The value of a formal job

Adolfo Fuentes*
University of Edinburgh

Job Market Paper
Latest version here

January 31, 2025

Abstract

Starting a career in a formal versus an informal job significantly impacts outcomes due to differences in job productivity and security, human capital accumulation, and offer arrival rates. In this paper, I attempt to measure the value of a formal job and assess the relative importance of each factor. I propose a frictional labor market model with two sectors, where firms are heterogeneous in productivity and job security, while workers search off- and on-the-job and accumulate human capital. I calibrate the model using survey and administrative data from Chile. I find that, on average, the value of a formal job is equivalent to a lump sum payment of 6.2 monthly minimum wages. Formal jobs lead, as a lower bound, to 7.8% higher earnings after five years. The main determinants of wages are productivity, human capital accumulation, and the minimum wage, while job security and offer arrival rates play secondary roles.

Keywords: Labor informality, job ladder, minimum wage, human capital accumulation.

*Email: a.i.fuentes@sms.ed.ac.uk. First version: August 2024. I am especially grateful to Axel Gottfries and Rafael Lopes de Melo for their continuous guidance and support throughout this project. I also thank Jesper Bagger, Andrew Clausen, Mike Elsby, Jan Grobosek, Jonna Olson, Victor Saldarriaga, Rachel Scarfe, the participants at the EALE Conference 2024, LM2C2-PUC Workshop 2024, SGPE Conference 2025, and seminars at the Institute for Fiscal Studies, Universidad Diego Portales, and the University of Edinburgh for their valuable comments. Finally, I am thankful for the financial support from the National Agency of Research and Development (ANID) of the Chilean Government for my doctoral studies. All errors are my own.

1 Introduction

Labor informality presents a pressing concern within developing economies. In these contexts, labor regulations do not apply to informal workers, exempting them from minimum wage requirements and limits on working hours, among other regulations (Ulyssea, 2020). The scale of these challenges is considerable, since labor informality significantly impacts 20 to 80 percent of the workforce in developing countries (Perry et al., 2007; La Porta & Shleifer, 2014), accounting for about 2 billion workers in the world (Bonnet et al., 2019) and one-third of total economic activity (Schneider & Enste, 2000; Elgin et al., 2021). Given the above, a natural question that arises is what is the difference in value between formal and informal jobs. Previous literature has established that formal jobs tend to be more productive (Ulyssea, 2018) and safe (La Porta & Shleifer, 2014), provide faster human capital accumulation (Jedwab et al., 2023), and have a higher exposure to formal offers than informal jobs (Meghir et al., 2015). The above implies that starting a career in one or the other affects outcomes such as wages, unemployment risk, human capital accumulation, and formality status. However, there is limited research on how the aspects mentioned above affect workers' outcomes over time, particularly regarding the value of formal employment and how it evolves as workers climb the job ladder.

To help close this gap in the literature, I propose a frictional search model with two sectors that differ in their institutional settings, particularly the existence of a minimum wage in the formal sector. In this model, firms vary in productivity and exogenous job destruction rates within each sector. Workers accumulate human capital while employed, at different rates in the formal and informal sector, and lose it when unemployed. Additionally, workers can search for jobs both off- and on-the-job.

This model will have several key characteristics. First, individuals can climb by moving to safer, more productive, or formal firms. This constitutes an extension of **?**. The job ladder will be slippery at the bottom because firms with high unemployment risk are concentrated there, making it more likely that workers who join these firms after a period of unemployment will return to unemployment. Second, the benefits of being employed at a firm are not limited to wages but also include the associated search gains, as the model incorporates competition between firms, following the sequential auction bargaining protocol introduced by Cahuc et al. (2006). This would lead safe and productive formal firms, presumably at the top of the job ladder, to internalize these gains and offer lower wages than informal firms. However, the minimum wage will restrict this behavior and instead will lead to formal firms paying higher wages than those observed in the informal sector. Third, the inclusion of a minimum wage will result in higher *de facto* bargaining power for workers employed in formal firms where the minimum wage is binding, as in Flinn (2006), but also will potentially exclude workers with low levels of human capital from formal jobs.

The data used to calibrate the model corresponds to Chilean data for the formal and informal sectors between 2010 and 2019. During this period, the informality rate remained relatively stable at around 28 percent of the labor force, while formal employment and unemployment accounted for approximately 65 percent and 7 percent, respectively (INE, 2019). The formal data consists of a 20 percent random sample from the administrative unemployment insurance records (*Seguro de Desempleo*). Data for both the formal and informal sectors is drawn from the National Employment Survey (*Encuesta Nacional de Empleo*), conducted quarterly,[1] and the Supplementary Income Survey (*Encuesta Suplementaria de Ingresos*) for the last quarter of each year. To calibrate the model, I match several moments from the data with those generated using simulated data from the model.

Once the model is calibrated, the value of formal and informal jobs can be directly inferred from the value functions of different jobs in equilibrium. The additional value provided by a formal job will depend on the corresponding informal counterpart selected. To conduct this analysis, I compare firms at the same relative position in both sectors. The model generates a distribution of the excess of value associated with moving to a formal job, allowing us to determine, for example, whether the greatest gains occur at the bottom, middle, or top of the job ladder and at what point during the career. When this analysis is performed across the entire distribution of firms, I find that, on average, the extra value provided by a formal job is equivalent to a lump sum payment of approximately 6.2 monthly minimum wages, or US$2,740, and that this value decreases with human capital.

How does the difference between a formal and informal job evolve over time? To answer this question, I simulate two groups of workers: one group initially receiving a formal offer and the other an informal one. Then, I compare their labor market outcomes over time. This simulation is conducted by allocating workers with the initial level of human capital to firms at the bottom, middle, and top of the job ladder. The results show that workers who start with a formal job have at least 7.8% higher earnings than those who begin with an informal job.

Once it is established that a formal job provides more value than informal jobs, it is relevant to study which components - productivity, unemployment risk, human capital accumulation, the arrival of offers, and the minimum wage - are key to providing this extra value. To assess this, I conduct counter-factual exercises in which the formal sector is assigned parameters equal to those estimated for the informal sector. I find that productivity, human capital accumulation, and the minimum wage are the most significant factors, accounting for 46, 28, and 22% changes in the wage difference, respectively. In contrast, the arrival of offers and unemployment risk are of secondary importance, with their effects on wages being approximately 13% and 4%, respectively.

[1] Centered in February, May, August, and November

The rest of the paper is organized as follows: The second section reviews the contributions of this paper to the existing literature. The third section provides an overview of the Chilean context and details the data used. The fourth section introduces the model. The fifth section addresses the aspects related to model estimation. The sixth section presents the calibration. The seventh section discusses three exercises designed to study the value of holding a formal job. Finally, the eighth section concludes the paper and outlines potential directions for future research.

2 Related literature

This paper takes as a basis the literature that studies labor informality using reduce-form and structural techniques, which has established several facts about the difference between formal and informal jobs. First, although there is considerable overlap in their productivity distributions, formal firms tend to be more productive than informal ones (Meghir et al., 2015; Allen et al., 2018; Ulyssea, 2018). This has been explained as the result of a selection process in which more productive firms hire more employees, increasing their likelihood of being discovered and fined by the government, which incentivizes them to operate formally (Dabla-Norris et al., 2008; Meghir et al., 2015; Haanwinckel & Soares, 2021; Ruggieri & Cisneros-Acevedo, 2023; Parente, 2024). Alternatively, it could be due to the fact that informal firms are managed by less skilled entrepreneurs and face limitations in capital and technological innovation (De Paula & Scheinkman, 2011; Fu et al., 2018). Second, informal firms tend to offer jobs with a higher risk of unemployment (Perry et al., 2007; La Porta & Shleifer, 2008; De Paula & Scheinkman, 2011; La Porta & Shleifer, 2014). Third, research shows that human capital accumulates more rapidly in the formal sector due to the availability of training and the acquisition of skills valuable to other employers (Lagakos et al., 2018; Bobba et al., 2021; Jedwab et al., 2023). Fourth, previous literature indicates that transition rates between formal and informal jobs vary according to worker characteristics (Magnac, 1991; Meghir et al., 2015; Allen et al., 2018; Ulyssea, 2018). Finally, the literature suggests that in the presence of labor market search frictions, informal firms create a negative externality for formal firms by making it more difficult for them to find workers, thereby increasing search costs (Amin & Okou, 2020; Ulyssea, 2018, 2020).

The contributions of this paper can be classified into two strands of literature. First, it addresses the literature that studies labor informality from a structural perspective. Previous research in this field has focused on the effects of several elements on informality, including the minimum wage (Parente, 2024), variations in human capital accumulation and educational investment across sectors (Bobba et al., 2021, 2022), firing costs, payroll, and taxes (Albrecht et al., 2009; Ulyssea, 2010; Bosch & Esteban-Pretel, 2012; Ordonez, 2014; Cisneros-Acevedo, 2022), complementarity between worker types (Haanwinckel & Soares, 2021), trade policy (Paz, 2014; Dix-Carneiro & Kovak, 2019;

Dix-Carneiro et al., 2021), intensive and extensive margins (Ulyssea, 2018), labor protection (Tejada, 2017), government regulation (Fugazza & Jacques, 2004; El Badaoui et al., 2010; Meghir et al., 2015), unemployment insurance (Bosch & Esteban-Pretel, 2015), and barriers to capital and investment (Amaral & Quintin, 2006; D'Erasmo & Boedo, 2012; Lopez-Martin, 2019). However, there is limited understanding of the career effects of starting in a formal versus an informal job and what drives the value of formal employment. This paper contributes to this literature by proposing a frictional search model capable of addressing this question, taking into account the most important characteristics of the sectors as identified in previous studies.

Conceptually, this paper extends this literature by including firms that can be different in three aspects: formality status, productivity, and exogenous destruction rate. While the first two components are common in the literature (Haanwinckel & Soares, 2021; Bobba et al., 2022; Parente, 2024), differences in exogenous destruction rates within each sector have not been extensively studied in the literature on labor informality. This concept was recently introduced by Pinheiro & Visschers (2015), and used by Jarosch (2023) to examine the consequences of job loss, as it creates a slippery job ladder at the bottom, where risky firms are located. Including heterogeneity in job security is relevant for several reasons: (i) Chilean formal firms experience annual destruction rates that vary by size, with differences of up to ten times (Arellano & Jimenez, 2016), (ii) It reduces the requirements to generate employment-to-employment transitions since it is not exclusively needed higher productivity to move to another firm, (iii) Generates additional heterogeneity in the effective rates of human capital accumulation inside each sector, (iv) As I will show later, the empirical destruction rates decline with tenure, something that it is not possible to explain in the traditional setting with a homogeneous destruction rate in each sector.

Also, this paper incorporates the between-employer competition framework introduced in Cahuc et al. (2006), and extended to include human capital accumulation in Bagger et al. (2014), into the analysis of informal markets. In this context, workers with the same levels of human capital and working in the same firm but having a different benchmark option will receive different wages, increasing the range of wages observed in the economy. The resulting wider wage distribution, and the possibility that more productive firms do not always pay higher wages than less productive ones, can serve as a natural explanation for the overlap in wage distributions between the formal and informal sectors, as documented in the literature (Maloney, 1999; Hsieh & Olken, 2014; Allen et al., 2018; Ulyssea, 2018).

Finally, this paper also contributes to the literature generating rich search models. Previous literature has focused, among others, on introducing between-employer competition (Postel-Vinay & Robin, 2002a,b; Cahuc et al., 2006), human capital accumulation (Bagger et al., 2014), unemployment policies (Lise et al., 2016), search intensity (Bagger & Lentz, 2019), and heterogeneity

in job security (Jarosch, 2023). However, in a two sectors model with the common setting of the previous papers, formal firms - being generally more productive, safer, and better in human capital accumulation and the arrival of offers - would likely pay lower wages than informal firms at the entry level of the job ladder, since in equilibrium they internalize the search gains provided to workers. However, this contradicts empirical evidence that shows that formal firms tend to offer higher wages across the entire wage distribution (INE, 2024). Then, this paper contributes to this strand of literature by introducing a minimum wage setting into the analysis. This inclusion is relevant since it will affect the extent to which best firms can pay low wages for those coming from unemployment, and also due to between-employer competition partially extending the influence of the minimum wage to informal firms.

The framework of this paper is closely related to Jarosch (2023), who studied individual wage dynamics in a frictional setting with human capital accumulation and heterogeneity in job security. The most important distinction is that this paper features two sectors, which are interconnected through on-the-job search, and the presence of a minimum wage in one of them.

3 Background and data

The period studied in this paper spans the years 2010 to 2019, which is relatively stable in terms of labor market indicators, particularly concerning unemployment and informality rates, which remained around 7 and 28%, respectively (Fuentes, 2019; INE, 2020).[2] The public sector fluctuates around 5 and 6 percent of the employment during this period, and it is relatively stable in terms of gender composition and its wage premium over the private sector (Fuentes & Vergara, 2018). Labor market participation at the national level was also relatively stable, averaging around 59%. However, we observe some internal shifts across gender. Men's participation slightly decreased from 72 to 70%, while women's participation significantly increased from 44 to 49% (INE, 2019). This change has been developing since the previous decade and can be primarily attributed to the expansion of tertiary education, cultural shifts, and the implementation of childcare policies during this period (Contreras & Plaza, 2007; Encina & Martínez, 2009; Contreras et al., 2012).

Regarding the minimum wage, it increased by roughly 50 percent in real terms between January 2010 and December 2019. Interestingly, this increase did not change either the ratio between the minimum and median wage in the economy, which was stable at around 0.7, or the proportion of workers who earn the minimum wage in the formal sector, which fluctuated around 15 percent according to the administrative records (Abud et al., 2022). The above evidence suggests that the

[2]Official data on informality has only been available since 2017; however, in another paper (Fuentes, 2024) I compare formal sector employment from survey data with administrative records, and show that the informality rate in the country remained constant throughout the decade of 2010.

minimum wage policy during this period was innocuous in affecting the aggregate formal status or the wage distribution in the formal sector.

Concerning the definition of an informal job, this will be closely tied to the institutional context of the labor market in question (Ulyssea, 2020), making it necessary to clarify what a formal job entails in Chile. The Chilean National Bureau of Statistics (INE) provides an official definition that includes criteria from the worker's and firm's perspectives (INE, 2017). For a job to be considered formal, it must involve a written contract, offer vacation time, pay at least the minimum wage, and contribute to social security on the worker's side. On the firm's side, the company must issue a document, typically a pay slip, detailing the worker's gross wage and deductions for social security contributions, taxes, and insurance. Additionally, the firm must be registered with the tax office.

Unfortunately, data corresponding to this definition has only been available since 2017. Therefore, I use an alternative definition of labor informality based on whether the firm contributes to social security, particularly the pension system. The pension system requires a contribution of 10 percent of the gross monthly wage, which is deposited into an individual account, along with an additional commission of approximately 1 percent that goes to the pension fund administrator. This information has been available since 2010, allowing the analysis of informality over the entire decade. This definition captures a combined condition involving the worker and the firm: to receive social security contributions in their accounts, workers must have a labor contract, and the firm must be registered in the social security system and, by extension, with the tax office. While this definition is less strict than the one used by the INE, it is widely employed in the literature on informality (Ulyssea, 2020).

The data used in this paper comes from three distinct datasets. First, I utilize administrative records from the unemployment insurance system (*Base del Seguro de Cesantía, SC*), which covers 20 percent of randomly selected workers in the formal sector, excluding the public sector.[3] The sample is drawn by randomly selecting 20 percent of workers from the most recent month of the dataset (June 2021) and reconstructing their labor histories back to 2002, when the unemployment insurance system began. This dataset includes IDs for both workers and firms. It also provides information on labor income at the individual and firm levels, as well as on gender, education, citizenship, age, region, and economic sector.

Individuals contribute to the unemployment insurance system for each formal job they have. Consequently, workers with multiple formal jobs appear in the monthly records. To address this, I

[3]Workers in the Chilean public sector do not have access to unemployment insurance, and therefore, they are not present in the data.

will focus on the primary employment for each individual in the analysis, where primary employment is defined as the job with the highest wage, among other criteria. The specific procedure for determining the primary employment is detailed in the Appendix.[4] As the dataset tracks individuals over time, it allows the identification of employment spells in the formal sector. A key assumption using the administrative data is that workers are unemployed between formal employment spells. This assumption facilitates the computation of human capital loss during unemployment periods. As will be shown later, this assumption is reasonable given the particularly low transition rates from the formal to the informal sector. After structuring the data as a panel, following the procedure outlined in the Appendix, the dataset will contain approximately 243 million observations from 2.2 million individuals.[5]

Other dataset is the panel data from the national employment survey (*Encuesta Nacional de Empleo, ENE*).[6] This survey tracks households over six quarters and collects information on employment status, age, gender, education, job tenure, region, economic sector, and occupation, though it does not include income data. Each worker is assigned a unique ID, allowing for tracking across the six quarters. While it is not possible to fully reconstruct individual labor histories, I can follow their employment history from their first observed job. Also, since I observe workers between quarters, I cannot distinguish between employment-to-employment transitions and workers who lost their jobs and then find another one between quarters. Then, I assume that workers who exhibit two different jobs in two consecutive quarters did an employment-to-employment transition. Also, as in the administrative records, I will only consider the main occupation, which is self-reported considering the occupation with the higher wage.[7] The survey is conducted monthly, with each dataset covering data from the past three months. I will focus on data centred in February, May, August, and November. There are 40 surveys from February 2010 to November 2019, each containing around 47,000 worker observations, resulting in approximately 1.88 million total observations.

To complement the employment survey data, I use the supplementary income survey (*Encuesta Suplementaria de Ingresos, ESI*), which is conducted during the last quarter of the year (around November).[8] This survey provides detailed information on various measures of labor income and uses an ID that matches the one in the employment survey, allowing for the linkage of income

[4] Appendix 1: Main job in administrative data

[5] Appendix 2: Administrative data to panel

[6] The data is publicly available at https://www.ine.gob.cl/estadisticas/sociales/mercado-laboral/ocupacion-y-desocupacion

[7] After filtering the data, only 3.4% of workers declare a secondary occupation. Of those, 53% declare working in that secondary occupation less than 10 hours per week, 83% declare less than 20 hours per week, and 96% declare less than 30 hours per week.

[8] The data is publicly available at https://www.ine.gob.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-suplementaria-de-ingresos

data with employment transitions. This linkage is crucial, as it is the only source of informal income data connected to employment information. This dataset provides approximately 423,000 observations.

I restrict the sample to individuals aged 25 to 64 in the labor force. Additionally, I will include only those workers with a self-reported job tenure of 10 years or less and who report working between 30 and 60 hours per week.[9] For unemployed individuals, I will limit the sample to those seeking a job for less than three years. Several filters were also applied to ensure consistency in the reported tenure and sector, with the details provided in the Appendix.[10]

Table 1: Summary statistics for employment: 2010-2019

	Overall	Men	Women	Tertiary	Non-tertiary
Share (%) - Formal	67.5	68.2	66.6	72.2	65.1
Share (%) - Informal	21.6	22.1	20.7	16.8	24.0
Share (%) - Unemployment	10.9	9.6	12.7	11.1	10.9
Mean Tenure (Months) - Formal	31.6	30.1	33.7	35.2	29.5
Mean Tenure (Months) - Informal	28.5	28.5	28.4	31.1	27.5
Mean Search (Months) - Unemployment	4.6	4.2	5.1	5.6	4.1
Share (%) - Men	57.1	-	-	51.1	60.2
Share (%) - Women	42.9	-	-	48.9	39.8
Share (%) - Tertiary education	34.1	30.5	38.8	-	-
Share (%) - Secondary or less	65.9	69.5	61.2	-	-
Observations	819,110	468,899	350,211	278,987	540,123

Source: Data from ENE surveys between 2010 and 2019.

Table 1 presents summary statistics for the sample used in the calibration, including the overall sample and breakdowns by gender and education level (tertiary education versus secondary education or less). Regarding employment and unemployment shares, nearly 68 percent of workers are employed in the formal sector, 21 percent in the informal sector, and the remaining 11 percent are unemployed. The shift in proportions compared to the entire dataset reflects the exclusion of partially employed individuals, which increases the relative unemployment share. Regarding gender differences, men are slightly more likely to be employed in both sectors, while women show a higher unemployment rate. Interestingly, workers with tertiary education also participate in the informal sector, though their participation is, as expected, lower than that of those with secondary education or less.

[9] The decision for tenure corresponds to drop observations on the top 26%. I make this decision to avoid considering workers unlikely to react to new offers and move to other jobs or receive unemployment shocks, like long term officials in the public sector, self-employed entrepreneurs, workers employed in a family business, among others. Regarding hours, after applying the other filters but before applying the filter on hours, 15% workers work less than 30 hours, and 5.5% work more than 60.

[10] Appendix 3: Filters to ensure consistency in tenure

The formal sector shows higher job tenure compared to the informal sector, which is expected as a consequence of the lower unemployment risk in the sector. When comparing genders, men tend to have lower tenure in the formal sector but experience fewer months of unemployment than women. When the sample is divided by education, workers with tertiary education degrees have longer tenures in both sectors, but they also experience longer periods of unemployment, a pattern documented previously in Fuentes (2019).

For the wage analysis, I use the employment dataset after applying the previously mentioned filters and merge it with wage information. I apply a series of adjustments to make the wages comparable to those in the administrative records.[11] Then calculate real wages using November 2019 as the baseline. Additionally, I focus on workers between the 6th and 95th percentiles of the wage distribution for each sector. Table 2 summarizes the income data. As shown, wages in the formal sector are higher than in the informal sector. Additionally, men earn higher wages than women in both sectors. In addition, there appears to be a wage premium associated with higher education in both sectors. It is also worth noting that these statistics for the formal sector align with those found in the administrative dataset. Summary statistics for tenure and wages in the administrative data can be found in the Appendix.[12]

Table 2: Summary statistics for income: 2010-2019

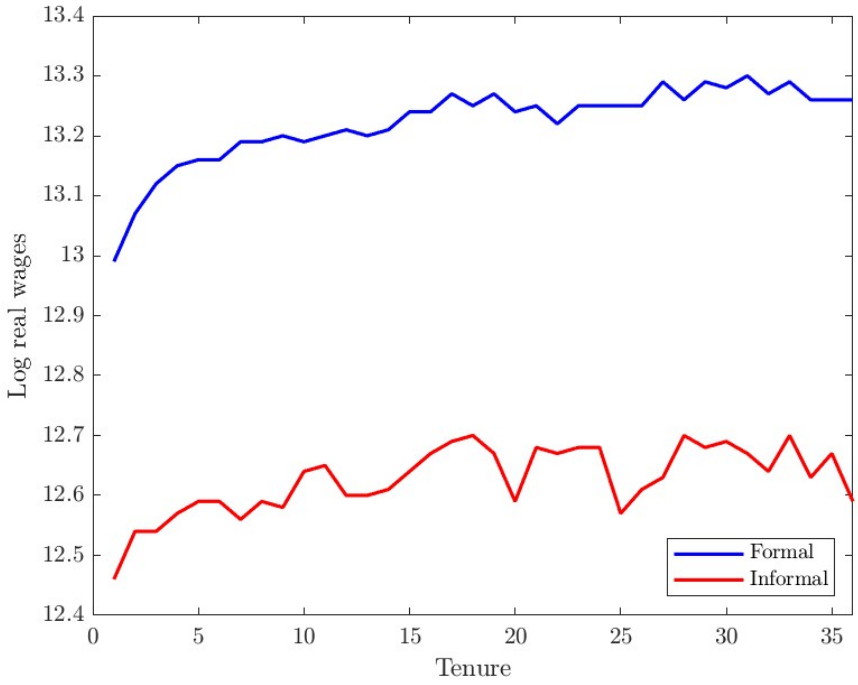
	Overall	Men	Women	Tertiary	Non-tertiary
Real log wage (Monthly) - Formal	13.23	13.29	13.16	13.58	13.05
Real log wage (Monthly) - Informal	12.61	12.67	12.52	12.89	12.53
Observations	164,259	95,754	68,505	53,872	110,387

Source: Data from ESI merged with ENE surveys between 2010 and 2019.

To conclude this section, I provide evidence of how wages and unemployment risk evolve with job tenure in the filtered dataset. Specifically, I calculate the average real log wage and employment-to-unemployment transition rates for each tenure level during the first three years. As shown in Figure 1, wages in the formal sector start higher than those in the informal sector, and wages in both sectors tend to increase with tenure. Wages in the formal sector grow approximately 27% during the three years, while wage growth is close to 12% in the informal sector. This difference can be explained through factors like faster human capital accumulation and better arrival of offers, allowing workers in the formal to re-bargain more frequently. An interesting observation is that even after three years, wages in the informal sector do not reach the levels seen in the formal sector during the first months of employment, which could be attributed to the minimum wage.

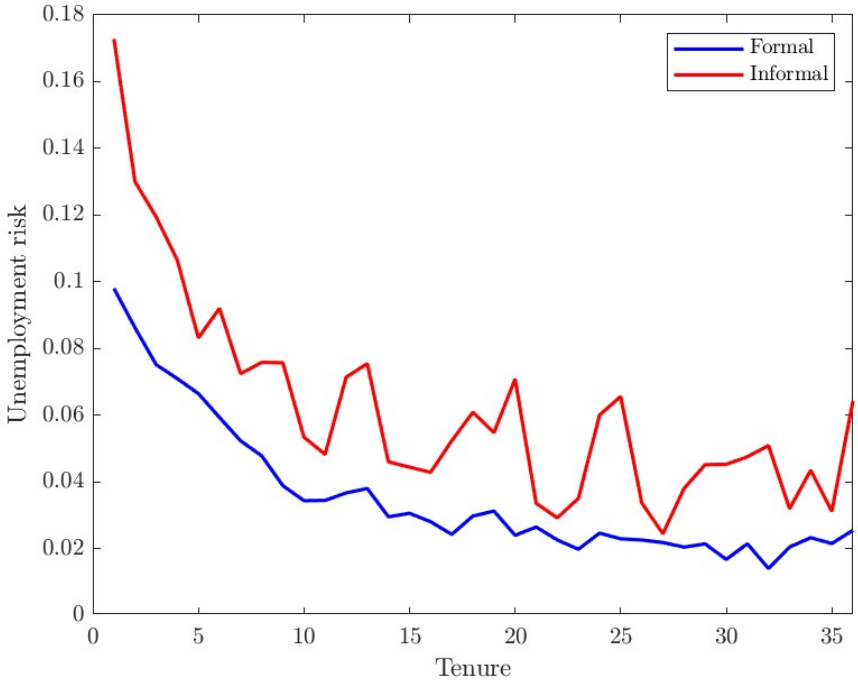
[11] Appendix 4: Adjustments to wages in income survey
[12] Appendix 5: Descriptive statistics administrative data

Figure 1: Evolution of real log wage per hour by tenure and sector (2010-2019)



Source: Data from ESI surveys between 2010 and 2019.

Figure 2: Employment to unemployment transitions by tenure and sector (2010-2019)



Source: Data from ENE surveys between 2010 and 2019.

Figure 2 illustrates the evolution of employment destruction rates across tenure in both sectors. As shown, informal jobs are consistently riskier than formal jobs, and both sectors experience a decline in unemployment risk with increasing tenure. However, this decrease in unemployment risk is more pronounced in the informal sector, decreasing from an unemployment risk of around 9.2% on average in the first year to an average of 4.3% during the third year, while these values for the formal sector are 5.8 and 2.1%, respectively. Also, an interesting feature in the data is that a worker with three years of tenure in the informal sector may face a lower unemployment risk than a newly hired worker in the formal sector. The above suggests that although formal firms are safer in the aggregate than informal firms, some informal firms are safer than formal firms at the base of the job ladder.

4 Model

4.1 Setting

Time is assumed to be discrete. Firms are characterized by their compliance status (formal or informal) and a vector $\theta = (\theta_y, \theta_\delta)$, where θ_y represents the firm's productivity and θ_δ denotes the exogenous job destruction rate associated with the firm. Workers search randomly for offers, samplying them from the same distribution, and can die with probability ρ. Workers can be in three possible states: unemployed (u), employed in the informal sector (i), and employed in the formal sector (f). Formal firms differ from the informal firms in two aspects: i) They sample their productivity and destruction rate parameters from different distributions, and ii) Formal firms must pay at least the minimum wage ($\overline{w}$).

When workers are unemployed, they receive a value z and they encounter job openings in formal and informal sector with probabilities λ_f and λ_i, respectively. When they are formally employed, they encounter job openings in formal and informal sectors with probabilities λ_{ff} and λ_{fi}. And, when they are informally employed, they encounter job openings in formal and informal sectors with probabilities λ_{if} and λ_{ii}.

Workers are endowed with a level of human capital h, which is observable to the firms and where $h \in \mathcal{H} = [\underline{h}, \overline{h}]$. Human capital is assumed to evolve following a process $G_e(h'|h)$ where $e \in \{f, i, u\}$. The above implies that human capital accumulation will differ in each possible state.

Once a worker and a firm form a match, output is determined by $p(\theta_y, h)$. If the match is dissolved, either because the worker transitions to another firm or due to exogenous destruction, the job opening holds no continuation value for the firm.

The timing of the model is as follows: First, the output $p(\theta_y, h)$ is observed and wages are paid. Second, workers may become unemployed due to the firm's exogenous destruction rate θ_δ. Third, based on their new state, human capital is updated according to the process $G_e(h'|h)$. Fourth, if workers remain employed, they may receive outside offers with sector-specific probabilities $(\lambda_{ff}, \lambda_{fi}, \lambda_{if}, \lambda_{ii})$ and can use these offers to renegotiate with their current firm or move to a new one. If workers are unemployed, they can receive offers from each sector with probabilities λ_f and λ_i.

In this model wages are restricted to fixed contracts, but can be renegotiated when workers receive an outside offer or experience an increase in their human capital. The wage determination for the informal sector follows the framework of Cahuc et al. (2006) as adapted in Jarosch (2023). In the formal sector, a modified version of this bargaining process is proposed to account for the minimum wage $(\overline{w})$.

4.2 Bargaining process

The bargaining process in this model has three main components. A first component includes the characteristics of the current firm in which a worker is employed, including compliance, productivity, and destruction rate denoted by c and θ. A second relevant component involves the characteristics of the best alternative offer received up to that point, or benchmark, which will be denoted by $\hat{c}$ and $\hat{\theta}$. The existence of a benchmark is relevant to generate the between-employer competition as presented in Cahuc et al. (2006) and Bagger et al. (2014). It is also important to note that the benchmark encompasses all possible firms in each sector, as well as unemployment. In this context, unemployment is treated as a "match" that provides an income of z and cannot be destroyed. Finally, the third component is the level of human capital h.

Let $W(c, \hat{c}, \theta, \hat{\theta}, h)$ denote the value of an employed worker in a firm with compliance status c, vector θ, and human capital h, and where $\hat{c}$ and $\hat{\theta}$ represent the sector and characteristics of the benchmark firm. Then $J(c, \hat{c}, \theta, \hat{\theta}, h)$ represents the value of a filled job, and $U(h)$ denote the value of unemployment.

With these components, it is possible to define the surplus of a match, $S(c, \theta, h)$, as:

$$S(c, \theta, h) \equiv max\{W(c, \hat{c}, \theta, \hat{\theta}, h) - U(h) + J(c, \hat{c}, \theta, \hat{\theta}, h), 0\}$$

It can be shown that this surplus depends solely on the current levels of θ and h, in addition to the compliance status c. The intuition behind this result is similar to that presented in Jarosch (2023): the benchmark firm's parameters determine the distribution of the surplus, not the total

amount of surplus generated. However, in this model with two sectors, the surplus also depends on the compliance status of the current firm. This is because the compliance status affects the probability of receiving future offers in each sector and because the minimum wage ($\overline{w}$) is applicable only in the formal sector.

If an unemployed worker with human capital h matches with an informal firm characterized by θ, the value of the employment match will be such that the worker receives a share α of the surplus, where α represents the worker's bargaining power. Then:

$$W(i, u, \theta, u, h) - U(h) = \alpha S(i, \theta, h)$$

This structure implies that workers will prefer firms with higher surpluses, as they will receive a greater amount when transitioning from unemployment to employment.

If an unemployed worker with human capital h matches with a formal firm, the same procedure applies as described above. However, if the worker's share α of the surplus is insufficient to meet the minimum wage, it is assumed that the worker's bargaining power increases to $\gamma(\theta, h)$. Here, $\gamma(\theta, h)$ represents the portion of the surplus that must be allocated to a worker with human capital h by a formal firm with characteristics θ to ensure that the minimum wage $\overline{w}$ is being paid. Then:

$$W(f, u, \theta, u, h) - U(h) = \max\{\alpha, \gamma(\theta, h)\} S(\theta, h)$$

Where the match will only occur if $0 < \gamma(\theta, h) < 1$.

This structure assumes that workers in both the formal and informal sectors will have the same bargaining power when the minimum wage is not binding. Although this assumption may not be evident, it is commonly used in the literature on labor informality to ensure that results are not skewed by differences in bargaining power across workers (Haanwinckel & Soares, 2021).

Next, consider a worker with human capital h currently employed at an informal firm with characteristics (i, θ), and using a benchmark firm with characteristics $(\hat{c}, \hat{\theta})$. If the worker remains with the firm (i, θ) rather than the benchmark firm $(\hat{c}, \hat{\theta})$, it implies that $S(i, \theta) > S(\hat{c}, \hat{\theta})$, otherwise, the worker would have chosen the benchmark firm. In this context, the between-employer competition implies that the current firm will outcompete the benchmark firm. To achieve this, the current firm will provide the worker with the entire surplus offered by the benchmark firm $S(\hat{c}, \hat{\theta})$, plus a share α of the surplus difference between the two firms $S(i, \theta, h) - S(\hat{c}, \hat{\theta}, h)$. Then:

$$W(i, \hat{c}, \theta, \hat{\theta}, h) - U(h) = S(\hat{c}, \hat{\theta}, h) + \alpha[S(i, \theta, h) - S(\hat{c}, \hat{\theta}, h)]$$

For a worker with human capital h currently employed at a formal firm with characteristics (f, θ), and using a benchmark firm with characteristics $(\hat{c}, \hat{\theta})$, the bargaining follows a similar logic, but we must account for the minimum wage $\overline{w}$. We need to consider two scenarios. In the first case, where the minimum wage is not binding, the worker receives:

$$NB = S(\hat{c}, \hat{\theta}, h) + \alpha[S(f, \theta, h) - S(\hat{c}, \hat{\theta}, h)]$$

In the second case, the minimum wage is binding, and therefore, the worker receives:

$$B = \gamma(\theta, h) S(f, \theta, h)$$

Then considering both cases, the value of being employed at a formal firm is the maximum of the two:

$$W(f, \hat{c}, \theta, \hat{\theta}, h) - U(h) = max\{NB, B\}$$

Regarding the decision process, consider a worker employed at a firm with characteristics (c, θ) and a benchmark $(\hat{c}, \hat{\theta})$ who receives an offer from a firm with characteristics (c', θ'). The procedure is as follows: If (c', θ') is a formal firm, we need to compute $\gamma(\theta', h)$. If $0 < \gamma(\theta', h) < 1$ the offer is considered feasible, otherwise the match is not possible. For a feasible offer, whether from the formal or informal sector, we have three cases to examine:

1. If $S(c', \theta', h) > S(c, \theta, h)$, the worker moves to θ'. The set of all options that result in a movement is defined as M. Then, in the case (c', θ', h) is informal the new value of employment is given by:

   $$W(i, c, \theta', \theta, h) - U(h) = S(c, \theta, h) + \alpha \left[S(i, \theta', h) - S(c, \theta, h)\right]$$

   And in the case (c', θ', h) is formal:

   $$W(f, c, \theta', \theta, h) - U(h) = max\{NB, B\}$$

   Where each condition is given by:

   $$\begin{aligned} NB &= S(c, \theta, h) + \alpha[S(f, \theta', h) - S(c, \theta, h)] \\ B &= \gamma(\theta', h) S(f, \theta', h) \end{aligned}$$

2. If $S(c, \theta, h) > S(c', \theta', h) > S(\hat{c}, \hat{\theta}, h)$ then the worker re-bargains, changing the benchmark from $(\hat{c}, \hat{\theta})$ to (c', θ'). The set of all options that result in a re-bargain is defined as RB.

Then, the value of employment will be given by:

$$\begin{aligned} W(i,c,\theta,\theta',h) - U(h) &= S(c,\theta',h) + \alpha\left[S(i,\theta,h) - S(c,\theta',h)\right] \\ W(f,c,\theta,\theta',h) - U(h) &= max\{NB, B\} \end{aligned}$$

Where each condition is given by:

$$\begin{aligned} NB &= S(c,\theta',h) + \alpha[S(f,\theta,h) - S(c,\theta',h)] \\ B &= \gamma(\theta,h)S(f,\theta,h) \end{aligned}$$

Note that this process means that whenever a worker re-bargains using a better benchmark in the informal sector, it results in an improved value of employment. However, this is not always the case for a worker in the formal sector. For example, if the minimum wage is binding for the worker and they receive an offer from another firm where the minimum wage is also binding, but this new firm offers a better benchmark, the worker would retain the current value of employment. In this situation, only the benchmark changes, with no impact on the value of employment.

3. If $S(c,\theta,h) > S(\hat{c},\hat{\theta},h) > S(c',\theta',h)$ then nothing happens, and the worker keeps the current firm and benchmark.

4.3 Value functions

Given the setting mentioned above, then the value of unemployment is given by:

$$\begin{aligned} U(h) = z + \beta(1-\rho)\int_{\mathcal{H}} \Bigg(\lambda_f \int_M \left[W(f,u,x,u,h') - U(h')\right] dF_f(x) \ ... \\ + \lambda_i \int_M \left[W(i,u,x,u,h') - U(h')\right] dF_i(x) + U(h') \Bigg) dG_u(h'|h) \end{aligned}$$

The intuition behind this equation is as follows: At the beginning of a period, the unemployed worker receives the value z. In the next period, there is a probability λ_f of receiving a formal job offer. Any offer that provides a value higher than the current value of unemployment (equivalent to a positive surplus) will be accepted. When a worker transitions from unemployment to employment, the benchmark is unemployment. Similarly, with probability λ_i, the worker may receive an informal offer, which follows the same process. Additionally, there is a possibility of not receiving any offer and remaining unemployed into the next period, where the worker faces a risk of losing human capital, as determined by the process $G_u(h'|h)$.

The value of being employed in the formal sector is given by:

$$W(f,\hat{c},\theta,\hat{\theta},h) = \max\{w(f,\hat{c},\theta,\hat{\theta},h),\overline{w}\} + \beta(1-\rho)\int_{\mathcal{H}}\Bigg[(1-\theta_\delta)\times ...$$
$$\Bigg\{\lambda_{ff}\left(\int_M W(f,f,x,\theta,h')dF_f(x) + \int_{RB} W(f,f,\theta,x,h')dF_f(x)\right) + ...$$
$$\lambda_{fi}\left(\int_M W(i,f,x,\theta,h')dF_i(x) + \int_{RB} W(f,i,\theta,x,h')dF_i(x)\right)\Bigg\} + ...$$
$$\left(1-\lambda_{ff}\int_{M\cup RB} dF_i(x) - \lambda_{fi}\int_{M\cup RB} dF_f(x)\right)\times W(f,\hat{c},\theta,\hat{\theta},h) + \theta_\delta U(h')\Bigg]dG_f(h'|h)$$

A worker employed in the formal sector will receive the higher of either the wage function or the minimum wage ($\overline{w}$). In the next period, if the job is not exogenously destroyed, the worker may receive an offer from either the formal or informal sector. Given such an offer, the worker can choose to move to a new job or use the offer to renegotiate their current condition. If no offer is received, the worker remains with the same firm in the following period. If the worker continues employed, there is a chance of increasing their human capital according to the updating process of the formal sector. However, if the job is exogenously destroyed, the worker begins the next period in unemployment.

Analogously, the value of employment in the informal sector is given by:

$$W(i,\hat{c},\theta,\hat{\theta},h) = w(i,\hat{c},\theta,\hat{\theta},h) + \beta(1-\rho)\int_{\mathcal{H}}\Bigg[(1-\theta_\delta)\times ...$$
$$\Bigg\{\lambda_{if}\left(\int_M W(f,i,x,\theta,h')dF_f(x) + \int_{RB} W(i,f,\theta,x,h')dF_f(x)\right) + ...$$
$$\lambda_{ii}\left(\int_M W(i,i,x,\theta,h')dF_i(x) + \int_{RB} W(i,i,\theta,x,h')dF_i(x)\right)\Bigg\} + ...$$
$$\left(1-\lambda_{fi}\int_{M\cup RB} dF_i(x) - \lambda_{ii}\int_{M\cup RB} dF_f(x)\right)\times W(i,\hat{c},\theta,\hat{\theta},h) + \theta_\delta U(h')\Bigg]dG_i(h'|h)$$

The primary difference between this value function and the one for the formal sector is that the probabilities of receiving offers shift from the pair $(\lambda_{ff},\lambda_{fi})$ to $(\lambda_{if},\lambda_{ii})$. Additionally, there is no minimum wage constraint in the informal sector.

Regarding firms, we have that the value of a formal filled vacancy is given by:

$$J(f,\hat{c},\theta,\hat{\theta},h) = p(\theta_y,s) - \max\{w(f,\hat{c},\theta,\hat{\theta},h),\overline{w}\} + \beta(1-\rho)...$$
$$\int_{\mathcal{H}}(1-\theta_\delta)\Bigg[\lambda_{ff}\int_{RB} J(f,f,\theta,x,h')dF_f(x) + \lambda_{fi}\int_{RB} J(f,i,\theta,x,h')dF_i(x) \ + \ ...$$
$$\left(1-\lambda_{ff}\int_{M\cup RB} dF_f(x) - \lambda_{if}\int_{M\cup RB} dF_i(x)\right)\times J(f,\hat{c},\theta,\hat{\theta},h')\Bigg]dG_f(h'|h)$$

In this scenario, a formal firm receives the value of production and pays the higher of either the unconstrained wage or the minimum wage. If the job is not exogenously destroyed in the next period, the worker may receive offers. If the worker chooses to move to a new job, the vacancy is terminated and holds no continuation value for the firm. The firm retains value from the filled vacancy only when the worker either uses the new offers to renegotiate or receives no offers at all. Additionally, if the worker remains employed, there is a probability that their level of human capital will increase.

Following the same reasoning, the value of a filled vacancy in the informal sector is given by:

$$
\begin{aligned}
&J(i,\hat{c},\theta,\hat{\theta},h) = p(\theta_y,s) - w(i,\hat{c},\theta,\hat{\theta},h) + \beta(1-\rho)... \\
&\int_S (1-\theta_\delta)\left[\lambda_{if}\int_{RB} J(i,f,\theta,x,h')dF_f(x) + \lambda_{ii}\int_{RB} J(i,i,\theta,x,h')dF_i(x) + ...\right. \\
&\left.\left(1-\lambda_{if}\int_{M\cup RB} dF_f(x) - \lambda_{ii}\int_{M\cup RB} dF_i(x)\right) \times J(i,\hat{c},\theta,\hat{\theta},h')\right) dG_i(h'|h)
\end{aligned}
$$

As with the value of employment, the key differences are the absence of a minimum wage and the probabilities of receiving offers from the informal sector.

Next, the last group of value functions needed in order to solve the model are the surplus functions. This functions represent the total gains obtain from the match are then are going to be divided between the firm and the worker, according to the bargaining rules. Following the definition of surplus, it can be shown that the surplus in the formal sector is given by[13]

$$
\begin{aligned}
&S(f,\theta,h) = p(\theta_y,s) - z + \beta(1-\rho)\int_H (1-\theta_\delta)\Big(S(f,\theta,h')... \\
&+\lambda_{ff}\int_M \Big[\max\{S(f,\theta,h') + \alpha[S(f,x,h') - S(f,\theta,h)], \gamma(x,h')S(f,x,h')\} - S(f,\theta,h')\Big] dF_f(x)... \\
&+\lambda_{fi}\int_M \alpha[S(i,x,h') - S(f,\theta,h')]dF_i(x)\Big)\, dG_f(h'|h)... \\
&-\beta(1-\rho)\int_H \left\{\lambda_f\int_M [\max\{\alpha,\gamma(\theta,h')\}S(f,x,h')dF_f(x) + \lambda_i\int_M \alpha S(i,x,h')dF_i(x)]\right\} dG_u(h'|h)... \\
&+\beta(1-\rho)\int_H U(h')dG_f(h'|h) - \beta(1-\rho)\int_H U(h')dG_u(h'|h)
\end{aligned}
$$

As mentioned earlier, the surplus does not depend on the benchmark values $(\hat{c},\hat{\theta})$.

Conceptually, first it captures the difference between the production at the current job and unemployment flows. Second, it shows the difference in gains from doing on-the-job versus off-the-job search. This difference will depend significantly on which state provides more access to offers and

[13] This procedure is shown in Appendix 6: Surplus proof

the quality of those offers. This process involves comparing the pair (λ_f, λ_i) with $(\lambda_{ff}, \lambda_{fi})$ and $(\lambda_{if}, \lambda_{ii})$. The intuition is that if the pair (λ_f, λ_i) strictly dominates the other two, workers will demand better jobs since accepting a job reduces their chances of receiving superior offers later on. Conversely, if (λ_f, λ_i) is strictly dominated by the other two, workers may accept lower-quality offers, recognizing that they will gain more search opportunities while employed.

Third, the surplus also accounts for human capital accumulation gains in the formal sector compared to unemployment. Since the literature indicates that human capital tends to erode during unemployment, this difference is likely to be positive.

Finally, the surplus only depends on the movement set (M) and not on the re-bargain set (RB). This result, again, is intuitive since re-bargaining does not increases the surplus, just the way in which is divided between parts.

Then, the surplus for the informal sector is given by:

$$
\begin{aligned}
&S(i,\theta,h) = p(\theta_y, s) - z + \beta(1-\rho)\int_H (1-\theta_\delta)\Bigg(S(i,\theta,h') + \lambda_{ii}\int_M \alpha[S(\tilde{i},x,h') - S(f,\theta,h')]dF_i(x)... \\
&+\lambda_{if}\int_M \Big[\max\{S(i,\theta,h') + \alpha[S(\tilde{f},x,h') - S(i,\theta,h)], \gamma(x,h')S(\tilde{f},x,h')\} - S(f,\theta,h')\Big]dF_f(x)\Bigg)dG_i(h'|h)... \\
&-\beta(1-\rho)\int_H \left\{\lambda_f \int_M [\max\{\alpha, \gamma(\theta,h')\}S(f,x,h')dF_f(x) + \lambda_i \int_M \alpha S(i,x,h')dF_i(x)]\right\} dG_u(h'|h)... \\
&+\beta(1-\rho)\int_H U(h')dG_i(h'|h) - \beta(1-\rho)\int_H U(h')dG_u(h'|h)
\end{aligned}
$$

4.4 Surplus properties and wages

These functional equations can be solved jointly with the bargaining protocol and the value of unemployment. A key property of the surpluses is that they summarize the value of a job by considering the three-dimensional ladder components: sector, productivity, and destruction rate. Given a set of model parameters, value function iteration can be used to determine the surpluses that generate a steady-state equilibrium in the model, allowing for the simulation of labor histories. However, there is a difference between the procedure outlined in Jarosch (2023) and the one presented here. In the former, wage computations are unnecessary for obtaining surpluses, whereas in this model, wage calculations are needed to ensure that formal firms are complying with the minimum wage requirement.

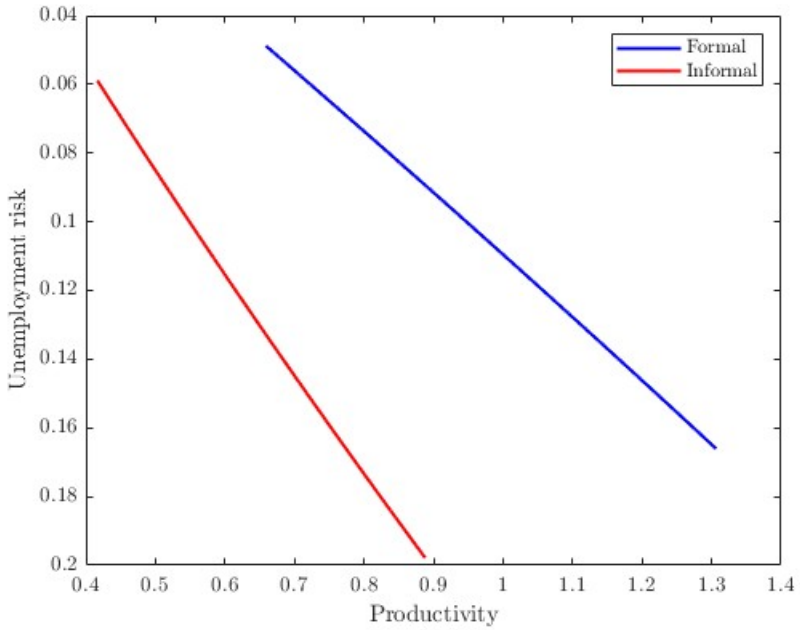
4.5 Indifference curves

To illustrate how the movement across firms works, Figure 3 shows the indifference curves using two examples from the calibrated model. As can be seen, there is a trade-off between unemploy-

ment risk and productivity, and the firms with the best surpluses are those with high levels of both variables, as expected.

To illustrate the job ladder component, Figure 4 shows a surface plot of the calibrated surpluses in each sector. Darkest colors imply a smaller surplus, while lighter colors are associated with higher surpluses. Here we can see that the job ladder is taller in the formal sector, which is as expected due the better characteristics of the formal sector. Notice that in the formal sector, the combination of the lowest unemployment risk and lowest productivity delivers a higher surplus than the combination of the higher productivity and the highest unemployment risk. The central intuition to explain this is that the lowest unemployment risk allows access to the search gains, e.g., allows the worker to do on-the-job search and improve their situation through bargaining.

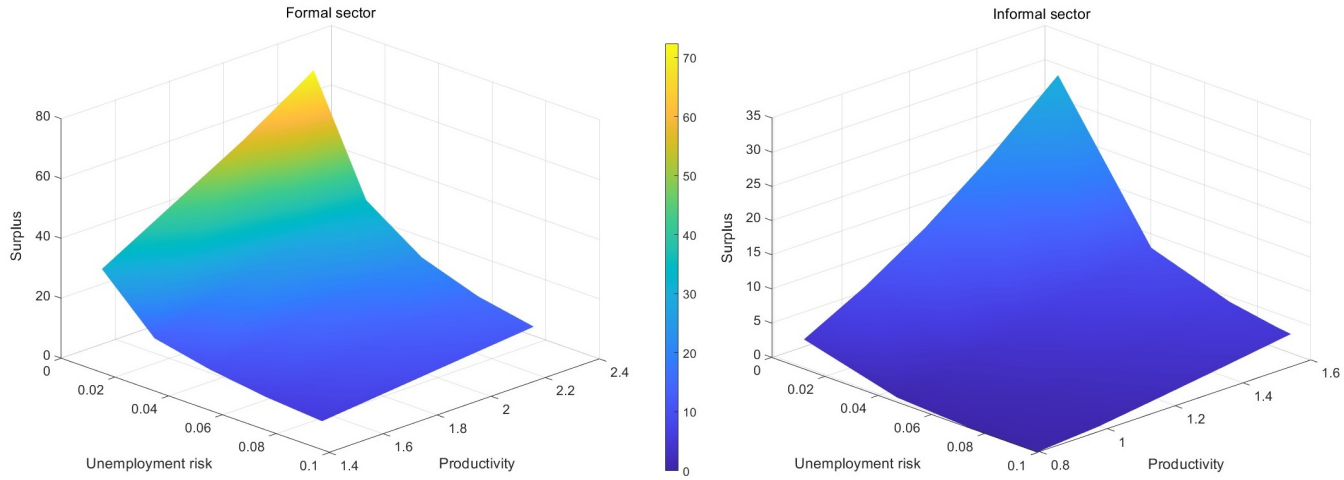
Figure 3: Indifference curves in both sectors



4.6 Relevant properties

This model presents several relevant characteristics. First, in the spirit of Jarosch (2023), the heterogeneity in productivity and unemployment risk across firms generates a job ladder, which base is particularly slippery since it is built from firms with high unemployment risk. Second, in this type of job ladder models there is no clear relationship between firm productivity and wages, since this relation depends on the bargaining power of workers and the potential search gains on-

Figure 4: Job ladder in both sectors



the-job (Cahuc et al., 2006).[14] However assuming that workers have the same bargaining power in both sectors and that the formal sector should provide higher search gains, the formal sector should exhibit lower wages than those observed in the informal sector. Nevertheless, the existence of a minimum wage in that sector restricts firms of taking this action and instead, forces them to grant a higher *de facto* bargaining power to workers as in Flinn (2006). This higher bargaining power will induce not only higher wages in the formal sector, which is something observed in the data, but also will make individuals more picky about taking an informal job, generating that informal firms will have to increase their wages as well to remain competitive.

5 Estimation

5.1 Functional assumptions and strategy

To solve the model, I will make some additional assumptions to ensure tractability. First, I assume there are only five productivity types and five destruction rate types in each sector. This creates 25 firm types per sector, resulting in 50 total firm types. The productivity and destruction rate for each firm will be randomly drawn from independent Beta distributions, with separate distributions for the formal and informal sectors. Regarding the Beta distributions for the destruction rates, these will be adjusted to deliver values between 0 and 0.2.[15] These four Beta distributions will be discretized. Regarding the output function, the functional form will be additively separable,

[14] If a worker has no bargaining power, firms will consistently pay the minimum amount possible, and therefore, more productive firms potentially can pay the lowest wages to workers coming from unemployment. Conversely, more productive firms will pay the highest wages when the worker has all the bargaining power.

[15] As I showed in Figure 2, the average destruction rates move between 0.17 and 0.04 for the informal sector and 0.1 and 0.02 for the formal sector.

$p_c(\theta, h) = \underline{p}_c + s + \theta_y$, where there will be two shifters, one for each sector.

Human capital will be discretized into five levels, i.e, $h \in \{1, 2, 3, 4, 5\}$. The updating of human capital will follow a Markov process, where workers have a probability ψ_f of increasing their human capital in one unit in the formal sector and a probability ψ_i in the informal sector. Additionally, unemployed workers will experience a decrease in their human capital in one unit with probability ψ_u.

With these assumptions, the set of parameters to estimate is given by Table 3.

Table 3: Parameters to estimate

Parameter	Description
λ_f	Formal offer arrival during unemployment
λ_i	Informal offer arrival during unemployment
λ_{ff}	Formal offer arrival during formal employment
λ_{fi}	Informal offer arrival during formal employment
λ_{if}	Formal offer arrival during informal employment
λ_{ii}	Informal offer arrival during informal employment
$\eta_{\delta f}, \mu_{\delta f}$	Unemployment risk distribution formal sector
$\eta_{\delta i}, \mu_{\delta i}$	Unemployment risk distribution informal sector
α	Bargaining power of workers
η_{yf}, μ_{yf}	Productivity distribution formal sector
η_{yi}, μ_{yi}	Productivity distribution informal sector
$\underline{p}_f$	Common output shifter formal sector
$\underline{p}_i$	Common output shifter informal sector
ψ_f	Skill accumulation during formal employment
ψ_i	Skill accumulation during informal employment
ψ_u	Skill depreciation during formal unemployment
z	Flow value of unemployment

For the estimation strategy, I will follow Jarosch (2023) and use the Simulated Method of Moments with indirect inference to estimate the model parameters on a monthly basis, including moments to match for the informal sector. The method minimizes the difference between the observed (or indirectly inferred) moments in the data and those generated by the model simulations. The model is estimated in a steady state, meaning that the flows across different states are balanced. The minimum wage will be normalized to 1.

5.2 Moments

The model has 21 parameters and is calibrated using 26 moments. The first two moments (1-2) capture unemployment-to-employment (UE) transitions into the formal and informal sectors. These moments correctly identify the arrival rates of offers, denoted by λ_f and λ_i.

The next four moments (3-6) are associated with the employment-to-employment (EE) transitions between sectors. These moments are monotonically related to the parameters λ_{ff}, λ_{fi}, λ_{if}, and λ_{ii}.

The next two moments (7-8) are related with the average job destruction rate in each sector. These moments inform the first shape parameter of the Beta distributions in each sector $\eta_{\delta f}$ and $\eta_{\delta i}$. To inform about the second shape parameter, I compute six moments (9-14) on how the unemployment risk changes through tenure in the labor market. To obtain this, I estimate the following regression for each sector:

$$I_{it}^{EU} = \alpha_0 + \sum_{\tau=1}^{\tau_{max}} \beta_\tau D_{it}^{\tau} + X_{it} + \varepsilon_{it}$$

Where I_{it}^{EU} is a binary variable indicating whether worker i transitioned to unemployment in period t. The variable D_{it}^{τ} is a binary variable indicating whether the worker has been employed at the same firm for D consecutive quarters. The coefficients β_τ reflect how job tenure reduces the destruction rate faced by workers. The matrix X includes a set of control variables such as gender dummies, month and year fixed effects, age group dummies (16-24, 25-34, 35-44, 45-54, and 55-64), education categories, and interactions between gender and those categories.

Then I will compute the average of the β_τ coefficients for the first, second, fourth and sixth year of tenure. Then, the moments used to calibrate the parameters $\mu_{\delta f}$ and $\mu_{\delta i}$ will be the differences between the averages of the last three groups and the first one:

$$\frac{1}{4}\sum_{\tau=5}^{8}\hat{\beta}_\tau - \frac{1}{4}\sum_{\tau=1}^{4}\hat{\beta}_\tau, \quad \frac{1}{4}\sum_{\tau=13}^{16}\hat{\beta}_\tau - \frac{1}{4}\sum_{\tau=1}^{4}\hat{\beta}_\tau, \quad \frac{1}{4}\sum_{\tau=21}^{24}\hat{\beta}_\tau - \frac{1}{4}\sum_{\tau=1}^{4}\hat{\beta}_\tau$$

Regarding wages, I will jointly calibrate the parameters α, $\underline{p}_f$, $\underline{p}_i$, and the productivity pairs in both sectors (η_{yf}, μ_{yf}) and (η_{yi}, μ_{yi}) using ten moments (15-24). The first pair are the mean log wages in each sector. The second and third pairs represent the differences between residulized log wages at the 90th and 50th percentiles, and the 50th and 10th percentiles, in both sectors. The obtain these residuals I estimate both in the real and simulated data the following regression:

$$w_{it} = \alpha_0 + \gamma_i + \sum_{y=1}^{9} \beta_i D_{it}^{y}$$

And then I compute the difference between the w_{it} and the estimated value. Where γ_i represent individual fixed effects and D_{it}^{y} is a set of binary variables identifying years.

The fourth pair captures the annual wage growth in each sector. The final pair reflects the wage gap in both sectors, defined as:

$$\mathrm{G}_c = 1 - \frac{\overline{w}_c^0}{\overline{w}_c}$$

Where $\overline{w}_c^0$ is the average wage for those starting a job after unemployment in the sector c, and $\overline{w}_c$ is the average wage in that sector.

Although all moments are influenced by the aforementioned parameters, some parameters affect certain moments more than others, which provides intuition about their relationships. The average log wages primarily inform the output shifters and the first shape parameter of both productivity distributions. The differences in residualized log wages across percentiles help identify the second shape parameters of the productivity distributions. Meanwhile, the wage growth rates and wage gaps in both sectors are informative about bargaining power and both parameters of the productivity distributions.

For the loss of human capital while unemployed, I will estimate the following regression:

$$w_{it}^0 = \alpha_0 + \gamma_1 \tau_{it}^u + \zeta_1 \overline{w}_i + \varepsilon_{it}$$

Where w_{it}^0 is the log wage after an unemployment spell, τ_{it} are the months of duration of the unemployment spell, and $\overline{w}$ is the average log wage of the worker. Then the moment $\hat{\gamma}_1$ (25) relates monotonically to the parameter ψ_u. It is important to note that, due to the nature of the data for the informal sector, it is not possible to accurately estimate this regression using data from that sector. However, since the model assumes that unemployment affects workers from both the formal and informal sectors equally, this parameter can be estimated using only administrative data from the formal sector.

For the accumulation of human capital, I will estimate the following regression:

$$w_{it}^t = \alpha_0 + \gamma_2 \tau_{it}^t + \zeta_2 w_{it} + \varepsilon_{it}$$

Where w_{it}^t is the log wage after t periods of continued tenure, τ^t is measuring those months, and w_{it} is the wage at the first period of employment in that firm. Then the moments γ_2 (26-27) identify the components ψ_f and ψ_i. Here, it is possible to estimate the parameters in both sectors with the data available.

Finally, for the parameter z, I will use the observed minimum wage in the informal sector as

moment (28) to calibrate it. The intuition is that when workers decide to take a job, they are forgoing the flow z, making this parameter relevant to determine the indifference point between working and not working. Since the model includes a minimum wage in the formal sector, that sector cannot be used to assess the opportunity cost of working. However, this is not the case of the informal sector, and therefore the minimum wage observed there is informative of z.

5.3 Firms in equilibrium

The algorithm used to solve the model can be found in the Appendix.[16] Although the model has a fixed number of firms, not all of them necessarily operate in equilibrium. This outcome can be driven by two factors. First, a formal firm may not participate in the market if, given the production parameters ($\underline{p}_f$, η_f, and μ_f), it is unable to pay at least the minimum wage. In the case of an informal firm, it may be absent from the market equilibrium if, given its production parameters ($\underline{p}_i$, η_i, and μ_i), it cannot offer a wage high enough to compensate unemployed workers for forgoing the flow z.

5.4 Simulation

Once the model has assigned a set of parameters, it will simulate labor histories for 10,000 workers over 420 months (representing 35 years). The workers will start unemployed with the lowest level of human capital and will receive offers and make employment decisions accordingly. The moments will be computed using the entire simulated dataset, aligning with the procedure applied to the real data. It is important to note that although the event of receiving an offer is random across simulations for the same worker, the firm associated with that offer is fixed across simulations. This means that if a worker, say W, receives an offer in the first period, it will come from the same firm in both simulations, even though W might not receive an offer at all in some simulations. This approach ensures that differences in the estimated parameters are not attributable to random variations in the simulations.[17]

6 Calibration

The moments used and the calibrated parameters can be seen in Table 4. First of all, it is important to mention that the overall calibration has a good fit. In general terms, the fit is very good for the informal sector, while it has some differences with the formal one.

Regarding the job finding rates, the model successfully replicates the moments observed in the data. The relation $\lambda_f > \lambda_i$, indicating that unemployed workers find it easier to secure a formal

[16] Appendix 6: Surplus proof

[17] More details about the simulation can be found in Appendix 7: Algorithm.

job than an informal one, and $\lambda_{if} < \lambda_f$, meaning it is easier to find a formal job while unemployed than when employed in the informal sector, implies a higher value of unemployment relative to being employed in the informal sector. Because of the above, workers entering the informal sector will require significant compensation to offset the reduced likelihood of transitioning to the formal sector. Concerning the employment-to-employment transitions, the model correctly fits three of the four transitions. The fact that $\lambda_{ff} > \lambda_f$ and $\lambda_{ii} > \lambda_i$ implies both sectors are more effective than unemployment in generating offers inside the sector. Also, we observe $\lambda_{fi} < \lambda_i$ and $\lambda_{if} < \lambda_f$, meaning that both sectors are less efficient that unemployment in generating offers in the other sector.

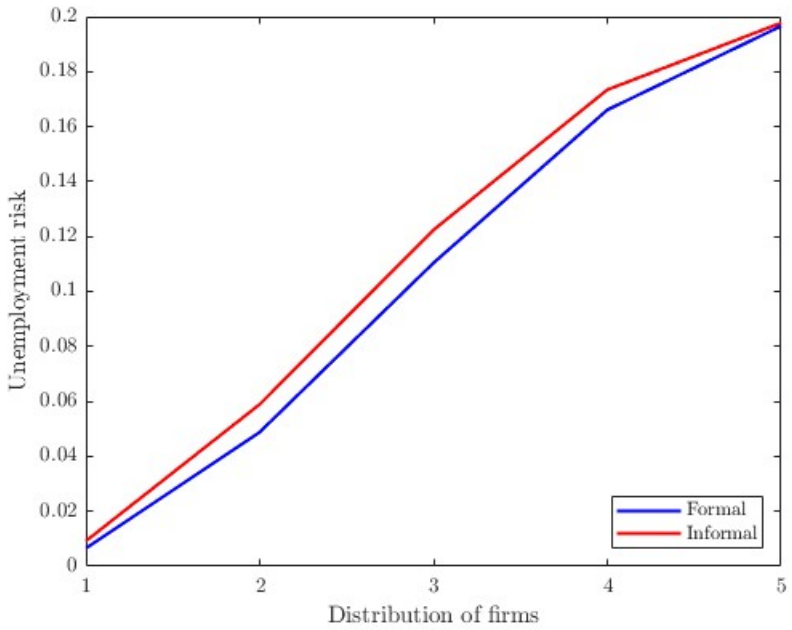
Table 4: Parameters and moments used to calibrate model

Moment	Target	Model	Parameter
Job finding rate - Formal	9.5%	9.5%	$\lambda_f = 0.122$
Job finding rate - Informal	5.7%	5.7%	$\lambda_i = 0.098$
E-E Transitions - FF	1.5%	1.5%	$\lambda_{ff} = 0.136$
E-E Transitions - FI	0.4%	0.3%	$\lambda_{fi} = 0.020$
E-E Transitions - IF	2.1%	1.7%	$\lambda_{if} = 0.066$
E-E Transitions - II	2.3%	1.3%	$\lambda_{ii} = 0.122$
Mean job loss - Formal	1.3%	2.2%	$\eta_{\delta,f} = 0.525$
Mean job loss - Informal	2.9%	2.9%	$\eta_{\delta,f} = 0.560$
Decay destruction - Formal (2 years)	-3.10%	-3.30%	$\mu_{\delta,f} = 0.469$
Decay destruction - Formal (4 years)	-3.78%	-4.51%	
Decay destruction - Formal (6 years)	-4.41%	-4.71%	
Decay destruction - Informal (2 years)	-4.26%	-3.39%	$\mu_{\delta,i} = 0.438$
Decay destruction - Informal (4 years)	-4.36%	-4.55%	
Decay destruction - Informal (6 years)	-5.39%	-4.62%	
Average log wages - Formal	13.21	13.08	$\underline{p}_f = 0.438$
Average log wages - Informal	12.62	12.58	$\underline{p}_i = 0.250$
$w_{90} - w_{50}$ - Formal	0.184	0.182	$\eta_{y,f} = 1.883$
$w_{50} - w_{10}$ - Formal	0.183	0.278	$\mu_{y,f} = 1.508$
$w_{90} - w_{50}$ - Informal	0.204	0.359	$\eta_{y,i} = 1.397$
$w_{50} - w_{10}$ - Informal	0.204	0.436	$\mu_{y,f} = 1.195$
Wage growth - Formal	9.1%	3.4%	
Wage growth - Informal	13.2%	12.4%	
Wage gap - Formal	25.0%	23.2%	$\alpha = 0.505$
Wage gap - Informal	14.9%	14.8%	
Human capital acc - Formal	0.17%	0.26%	$\psi_f = 0.180$
Human capital acc - Informal	0.15%	0.16%	$\psi_i = 0.032$
Human capital destruction	-0.25%	-0.27%	$\psi_u = 0.101$
Minimum wage - Informal sector	11.57	10.12	$z = 12.59$

Concerning the moments associated with the destruction rate distributions, the model overestimates the average job loss for the formal sector. However, it correctly captures that the mean destruction rate is significantly higher in the informal sector than in the formal sector. Also, the

mean rate for the informal sector is correctly matched. As for the decline in unemployment risk, the model shows that the risk decreases at a similar rate in both sectors, correctly matching the moments in the formal sector but underestimating the decay in the informal sector by one percentage point.

Figure 5: Distribution of destruction rates across types of firms



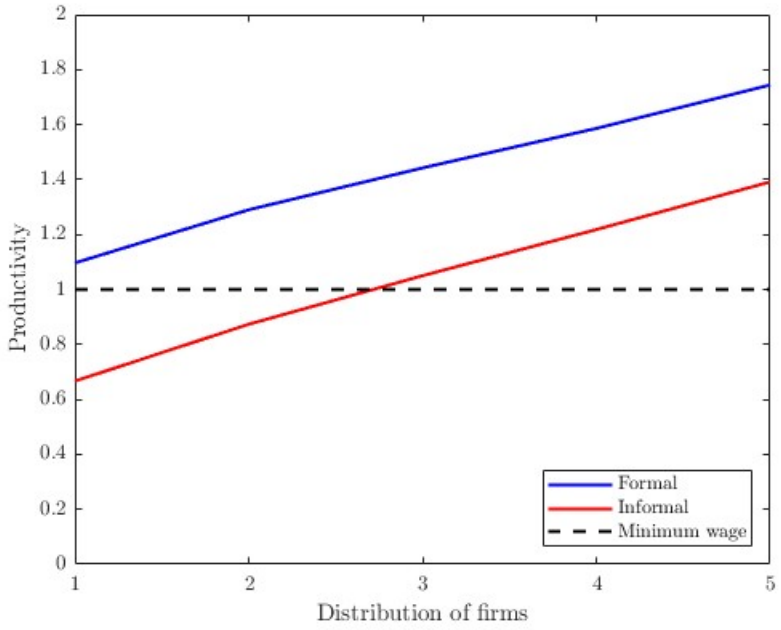
The unemployment risk distribution in both sectors is fairly similar, as illustrated in Figure 5, which depicts the discretized Beta distributions for unemployment risk in both sectors. This suggests that the differences between formal and informal firms are not primarily driven by employment stability. This finding aligns with Arellano & Jimenez (2016), who show considerable dispersion in the annual destruction rates of firms within the formal sector, ranging from 1.3 to 12.0 percent in 2013.

For the wage-related moments, the model shows a relatively good fit for average wages. In terms of residualized wages, the model overestimates the percentile differences, although it is relevant to keep in mind that those are residual wages. Regarding annual wage growth in both sectors, the model replicates the qualitative trend that wage growth is higher in the informal sector, though the formal sector is underestimated. For wage gaps, the model correctly matches the moments.

Regarding the parameters associated with productivity, the output shifter in the formal sector

($\underline{p}_f$) is considerably higher than in the informal sector ($\underline{p}_i$). As we can see in Figure 6, the productivity in the formal sector is consistently higher than in the informal sector, but they share the degree of dispersion. The estimated bargaining power is relatively low compared to what is typically found in the literature (for example, Jarosch (2023) found $\alpha = 0.92$ for Germany). Unfortunately, there are no comparable studies on bargaining power in Chile to further contextualize this result. Also, although all the firms' productivity is higher than the minimum wage in the formal sector, this does not imply that the minimum wage is not binding for some firms. The above is because firms at the bottom of the job ladder will have significant search gains. Since these gains are part of the surplus, it is not evident that they also have enough space to pay the minimum wage.

Figure 6: Distribution of productivity across types of firms



Concerning the human capital upgrading process, the model accurately matches the moments observed in the data for the informal sector and unemployment, although it overestimates the moment for the formal sector. As expected, $\psi_f > \psi_i$, indicating that human capital accumulation is faster in the formal sector, which aligns with the literature (Bobba et al., 2021). This shows that an advantage of working in the formal sector is the access to a faster rate of human capital accumulation. The human capital destruction rate, ψ_u, is greater than accumulation in the informal sector.

The model underestimates the minimum wage observed in the informal sector. About the minimum wage observed in the informal sector, the model underestimates this value. A value z (12.59) is higher than the minimum wage observed in the informal sector in the model (10.12) implies that workers are willing to accept wages below their unemployment flow because they anticipate

accumulating human capital and receiving better offers in the future.

Finally, Table 5 presents the distribution of workers in both states, according to the data and the model. The model accurately predicts the number of workers in each state.

Table 5: Distribution of workers across states

State	Data	Model
Employment - Formal	67.5%	67.8%
Employment - Informal	21.6%	21.3%
Unemployment	10.9%	10.9%

7 The value of a formal job

With the calibrated model, it is possible to estimate the value of a formal job. Since the question of what constitutes the value of a formal job can be interpreted in various ways, this section will explore different approaches to addressing it.

First, I will take a theoretical approach by comparing the value functions for workers employed in both sectors. Second, I will run a simulation where half of the workers are guaranteed an offer from a formal firm, while the other half receive an offer from an informal firm, and I will compare labor outcomes 1, 5, and 10 years later. Lastly, I will present an exercise to understand the relative importance of different aspects of the formal and informal sector in the total earnings.

7.1 Differences in value functions

Using the value functions for employed workers provides a direct way to compare the value of formal and informal jobs. However, a few important considerations must be kept in mind. The results are presented in units of the minimum wage in 2019.[18] Second, these value functions apply to infinitely lived individuals, meaning they reflect the present discounted value of an infinite stream of wages. Third, the selection of firms to compare is relevant. For this exercise, I will compare the value of being employed at a formal and an informal firm that occupy the same relative position in their respective distributions. This approach is preferable to forcing an informal firm to operate formally, as a significant proportion of these firms would exit the market if required to pay the minimum wage.

However, even in the scenario described, it is not clear which value functions to compare, as the value of being employed at a firm depends on the benchmark used in bargaining. To sim-

[18]The value was CLP $301,000 at that time, which is equivalent to approximately USD $435 today.

plify the exercise, I will compare the value of being employed between firms when the benchmark is unemployment. This approach is chosen for two reasons. First, the value of being employed with a benchmark is not always valid, as it requires the benchmark to have a lower surplus than the current firm. This means that firms with higher surpluses will offer fewer comparable value functions. Second, since on-the-job offer arrivals are relatively infrequent, most workers in both sectors will initially use unemployment as their benchmark during the early periods of employment.

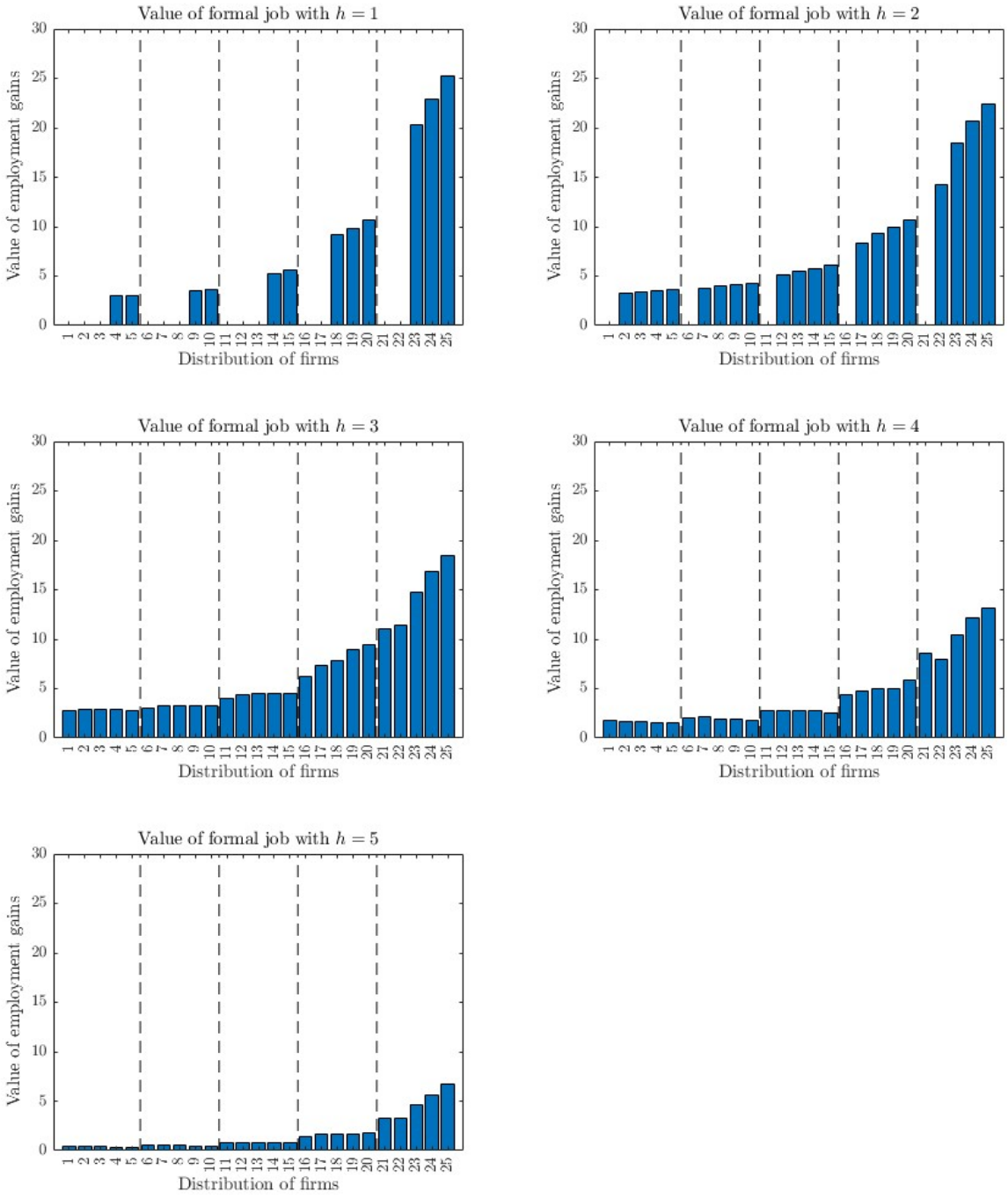
Since the value of being employed depends on human capital levels, I calculate the differences in value for each level. The results for the first level of human capital are shown in Figure 7. This figure is structured to display $W(f, \theta, u, u, h) - W(i, \theta, u, u, h)$ for all θ and h. There are 25 possible pairs $(\theta_y, \theta_\delta)$. The figure arranges firms from those with higher unemployment risk to the safer ones. Within each risk group, the first firm is the least productive, and the last is the most productive. If the corresponding informal firm is not in the market, I assign a value of 0 for the difference.

Workers with higher levels of human capital will provide the firms with more output, making them more likely to be in the market. For example, for the first level of human capital, firms in positions 1-3, 6-8, 11-13, 16-17, and 21-22,[19] are not valid hirings in the formal or informal sector. However, when we increase human capital to the second level, only the firms in the first level of productivity do not have a positive surplus in at least one of the sectors. From the third level of human capital onwards, all the firms can hire workers.

First, the figure shows that transitioning from an informal to a formal firm increases the value of employment. This increase is driven by access to higher productivity, safer jobs, faster human capital accumulation, and a different structure of arrival of offers in the formal sector.

[19] Which corresponds to all firms in the first two productivity levels and three firms in the third level

Figure 7: Gains of moving to a formal job for workers by level of human capital



Second, the gains from holding a formal job decrease as human capital levels increase. This occurs because workers at higher levels on the human capital ladder derive less incremental value from the faster accumulation of human capital. To illustrate this, I will compute the average gains of moving to a formal job across the five levels of human capital. Table 6 shows the results. The gains from holding a formal job are on average 6.2 monthly minimum wages, approximately USD$ 2,740. The gains decrease consistently from a value of 10.19 minimum wages (USD$ 4,490) for the first level of human capital, up to 1.62 minimum wages (USD$ 715)

Table 6: Average gains of having a formal job by level of human capital

Level of human capital (h)	Average gains of a formal job	Relative to $h = 1$
1	10.19	1.00
2	8.32	0.82
3	6.59	0.65
4	4.35	0.43
5	1.62	0.16
Average	6.21	0.61

7.2 Simulations

Regarding the simulations, the procedure will be as follows: I will simulate an economy with 10,000 individuals, starting as unemployed with the lowest level of human capital. Half of these individuals will receive an offer from a formal firm, while the other half will receive an informal offer. Offers will come from firms that are in equilibrium and will be at the same relative position in the joint distribution of productivity and destruction rate across sectors. I will consider three cases to represent the bottom (percentile 16), median (percentile 56) and top firms (percentile 92) of this joint distribution. Since these firms have positive surpluses, workers will accept the offers. I will then compute the total earnings, average wage, and unemployment periods after up to 5 years after workers received the initial offer.

Figure 8 shows the evolution of the total earnings by year and the firm type assigned at the beginning. First, workers who started in a better position on the job ladder consistently have higher earnings than those who started in lower positions. Second, we can see that earnings for those who started in the informal sector are relatively close, with slight differences for those who started at the bottom. This is also the case for firms at the formal sector's bottom and median of the job ladder. However, workers who start at firms at the top of the job ladder generate a significant difference in total earnings.

Figure 8: Gains of moving to a formal job for workers by level of human capital

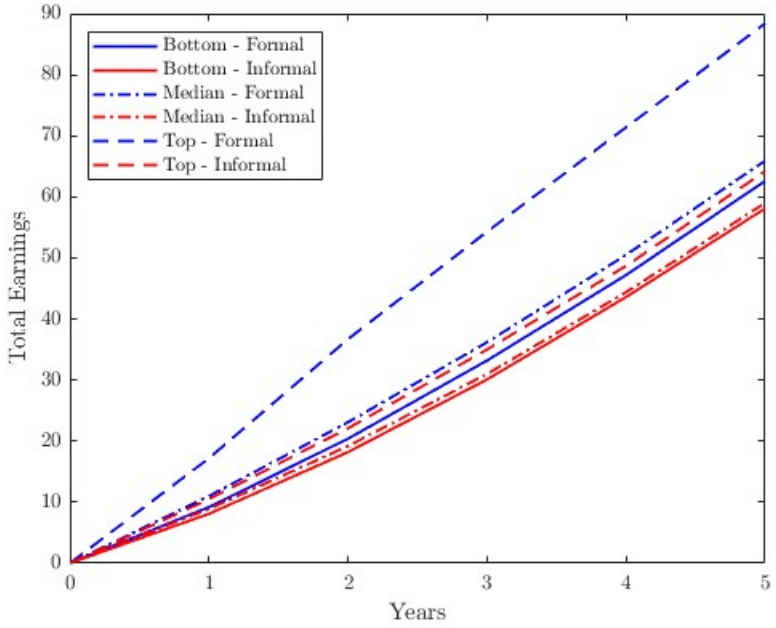


Table 7 provides average earnings, wages and unemployment periods per year and firm type. Earnings and average wages are measured in units of the minimum wage. For firms at the bottom of the productivity distribution, workers receive around 13% higher earnings in the formal sector during the first year, and it decreases up to 7.8% in the fifth year. Also, for the first years the difference in unemployment periods is practically negligible, as the destruction rates of both firms are similar. In addition, informal firms at the bottom of the job ladder pay on average more than the minimum wage. Given that the search gains in these firms are low, because of the high unemployment risk, they must compensate workers paying higher wages.

For the median firms, the differences in earnings are around 23.7% in the first year, and this decreases up to 11.7% in the fifth year. We can notice differences in average wages and unemployment periods at the median firm, especially in the first years. The workers who started in the informal sector have, on average, lower wages than those allocated in the informal firm at the bottom of the job ladder. This is due to the compensating differential mentioned above. However, since the median informal firm faces a lower unemployment risk, the average earnings are higher than the informal firm at the bottom.

For the firms at the top, the differences in earnings are substantial, being 63.5% in the first year and 37.7% in the fifth year. Here we can see that the unemployment risk is significantly low, with unemployment periods close to 10% of those observed in the first years of those allocated to the median formal firm. For the informal firms at the top, the earnings are higher than for the

informal firms at the median but notice again that average wages are lower than in that scenario, and the higher earnings are due to lower unemployment risk.

Table 7: Labor outcomes from simulation for first 5 years

Formal - Perc. 15% (Bottom)	1 Year	2 Years	3 Years	4 Years	5 Years
Earnings	9.18	20.41	33.24	47.17	62.60
Average wage	1.19	1.27	1.31	1.34	1.36
Unemployment periods	4.29	7.93	10.61	12.83	14.08
Informal - Perc. 15% (Bottom)					
Earnings	8.12	18.18	30.06	43.59	58.07
Average wage	1.06	1.16	1.23	1.27	1.31
Unemployment periods	4.32	8.32	11.57	13.72	15.63
Differences in Earnings	13.0%	12.2%	10.6%	8.2%	7.8%

Formal - Perc. 56% (Median)	1 Year	2 Years	3 Years	4 Years	5 Years
Earnings	11.02	23.06	36.19	50.53	65.93
Average wage	1.24	1.31	1.34	1.37	1.38
Unemployment periods	3.11	6.36	8.91	10.99	12.24
Informal - Perc. 56% (Median)					
Earnings	8.91	19.16	30.96	44.36	59.03
Average wage	1.03	1.14	1.21	1.25	1.30
Unemployment periods	3.37	7.16	10.31	12.51	14.43
Differences in Earnings	23.7%	20.4%	16.9%	13.9%	11.7%

Formal - Perc. 92% (Top)	1 Year	2 Years	3 Years	4 Years	5 Years
Earnings	17.22	36.70	54.27	71.38	88.36
Average wage	1.47	1.58	1.57	1.56	1.55
Unemployment periods	0.29	0.77	1.44	2.18	2.95
Informal - Perc. 92% (Top)					
Earnings	10.54	22.07	34.96	48.68	64.19
Average wage	0.91	0.98	1.05	1.11	1.18
Unemployment periods	0.44	1.43	2.69	4.05	5.41
Differences in Earnings	63.5%	66.2%	55.2%	46.6%	37.7%

Overall, these results suggests that starting a career in a formal job has substantial value, as the differences in outcomes persist over the medium term, with a minimum difference around 7.6% in total earnings after 5 years.

7.3 Decomposition

An interesting question regarding sectoral earnings differences is what fraction is due to wage disparities, unemployment periods, and human capital accumulation. Using the calibrated model, I can address this by simulating an economy where half the workers receive offers from a formal firm in a specific relative position, and the other half receives offers from an informal firm in the

same relative position. I will then calculate the total earnings difference (ΔEarnings), which is equal to:

$$\Delta\text{Earnings} = \text{Earnings}_f - \text{Earnings}_i$$

Here, Earnings_f represents the total earnings for those who started in the formal sector, and Earnings_i for those who started in the informal sector. Next, note that the model can be simulated again by setting $\psi_f = \psi_i = 0.032$, meaning human capital accumulation in the formal sector is fixed to match that in the informal sector. This allows to compute a new total earnings difference:

$$\Delta\text{Earnings}_{sh} = \text{Earnings}_{f,sh} - \text{Earnings}_{i,sh}$$

Here, $\Delta\text{Earnings}_{sh}$ represents the earnings difference when human capital accumulation is the same in both sectors. Thus, $\Delta\text{Earnings} - \Delta\text{Earnings}_{sh}$ indicates how much of the earnings gap is due to human capital differences. I use this method because changes in human capital affect not only wage growth but also unemployment periods in both sectors. This is because less productive firms require workers with medium or high human capital to generate positive surplus, making those with lower human capital more likely to experience unemployment.

Then, the earnings difference between sectors with equal levels of human capital accumulation can be decomposed as follows:

$$\Delta\text{Earnings}_{sh} = w_{f,sh} \times t_{f,sh} - w_{i,sh} \times t_{i,sh}$$

Here, $w_{f,sh}$ represents the average wage for those who started with a formal job, and $t_{f,sh}$ is the average time they are employed. Using the same procedure as the Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973), this can be expressed as:

$$\Delta\text{Earnings}_{sh} = (w_{f,sh} - w_{i,sh}) \times t_{f,sh} - w_{i,sh} \times (t_{f,sh} - t_{i,sh})$$

The first term reflects how wage differences drive the earnings gap, while the second term shows the contribution of differing employment periods.

The results of this decomposition are presented in Table 8. The relative importance of each component changes through the years, and it can be even negative.[20] Then, for workers assigned to the bottom of the distribution, higher wages are the main component to explain the differences between workers during the first years while human capital is a second relevant component, but

[20] In the specific case of workers assigned to the formal firm at the bottom, this is because, in the simulation, they have more periods of unemployment.

in the later years, this ranking is reverted.

For workers assigned at the median firm, the decomposition provides the same effect. At the beginning wages are the most relevant component, but over time they become the second after human capital. Regarding those assigned at the top of the job ladder, wages never lose their position as the most important component to explain earnings.

Table 8: Decomposition effects by percentile and year

Perc. 15 (Bottom) - Differences (%)	1 Year	2 Years	3 Years	4 Years	5 Years
Human capital	39.9	60.4	58.0	50.8	54.3
Wages	72.4	48.5	37.9	45.5	33.0
Employment	-12.3	-8.9	4.1	3.7	12.7
Perc. 56 (Median) - Differences (%)	**1 Year**	**2 Years**	**3 Years**	**4 Years**	**5 Years**
Human capital	27.8	45.2	49.3	45.3	45.5
Wages	65.7	48.8	42.0	46.1	38.7
Employment	6.5	6.0	8.7	8.6	15.8
Perc. 92 (Top) - Differences (%)	**1 Year**	**2 Years**	**3 Years**	**4 Years**	**5 Years**
Human capital	34.7	37.0	29.9	20.9	13.8
Wages	63.4	59.2	64.3	71.5	76.3
Employment	1.8	3.8	5.8	7.7	9.9

7.4 Counterfactuals

The final exercise involves recreating the simulations for the median firm while altering various parameters to understand their impact on total earnings after 5 years. The main differences between the formal and informal sectors can be attributed to five factors: different productivity distributions, represented by the pairs (η_{yf}, μ_{yf}) and (η_{yi}, μ_{yi}) and the productivity shifters $\underline{p}_f$ and $\underline{p}_i$; varying rates of job offers on-the-job, determined by λ_{ff}, λ_{fi}, λ_{if}, and λ_{ii}; different destruction rates, represented by the pairs $(\eta_{\delta f}, \mu_{\delta f})$ and $(\eta_{\delta i}, \mu_{\delta i})$; different rates of human capital accumulation, given by ψ_f and ψ_i : and finally, the existence of the minimum wage in the formal sector.

For each exercise, I will compute a new equilibrium of the model by imposing that one group of parameters associated with the formal sector be equal to their counterparts in the informal sector. In the case of the minimum wage, I will impose it equal to 0. After obtaining this new equilibrium, I will repeat the simulation for workers starting in the median firm and compute their total earnings after 5 years.

The results of this exercise are presented in Table 9. As can be seen, total earnings decrease for

those who receive an offer in the formal sector independently of the group of parameters changed. This occurs because the informal sector's parameters are generally less favorable than those in the formal sector. Interestingly, workers who receive an informal offer also experience declining earnings. This result is driven by poorer outcomes when transitioning to the formal sector and a drop in wages within the informal sector. The intuition behind this is this changes in the parameters reduce the need to compensate workers in the informal sector.

After performing the exercises, it is clear that the primary differentiating parameters between the sectors are those related to the productivity distributions, accumulation of human capital, and the minimum wage. When these parameters are equalized, earnings in the formal sector decrease by nearly 46%, while in the informal sector, they decline by 32%. The second most significant factor is human capital accumulation, and when these parameters are equalized, earnings in the formal sector decrease by nearly 28%, while in the informal sector, they decline by 26%. The third relevant parameter is the minimum wage, and when it is set to 0, earnings in the formal and informal sector decrease 22 and 23%, respectively.

On the other hand, the arrival of offer on-the-job and job security are more of second order components. When arrival of offers is equalized, we observe a decrease in earnings of about 12.9% in the formal sector, and 10.8% in the informal sector. And when we equalize the job security distribution, earnings are reduced in only 4.2 and 3.3%, respectively.

Table 9: Total earnings after changes in structural parameters

Scenario	Formal		Informal	
Baseline	65.93		59.03	
Change	Earnings	Diff.	Earnings	Diff.
Same productivity	35.59	-46.0%	40.01	-32.2%
Same human capital acc.	47.53	-27.9%	43.67	-26.0%
No minimum wage	51.54	-21.8%	45.36	-23.2%
Same transitions	57.42	-12.9%	52.66	-10.8%
Same job security	63.19	-4.2%	57.07	-3.3%

8 Conclusion and further research

The literature on labor informality has explored various factors that distinguish the formal sector from the informal one. Key characteristics differentiating formal and informal firms include varying levels of productivity, unemployment risk, and faster human capital accumulation in the formal sector. However, there remains a gap in understanding how these components, when combined, influence workers' career trajectories.

To answer this question, I develop a novel model with search frictions that incorporates a three-dimensional job ladder (sector, productivity, and unemployment risk). In this model, workers accumulate human capital at different rates depending on the sector and engage in both off- and on-the-job search. Additionally, formal sector firms are required to pay their workers at least the minimum wage.

I calibrate the model using both administrative and survey data from Chile between 2010 and 2019. During this period, Chile exhibited stable employment trends across the formal and informal sectors. Overall, the model achieves a reasonable fit, though certain components, particularly those related to the arrival of on-the-job offers and job destruction rates, require further refinement.

With the calibrated model, I conduct three exercises to explore the value of a formal job. In the first exercise, I show that the difference in value functions is on average 6.2 monthly minimum wages, although there is a significant heterogeneity across the firm distribution. Additionally, the gains from holding a formal job decrease as individuals' levels of human capital increase. In the second exercise, I show that receiving a formal job offer at the start of one's career has medium-term consequences for total earnings. Specifically, I find that the lower bound of this effect results in a 7.6% increase in earnings five years later for those who begin their careers in the formal sector.

Next, I analyze the relative importance of the components that distinguish the two sectors. The results show that differences in productivity parameters play a central role in explaining the observed earnings gap between the formal and informal sectors, contributing to nearly a 46% difference in total earnings in the formal sector. This is followed by the variation in human capital accumulation, which accounts for a 28% difference in total earnings. The minimum wage is also a relevant component, generating differences of a 22%. Additionally, I find that components related to the arrival of job offers, and unemployment risk are of secondary importance, as their counterfactuals result in smaller decreases in total earnings: 12.9% and 4.2%, respectively.

This paper opens a research agenda that can be expanded in several directions. Using this model and its simulations, it is possible to study the consequences of job loss in the informal sec-

tor, an analysis that is not feasible with current data due to the need for long-term observations of workers in the informal sector. Another potential application of this model is to explore the effects of changes in the minimum wage on both informality levels and the value of having a formal job. Currently, the model operates under a partial equilibrium framework and does not account for firm reactions to minimum wage changes. However, it is feasible to incorporate general equilibrium features by assuming a matching function, as presented in Flinn (2006).

Finally, as shown in Table 1, there is significant heterogeneity across gender and education, which the current model does not account for. One way to address this is by calibrating the model separately for each group, with the assumption that they participate in distinct labor markets. With these new calibrations, it would be possible to replicate the three main exercises and better understand how the value of a formal job differs for these workers.

References

Abud, M. J., Cases, R., Fuentes, A., & Vergara, R. (2022). *Salario mínimo en Chile: Un aporte a la discusión* (Tech. Rep.). Centro de Estudios Publicos.

Albrecht, J., Navarro, L., & Vroman, S. (2009). The effects of labour market policies in an economy with an informal sector. *The Economic Journal*, *119*(539), 1105–1129.

Allen, J., Nataraj, S., & Schipper, T. C. (2018). Strict duality and overlapping productivity distributions between formal and informal firms. *Journal of Development Economics*, *135*, 534–554.

Amaral, P. S., & Quintin, E. (2006). A competitive model of the informal sector. *Journal of monetary Economics*, *53*(7), 1541–1553.

Amin, M., & Okou, C. (2020). Casting a shadow: Productivity of formal firms and informality. *Review of Development Economics*, *24*(4), 1610–1630.

Arellano, P., & Jimenez, E. (2016). *Dinámica empresarial. brechas regionales y sectoriales de las pymes en chile (2005-2014)* (Tech. Rep.). Ministerio de Economía, Gobierno de Chile.

Bagger, J., Fontaine, F., Postel-Vinay, F., & Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, *104*(6), 1551–1596.

Bagger, J., & Lentz, R. (2019). An empirical model of wage dispersion with sorting. *The Review of Economic Studies*, *86*(1), 153–190.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436–455.

Bobba, M., Flabbi, L., & Levy, S. (2022). Labor market search, informality, and schooling investments. *International Economic Review*, *63*(1), 211-259.

Bobba, M., Flabbi, L., Levy, S., & Tejada, M. (2021). Labor market search, informality, and on-the-job human capital accumulation. *Journal of Econometrics*, *223*(2), 433–453.

Bonnet, F., Vanek, J., & Chen, M. (2019). Women and men in the informal economy: A statistical brief. *International Labour Office, Geneva*, *20*, 22.

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

Bosch, M., & Esteban-Pretel, J. (2015). The labor market effects of introducing unemployment benefits in an economy with high informality. *European Economic Review*, *75*, 1–17.

Cahuc, P., Postel-Vinay, F., & Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, *74*(2), 323–364.

Cases, R., & Fuentes, A. (2024). How substitutable are high-skilled workers? The case of expansion of tertiary education in Chile. *Working paper*.

Cisneros-Acevedo, C. (2022). Unfolding trade effect in two margins of informality. the peruvian case. *The World Bank Economic Review*, *36*(1), 141–170.

Contreras, D., Hurtado, A., & Sara, M. F. (2012). *La excepción chilena y las percepciones de género en la participación laboral femenina* (Tech. Rep.). Departamento de Economía, Universidad de Chile.

Contreras, D., & Plaza, G. (2007). *Participación laboral femenina en Chile.¿Cuánto importan los factores culturales?* (Tech. Rep.). Facultad de Economía y Negocios, Universidad de Chile.

Dabla-Norris, E., Gradstein, M., & Inchauste, G. (2008). What causes firms to hide output? The determinants of informality. *Journal of Development Economics*, *85*(1-2), 1–27.

De Paula, A., & Scheinkman, J. A. (2011). The informal sector: An equilibrium model and some empirical evidence from brazil. *Review of Income and Wealth*, *57*, S8–S26.

D'Erasmo, P. N., & Boedo, H. J. M. (2012). Financial structure, informality and development. *Journal of Monetary Economics*, *59*(3), 286–302.

Dix-Carneiro, R., Goldberg, P., Meghir, C., & Ulyssea, G. (2021). Trade and informality in the presence of labor market frictions and regulations. *Economic Research Initiatives at Duke (ERID) Working Paper*(302).

Dix-Carneiro, R., & Kovak, B. K. (2019). Margins of labor market adjustment to trade. *Journal of International Economics*, *117*, 125–142.

El Badaoui, E., Strobl, E., & Walsh, F. (2010). The formal sector wage premium and firm size. *Journal of development economics*, *91*(1), 37–47.

Elgin, C., Kose, M. A., Ohnsorge, F., & Yu, S. (2021). Understanding informality. *CAMA Working Paper*.

Encina, J., & Martínez, C. (2009). *Efecto de una mayor cobertura de salas cuna en la participación laboral femenina: Evidencia de Chile* (Tech. Rep.). Facultad de Economía y Negocios, Universidad de Chile.

Flinn, C. J. (2006). Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates. *Econometrica*, *74*(4), 1013–1062.

Fu, X., Mohnen, P., & Zanello, G. (2018). Innovation and productivity in formal and informal firms in ghana. *Technological Forecasting and Social Change*, *131*, 315–325.

Fuentes, A. (2019). *Caracterización del desempleo en Chile* (Tech. Rep.). Centro de Estudios Públicos.

Fuentes, A. (2024). When education is not enough: Explaining labor informality inertia in Chile. *Working paper*.

Fuentes, A., & Vergara, R. (2018). Premio salarial en el sector público: Evolución en el periodo 1990-2017. *Estudios públicos*(152).

Fuentes, A., & Vergara, R. (2021). Impuestos a la renta de personas en Chile: simulaciones siguiendo esquemas de otros países OCDE. *Estudios Públicos*, 69-111.

Fugazza, M., & Jacques, J.-F. (2004). Labor market institutions, taxation and the underground economy. *Journal of Public Economics*, *88*(1-2), 395–418.

Haanwinckel, D., & Soares, R. R. (2021). Workforce composition, productivity, and labour regulations in a compensating differentials theory of informality. *The Review of Economic Studies*, *88*(6), 2970–3010.

Hsieh, C.-T., & Olken, B. A. (2014). The missing "missing middle". *Journal of Economic Perspectives*, *28*(3), 89–108.

INE. (2017). *Boletín estadístico: Informalidad laboral. Edición 1* (Tech. Rep.). Instituto Nacional de Estadísticas, Chile.

INE. (2019). *Boletín estadístico: Empleo trimestral. Edición 256* (Tech. Rep.). Instituto Nacional de Estadísticas, Chile.

INE. (2020). *Boletín estadístico: Informalidad laboral. Edición 9* (Tech. Rep.). Instituto Nacional de Estadísticas, Chile.

INE. (2024). *Nota complementaria: Encuesta suplementaria de ingresos 2023* (Tech. Rep.). Instituto Nacional de Estadísticas, Chile.

Jarosch, G. (2023). Searching for job security and the consequences of job loss. *Econometrica*, *91*(3), 903–942.

Jedwab, R., Romer, P., Islam, A. M., & Samaniego, R. (2023). Human capital accumulation at work: Estimates for the world and implications for development. *American Economic Journal: Macroeconomics*, *15*(3), 191–223.

Lagakos, D., Moll, B., Porzio, T., Qian, N., & Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, *126*(2), 797–849.

La Porta, R., & Shleifer, A. (2008). *The unofficial economy and economic development* (Tech. Rep.). National Bureau of Economic Research.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, *28*(3), 109–126.

Lise, J., Meghir, C., & Robin, J.-M. (2016). Matching, sorting and wages. *Review of Economic Dynamics*, *19*, 63–87.

Lopez-Martin, B. (2019). Informal sector misallocation. *Macroeconomic Dynamics*, *23*(8), 3065–3098.

Magnac, T. (1991). Segmented or competitive labor markets. *Econometrica*, *59*(1), 165–187.

Maloney, W. F. (1999). Does informality imply segmentation in urban labor markets? evidence from sectoral transitions in mexico. *The World Bank Economic Review*, *13*(2), 275–302.

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

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.

Ordonez, J. C. L. (2014). Tax collection, the informal sector, and productivity. *Review of Economic Dynamics*, *17*(2), 262–286.

Parente, R. (2024). Minimum wages, inequality, and the informal sector.

Paz, L. S. (2014). The impacts of trade liberalization on informal labor markets: A theoretical and empirical evaluation of the brazilian case. *Journal of International Economics*, *92*(2), 330–348.

Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and exclusion*. World Bank Publications.

Pinheiro, R., & Visschers, L. (2015). Unemployment risk and wage differentials. *Journal of Economic Theory*, *157*, 397–424.

Postel-Vinay, F., & Robin, J.-M. (2002a). The distribution of earnings in an equilibrium search model with state-dependent offers and counteroffers. *International Economic Review*, *43*(4), 989–1016.

Postel-Vinay, F., & Robin, J.-M. (2002b). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, *70*(6), 2295–2350.

Ruggieri, A., & Cisneros-Acevedo, C. (2023). Firms, policies, informality, and the labor market. *Available at SSRN 4638545*.

Schneider, F., & Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, *38*(1), 77–114.

Tejada, M. (2017). Dual labor markets and labor protection in an estimated search and matching model. *Labour Economics*, *46*, 26–46.

Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, *91*(1), 87–99.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from brazil. *American Economic Review*, *108*(8), 2015–2047.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, *12*, 525–546.

9 Appendix

9.1 Appendix 1: Main job in administrative data

To identify the main job in the administrative records, I follow the procedure outlined in Cases & Fuentes (2024). This procedure involves the following steps: First, identify workers with multiple jobs in the same period, which can be determined because formal workers must pay unemployment insurance for each formal job within a month. For these workers, the primary criterion is to select the job with the highest consistent payment during the overlap period. If multiple jobs have similar payments, the next criterion is to choose the job with the longest tenure duration. Any remaining cases with duplicate entries are dropped, as it is not possible to definitively identify the main job in such instances.

9.2 Appendix 2: Administrative data to panel

To create the panel used in this paper, follow these steps: First, download the 3, 5, and 12% samples of the administrative records from the *Superintendencia de Pensiones* website.[21] The documents consist of a series of .csv files containing data on workers, their individual accounts, requests for unemployment insurance, associated payments, wages, and information on rejected requests. For this paper, only the data on workers and their wages is utilized.

The second step involves merging the workers' data with wage information using the common ID present in both databases. This merge will yield a dataset that includes the worker ID, firm ID, and the wages earned from each job, among other details.

Third, apply the procedure outlined in Appendix 1 to retain only one observation per worker. The remaining data will include information about wages and workers. To compute wage losses due to unemployment, I will create empty observations to indicate unemployment periods between employment spells. I will not add further employment spells beyond the last recorded employment, assuming the worker has retired.

The fourth step involves removing data with potential errors. Specifically, I will drop all observations where the database records a worker as employed but no wage data is available. Additionally, observations with wages below 0.75 minimum wages will be removed, as these are likely indicative of part-time work and to ensure consistency with the procedure outlined in Abud et al. (2022).

[21] The website can be accessed through www.spensiones.cl

Finally, adjust all wages for inflation using the CPI index, covering the period from January 2010 to December 2019.

9.3 Appendix 3: Filters to ensure consistency in tenure

After applying the filters mentioned in the main text, an inconsistency in tenure across survey waves may be observed. This inconsistency arises because there is no process ensuring consistency of tenure reporting between waves; respondents may report more or less tenure on their current job compared to previous waves. To address this issue, I will follow the procedure outlined below.

First, I compute the tenure for each individual in the first wave by calculating the difference in months between the date they report starting the job and the survey date. Next, I compare this tenure with the tenure reported in subsequent waves. If a worker reports being in the same sector in the following wave and indicates a tenure greater than the previous tenure plus three months (since the waves are quarterly), I will adjust their tenure to match the tenure calculated in the first wave plus three months. This procedure will be repeated for each subsequent wave, using the adjusted tenure from the previous wave as the basis. This approach ensures consistency in tenure measures for the specific set of workers.

Second, I will discard several sets of observations associated with workers who have inconsistent labor histories. This includes workers who report being employed in two different firms within the same sector between waves, but where the reported tenure in the first wave exceeds that in the second wave (indicating a firm change), and the tenure reported in the second wave is more than three months, creating an inconsistency. Additionally, I will eliminate observations where there is a sector change (from formal to informal or vice versa) but the reported tenure in the second wave is either higher than in the first wave or exceeds three months.

9.4 Appendix 4: Adjustments to wages in income survey

A significant difference between the wages reported in the income survey and those in the administrative records is that the survey wages are net of taxes, while the administrative records show gross wages. This means that one of the wage sets must be adjusted for comparison. In this case, I will adjust the survey wages to convert them into gross wages. To do so, I will use a simplified version of the procedure from Fuentes & Vergara (2021). The steps are as follows: .

1. First, identify whether the employer is paying social security contributions, which can be found directly in the ESI database. If the employer is not paying taxes, we need to verify if the employer is issuing a receipt for the services. If so, the worker's gross wage is calculated

as follows:

$$\text{Gross Wage} = \frac{\text{Declared Wage}}{1 - 0.1}$$

This adjustment accounts for the 10 percent mandatory withholding that workers using receipts must pay to the government. This withholding serves as a savings account to assist workers with tax payments if necessary.

2. If the employer is not paying social security contributions and the worker is not issuing receipts for services provided, then the declared wage is equal to the gross wage.

$$\text{Gross Wage} = \text{Declared Wage}$$

3. If the employer is paying social security contributions, the worker's wage is reduced by that amount and also by applicable taxes. In the Chilean context, the order of deductions is: first social security contributions, followed by taxes on the remaining amount. Therefore, we must first compute the wage before taxes and then account for the social security contributions. To obtain the wage before taxes, we need to consult the tax table for each year. For example, Table 9 shows the tax structure for November 2019.

Table 10: Tax structure for those paying social security contributions. November 2019

Before Tax Income				
From	Until	Tax rate	Allowance	Effective tax rate
-.-	$ 664,591	Exent	-.-	Exent
$ 664,591	$ 1,476,870	4%	$ 26,583.66	2.20%
$ 1,476,870	$ 2,461,450	8%	$ 85,658.46	4.52%
$ 2,461,450	$ 3,446,030	13.5%	$ 221,038.21	7.09%
$ 3,446,030	$ 4,430,610	23%	$ 548,411.06	10.62%
$ 4,430,610	$ 5,907,480	30.4%	$ 876,276.20	15.57%
$ 5,907,480	-.-	35%	$ 1,148,020,28	>15.57%

For example, if a worker has a net income of $1,250,000, this means they paid taxes on 4% of their income and received $26,584 in allowances. The income before taxes can be calculated as follows:

$$\text{Income before taxes} = \frac{\text{Net Income} - \text{Allowance}}{1 - \text{Tax rate}}$$

In this example, the income before taxes is $1,274,392.

Next, we must account for social security payments, which primarily consist of contributions to the health system (7%) and pensions (11.14%).[22] The gross wage can then be calculated

[22]The last value considers the mandatory payment of 10% plus the average commission of 1.14%

as follows:

$$\text{Gross Wage} = \frac{\text{Wage Before Taxes}}{1 - 0.07 - 0.1114}$$

Therefore, the gross wage of a worker with an income of $1,274,392 before taxes is $1,556,794. This wage is nearly 25 percent higher than what was declared in the survey, highlighting the importance of this adjustment for wage comparisons.

This process is replicated for each year using the tax table announced for November. While this choice may seem arbitrary, it aligns with the fact that the income survey is conducted in the last quarter of each year. Tax tables are available on the Chilean Tax Office's website.[23]

9.5 Appendix 5: Descriptive statistics administrative data

Table 10 presents descriptive statistics for the formal sector using administrative data. Notably, there are only small differences between the total averages of log wages and tenure in the administrative and survey data. The proportion of women is also roughly consistent with the survey data. However, there is a significant discrepancy in the proportion of workers with tertiary education between the administrative datasets and the surveys. This discrepancy arises because the administrative records capture the level of education workers had at the time they first received formal income. As many workers with tertiary education held formal jobs during their studies, the administrative data tends to underestimate the level of education in the formal sector.

Table 11: Summary statistics - Administrative data (2010-2019)

Year	Average log wages	Average job tenure	Prop. Women	Prop. Tertiary
2010	13.0	25.2	36.0	16.2
2011	13.0	26.6	36.2	16.1
2012	13.1	27.8	36.7	15.9
2013	13.2	29.9	37.5	15.7
2014	13.2	32.4	38.2	15.6
2015	13.2	34.4	38.6	15.3
2016	13.3	36.3	38.9	15.0
2017	13.3	38.2	39.1	14.8
2018	13.3	38.8	39.1	14.5
2019	13.3	40.2	39.2	14.2
Total Average	13.2	33.0	37.9	15.3

[23] https://www.sii.cl/valores_y_fechas/impuesto_2da_categoria/impuesto2024.htm

9.6 Appendix 6: Surplus proof

The surplus in the formal sector is by definition:

$$S(f,\theta,h) = W(f,\hat{c},\theta,\hat{\theta},h) - U(h) + J(f,\hat{c},\theta,\hat{\theta},h)$$

Replacing the definitions of each element we have:

$$\begin{aligned}
&S(f,\theta,h) = \max\{w(f,\hat{c},\theta,\hat{\theta},h),\overline{w}\} + \beta(1-\rho)\int_{\mathcal{H}}[(1-\theta_\delta)\times ... \\
&\left(\lambda_{ff}\left(\int_M W(f,f,x,\theta,h')dF_f(x) + \int_{RB} W(f,f,\theta,x,h')dF_f(x)\right) + ... \right. \\
&\left.\lambda_{fi}\left(\int_M W(i,f,x,\theta,h')dF_i(x) + \int_{RB} W(f,i,\theta,x,h')dF_i(x)\right)\right) + ... \\
&\left(1-\lambda_{ff}\int_{M\cup RB} dF_i(x) - \lambda_{fi}\int_{M\cup RB} dF_f(x)\right)\times W(f,\hat{c},\theta,\hat{\theta},h) + \theta_\delta U(h')]\, dG_f(h'|h) \\
&-z - \beta(1-\rho)\int_{\mathcal{H}}\left(\lambda_f\int_M [W(f,u,x,u,h') - U(h')]\, dF_f(x) ... \right. \\
&\left. + \lambda_i\int_M [W(i,u,x,u,h') - U(h')]\, dF_i(x) + U(h')\right) dG_u(h'|h) \\
&+p(\theta_y,s) - \max\{w(f,\hat{c},\theta,\hat{\theta},h),\overline{w}\} + ... \\
&\beta(1-\rho)\int_{\mathcal{H}}(1-\theta_\delta)\left(\lambda_{ff}\int_{RB} J(f,f,\theta,x,h')dF_f(x) + \lambda_{fi}\int_{RB} J(f,i,\theta,x,h')dF_i(x) \ + \ ... \right. \\
&\left.\left(1-\lambda_{ff}\int_{M+RB} dF_f(x) - \lambda_{if}\int_{M+RB} dF_i(x)\right)\times J(f,\hat{c},\theta,\hat{\theta},h')\right) dG_f(h'|h)
\end{aligned}$$

Then, we can simply the wages and group the production and unemployment flows. The intuition for the surplus not depending on wages is due the fact that wages are just a distribution of the surplus across the parties. Next, note that we can group the value functions for workers and filled vacancies associated with the re-bargaining and not moving:

$$\begin{aligned}
&S(f,\theta,h) = p(\theta_y,s) - z + \beta(1-\rho)\int_{\mathcal{H}}[(1-\theta_\delta)\times ... \\
&\left(\lambda_{ff}\left(\int_M W(f,f,x,\theta,h')dF_f(x) + \int_{RB} W(f,f,\theta,x,h')dF_f(x)\right) + ... \right. \\
&\left.\lambda_{fi}\left(\int_M W(i,f,x,\theta,h')dF_i(x) + \int_{RB} W(f,i,\theta,x,h')dF_i(x)\right)\right) + ... \\
&\left(1-\lambda_{ff}\int_{M\cup RB} dF_i(x) - \lambda_{fi}\int_{M\cup RB} dF_f(x)\right)\times W(f,\hat{c},\theta,\hat{\theta},h) + \theta_\delta U(h')]\, dG_f(h'|h) \\
&-\beta(1-\rho)\int_{\mathcal{H}}\left(\lambda_f\int_M [W(f,u,x,u,h') - U(h')]\, dF_f(x) ... \right.
\end{aligned}$$

$$+\lambda_i \int_M \left[W(i,u,x,u,h') - U(h')\right] dF_i(x) + U(h')\Big) dG_u(h'|h)$$
$$+\beta(1-\rho)\int_{\mathcal{H}}(1-\theta_\delta)\left(\lambda_{ff}\int_{RB} J(f,f,\theta,x,h')dF_f(x) + \lambda_{fi}\int_{RB} J(f,i,\theta,x,h')dF_i(x) \; + \; ...\right.$$
$$\left.\left(1-\lambda_{ff}\int_{M+RB} dF_f(x) - \lambda_{if}\int_{M+RB} dF_i(x)\right) \times J(f,\hat{c},\theta,\hat{\theta},h')\right) dG_f(h'|h)$$

Then we will take out $U(h')$, and add and subtract $\beta(1-\rho)\int_{\mathcal{H}} U(h')dG_f h'|h$:

$$S(f,\theta,h) = p(\theta_y,s) - z + \beta(1-\rho)\int_{\mathcal{H}}[(1-\theta_\delta)\times ...$$
$$\left(\lambda_{ff}\left(\int_M W(f,f,x,\theta,h')dF_f(x) + \int_{RB}\left[W(f,f,\theta,x,h') + J(f,f,\theta,x,h')\right]dF_f(x)\right) + ...\right.$$
$$\left.\lambda_{fi}\left(\int_M W(i,f,x,\theta,h')dF_i(x) + \int_{RB}\left[W(f,i,\theta,x,h') + J(f,i,\theta,x,h')\right]dF_i(x)\right)\right) + ...$$
$$\left(1-\lambda_{ff}\int_{M\cup RB} dF_i(x) - \lambda_{fi}\int_{M\cup RB} dF_f(x)\right)\times\Big[W(f,\hat{c},\theta,\hat{\theta},h) + J(f,\hat{c},\theta,\hat{\theta},h)\Big]$$
$$+\theta_\delta U(h')\big]\, dG_f(h'|h) - \beta(1-\rho)\int_{\mathcal{H}}\left(\lambda_f\int_M\left[W(f,u,x,u,h') - U(h')\right]dF_f(x)\; ...\right.$$
$$\left.+\lambda_i\int_M\left[W(i,u,x,u,h') - U(h')\right]dF_i(x) + U(h')\right) dG_u(h'|h) + ...$$
$$\beta(1-\rho)\int_{\mathcal{H}} U(h')dG_f h'|h - \beta(1-\rho)\int_{\mathcal{H}} U(h')dG_f h'|h$$

Now, notice that the highlighted unemployment value can be decomposed into pieces to fill all the highlighted terms:

$$S(f,\theta,h) = p(\theta_y,s) - z + \beta(1-\rho)\int_{\mathcal{H}}[(1-\theta_\delta)\times ...$$
$$\left(\lambda_{ff}\left(\int_M W(f,f,x,\theta,h')dF_f(x) + \int_{RB}\left[W(f,f,\theta,x,h') + J(f,f,\theta,x,h')\right]dF_f(x)\right) + ...\right.$$
$$\left.\lambda_{fi}\left(\int_M W(i,f,x,\theta,h')dF_i(x) + \int_{RB}\left[W(f,i,\theta,x,h') + J(f,i,\theta,x,h')\right]dF_i(x)\right)\right) + ...$$
$$\left(1-\lambda_{ff}\int_{M\cup RB} dF_i(x) - \lambda_{fi}\int_{M\cup RB} dF_f(x)\right)\times\Big[W(f,\hat{c},\theta,\hat{\theta},h) + J(f,\hat{c},\theta,\hat{\theta},h)\Big]$$
$$+\theta_\delta U(h')\big]\, dG_f(h'|h) - \beta(1-\rho)\int_{\mathcal{H}}\left(\lambda_f\int_M\left[W(f,u,x,u,h') - U(h')\right]dF_f(x)\; ...\right.$$
$$\left.+\lambda_i\int_M\left[W(i,u,x,u,h') - U(h')\right]dF_i(x)\right) dG_u(h'|h) - \beta(1-\rho)\int_{\mathcal{H}} U(h')dG_u(h'|h)$$
$$+\beta(1-\rho)\int_{\mathcal{H}} U(h')dG_f(h'|h) - \beta(1-\rho)\int_{\mathcal{H}} U(h')dG_f(h'|h)$$

Then, we will obtain this expression:

$$
\begin{aligned}
&S(f,\theta,h) = p(\theta_y,s) - z + \beta(1-\rho)\int_{\mathcal{H}}\left[(1-\theta_\delta)\times ...\right.\\
&\left(\lambda_{ff}\left(\int_M \left[W(f,f,x,\theta,h') - U(h'|h)\right]dF_f(x)\right.\right.\\
&\left.+\int_{RB}\left[W(f,f,\theta,x,h') - U(h'|h) + J(f,f,\theta,x,h')\right]dF_f(x)\right) + ...\\
&\lambda_{fi}\left(\int_M \left[W(i,f,x,\theta,h') - U(h'|h)\right]dF_i(x) + ...\right.\\
&\left.\left.\int_{RB}\left[W(f,i,\theta,x,h') - U(h'|h) + J(f,i,\theta,x,h')\right]dF_i(x)\right)\right) + ...\\
&\left(1-\lambda_{ff}\int_{M\cup RB}dF_i(x) - \lambda_{fi}\int_{M\cup RB}dF_f(x)\right)\times\left[W(f,\hat{c},\theta,\hat{\theta},h) - U(h'|h) + J(f,\hat{c},\theta,\hat{\theta},h)\right]\\
&dG_f(h'|h) - \beta(1-\rho)\int_{\mathcal{H}}\left(\lambda_f\int_M\left[W(f,u,x,u,h') - U(h')\right]dF_f(x)\ ...\right.\\
&\left.+\lambda_i\int_M\left[W(i,u,x,u,h') - U(h')\right]dF_i(x)\right)dG_u(h'|h) - \beta(1-\rho)\int_{\mathcal{H}}U(h')dG_u(h'|h)\\
&+\beta(1-\rho)\int_{\mathcal{H}}U(h')dG_f(h'|h)
\end{aligned}
$$

Now it is possible to apply the definition of surplus again and also use the bargaining rule to replace the highlighted terms by expressions in terms of the surplus:

$$
\begin{aligned}
&S(f,\theta,h) = p(\theta_y,s) - z + \beta(1-\rho)\int_{\mathcal{H}}\left[(1-\theta_\delta)\times ...\right.\\
&\left(\lambda_{ff}\left(\int_M \left[\max\{S(f,\theta,h') + \alpha[S(f,x,h') - S(f,\theta,h')], \gamma(x,h)S(f,x,h)\}\right]dF_f(x)\right.\right.\\
&\left.+\int_{RB}\left[S(f,\theta,h')\right]dF_f(x)\right) + ...\\
&\lambda_{fi}\left(\int_M \left[S(f,\theta,h') + \alpha\left[S(i,x,h') - S(f,\theta,h')\right]\right]dF_i(x) + \int_{RB}S(f,\theta,h')dF_i(x)\right)\Big) + ...\\
&\left(1-\lambda_{ff}\int_{M\cup RB}dF_i(x) - \lambda_{fi}\int_{M\cup RB}dF_f(x)\right)\times S(f,\theta,h')\Big]dG_f(h'|h)\\
&-\beta(1-\rho)\int_{\mathcal{H}}\left(\lambda_f\int_M\max\{\gamma(x,h'),\alpha\}S(f,x,h')dF_f(x)\ ...\right.\\
&\left.+\lambda_i\int_M S(i,x,h')dF_i(x)\right)dG_u(h'|h) - \beta(1-\rho)\int_{\mathcal{H}}U(h')dG_u(h'|h)\\
&+\beta(1-\rho)\int_{\mathcal{H}}U(h')dG_f(h'|h)
\end{aligned}
$$

Finally, we can simply common terms:

$$
\begin{aligned}
&S(f,\theta,h) = p(\theta_y, s) - z + \beta(1-\rho)\int_H (1-\theta_\delta)\left(S(f,\theta,h') + ...\right.\\
&\lambda_{ff}\int_M [\max\{S(f,\theta,h') + \alpha[S(f,x,h') - S(f,\theta,h)], \gamma(x,h')S(f,x,h')\}...\\
&\left.-S(f,\theta,h')\right] dF_f(x) + \lambda_{fi}\int_M \left[\alpha[S(i,x,h') - S(f,\theta,h')]\right] dF_i(x)\Bigg) dG_f(h'|h))...\\
&-\beta(1-\rho)\int_H \Bigg\{\lambda_f \int_M [\max\{\gamma(\theta,h')S(f,x,h')dF_f(x)\} + ...\\
&\lambda_i \int_M \alpha S(i,x,h')dF_i(x)]\Bigg\} dG_u(h'|h) + ...\\
&\beta(1-\rho)\int_H U(h')dG_f(h'|h) - \beta(1-\rho)\int_H U(h')dG_u(h'|h)
\end{aligned}
$$

Notice that the surplus only depends on movements, and given that movements are done comparing the surpluses between the actual and the new firms, the benchmark firms to not play a role there, and because of this, the surplus can be written only as a function of the current firm parameters. The structure to obtain the informal surplus is analogous, with the main differences being the minimum wage and the arrival of offers on-the-job.

9.7 Appendix 7: Algorithm

I solve the model using value function iteration. Specifically, solving the model involves determining the values of firm surpluses and wages for every possible combination of current and benchmark firms. The presence of a minimum wage in the formal sector adds complexity, requiring the computation of the values of the vector $\gamma(\theta, h)$

To solve the model, the surpluses in both sectors must be determined simultaneously. Due to the interdependence between sectors, the minimum wage in the formal sector can influence informal firms, particularly those that are similar to formal firms where the minimum wage is binding. Consequently, it is necessary to estimate the extent to which the minimum wage is binding across formal firms before beginning the solution process. To address this, I propose the following algorithm:

1.1 Guess $\gamma(\theta, h)$

The initial guess I use is 0 for all the firms. Using α is also consistent with this start.

2.1 Guess $S_f(\theta, h)$

The first guess is just a randomly generated matrix with positive values between 0 and 1.

3.1 Guess $S_i(\theta, h)$

The first guess is just a randomly generated matrix with positive values between 0 and 1.

3.2 Compute $U(h)$

This is done using the equation for the value of unemployment.

3.3 $S_i(\theta, h) = \max\{S_i(\theta, h), 0\}$

This is due the definition of the surplus.

3.4 Check and update $S_i(\theta, h)$ until convergence.

2.2 Compute $w(f, \hat{\tau}, \theta, \hat{\theta}, h)$ using $J(\cdot)$

This step implies computing the wage for each feasible combination of current and benchmark firm.

2.3 Update $\gamma(\theta, h) = \max(\max(Aux(f, :, \theta, :, h)))$

Here, I impose $\gamma(\theta, h)$ as the maximum γ observed across all possible benchmarks. The rationale is to associate a firm with the highest γ required to meet the minimum wage, ensuring that the firm can pay it. While it is possible to relax this assumption and consider case-by-case settings, doing so would significantly impact the efficiency of the γ iteration process. Additionally, visual inspection of the different γ values for each firm reveals that they are relatively similar in the different benchmarks..

2.4 $S_f(\theta, h) = \max\{S_f(\theta, h), 0\}$

This is just apply the definition of surplus of a formal firm

2.5 $S_f(\theta, h) = S_f(\theta, h) \times (\gamma(\theta, h) \leq 1)$

Cancel the surplus if the minimum wage is binding.

2.6 Check and update $S_f(\theta, h)$ until convergence.

1.2 Check and update $\gamma(\theta, h)$ until convergence.

Finally, it is important to account for a margin of error in the transitions between firms. Since the surpluses are determined through value function iteration and there is an interconnection between sectors, high precision is required to ensure accurate transitions between firms. To address this, I allow for an additional 0.01 points of surplus as a margin to allow for a job transition.

9.8 Appendix 8: Simulation details

To conduct the simulations, I follow these steps. First, I solve the model using the calibrated parameters. Second, I generate 10 random vectors with dimensions 420 realizations (35 years) each using a uniform distribution. Each vector represents the following components:

1-2 Arrival of formal and informal offers: Take the value 1 if the random realization is below λ_f and λ_i, respectively, and 0 otherwise.

3-6 Arrival of on-the-job offers: Take the value 1 if the random realization is below λ_{ff} and λ_{fi} for the formal sector, and below λ_{if} and λ_{ii} for the informal sector, and 0 otherwise.

7-9 Update of human capital: Take the value 1 if the random realization is below ψ_f, ψ_i and ψ_u, respectively, and 0 otherwise.

10 Firm associated (only for offers): Assign a value between 1 and 25 following a proportional distribution between 0 and 1.

Note that these random vectors are fixed across simulations. This is necessary to prevent variations in the results due to the randomness of the shocks generation.

Once the random realizations are generated, we can begin simulating the model. In the first period, individuals immediately receive offers, as they start unemployed. If they accept an offer, the firm's destruction parameter (θ_δ) is incorporated into their employment history. Also, use the wage function to compute the relevant wage for the worker. Next, another random vector is generated to simulate unemployment shocks, which will occur if the random realization is below θ_δ. Also workers will incorporate unemployment as the benchmark.

Once workers begin receiving on-the-job offers, they will initially accept any firm as their benchmark, as their current benchmark value is unemployment, equivalent to zero. After updating the benchmark, they will use the wage function to compute the new relevant wage. From that point on, whenever they receive an on-the-job offer, they will compare the surplus of the new offer with that of their current firm and the benchmark to decide their next move. Each time a worker switches to a new firm, their previous firm will then serve as the updated benchmark.

If workers experience an unemployment shock, they will exit both their current firm and benchmark, entering the unemployment pool and potentially facing human capital de-accumulation shocks.

Also, workers can increase their levels of human capital if they survive the destruction realization.

After completing these steps, the labor history of one worker is generated. To simulate the entire economy, I repeat this procedure for 10,000 workers, creating a comprehensive simulated dataset.