliver the result, where the overlap is assured by enough variation in the distributions and by the plausible constraint that the highest acceptable informal wage is higher than the mandatory minimum wage w_M.

Fourth, the formal accepted wage distribution should exhibit a spike at exactly the minimum wage. As explained in Section 2.3, this occurs in the model because all the meetings with match value x such that $\underline{x}_M \leq x < \overline{x}_M$ will be accepted and will pay exactly w_M.

Fifth, wages of informal jobs should also depend on the mandatory minimum wage. The result is clear when looking at wage schedule (13). The schedule represents the wage paid to an informal salaried worker and it is a function of the parameters α, c, ρ; the match-specific productivity x; and the value of unemployment U. This last component shows up in the wage schedule as a result of bargaining since it corresponds to the outside option of the worker. But being the outside option of the worker defined in (1), it does also represents the present discounted value of participating in the labor market defined by all the parameters of the model, including the mandatory minimum wage w_M. As a result, any changes in a binding mandatory minimum wage will affect the value of U and, in turn, the value of $w_0(x)$ paid to salaried informal workers. Intuitively: a worker bargaining with an employer for an informal job knows that he could go back and search and find a formal jobs subject to the minimum wage constraint. Therefore, different levels of minimum wage impact his bargaining power with the employer.

3 Estimation

We estimate the model using the maximum likelihood method, exploiting the empirical implications of the model in terms of the distributions of the data. The theoretical framework provides specific predictions about the joint distribution of labor market states, durations in each state, and wage/earnings outcomes. By mapping these theoretical distributions to the observed data, we can recover the structural parameters that govern worker behavior and labor market dynamics.

3.1 Data

The data requirements for the empirical implementation assume access to a representative sample of N individual workers, indexed by $j = 1, 2, 3...N$. For each worker in the sample, three types of information are required to estimate the model parameters. First, we need data on labor market states, captured by indicator functions $\{I[j \in u], I[j \in e_1], I[j \in e_0], I[j \in s]\}$, where u, e_1, e_0, and s represent unemployment, formal employment, informal employment, and self-employment, respectively. Second, duration data are also necessary, denoted by variables $\{t_{j,u}, t_{j,e_1}, t_{j,e_0}, t_{j,s}\}$, where $t_{j,i}$ represents the complete or ongoing duration in months for worker j in labor market state i. Finally, labor income information is also required and is captured through variables $\{w_j, y_j\}$, where w_j denotes the observed wage for worker j in salaried employment (either formal or informal) and y_j represents the net earnings for worker j in self-employment. Together, these data elements provide the necessary information to identify the model's structural parameters through the estimation procedure outlined below.

We complement our data with demand-side information; specifically, we use job posting data from vacancies indexed in Google Jobs for Colombia and Mexico. This database is compiled by the *Labor Observatory* at the Inter-American Development Bank (IDB).[8] The vacancy data are categorized, among other dimensions, by the Degree of Qualifications Required for the Position. Since our sample consists of young unskilled workers, we focus solely on job postings that demand little or no skill to perform the job.

3.2 Supply Side Parameters: Likelihood Estimation

Given the model presented in Section 2, it is possible to write the contribution to the likelihood of the information contained in the data set described in Section 3.1. The likelihood function captures the probability of observing the data given the model parameters. We construct it by considering the contribution of each labor market state separately, accounting for both duration and income information where applicable. The key idea is that the model's structural equations determine the distribution of observable outcomes in each state.

3.2.1 Unemployment Duration Contribution

To write the contribution of unemployed individuals, we begin by defining the total hazard rate out of unemployment, which represents the rate at which unemployed workers transition to any employment state. This hazard rate is denoted by ζ_u and is the sum of transition rates to each employment state:

$$\zeta_u = \zeta_{u\to e_1} + \zeta_{u\to e_0} + \zeta_{u\to s}$$

The individual components of the last equation represent the transition rates from unemployment to formal salaried employment $\zeta_{u\to e_1} = \lambda_1 \tilde{G}_1(\max\{x_1^*, \underline{x}_M\})$, to informal salaried employment $\zeta_{u\to e_0} = \lambda_0 \tilde{G}_0(x_0^*)$, and to self-employment $(\zeta_{u\to s} = \gamma \tilde{R}(y^*))$, where $\tilde{G}(x) = 1 - G(x)$ and $\tilde{R}(y) = 1 - R(y)$.

Using the reservation values from the worker's optimization problem, we can express the hazard

[8] See the website of the project for details: `https://observatoriolaboral.iadb.org/en/vacantes/`

rate out of unemployment as:

$$\zeta_u = \lambda_1 \tilde{G}_1((1+t)\max\{[\rho U - \beta_1], w_M\}) + \lambda_0 \tilde{G}_0(\rho U + c) + \gamma \tilde{R}(\rho U - \beta_s)$$

This expression incorporates the worker's optimal decision rules through the reservation values. The first term reflects transitions to formal employment, where the reservation productivity depends on whether the minimum wage is binding. The second term captures transitions to informal employment given the cost of being caught working informally, and the third term represents transitions to self-employment.

Given that, conditional on the structure of the model, the hazard rate ζ_u is constant (i.e. independent of duration), unemployment durations follow an exponential distribution. Additionally, the unemployment duration data is only observed for those that are currently unemployed. Therefore the likelihood contribution for an unemployed worker with duration $t_{j,u}$ is:

$$f_{j,u}(t_{j,u}, j \in u) = \zeta_u \exp(-\zeta_u t_{j,u}) \Pr[j \in u] \tag{21}$$

Equation (21) represents the joint probability of observing the duration and that the current state is unemployment.

3.2.2 Employment Contributions

Workers working as employees provide information through both their durations in the job and their wages. We model these contributions separately and then combine them in the employment likelihood contribution.

Duration Component: For salaried employment of type f (where $f = 1$ for formal and $f = 0$ for informal), the hazard rate out of employment is simply the exogenous termination rate:

$$\zeta_{e_f} = \eta_f, \quad f = 1, 0$$

Since job termination is exogenous and independent of duration, employment durations also follow exponential distributions with parameters η_f.

Wage Component for Informal Workers: The wage contribution requires mapping from the underlying productivity distribution to the observed wage distribution, accounting for the selection process inherent in the worker's optimal decisions. For informal workers, we have to consider that only wages above the reservation wage are accepted. The conditional wage density for informal workers, given that the wage exceeds the reservation value, is:

$$f_{j,e_0}(w_j | w_j \geq \rho U, j \in e_0) = \frac{1}{\alpha} \frac{g_0\left(\frac{w_j + \alpha c - (1-\alpha)\rho U}{\alpha}\right)}{\tilde{G}_0(\rho U + c)}$$

This expression maps the productivity density $g_0(\cdot)$ to the wage space using the wage equation, while the denominator $\tilde{G}_0(\rho U + c)$ normalizes for the truncation at the reservation productivity level.

Combining duration and wage information, the complete likelihood contribution for an informal employee is:

$$f_{j,e_0}(w_j, w_j \geq \rho U, t_{j,e_0}, j \in e_0) = f_{j,e_0}(w_j | w_j \geq \rho U)\eta_0 \exp(-\eta_0 t_{j,e_0}) \Pr[j \in e_0] \tag{22}$$

where again this expression represents the joint probability of observing the duration and the wage for those currently in the informal employment state.

Wage Component for Formal Workers: Formal employment presents additional complexity due to the minimum wage constraint. This creates a discontinuity in the wage distribution: some workers earn exactly the minimum wage, while others earn above it. We must account for both cases in the likelihood.

First, we derive the probability that a formal worker earns exactly the minimum wage. This occurs when the worker's productivity would imply a wage below the minimum, but the minimum wage constraint binds:

$$\begin{aligned}\Pr[w_j = w_M | e_1] &= \frac{\tilde{G}_1(\underline{x}_M) - \tilde{G}_1(\bar{x}_M)}{\tilde{G}_1(\underline{x}_M)} \\ &= \frac{\tilde{G}_1((1+t)w_M) - \tilde{G}_1\left(\frac{(1+t)}{\alpha}\{w_M + (1-\alpha)[\rho U - \beta_1]\}\right)}{\tilde{G}_1((1+t)w_M)}\end{aligned}$$

The numerator of the last expression captures the mass of workers whose productivity falls in the range where the minimum wage is paid, while the denominator normalizes by the total probability of accepting formal employment.

The joint likelihood contribution for workers earning exactly the minimum wage accounts for the possibility that the minimum wage may not be binding:

$$\Pr[w_j = w_M, t_{j,e_1}, j \in e_1] = \begin{cases} 0 & \text{if } w_M < \rho U - \beta_1 \\ \Pr[w_j = w_M | e_1]\eta_1 \exp(-\eta_1 t_{j,e_1}) \Pr[j \in e_1] & \text{otherwise} \end{cases}$$

For workers earning above the minimum wage, we need the conditional wage density truncated at the effective lower bound. This bound is either the reservation wage when the minimum wage is not binding or the minimum wage when it is. Therefore:

$$f_{j,e_1}(w_j | w_j > \max\{[\rho U - \beta_1], w_M\}, j \in e_1) = \frac{1+t}{\alpha} \frac{g_1\left(\frac{(1+t)(w_j - (1-\alpha)[\rho U - \beta_1])}{\alpha}\right)}{\tilde{G}_1((1+t)\max\{[\rho U - \beta_1], w_M\})}$$

This can be expressed more precisely as a piecewise function that depends on whether the minimum wage binds:

$$f_{j,e_1}(w_j|w_j > \max\{\rho U - \beta_1, w_M\}, j \in e_1) = \begin{cases} \frac{1+t}{\alpha} \frac{g_1\left(\frac{(1+t)(w_j-(1-\alpha)[\rho U-\beta_1])}{\alpha}\right)}{\tilde{G}_1((1+t)[\rho U-\beta_1])} & \text{if } w_M < \rho U - \beta_1 \\ \frac{1+t}{\alpha} \frac{g_1\left(\frac{(1+t)(w_j-(1-\alpha)[\rho U-\beta_1])}{\alpha}\right)}{\tilde{G}_1((1+t)w_M)} & \text{otherwise} \end{cases}$$

The complete likelihood contribution for formal workers earning above the minimum wage is:

$$\begin{aligned} & f_{j,e_1}(w_j, w_j > \max\{[\rho U - \beta_1], w_M\}, t_{j,e_1}, j \in e_1) \\ & = f_{j,e_1}(w_j|w_j > \max\{[\rho U - \beta_1], w_M\}, j \in e_1) \times \eta_1 \exp(-\eta_1 t_{j,e_1}) \Pr[j \in e_1] \end{aligned} \tag{23}$$

where, as before, this expression represents the joint probability of observing the duration and wage for workers currently in formal employment.

3.2.3 Self-Employment Contribution

Self-employed workers provide information through both their duration in self-employment and their net earnings. Unlike salaried workers, there is no wage equation linking productivity to earnings for the self-employed, which simplifies the analysis considerably.

The hazard rate out of self-employment is simply the exogenous termination rate:

$$\zeta_s = \delta$$

The earnings component reflects the endogenous selection process in which workers enter self-employment only if their earnings potential exceeds their reservation value. The conditional earnings density is:

$$f_{j,s}(y_j|y_j \geq \rho U - \beta_s, j \in s) = \frac{r(y_j)}{\tilde{R}(\rho U - \beta_s)}$$

where $r(y_j)$ is the density of net earnings and the denominator accounts for the truncation at the reservation level.

The complete likelihood contribution for self-employed workers is:

$$f_{j,s}(y_j, y_j \geq \rho U - \beta_s, t_{j,s}, j \in s) = f_{j,s}(y_j|y_j \geq \rho U - \beta_s, j \in s)\zeta_s \exp(-\zeta_s t_{j,s}) \Pr[j \in s] \tag{24}$$

Analogous to the previous cases, this expression gives the joint probability of observing the duration and earnings for workers in self-employment.

3.2.4 Steady-State Labor Market Probabilities

The likelihood contributions in equations (21) to (24) require the endogenous probabilities of observing individuals in each labor market state. These probabilities come from the steady-state equilibrium of the model in which flows into and out of each state must balance.

As stated before, in steady state the flows into and out of each state must be equal, which conditional on the model gives us the following system of equations:

$$
\begin{aligned}
\lambda_1 \tilde{G}_1((1+t)\max\{[\rho U - \beta_1], w_M\}) \cdot u &= \eta_1 e_1 \\
\lambda_0 \tilde{G}_0(\rho U + c) \cdot u &= \eta_0 e_0 \\
\gamma \tilde{R}(\rho U - \beta_s) \cdot u &= \delta s \\
u + e_1 + e_0 + s &= 1
\end{aligned}
$$

The first three equations represent flow balance: the left-hand side captures the flow from unemployment into each employment state, while the right-hand side represents the flow out of each employment state back to unemployment. The final equation is a normalization constraint ensuring that the state probabilities sum to one (which is equivalent to normalizing the total labor force to 1). Solving this linear system yields the steady-state probabilities: $\Pr[u] = u$, $\Pr[e_1] = e_1$, $\Pr[e_0] = e_0$, and $\Pr[s] = s$.

3.2.5 Complete Likelihood Function

Combining all individual contributions from equations (21) to (24), the complete log-likelihood function is:

$$
\begin{aligned}
\ln L(\Theta; t, w, y) = &\sum_{j \in u} \ln f_{j,u}(t_{j,u}, j \in u) + \sum_{j \in e_1} I[w_j = w_M] \ln \Pr[w_j = w_M, t_{j,e_1}, j \in e_1] \\
&+ \sum_{j \in e_1} I[w_j > w_M] \ln f_{j,e_1}(w_j, w_j > \max\{[\rho U - \beta_1], w_M\}, t_{j,e_1}, j \in e_1) \\
&+ \sum_{j \in e_0} \ln f_{j,e_0}(w_j, w_j \geq \rho U, t_{j,e_0}, j \in e_0) \\
&+ \sum_{j \in s} \ln f_{j,s}(y_j, y_j \geq \rho U - \beta_s, t_{j,s}, j \in s)
\end{aligned} \tag{25}
$$

In the log-likelihood function above, the parameter space of the supply side of the model is defined as $\Theta^S = \{\rho U, \lambda_1, \lambda_0, \gamma, \eta_1, \eta_0, \delta, \mu_1, \sigma_1, \mu_0, \sigma_0, \mu_y, \sigma_y, \beta_1, \beta_s, c\}$, assuming that $G_1(x)$, $G_0(x)$, and $R(y)$ are log-normal distributions with parameters (μ_1, σ_1), (μ_0, σ_0), and (μ_y, σ_y), respectively.

It is important to note that ρU, λ_1, and λ_0 are treated as parameters in the likelihood estimation, despite being endogenous in the full equilibrium model. This reparameterization is convenient because it facilitates estimation and can be used because the link between the supply and demand sides of the model is completely characterized by $\lambda_1 = \pi_1 \lambda$ and $\lambda_0 = (1 - \pi_1)\lambda$, as described below. In addition, U can be uniquely mapped to β_u using the value function in Equation (1).

Real-world data often contains measurement error, particularly for income variables, which we incorporate by modeling observed wages and earnings as the product of true values and a multiplicative error term. Let observed wages be related to true wages through $w^o = w \times \epsilon$, where ϵ is the measurement error and this specification implies that $\epsilon = \frac{w^o}{w}$.

The measurement error process is governed by three key assumptions that ensure a tractable estimation while maintaining economic realism. First, we assume state independence, meaning that the measurement error does not vary systematically across labor market states, ensuring that any bias in wage reporting is consistent regardless of whether workers are in formal employment, informal employment, or self-employment. Second, we assume that the error term follows a log-normal distribution such that $\epsilon \sim \text{LogNormal}(\mu_\epsilon, \sigma_\epsilon^2)$, which provides flexibility while maintaining the positive support required for multiplicative errors. Third, we impose unbiased measurement, requiring that the conditional expectation of observed wages equals true wages, formally $E[w^o|w] = w$, which implies $E[\epsilon|w] = 1$. This third assumption places a restriction on the parameters of the error distribution because, for a log-normal random variable, the mean is $\exp(\mu_\epsilon + \sigma_\epsilon^2/2)$, and setting this equal to one gives us $\exp(\mu_\epsilon + \sigma_\epsilon^2/2) = 1$. Taking logarithms and solving for μ_ϵ yields $\mu_\epsilon = -\frac{\sigma_\epsilon^2}{2}$, which means that we only need to estimate one measurement error parameter, as we can express $\mu_\epsilon = -0.5\sigma_\epsilon^2$.

The likelihood contributions of wages and net earnings must be modified to account for the measurement error, such that for each state $i \in \{e_1, e_0, s\}$, the observed likelihood becomes:

$$f_{j,i}^o = \int \frac{1}{w} v\left(\frac{w_j^o}{w}\right) f_{j,i}(w|\cdot)dw$$

where $v(\cdot)$ is the density function of the log-normal measurement error, and $f_{j,i}(w|\cdot)$ is the theoretical density of true wages/earnings conditional on the relevant truncation. This integral represents the probability of observing wage w_j^o given the distribution of true wages and the measurement-error process. This completes the specification of the likelihood function required for the maximum likelihood estimation of the theoretical model.

3.3 Demand Side Parameters

From the estimation of the supply side of the model, we obtain the arrival rates of jobs, λ_1 and λ_0. These estimates, combined with information on job postings, are sufficient to recover the distribution of formal vacancies π_1 and a matching function characterized by one parameter. Specifically, we have $\lambda_1 = \pi_1\lambda$ and $\lambda_0 = (1-\pi_1)\lambda$ and therefore the proportion of formal vacancies is

$$\pi_1 = \frac{\lambda_1}{\lambda_1 + \lambda_0} \tag{26}$$

Furthermore, assuming a one-parameter Cobb-Douglas matching function $Q(v, u) = v^\phi u^{1-\phi}$, with ϕ representing the matching elasticity, the aggregated rate at which jobs arrive can be written as $\lambda = \theta^\phi$. Here $\theta = \frac{v}{u}$ is the market tightness, thereby the matching elasticity can be recovered using the estimated parameters λ_1 and π_1 as

$$\phi = \frac{\log \lambda_1 - \log \pi_1}{\log v - \log u} \tag{27}$$

where u is the unemployment rate in the population of young unskilled workers and v is the number of job posts for the skill level considered here.

The next parameter is the cost of posting vacancies κ. With the matching function defined earlier, the rate at which firms meet workers is $\varphi = \theta^{\phi-1}$. Therefore, the free-entry condition for informal employment, together with the estimated parameters of the supply side (ρU, μ_0, σ_0) and demand side (ϕ), can be used to recover the cost of posting vacancies:

$$\kappa = \theta^{\phi-1}(1-\alpha) \int_{\rho U + c} \left[\frac{x - c - \rho U}{\rho + \eta_0} \right] dG_0(x) \tag{28}$$

where in Equation (28), we have used the definitions of the wage equation and reservation productivity.

To complete the set of parameters on the demand side of the model, we need to recover the distribution of the additional cost of posting vacancies faced by firms that want to post a formal vacancy $F(\nu)$. Given that we do not have any information related to the posting process, we simply use the structure of the model to recover the distribution, which makes it possible to identify a distribution characterized by one parameter. Specifically, we assume that ν follows an exponential distribution with the parameter ψ, such that $F(\nu) = 1 + \exp(-\psi\nu)$. Given equation (10) and π_1 pinned down as described above, the parameter ψ can be recovered as

$$\psi = \frac{-\log(\pi_1 - 1)}{\nu^*} \tag{29}$$

Overall, the parameter space on the demand side of the model is $\Theta^D = \{\phi, \kappa, \psi\}$.

4 Preliminary Results on Colombia and Mexico

4.1 Estimation Sample

We apply the estimation strategy described above for two countries with high informality: Colombia and Mexico. For the case of Colombia we use the *Gran Encuesta Integrada de Hogares (GEIH) 2023*, while for the case of Mexico we use the *Encuesta Nacional de Ocupación y Empleo (ENOE) 2023*. In this section, we describe the definition of the sample for both countries that allows us to identify the parameters of the model and, therefore, to analyze the minimum wage effects on labor market outcomes.

The sample is constructed with specific restrictions to ensure consistency with the assumptions of the model while ensuring that we capture a relevant part of the population for which the minimum wage is an issue. Therefore, we focus exclusively on full-time workers who work 48 hours per week, examining all household members rather than limiting the analysis to household heads. To maintain some degree of homogeneity in our sample, we restrict our attention to male workers between 18 and 30 years old who have completed at most a high school education. As stated before, these restrictions allow us to analyze a relatively homogeneous group, where minimum wage effects are most likely to be observable and significant.

A critical component of our methodology is the precise identification of jobs that earn the minimum wage. Our approach begins by examining the distribution of monthly nominal wages for full-time workers, where we observe a clear spike at the minimum wage.

In Colombia, the monthly minimum wage is 1,160,000 pesos. The spike at the minimum wage represents a substantial 44% of our sample, corresponding to 3,249 observations, providing strong evidence of minimum wage compliance and binding constraints. Given the tendency of survey respondents to round their wage reports, we carefully examine the distribution around the minimum wage spike to identify any clustering patterns. While there is no significant concentration of wages in round numbers immediately surrounding the minimum wage, we do observe notable spikes at 1,100,000 pesos (1.2% of the sample, 89 observations) and 1,200,000 pesos (8.3% of the sample, 613 observations) wages. This pattern is consistent with the general rounding behavior observed throughout the wage distribution, where round numbers tend to attract responses regardless of their proximity to the minimum wage. All other wages below the minimum wage (except for these spikes), which represent 6.8% of our sample (501 observations), are left to be captured by the measurement error (assuming compliance). The complete distribution of observed wages (for formal and informal employees) and self-employment income are presented in Figure 1a.

For the case of Mexico, there are almost no observations at the minimum wage (only three observations or 0.01%). This could be explained by two factors: first, the minimum wage is daily, and workers have to make monthly conversions at the time of the survey; second, there are multiple minimum wages by type of occupation, creating potential confusion about the applicable rate. While there is no significant concentration of wages in round numbers immediately surrounding the minimum wage, we do observe notable spikes at 6,000 pesos (4.5% of the sample, 106 observations), 6,020 pesos (3.13% of the sample, 72 observations), 6,450 pesos (8.5% of the sample, 198 observations), and 6,880 pesos (4.5% of the sample, 105 observations). This pattern is consistent with the general rounding behavior and potential miscalculations arising from the daily wage structure and occupational variations in minimum wage levels. All other wages below 6,000 pesos, which represent 12.7% of our sample (297 observations), are left to be captured by the measurement error under the assumption of minimum-wage compliance. The complete distribution of observed wages for formal and informal employees and self-employment income is shown in Figure 1b.

For our empirical analysis, we convert all monetary variables to hourly wages expressed in 2023 US dollars to facilitate interpretation and comparison across countries. This conversion employs the average exchange rate from December 2023, which stood at 3,952.569 pesos per US dollar for

Colombia and 17.186 pesos per US dollar for Mexico.

Our estimation strategy also involves calibrating a subset of model parameters, a decision driven by both empirical identification challenges and data limitations.

Two key macroeconomic parameters are calibrated based on institutional knowledge of the Colombian and Mexican economies. The discount rate ρ cannot be separately identified within our model framework, while the payroll contribution rate t can be observed directly from public finance data. Following OECD (2023), we set $t = 0.16$ for Colombia and $t = 0.33$ for Mexico. The discount rate is set at $\rho = 0.10$ for both Colombia and Mexico, following the recommendations for Latin American economies provided by multilateral development banks as documented by Moore et al. (2020).

The cost of informality parameter c is calibrated at 10% of the informal average wage observed in the samples for both countries. This calibration draws on the estimation results from Bobba et al. (2022), providing a theoretically grounded starting point for this parameter.

Finally, our approach incorporates the standard assumptions from the search and matching literature regarding bargaining power. We impose the common assumption of symmetric bargaining power, setting the parameter α to 0.5 for both countries. While this assumption is standard in the literature, the lack of access to firm-side data does not allow us to identify this parameter in the estimation.

4.2 Estimation Results

Table 1 reports the point estimates of the structural parameters of the model for Colombia and Mexico. The top panel reports maximum likelihood estimates while the bottom panel reports calibrated values. The arrival rate estimates imply a higher arrival rate of offers for employee jobs (λ_1 and λ_0) than for self-employment opportunities (γ) in both countries, though this difference is less pronounced in Colombia than in Mexico. The formality composition of employee offers varies between 61 and 66% in the two countries.

The distributions of wage offers and self-employment opportunities confirm higher average wages for formal employees in both countries, but they also exhibit substantial overlap between formal and informal wage offers and between wage offers and (net) labor income in self-employment across both labor markets. These results are a direct consequence of what is observed in Figure 1.

Termination rates η_1 and η_0 confirm the higher instability of informal employee jobs compared to formal ones in Colombia, while in Mexico the termination rates are similar between formal and informal employment; self-employment jobs are estimated to be more stable in both Colombia and Mexico (see δ).

The flow cost of posting employee vacancies κ differs considerably between countries: in Colombia, to post a vacancy costs around half of what it costs in Mexico. In turn, the estimation of the distribution of the additional cost of posting a formal vacancy shows that this cost is also higher in Mexico than in Colombia (on average 60% more). These results suggest different regulatory and institutional environments affecting vacancy creation costs. The matching elasticity ϕ is similar across both countries: around 0.4.

Table 2 reports the model fit, which is extremely good along all dimensions and for both countries.

4.3 Policy Experiments

Figures 2 through 6 report the results of varying the level of the mandatory minimum wage from 50% lower than the 2023 value to 50% higher for both countries. For each minimum wage value, we recompute the equilibrium and simulate labor market careers under the new equilibrium and minimum wage regimes. We present two versions of these results. On the one hand, the equilibrium for each value of the minimum wage includes the endogenous adjustments made by all agents in the economy, including the firms' different posting behavior. On the other hand, we fix the posting behavior of the firms at the estimated values (i.e., we fix the arrival rates λ_1 and λ_0) and include only the endogenous responses that workers make in choosing their occupations.

Figure 2 reports the impact on the endogenous proportion of workers in the four labor market states for Colombia and Mexico. For Colombia, the changes are asymmetric around the current value (denoted by the vertical dashed line), with the proportions of formal and informal employment being more elastic when the minimum wage increases than when it decreases. The same asymmetric behavior is observed for the proportion of informal self-employment and unemployment, although it is considerably less pronounced than for formal employment. The proportions in all other labor market states increase when the minimum wage increases, suggesting some degree of substitution from formal employment toward informal employment, self-employment, and unemployment. The substitution for informal employment is particularly relevant; it almost mirrors the reduction in formal employment. When we fix the arrival rates, formal employment is considerably less elastic and the substitution is almost evenly distributed across informal employment, self-employment, and unemployment.

Mexico exhibits a similar pattern and comparable elasticities in the behavior of formal and informal employment in Colombia. However, notable differences exist in the distribution of labor market states between the two countries. In Colombia, informal employment (19%) is lower than both informal self-employment (26%) and unemployment (21%), whereas in Mexico, informal employment is significantly higher (28%) than both informal self-employment (5%) and unemployment (15%). Additionally, formal employment is considerably more prevalent in Mexico (50%) than Colombia (34%). Informal self-employment shows minimal responsiveness to increases in minimum wage. At very high minimum wage levels, formal employment nearly disappears in Colombia, while it remains above approximately 10% in Mexico, indicating that the minimum wage imposes different binding constraints on the two labor markets. As Figure 3 shows, in the benchmark Colombia has almost 50% of formal employment earning the minimum wage, while Mexico has just above 20%. Moreover, for the highest values of minimum wage in the experiment, Colombia has almost all its formal employment earning the minimum wage.

Figure 4 report the impact on employee wages and self-employment incomes for Colombia and Mexico. Again, the most pronounced effects occur in the formal employee sector, where average wages increase substantially with minimum wage increases in both countries. This impact stems from three distinct sources: the direct impact on low wages (workers previously earning below the new minimum), the equilibrium impact on wages above the minimum wage (through general equilibrium effects on the wage distribution), and the composition effect induced by changes in the reservation productivity value required to work formally (altering the mix of workers in formal employment). In contrast, average wages of informal employees drop with high values of the minimum wage in both countries. For Colombia, this decline is more pronounced. The reduction in wages is explained by the composition effect induced by changes in the reservation productivity value, as higher minimum wages encourage more productive workers to seek formal employment, leaving a pool of lower-productivity workers in the informal sector and thereby reducing average informal wages (see Figure 5). Self-employment income does not respond to minimum wage changes in either country. When we fix the arrival rates, the average wages of informal employees actually increase with high minimum wage values in both countries, although the effect is very small. This suggests that the equilibrium effect on the demand side, that is, the posting process, is important in explaining the composition effect.

Figure 6 shows how the demand side is adjusted in Colombia and Mexico. This figure presents the arrival rate of jobs to workers, composition of vacancies, and meeting rates at which firms meet workers. All the cases in these figures refer to the results with firms endogenously adjusting their posting behavior. The results are qualitatively similar between countries, although the magnitudes differ. The proportion of posted formal employee vacancies fall with the minimum wage, but the overall number of vacancies posted does not, increasing market tightness. The reduction in the expected value of formal jobs makes it less attractive to post this type of vacancy and more attractive to post informal ones. Consequently, the arrival rate of formal jobs drops as the effect of the fall in the proportion of formal vacancies dominates. In contrast, the arrival rate of informal jobs increases, explained by the change in composition of vacancies in the economy from formal to informal and by a less tight labor market for workers. Meanwhile, meeting rates drop, making it harder for firms to meet workers. This is due to a competition effect: more informal vacancies are posted, intensifying competition among firms for the same pool of workers.

Figure 7 shows the unemployment values for Colombia and Mexico. For Colombia, the unemployment value decreases for higher minimum wage values. Since the value of unemployment also represents the worker's outside option, its decrease means a reduction in workers' bargaining power and their share of the surplus. An additional interesting feature of the Colombian case is that over the range of policy experiments we run, the value of unemployment is non-monotonic, increasing at low levels and decreasing at high levels. The case of Mexico is different: the value of unemployment is increasing, but within a very narrow range, which implies that Mexico is probably in the first part of the non-monotonic relationship observed in Colombia. Since value unemployment also represents the present discounted value of participating in the market for workers, these results imply that workers' welfare is non-monotonic at the minimum wage level, with potential welfare gains at low minimum wage levels but welfare losses at high levels, particularly evident in the Colombian labor market.

Finally, Figure 7 also shows the total level of production generated by the labor market (i.e., the sum of all match-specific productivities x and all self-employment productivities y). Again, the results for Colombia and Mexico are similar. The total production level decreases monotonically with minimum wage levels. These results, combined with those found for the value of unemployment, suggest a nuanced discussion of the optimal minimum wage level. If the bite of the minimum wage is high, it may be positive if the objective function is workers' welfare, but it may well be zero if the objective function is the GDP level.

When fixing the arrival rates, we find significantly different outcomes. The unemployment value appears to be non-monotonic, yet it tends to increase over a wide range of higher minimum wage levels, with a particularly sharp increase in Colombia. The total level of production is increasing in Mexico and non-decreasing in Colombia for a wide range of higher minimum wage values. These results suggest that the equilibrium effect on the demand side, specifically the posting process, is important for explaining the composition effect.

5 Concluding Remarks

This paper develops a search and matching model that captures the interaction between a labor market with high informality and binding minimum wage policies. By incorporating institutional parameters, such as payroll contributions and minimum wage constraints, alongside labor market frictions and occupational choices, including formal, informal, and self-employment states, the model is successful in reproducing the key empirical patterns observed in Latin American labor markets.

The results for Colombia and Mexico show that the model is able to fit very well the observed dynamics of the labor market, wage patterns and distributions of labor market states. It can also account for the large overlaps between formal and informal wages and for the important bit of the minimum wage. In the model, the costs of job posting and the process of vacancy creation are important for shaping labor market outcomes and creating formal jobs.

Policy simulations show how different minimum wage levels affect the workers. Small cuts in minimum wage can help workers, but large cuts or increases can cause adverse outcomes, including moving from formal jobs to informal work or self-employment. This change highlights the trade-offs policymakers face when protecting low-wage workers and maintaining formal jobs. In Colombia and, to a lesser extent, in Mexico, the link between minimum wage levels, welfare, and production is not straightforward. Worker welfare and total production react differently to changes in the minimum wage; therefore, optimal policy design requires balancing these competing objectives. Endogenizing firms' vacancy-posting behavior is critical to capturing labor market adjustments, as it drives the composition of vacancies, arrival rates of job offers, and overall market tightness. Ignoring these demand-side responses can lead to very different conclusions regarding the minimum wage impacts.

References

Acemoglu, Daron, "Good Jobs versus Bad Jobs," *Journal of Labor Economics*, 2001, *19* (1), 1–21.

Albrecht, James, Lucas Navarro, and Susan Vroman, "The Effects of Labour Market Policies in an Economy with an Informal Sector," *Economic Journal*, 07 2009, *119* (539), 1105–1129.

Bianchi, Milo and Matteo Bobba, "Liquidity, Risk, and Occupational Choices," *Review of Economic Studies*, 2013, *80* (2), 491–511.

Bobba, Matteo, Luca Flabbi, and Santiago Levy, "Labor market search, informality, and schooling investments," *International Economic Review*, 2022, *63* (1), 211–259.

______ , ______ , ______ , and Mauricio Tejada, "Labor market search, informality, and on-the-job human capital accumulation," *Journal of Econometrics*, 2021, *223* (2), 433–453.

Bosch, Mariano and Julen Esteban-Pretel, "Job creation and job destruction in the presence of informal markets," *Journal of Development Economics*, 2012, *98* (2), 270–286.

______ and William F Maloney, "Comparative analysis of labor market dynamics using Markov processes: An application to informality," *Labour Economics*, 2010, *17* (4), 621–631.

Conti, Gabriella, Rita Ginja, and Renata Narita, "The value of health insurance: a household job search approach," 2018.

Dey, Matthew and Christopher Flinn, "Household search and health insurance coverage," *Journal of Econometrics*, 2008, *145* (1-2), 43–63.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman, "Labor market dynamics and development," *The Quarterly Journal of Economics*, 2023, *138* (4), 2287–2325.

Eckstein, Zvi and Gerard J. van den Berg, "Empirical labor search: A survey," *Journal of Econometrics*, February 2007, *136* (2), 531–564.

______ and Kenneth I Wolpin, "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model," *Review of Economic Studies*, April 1995, *62* (2), 263–86.

______ , Suqin Ge, and Barbara Petrongolo, "Job and wage mobility with minimum wages and imperfect compliance," *Journal of Applied Econometrics*, 2011, *26* (4), 580–612.

Falco, Paolo, William F Maloney, Bob Rijkers, and Mauricio Sarrias, "Heterogeneity in subjective wellbeing: An application to occupational allocation in Africa," *Journal of Economic Behavior & Organization*, 2015, *111*, 137–153.

Fields, Gary S., "Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs," *Journal of Development Economics*, June 1975, *2* (2), 165–187.

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

Flabbi, Luca, "Implications of Minimum Wages Policies in Labor Markets with High Informality and Frictions.," Technical Report, Background paper for the UNDP Regional Human

Development Report for Latin America and the Caribbean 2021.

——— **and James Mabli**, "Household Search or Individual Search: Does It Matter?," *Journal of Labor Economics*, 2018, *35* (1).

Flinn, C., "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining and Endogenous Contact Rates," *Econometrica*, 2006, *73*, 1013–1062.

Flinn, Christopher and Joseph Mullins, "Firms'choices of wage-setting protocols," Technical Report, Discussion paper, New York University 2021.

Flinn, Christopher J, *The minimum wage and labor market outcomes*, MIT press, 2011.

Maloney, William F, "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico," *World Bank Economic Review*, May 1999, *13* (2), 275–302.

McKenzie, David J and Christopher Woodruff, "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises," *Economic Development and Cultural Change*, October 2006, *55* (1), 3–42.

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

Moore, Mark A., Anthony E. Boardman, and Aidan R. Vining, "Social Discount Rates for Seventeen Latin American Countries: Theory and Parameter Estimation," *Public Finance Review*, January 2020, *48* (1), 43–71.

OECD, *Taxing Wages 2023: Indexation of Labour Taxation and Benefits in OECD Countries*, OECD, 2023.

Rogerson, Richard, Robert Shimer, and Randall Wright, "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, December 2005, *43* (4), 959–988.

Table 1: Estimated Parameters

Parameter	Colombia		Mexico	
	Estimate	Standard Error	Estimate	Standard Error
Dynamics				
λ_1	0.078	(0.0058)	0.090	(0.0066)
λ_0	0.050	(0.0030)	0.046	(0.0034)
γ	0.033	(0.0017)	0.005	(0.0005)
η_1	0.049	(0.0008)	0.029	(0.0007)
η_0	0.056	(0.0013)	0.028	(0.0009)
δ	0.027	(0.0004)	0.016	(0.0009)
Productivity Distribution				
μ_y	-0.069	(0.0532)	0.966	(0.0529)
σ_y	0.462	(0.0133)	0.444	(0.0572)
μ_1	1.202	(0.6065)	1.464	(0.5444)
σ_1	0.255	(0.0786)	0.395	(0.1080)
μ_0	0.815	(0.0808)	1.092	(0.0151)
σ_0	0.349	(0.0288)	0.309	(0.0146)
Utility				
β_u	-3.542	(23.5913)	-3.762	(11.6304)
β_1	0.303	(17.0538)	0.000	(5.0459)
β_s	0.321	(0.1902)	0.445	(0.3342)
ρU	0.342	(0.1764)	1.004	(0.0125)
Demand Side				
ϕ	0.429	(0.0115)	0.399	(0.0136)
π_1	0.610	(0.0168)	0.664	(0.0144)
κ	234.141	(41.3327)	532.753	(37.6598)
ψ	0.005	(0.008)	0.003	(0.0013)
Measurement Error				
σ_ϵ	0.043	(0.0247)	0.195	(0.0068)
Fixed Parameters				
w_M	1.456		1.796	
t	0.160		0.330	
c	0.130		0.197	
α	0.500		0.500	
θ	0.0082		0.0067	
$LogL$	53847.0		11435.0	

NOTE: Bootstrap standard errors in parenthesis. Following Bobba et al. (2022), c is calibrated as 10% of the average wage of informal employees.

Table 2: Model Fit

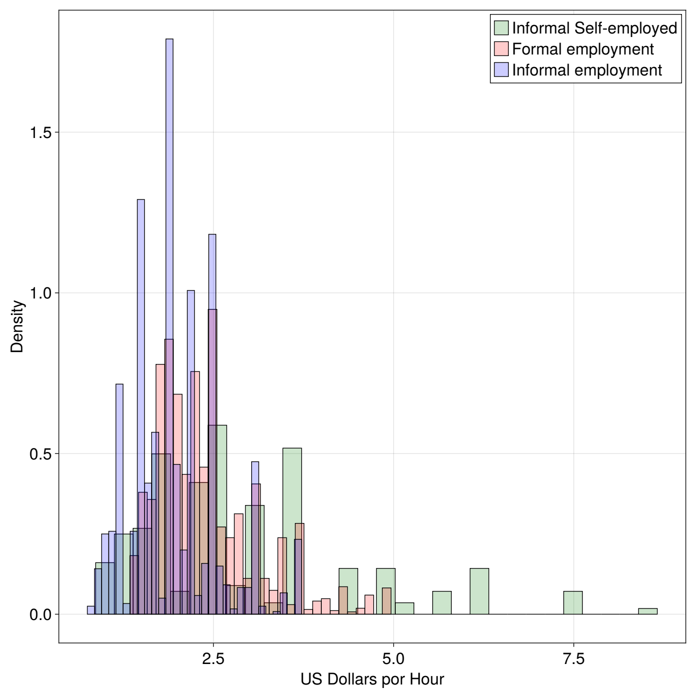
Statistic	Colombia		Mexico	
	Model	Data	Model	Data
u	0.211	0.254	0.171	0.153
s	0.263	0.253	0.049	0.050
e_1	0.337	0.317	0.504	0.516
e_0	0.190	0.176	0.275	0.281
$P[w = w_M \mid e_1]$	0.500	0.497	0.231	0.246
$E[w \mid s]$	1.041	1.079	2.927	2.897
$SD[w \mid s]$	0.511	0.460	1.496	1.472
$E[w \mid e_1]$	1.627	1.628	2.370	2.401
$SD[w \mid e_1]$	0.275	0.276	0.791	0.719
$E[w \mid e_1, w > w_M]$	1.801	1.810	2.539	2.589
$SD[w \mid e_1, w > w_M]$	0.304	0.302	0.827	0.681
$E[w \mid e_0]$	1.311	1.302	1.976	1.966
$SD[w \mid e_0]$	0.438	0.366	0.642	0.609
$E[t \mid u]$	6.217	5.125	7.444	8.331
$E[t \mid s]$	37.426	37.426	64.051	64.051
$E[t \mid e_1]$	20.555	20.555	35.031	35.031
$E[t \mid e_0]$	18.006	18.006	35.331	35.331

NOTE: Based on 50.000 simulated individuals.

Figure 1: Wages Distributions by Labor Market State

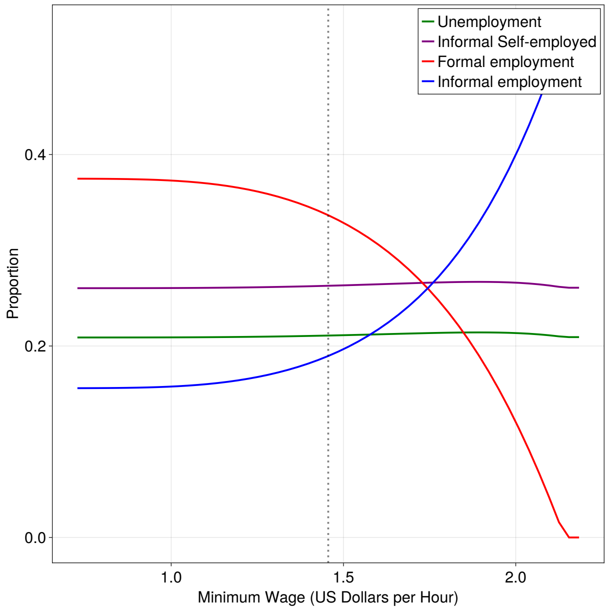
figs/col_wages_distributions_hist.pdf

(a) Colombia

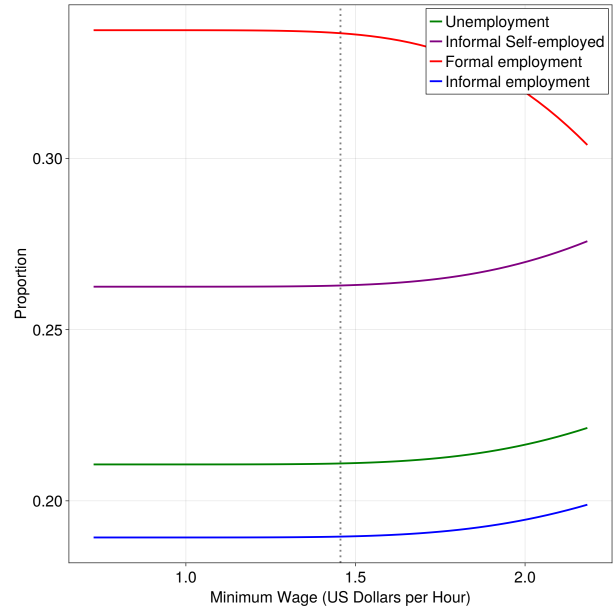


(b) Mexico

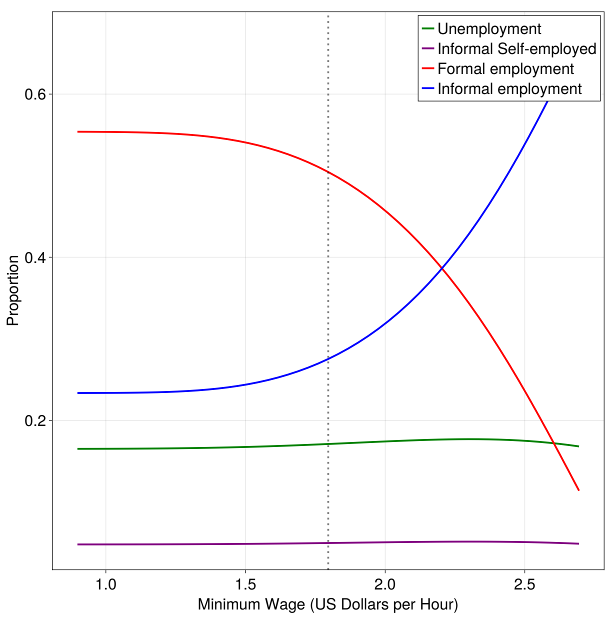
Figure 2: Labor Market States under Different Counterfactual Scenarios



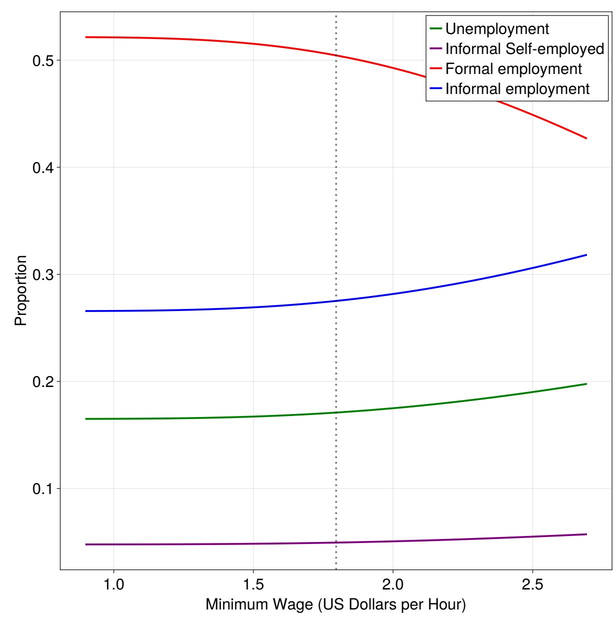
(a) Colombia (Endogenous Vacancy Creation)



(b) Colombia (Exogenous Vacancy Creation)

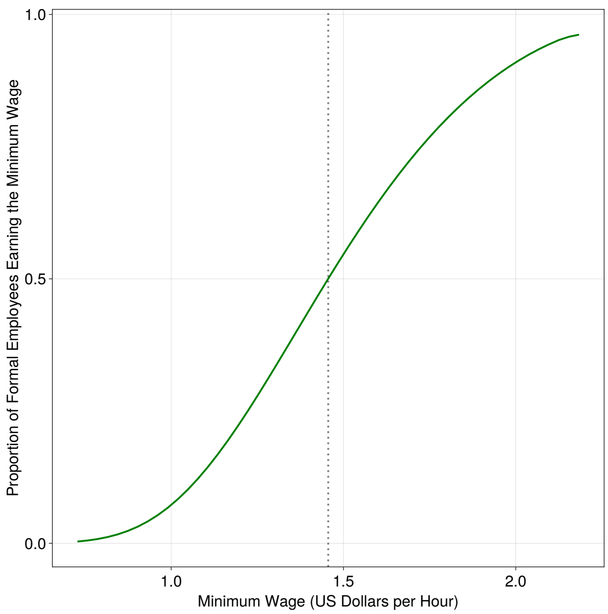


(c) Mexico (Endogenous Vacancy Creation)

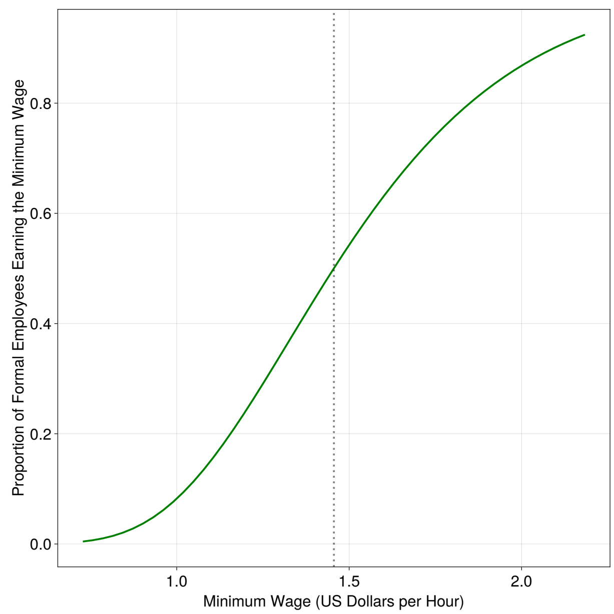


(d) Mexico (Exogenous Vacancy Creation)

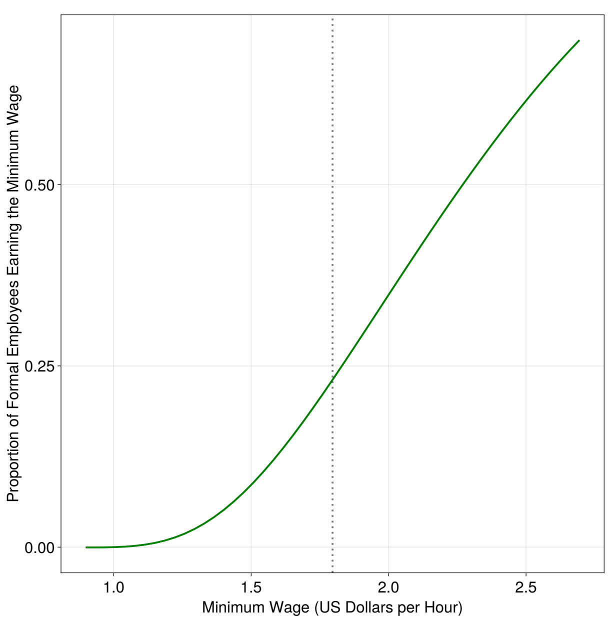
Figure 3: Proportion of Formal Employees Earning the Minimum Wage



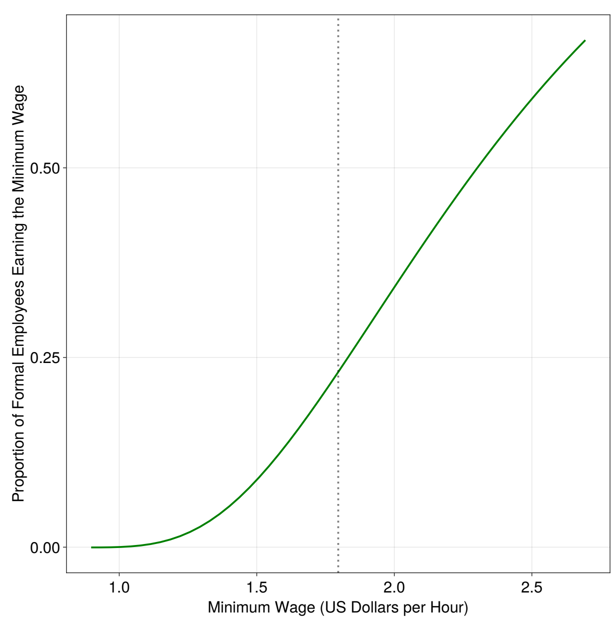
(a) Colombia (Endogenous Vacancy Creation)



(b) Colombia (Exogenous Vacancy Creation)

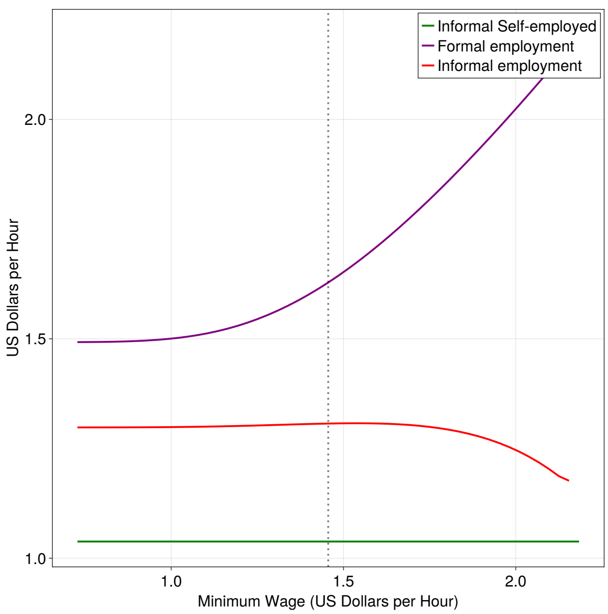


(c) Mexico (Endogenous Vacancy Creation)

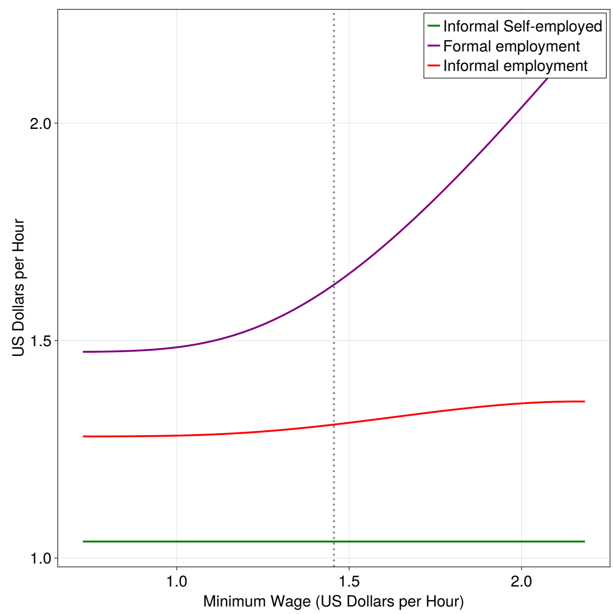


(d) Mexico (Exogenous Vacancy Creation)

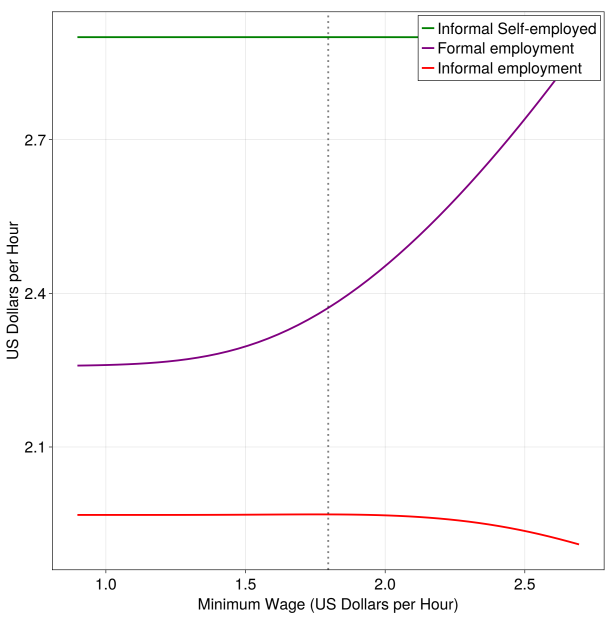
Figure 4: Average Wages by Labor Market State under Different Counterfactual Scenarios



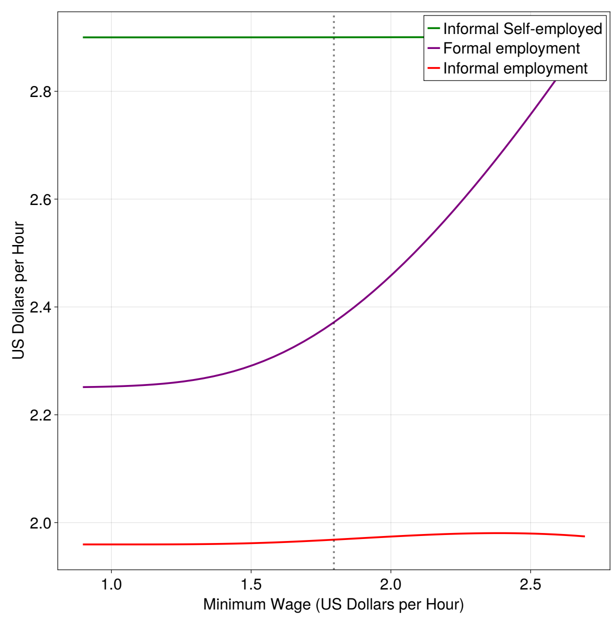
(a) Colombia (Endogenous Vacancy Creation)



(b) Colombia (Exogenous Vacancy Creation)

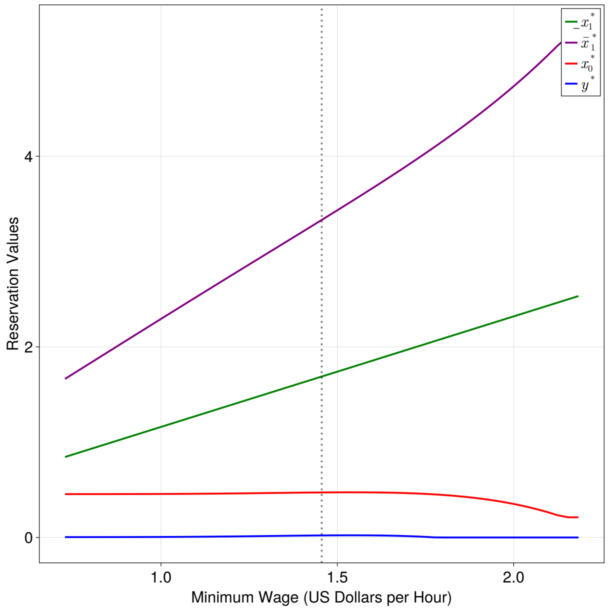


(c) Mexico (Endogenous Vacancy Creation)

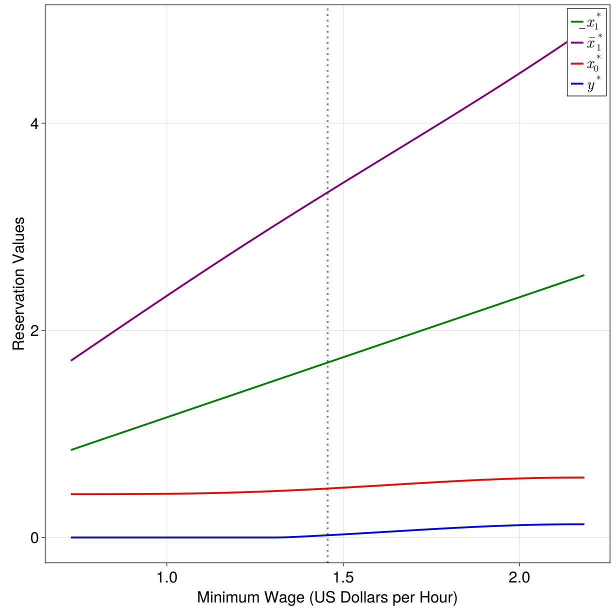


(d) Mexico (Exogenous Vacancy Creation)

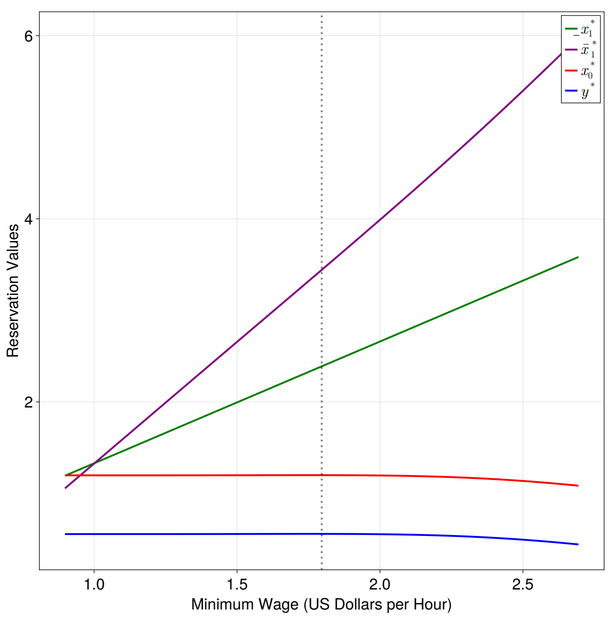
Figure 5: Reservation Productivities under Different Counterfactual Scenarios



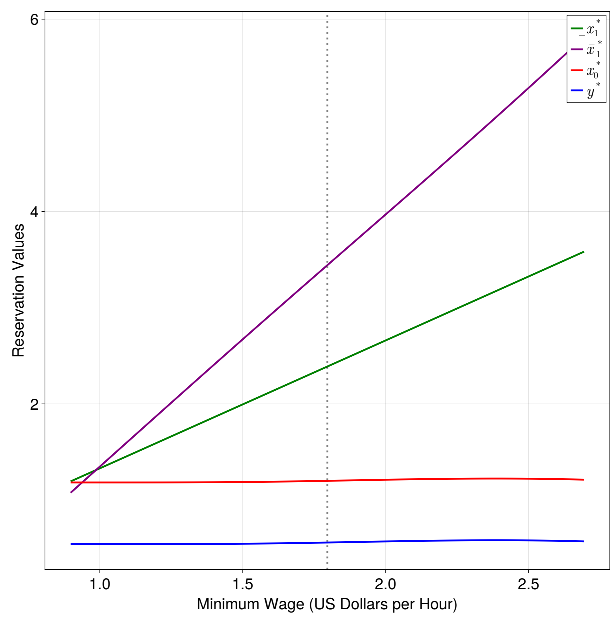
(a) Colombia (Endogenous Vacancy Creation)



(b) Colombia (Exogenous Vacancy Creation)

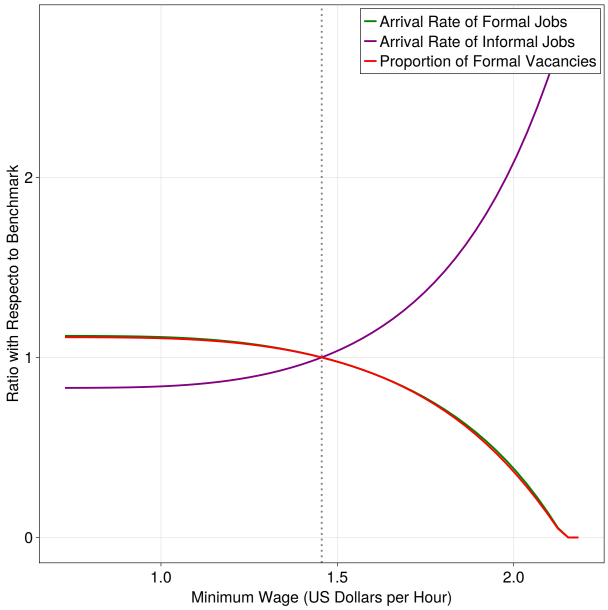


(c) Mexico (Endogenous Vacancy Creation)

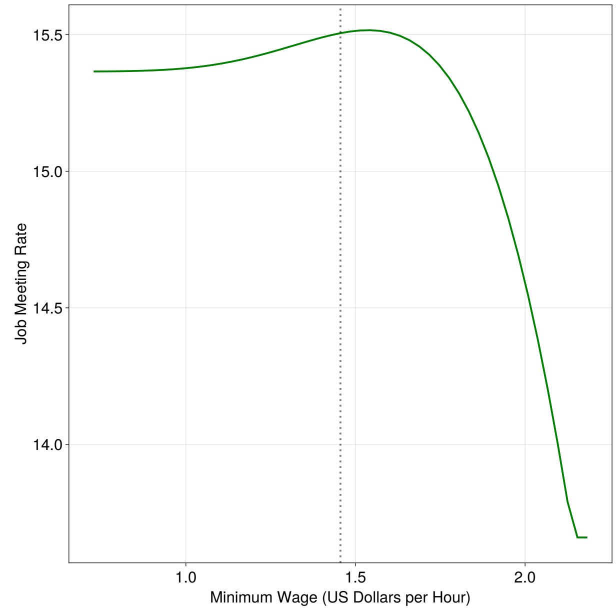


(d) Mexico (Exogenous Vacancy Creation)

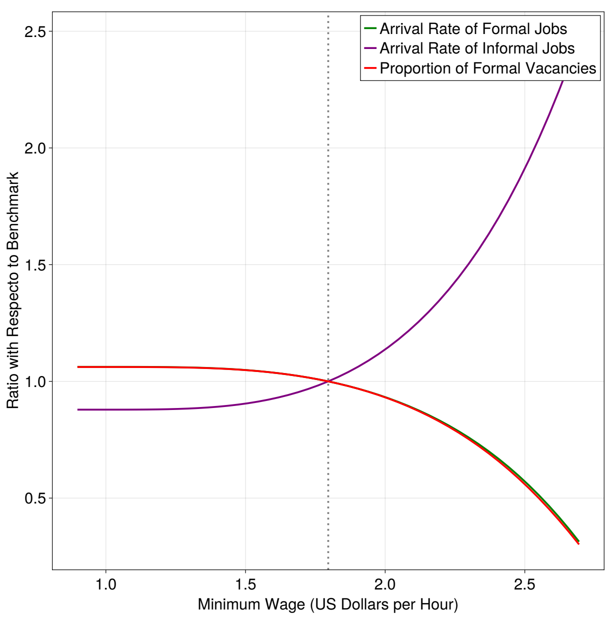
Figure 6: Arrival Rates and Job Meeting Rates under Different Counterfactual Scenarios



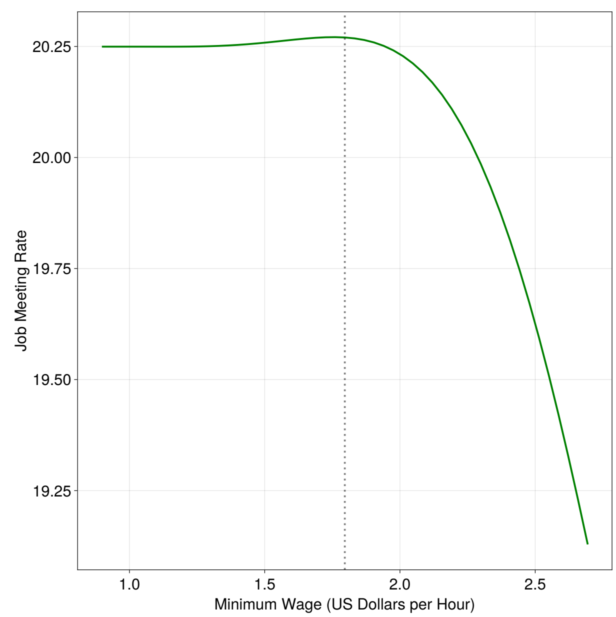
(a) Colombia (Supply Side)



(b) Colombia (Demand Side)

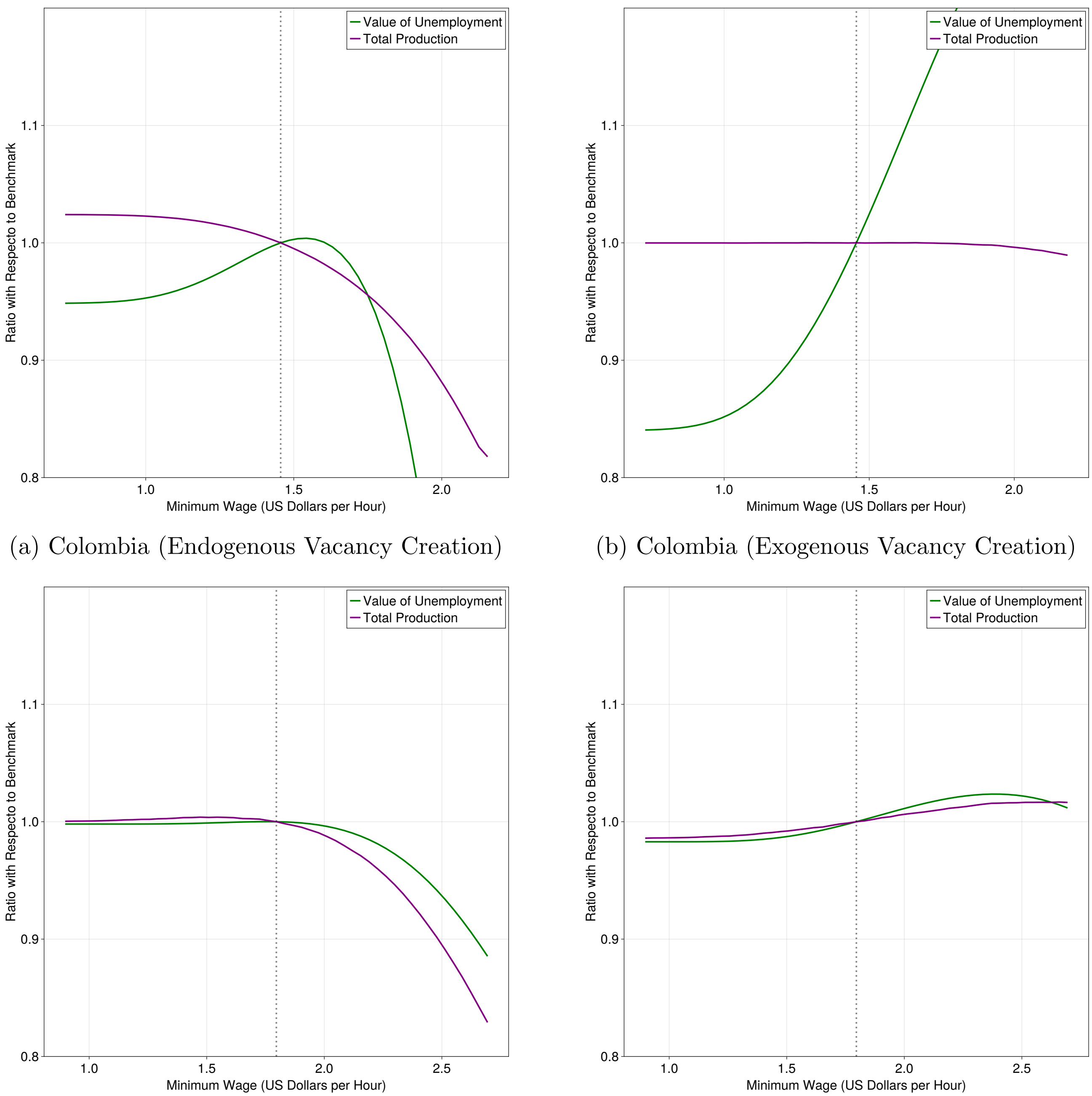


(c) Mexico (Supply Side)



(d) Mexico (Demand Side)

Figure 7: Life Time Value of Unemployment and Total Productivity under Different Counterfactual Scenarios



(a) Colombia (Endogenous Vacancy Creation)

(b) Colombia (Exogenous Vacancy Creation)

(c) Mexico (Endogenous Vacancy Creation)

(d) Mexico (Exogenous Vacancy Creation)