Abstract

This paper analyzes how informality responds to the quality of labor enforcement and the bundle of benefits that the formal workers receive. In the model, in a general equilibrium framework, the government maximizes the workers' utility subject to a budget constraint, choosing the level of labor contributions' enforcement and the bundle of benefits that the formal worker enjoys. Then, there are a representative firm chooses the share of formal and informal workers that they want to hire; and workers offering a percentage of their work time formally and informally. I estimate the main parameters of the model, the production function, the quality of government enforcement and the quality of benefits for five countries: Argentina, Brazil, Colombia, Peru, and Uruguay. Differences in the quality functions of the government enforcement and benefits are found, as well as in the fines established to enforce labor regulations.

JEL codes: E26, H26, H53, O17, O54.

Keywords: Informality, evasion, labor regulation, government enforcement, Latin America.

*I am grateful for very helpful comments by Russell Cooper, Andrea Ichino, Jérôme Adda, Marco Manacorda, Giovanni Pica, Friedrich Schneider, Catalina Granda Carvajal, Alvaro Forteza, German Cubas and Julio Garin. This paper has a previous version: Informality and government enforcement.
1 Introduction

Informal work is a standard characteristic in developing economies and governments play a crucial role in it. On the one hand, governments fight against this practice, as the lack of social contributions implies difficulties both in collecting other taxes from firms and workers and in providing benefits and public goods. However, at the same time, as informality is a source of flexibility in the low phase of the economic cycle, the allowance of this practice mitigates the costs of an economic recession (Ulyssea, 2010; Loayza and Rigolini, 2011; Meghir et al., 2015) and, simultaneously it is the entry gate to the labor market for younger individuals (Cruces et al., 2012). Thus, understanding how collection enforcement and the provision of benefits affect informality levels are key points to shape policies in developing countries.

Public policy can mainly affect informality through three different channels: first, revenue agencies collect taxes and contributions from formal workers; second, regulators monitor firms by checking; and third, the administration provides benefits to formal workers through the health system, pensions, and unemployment benefits (Almeida and Carneiro, 2009, 2012; Bergolo and Cruces, 2014; Rocha et al., 2018; Ulyssea and Ponczek, 2018). These three assignments are carried out to different degrees and lead to varying levels of enforcement for firms and workers. This paper aims to shed light on how the labor market responds to the quality of government enforcement and the quality of the bundle of benefits provided by public institutions. Then, I compared the level of informality, estimating the shape of the production function, the quality of government enforcement and the bundle of benefits offered in several countries in Latin American, a region where informality is about 50% of salaried workers are employed informally (Portes et al., 1989; Schneider, 2012).

My model introduces a novel general equilibrium framework, where households, firms and government optimize the level of informality, government enforcement and the benefits that formal workers receive. Additionally, I allow a loose definition of dualism through the definition of the production function of the representative firm which includes formal and informal workers, where the level of substitutability is specifically estimated. As in the intensive margin of Ulyssea (2018), firms hire both formal and informal workers contemporaneously, but in my model both workers enter in the production function with different levels of substitutability by education. If both inputs are perfect substitutes, the optimal is the corner solution and inputs can be treated as only one. But, if the level of substitutability is lower, both inputs coexist in the production function.

All three agents in the economy face costs and benefits by being either in formality or informality. Firstly, in formality the employees are obliged to pay contributions and taxes, but they have the right to receive benefits in the present and in the future (such as the right to be covered by the health system, enjoy holidays, receive some extra payments and a pension upon retirement). However, some workers do not value some of these benefits, because the services which are provided are of poor quality, or the government commitment is too weak i.e. they believe that in the future a form of survival pension will be available for everyone. Additionally, informality is attractive for some workers because it is a more flexible sector, which allows easier entrance for unemployed workers or for those who want to acquire experience without a strict or full-time work schedule.

There are three strands in the literature that try to address the nature of informality. The first one has extensively claimed that there are two separate segmented markets, which have different rules and have been related with low and high productivity sectors (La-Porta and Shleifer, 2014). This concept has been discussed by the empirical literature using data from Mexico, Colombia, Argentina and Uruguay (Magneac, 1991; Maloney, 2004; Pratap and Quintin, 2006; Bucheli and Ceni, 2010). In these cases, the evidence seems to suggest that workers are the ones who decide whether to be formal or informal employees. Additionally, Meghir et al. (2015) discuss this point by using firm-level data and shows an overlapping zone of formal and informal firms according to size and productivity. The second strand focuses on considering a unique labor market where workers and firms can always decide to be either in formality or informality. Finally, the third one proposes a moderate

---

1Salaried workers are those who have a salary paid by an employer and the definition of informal workers are those who are not covered by labor regulations, such as taxation, the right to public health care and the right to a pension income upon retirement.
dualism, which has been considered in most of the recent theoretical papers. In this case only those workers that are above some productivity (education) threshold can decide.

This third strand is used by the major part of the literature both at the firm and worker levels. Traditionally the literature focuses on the extensive margin of the informality, identifying the main features that determine the probability of being either formality or informality: characteristics such as managerial ability or physical capital (Galiani and Weinschelbaum, 2007; Amaral and Quintin, 2006; De-Paula and Scheinkman, 2011). Working with macro models Ulyssea (2018) and Meghir et al. (2015) exploit heterogeneity among firms, focusing on the intensive margin. Ulyssea (2018) separates the entrance or extensive margin (firms decide to be either formal or not) from worker hiring or the intensive margin (those formal firms decide the number of formal and informal workers), to analyze the effect of a policy in each margin. In my paper, I mainly consider the intensive margin, in terms of a representative firm deciding how to split the hours among formal and informal workers.

At the worker level the main controversial point is whether or not the worker’s decision plays a role in the equilibrium. The dual market assumes a negative answer, but the evidence, especially in the middle and high educated workers, indicate that the worker at least moderates it. The central argument against this dualism is the mobility between these sectors, which is clearly observed in these economies. If workers move between sectors it means that their intrinsic characteristics are not so different, so it can be considered as the same input in the production function. Mobility is a permanent feature of this phenomenon (Fields, 2011). The informal sector can be more attractive than the formal one for the unemployed workers, and the formal sector is the more stable one (Ceni, 2017).

Additionally, work incentives could be changed by some policies designed to improve the life quality of the poorest population, such as health insurance or Conditional Cash Transfer programs. These programs have a formal income threshold as a condition, leading to less formalization among the beneficiaries (World-Bank, 2010; Bergolo and Cruces, 2014).

Secondly, firms in informality do not pay any contributions and taxes because they are not monitored by the government or the fines are too low. Moreover, in Latin America there is no social punishment because this is not viewed as a crime even for the high educated workers. However, by being in the formal sector workers benefit from government protection against possible abuse from criminal activities, such as blackmail (Loayza et al., 2009).

Finally, the role of government emerges as crucial in both academic and political discussions about informality. Traditionally in the public economics literature the government decides the level of enforcement through a cost-benefit problem (Becker, 1968; Allingham and Sandmo, 1972). In this paper, the government decides about the level of enforcement and benefits, and the enforcement definition is in line with the literature of social norms (Posner, 1997; Besley et al., 2015; Benabou and Tirole, 2011; Acemoglu and Jackson, 2014). The government decides on household utility, thus informality and low levels of enforcement could be desirable independently of cost functions. Theoretically, this is one of the main contributions of this paper in the same fashion as Forteza and Noboa (2014) who consider a benevolent government that maximize a welfare function and prefers discretion to signing a commitment in an insurance environment.

This model allows me to explore what the underlying differences are (e.g. in enforcement technology, the ability to provide public benefits, the production technology, the skill composition of the labor force) that jointly explain the choice of governments on the level of government enforcement and benefits, and the reaction of firms and workers in terms of operating in formality/informality. Theoretically, endogenous enforcement is the main contribution of my model. I can subsequently compare the informality levels in different countries given the quality of the bundle of benefits, the costs, fines and quality of the government enforcement.

The main objective of this paper is to explain the heterogeneity in informality levels, given that this phenomenon is present in all countries in Latin America, and explore the features that could explain these differences, focusing on costs, benefits and enforcement. I estimate the main parameters that characterize each country (Argentina, Brazil, Colombia, Peru and Uruguay) in order to assess the differences in the informality
levels. The main results show that the model captures the heterogeneity among countries following the ranking and the differences in education.

My analysis provides insights to assess which policy is more suitable, including changes in enforcement quality. Government can decide which measure is better in each situation to decrease informality, either through high penalty fines or structural changes in quality functions. This second policy, in contrast to the first one, requires a profound change not only in the government agency, but also in household’s behavior.

This paper is structured as follows. In Section 2 I describe the data and the main variables of the paper, Section 3 provides the model, with one period and two types of workers with different education levels, Section 4 presents the main results of the estimation and the experiments, and finally, in Section 5 I present the main conclusions.

2 Data and main variables

I use data from five countries with different levels of informality: Argentina, Brazil, Colombia, Peru and Uruguay. In order to estimate the production function I use data from the National Accounts and the household surveys for each one. There are series of National Accounts publicly available in the National Central Banks’ websites. Households surveys generally have a socioeconomic purpose and they are crucial in identifying workers in different productive sectors in the economy. Identifying the formal workers\(^2\) is done directly by asking if the employer pays the contribution in order to obtain the right to a pension in retirement. The high educated workers are identified as those who declare that they completed high school (completed high school and higher), and the low educated all the other ones (uncompleted high school and lower).

I estimate the informality by education and sector quarterly, and the GDP for each sector quarterly. In this way the data-base to estimate the production function has 1,162 observations (414 for Argentina, 88 for Brazil, 108 for Colombia, 240 for Peru and 312 for Uruguay).

For Argentina, I use the Permanent Household Survey (EPH in Spanish) carried out by the National Institute of Statistics and Census (INDEC in Spanish) for the period 1995-2010. The sample is restricted to the urban regions, covering 28 large urban centers where 70% of the urban population of Argentina live\(^3\). In this case I have four observations per year for a long period.

For Brazil, I use the Continuous Household Survey (PNAD)\(^4\), conducted by IBGE\(^5\) in September of each year between 1996 and 2007. The survey is only carried out in September, so I only have one observation per year.

In the case of Colombia, I use the Continuous Household Survey (ECH) between 2002 and 2005 and the Large Integrated Household Survey (GEIH) between 2007 and 2010, conducted by the National Bureau of Statistics (DANE). The question about the social contribution is only present in the second quarter in the ECH and in the first half of the year in the GEIH, so the number of observations in this case is also limited.

I use the National Household Survey (ENAHO) of Peru carried out by the National Institute of Statistics and Informatics (INEI) in the period 2001-2010. The sample includes all urban and rural areas in the country, and social contribution questions are included in all quarters.

In the case of Uruguay, I use the Continuous Household Survey (ECH) conducted by the National Statistics Institute (INE), between 1997 and 2010 throughout the whole year. The ECH is a survey carried out in urban areas between 1997 and 2005, where more than 90% of the Uruguayan population lives, so the survey gives a good representation of the country. From 2006 the survey includes rural areas as well.

---

\(^2\) Note that, informal workers are those who are not covered by labor regulation, such as taxation, the right to the health system and the right to a pension income in retirement.

\(^3\) Urban population accounts for the 90% for the total population of Argentina, so the survey gives a good representation of the country.

\(^4\) Pesquisa Nacional por Amostra de Domicílios

\(^5\) Instituto Brasileiro de Geografia e Estatística
In this paper I will only measure the informality of those who declare that their employer does not pay the necessary contributions in order to give them the right to a pension in old age\textsuperscript{6}. The decision to use this definition is due to the fact that this question is present in all the household surveys and its consequences have been widely analyzed in the literature (Holzmann and Takayama, 2009; Joubert, 2012; Ceni, 2017).

Finally, to estimate government enforcement and the quality of the benefits that the formal workers receive, I use some indicators collected by the InterAmerican Development Bank. In particular for the quality of government enforcement I use two indicators: compliance with the law\textsuperscript{7} and confidence in the judiciary system\textsuperscript{8}. For the benefits that the government provides, I use the citizens’ perception of the taxes being well spent\textsuperscript{9}.

2.1 Some facts from the data

In this section, I analyze some general facts from informality in the Latinamerican countries and how the public policy interact with the informal work. I compare the levels of informality in the countries in the five countries and there are clearly heterogeneous using the pension rights definition, and it is also clearly heterogeneous if other benefits in the definition, such as health benefits, the 13th salary, and the holidays right, as is shown in Table 1. The rates are from to below to 20% to about 50% in all the definitions. In Table 2 shows the there are informal work in all educative levels and those figures are also extremely heterogeneous between different countries.\textsuperscript{10}

<table>
<thead>
<tr>
<th>Country</th>
<th>Pensions</th>
<th>Health</th>
<th>13th month</th>
<th>Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>36.1%</td>
<td>35.9%</td>
<td>34.6%</td>
<td>34.4%</td>
</tr>
<tr>
<td>Brazil</td>
<td>21.6%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Colombia</td>
<td>42.7%</td>
<td>42.4%</td>
<td>53.7%</td>
<td>28.2%</td>
</tr>
<tr>
<td>Peru</td>
<td>51.6%</td>
<td>51.3%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Uruguay</td>
<td>19.4%</td>
<td>21.5%</td>
<td>25.3%</td>
<td>-</td>
</tr>
</tbody>
</table>

Source: CEDLAS.

<table>
<thead>
<tr>
<th>Country</th>
<th>Mean</th>
<th>Low education</th>
<th>High education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>36.1%</td>
<td>39.7%</td>
<td>15.2%</td>
</tr>
<tr>
<td>Brazil</td>
<td>21.6%</td>
<td>26.8%</td>
<td>14.3%</td>
</tr>
<tr>
<td>Colombia</td>
<td>42.7%</td>
<td>46.9%</td>
<td>14.8%</td>
</tr>
<tr>
<td>Peru</td>
<td>51.6%</td>
<td>66.8%</td>
<td>25.2%</td>
</tr>
<tr>
<td>Uruguay</td>
<td>19.4%</td>
<td>25.0%</td>
<td>6.5%</td>
</tr>
</tbody>
</table>

Source: CEDLAS, World Bank.

In line with the idea of the workers move from one sector to other constantly, there are a significant number of annual changes among sectors in Argentina as shown in the figures corresponding to the years between 2003 and 2011 (Table 3). However, the formal work is more stable than the informal or being at unemployment; there are about 7% of formal workers that would be informal in the next year.

The public policy can be summarized by three main channels of action: collecting taxes, providing benefits and monitoring and regulating the economy. Empirical analysis shows an ambiguous relation between the unofficial economy and the level of taxation, and a positive relation with corruption using a large cross-country

\textsuperscript{6}The worker declaration as a source of measure is the common use in the literature, and mismeasurement is not considered as a problem as informality is not viewed as a crime.

\textsuperscript{7}This indicator represents the percentage of those surveyed who respond that they believe that citizens comply with the law very much or a fair amount. Source: Latinobarometer.

\textsuperscript{8}It measures the percentage of firms that agree with the statement: I am confident that the judicial system will enforce my contractual and property rights in business disputes. Source: World Bank.

\textsuperscript{9}This indicator represents the percentage of answers to the question: Do you believe that the government spends your tax dollars well?

\textsuperscript{10}High educated workers are defined as those who at least finished high school, and the low educated workers as those who did not.
Table 3: Probability of changing sector in Argentina, based on the multinomial model 2003-2010 (only men).

<table>
<thead>
<tr>
<th>Probability of yearly change of sector</th>
<th>Unemployment</th>
<th>Formal</th>
<th>Informal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment (-1)</td>
<td>0.337</td>
<td>0.254</td>
<td>0.409</td>
</tr>
<tr>
<td>Formal (-1)</td>
<td>0.024</td>
<td>0.909</td>
<td>0.067</td>
</tr>
<tr>
<td>Informal (-1)</td>
<td>0.086</td>
<td>0.240</td>
<td>0.674</td>
</tr>
</tbody>
</table>

data-base of entrepreneurs. Higher taxation means stronger institutions and lower informality, but if taxes can be an entry barrier. Johnson et al. (1998), using simple OLS regression, find a positive relation between regulation bureaucracy, tax burden and corruption with higher unofficial activities. Friedman et al. (2000) go further with a larger number of countries and find that the taxes have a negative effect on the unofficial economy, and it is the corruption and the bureaucracy which have a positive effect.

First, in Figure 1 we can observe the relationship between informality and the social contributions (taxes) which are paid by the employee and the employer. Due to the legislation it is not possible to evade one of those contributions independently, then both might have a role in informality decisions. There is a slightly negative relation between them, i.e. countries with a high level of contribution have less informality. This relation is observed both with the total, employer and employee contribution, and it could be interpreted as part of the institutional framework. The countries in the region with stronger institutions are those which have lower levels of informality and higher levels of social contributions.

Second, I explore the relation between informality and the quality of the benefits. Workers can decide not to work formally because the benefits that they can enjoy are so poor that it is no advantage to them. The quality of the benefits is a sign of good government practices, low corruption and strong institutions. To measure this feature I use an indicator about the citizens perceptions that taxes are well spent by the government. Figure 2 establishes that there is a clear negative relation between the quality of benefits and informality.

Finally, I consider the relation between informality and government enforcement, measured by citizen compliance with the law (Figure 3). There is a clear negative relation between them, which is in line with the literature on the relation between informality and institutional quality.

These three figures show that two of the three relations between government role and informality appear clearly in the cross-country data, and there is room for action. The benefits that the workers receive and the
level of government enforcement in the labor market seem to play an important role in the level of informality, and in the process of formalization. Thus, in the model I decide to work with these last two roles (enforcement and provide benefits) and leave the level of taxation as exogenous.

3 Model

I develop a simple model to analyze the relation between the levels of informality by education \((\theta_L^{i}, \theta_H^{i})\) and the quality of government enforcement \((q(c))\) and the quality of benefits \((K(\kappa))\). The households and firms decide on the level of informality (formality), and simultaneously there is a government choosing the level of enforcement in the labor market, and the benefits that the formal workers would enjoy.
3.1 Optimization problem: households and firms

**Representative household problem:** In a representative household, there is a continuous of $x$ workers with a low level of education ($s = L$), and $1 - x$ with a high level ($s = H$). This household maximizes its consumption by deciding the share of informal ($\theta^i_s$) and formal ($\theta^f_s$) work for each level of education $s$. There is no utility of leisure, the worker decides how to split their total number of hours ($x$ and $1 - x$) between formality and informality. The consumption is determined by the revenues from formal and informal work, a lump sum tax ($T$) and the profit from firms (II) which is fixed at zero. The formal revenues are the wage of formal hours $\omega^f_s \theta^f_s$ and they also receive $(K(\kappa))$ representing the quality of the benefits that the formal workers enjoy. The informal revenues are the informal wage $\omega^i_s \theta^i_s$ for the hours worked in informality, and there is also a share $(\phi_2 q(e))$ of this total wage which is lost. This loss depends on the quality of the enforcement $(q(e))$ and a parameter $\phi_2$, and represents a market imperfection in the informal labor market.

$$\max_{C, \{\theta^i_s, \theta^f_i\}} U(C)$$  \hspace{1cm} (1)

Subject to

$$C = \sum_s \left[ \omega^i_s \theta^i_s \left( 1 - \phi_2 q(e) \right) + \omega^f_s \theta^f_s \left( 1 + K(\kappa) \right) \right] - T + \Pi$$  \hspace{1cm} (2)

The maximum of low skill hours is $x$:

$$\theta^i_L + \theta^f_L \leq x$$  \hspace{1cm} (3)

and the maximum of high skill hours is $1 - x$:

$$\theta^i_H + \theta^f_H \leq 1 - x$$  \hspace{1cm} (4)

**Lagrangian function:**

$$L = U(C) + \lambda^H \left[ C - \sum_s \left[ \omega^i_s \theta^i_s \left( 1 - \phi_2 q(e) \right) - \omega^f_s \theta^f_s \left( 1 + K(\kappa) \right) \right] + T - \Pi \right]$$

$$+ \lambda^H \left( \theta^i_L + \theta^f_L - x \right) + \lambda^H \left( \theta^i_H + \theta^f_H - (1 - x) \right)$$  \hspace{1cm} (5)

The interior solution to the household problem ($\theta^i_s \neq 0$) implies workers offering formal and informal hours in the labor market, then wages in informality after the market imperfection loss is equal to the formal wage plus the benefits:

$$\theta^i_L \frac{\partial L}{\partial \theta^i_L} = 0 : \hspace{0.5cm} \theta^i_L \left( - \lambda^H \left( \omega^i_s \left( 1 - \phi_2 q(e) \right) \right) + \lambda^H \right) = 0$$  \hspace{1cm} (6)

$$\theta^i_L \frac{\partial L}{\partial \theta^i_L} = 0 : \hspace{0.5cm} \theta^i_L \left( - \lambda^H \left( \omega^i_s \left( 1 + K(\kappa) \right) \right) + \lambda^H \right) = 0$$  \hspace{1cm} (7)

$$\theta^i_H \frac{\partial L}{\partial \theta^i_H} = 0 : \hspace{0.5cm} \theta^i_H \left( - \lambda^H \left( \omega^i_s \left( 1 - \phi_2 q(e) \right) \right) + \lambda^H \right) = 0$$  \hspace{1cm} (8)

$$\theta^i_H \frac{\partial L}{\partial \theta^i_H} = 0 : \hspace{0.5cm} \theta^i_H \left( - \lambda^H \left( \omega^i_s \left( 1 + K(\kappa) \right) \right) + \lambda^H \right) = 0$$  \hspace{1cm} (9)

**Representative firm problem:** A representative firm, decides to hire a share of informal $l^i_s$ and formal workers $l^f_s$ for each educative level $s$. The firm pays the formal workers $\omega^f_s l^f_s$ plus taxes $\tau$, I am considering that the net wage (after taxes) and the taxes are only paid by the firm. Informal workers receive $\omega^i_s l^i_s$ but the firm
faces a proportional fine $\phi_1 q(e)$ if that job is monitored.

$$\max_{l_s^f, l_s^f} \Pi = y(l_s^f, l_s^f) - \sum_s \left( (1 + \phi_1 q(e)) \omega_s^f l_s^f + \omega_s^l l_s^l (1 + \tau) \right)$$  \hspace{1cm} (10)

F.O.C.

$$l_s^f : \frac{\partial y(l_s^f, l_s^f)}{\partial l_s^f} - \left( \omega_s^f (1 + \phi_1 q(e)) \right) = 0$$  \hspace{1cm} (11)

$$l_s^l : \frac{\partial y(l_s^f, l_s^f)}{\partial l_s^l} - \left( \omega_s^l (1 + \tau) \right) = 0$$  \hspace{1cm} (12)

The market clearing condition equalizes the share of hours in formality and informality for each level of education, which the firm demands and the worker supplies:

$$l_s^f = \theta_s^f \text{ ; } l_s^l = \theta_s^l$$  \hspace{1cm} (13)

The definition of the production function is one of the contributions of this paper. In the literature most of the papers either introduce the formality and informality as substitutes as in Ihrig and Moe (2004), or treat them as complements modeled in a Cobb-Douglass framework. The functional form which I choose is the CES function as in Dolado et al. (2001); Giuliodori and Stucchi (2010) and Cappellari et al. (2012) who model the coexistence of temporary and permanent workers to reflect the fact that there are two types of workers who are not different in essence, but contractually. In the informality literature, Ulyssea (2010) also presents a model with a CES production function with formal and informal intermediate goods. The CES function allows me to introduce the loose form of market duality, if formal and informal are perfect substitutes the solution tends to be a corner solution. However, if there is an imperfect substitution, formality and informality coexist in the production function.

**Remark 1** The CES production function of the representative firm including contemporaneously formal and informal workers captures the market duality through the level of substitutability $\delta_j$. If inputs have high substitutability, it shows that both inputs are more similar than in the case when the parameter goes to the complementarity. The production function is:

- $y = \gamma l_s^f l_s^l$

- $l_H = \left[ \psi_1(l_H)^{-\delta_1} + (1 - \psi_1)(l_H)^{-\delta_1} \right]^{-\frac{\psi_1}{\delta_1}} \text{ ; } l_L = \left[ \psi_2(l_L)^{-\delta_2} + (1 - \psi_2)(l_L)^{-\delta_2} \right]^{-\frac{\psi_2}{\delta_2}}$

The level of substitutability is determined by $\delta_j$, if it is close to -1, both inputs are perfect substitutes. Conversely if both inputs are complements, $\delta_j \rightarrow \infty$.

From the F.O.C. of the firm problem

$$\gamma \psi_2 \rho_2 \theta_2 \psi_2 \theta_2^{-1} \left( \psi_2 (\theta_L)^{-\delta_2} + (1 - \psi_2) (\theta_L)^{-\delta_2} \right)^{-\frac{\psi_2 + \delta_2}{\delta_2}} \left( 1 - \psi_2 \right) (\theta_L)^{-\delta_2 - 1}$$

$$= \omega_L^l (1 + \phi_1 q(e))$$

$$\gamma \psi_2 \rho_2 \theta_2 \psi_2 \theta_2^{-1} \left( \psi_2 (\theta_L)^{-\delta_2} + (1 - \psi_2) (\theta_L)^{-\delta_2} \right)^{-\frac{\psi_2 + \delta_2}{\delta_2}} \psi_2 (\theta_L)^{-\delta_2 - 1}$$

$$= \omega_L^l (1 + \tau)$$

$$\gamma \psi_1 \rho_1 \theta_1 \psi_1 \theta_1^{-1} \theta_1 \psi_1 \theta_1^{-1} \left( \psi_1 (\theta_H)^{-\delta_1} + (1 - \psi_1) (\theta_H)^{-\delta_1} \right)^{-\frac{\psi_1 + \delta_1}{\delta_1}} \left( 1 - \psi_1 \right) (\theta_H)^{-\delta_1 - 1}$$

$$= \omega_H^l (1 + \phi_1 q(e))$$

$$\gamma \psi_1 \rho_1 \theta_1 \psi_1 \theta_1^{-1} \theta_1 \psi_1 \theta_1^{-1} \left( \psi_1 (\theta_H)^{-\delta_1} + (1 - \psi_1) (\theta_H)^{-\delta_1} \right)^{-\frac{\psi_1 + \delta_1}{\delta_1}} \psi_1 (\theta_H)^{-\delta_1 - 1}$$

$$= \omega_H^l (1 + \tau)$$
3.2 Equilibrium: households and firms

I will focus only on the interior solution ($\theta^*_e \neq 0$) firms where both formal and informal workers coexist. The relative informal wages depend positively on the quality of benefits and the quality of government enforcement:

$$\frac{\omega^*_L}{\omega^*_F} = \frac{(1 + K(\kappa))}{(1 - \phi_2 q(e))}$$
$$\frac{\omega^*_I}{\omega^*_F} = \frac{(1 + K(\kappa))}{(1 - \phi_2 q(e))}$$

The relative size of the informal sector depends on the relative wage, the relation between fines from being in informality ($\phi_1 q(e)$), and the contribution paid by the formal employer ($\tau$), while the differences in educative levels is given by the formal shares ($\psi_1$ and $\psi_2$) and the level of substitutability ($\delta_1$ and $\delta_2$) of the production function.

$$\frac{\theta^*_I}{\theta^*_F} = \frac{\psi_2}{1 - \psi_2} \frac{(1 + K(\kappa))(1 + \phi_1 q(e))}{1 + \tau}$$
$$\frac{\theta^*_I}{\theta^*_F} = \frac{\psi_1}{1 - \psi_1} \frac{(1 + K(\kappa))(1 + \phi_1 q(e))}{1 + \tau}$$

The share of informality in both educative levels in equilibrium depend negatively on the level of enforcement $e$ (Equations 16) and the benefits $\kappa$ (Equations 17).

$$\frac{\partial \theta^*_I}{\partial e} = -A_1 \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} (\phi_1 + \phi_2) \frac{\partial q}{\partial e} < 0; \quad \frac{\partial \theta^*_I}{\partial \kappa} = -A_2 \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} (\phi_1 + \phi_2) \frac{\partial q}{\partial e} < 0$$

$$\frac{\partial \theta^*_I}{\partial \kappa} = -A_1 (1 + \phi_1 q(e)) \frac{\partial K(\kappa)}{\partial \kappa} < 0; \quad \frac{\partial \theta^*_I}{\partial \kappa} = -A_2 (1 + \phi_1 q(e)) \frac{\partial K(\kappa)}{\partial \kappa} < 0$$

$$A_j = \frac{1}{\delta_j + 1} \left[ \psi_j \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \frac{1 + \phi_1 q(e)}{1 + \tau} \right] \frac{\psi_j}{1 - \psi_j} \frac{1}{1 - \psi_j (1 - \phi_2 q(e))(1 + \tau)} > 0 \quad j \{1, 2\}$$

3.3 Optimization problem: Government

The government maximizes the consumer’s utility by choosing benefits $\kappa$ and the level of enforcement $e$ as a carrot and a stick. For workers in formality there is a carrot which is benefits $K(\kappa)$, and there is a stick for those in informality which is the quality of government enforcement ($q(e)$) and the level of fines $\phi_1$.

The government equalizes the resources (from the social contributions in the formal jobs $\tau$, the fines in the informal jobs $\phi_1 q(e)$ and a lump sum tax ($T$) with the spending $e$, e.g. the cost of the government enforcement and providing benefits which is given by function $B(e, \kappa)$ (Equation 20).

Remark 2 This maximization allows the government to choose the level of enforcement and benefits such that informality can exist in the labor market in its optimal choices. This possibility goes in the direction that some level of informality can be allowed by the government through the relaxation of the quality of the government enforcement and benefits. This point could be controversial if informality is perceived as a purely negative feature in the economy, although in this paper I use an agonistical approach to informality.

$$\max_{e, \kappa} U(C)$$

10
Subject to

\[ C = \sum_s \left[ \omega^i_s \theta^i_s \left( 1 - \phi_2 q(e) \right) + \omega^f_s \theta^f_s \left( 1 + K(\kappa) \right) \right] - T + \Pi \] (19)

\[ \sum_s \left[ \tau \omega^i_s \theta^i_s + \phi_1 q(e) \theta^i_s \right] + T = B(e, \kappa) \quad s = \{L, H\} \] (20)

F.O.C.

\[ e : \quad \frac{\partial U}{\partial C} \left( \sum_s \omega^i_s \theta^i_s (\phi_1 - \phi_2) \frac{\partial q(e)}{\partial e} - \frac{\partial B(e, \kappa)}{\partial e} \right) = 0 \] (21)

\[ \kappa : \quad \frac{\partial U}{\partial C} \left( \sum_s \omega^f_s \theta^f_s \frac{\partial K(\kappa)}{\partial \kappa} - \frac{\partial B(e, \kappa)}{\partial \kappa} \right) = 0 \] (22)

### 3.4 Equilibrium: Government

The functional forms which are chosen to estimate this model are such that the quality functions (enforcement \( e \), and benefits \( \kappa \)) are increasing and concave where \( a \) and \( d \) are the quality parameters. The cost function of government enforcement and benefits is quadratic

\[ q(e) = a \sqrt{e} \quad K(\kappa) = d \sqrt{\kappa} \] (23)

\[ B(e, \kappa) = b_1 e^2 + b_2 \kappa^2 + b_3 \]

The equilibrium enforcement and benefits are:

\[ e^* = \left( \frac{a (\phi_1 - \phi_2) (\omega^L \theta^L + \omega^H \theta^H)}{4 b_1} \right)^{\frac{1}{2}} \] (24)

The level of enforcement \( e \) depends positively on the mass of salaries in the informality \( (\omega^i \theta^i_L + \omega^i \theta^i_H) \), the ratio of the quality and cost function parameters \( \frac{a}{b_1} \), and the term \( \phi_1 - \phi_2 \), which is the difference between the fines which the firms pay for any informal job monitored and the worker loss in the informality. If the fines were equal to the workers’ losses, the level of government enforcement would be zero because the effect of the informalidad is solved within the market, and the action of the government is not necessary.

The level of benefits \( \kappa \) that the government chooses depends positively on the mass of salaries in the formality \( (\omega^f \theta^f_L + \omega^f \theta^f_H) \), and the ratio of the quality and cost function parameters \( \frac{d}{b_2} \).

\[ \kappa^* = \left( \frac{d (\omega^L \theta^L + \omega^H \theta^H)}{4 b_2} \right)^{\frac{1}{2}} \] (25)

The level of taxes is given by the equalized budget constraint of the government:

\[ T^* = \sum_s B(e^*, \kappa^*) - \tau \omega^i \theta^i H_s \phi_1 \omega^i \theta^i s q(e^*) \] (26)

**Definition 1** Given the set of parameters, there is a unique equilibrium which determines the level of informality in each level of education \( (\theta^i_s) \) working, the level of government enforcement over informal jobs \( (e^*) \), the benefits that the government brings to the formal ones \( (\kappa^*) \), and the lump sum taxes collected from the households \( (T^*) \).

### 3.5 Comparative statics

In this section, I analyze the effect of the exogenous parameters in the main model equilibrium outcomes: informality levels, enforcement and benefits \( (\theta^i_s, e \text{ and } \kappa) \). I focus not only on the comparative statics with
respect to single parameters\textsuperscript{11}, but also considering the effect of a couple of them that allows me compare the magnitude of each relevant pairs effects.\textsuperscript{12}

In the set of equations in 27, I analyze the comparative statics of the equilibrium relative size of the informality ($\theta^*_i$) with respect to the fines ($\phi_1$) and informal market imperfection or informal wage loss ($\phi_2$), and the quality parameters. These shares depend negatively on the fines that the firm has to pay if the informal job is monitored ($\phi_1$) but it is uncertain on the share of informal wage that the worker loses if their job is monitored ($\phi_2$). To compare these effects, Figure C.1 shows the simulated performance of a grid of $\phi_1$ and $\phi_2$, the informality is more sensitive with a change of $\phi_2$ than $\phi_1$, and there is a predominance of a positive effect of market imperfections in informality.

Additionally, I consider the comparative statics of the informality shares respect to the quality function parameters ($a$ and $d$). These shares decrease when the parameters of the quality of enforcement and benefits function ($a$ and $d$) increase. The effect of $a$ on the informality is higher than the effect of $d$, as is shown in Figure C.3, in the model the enforcement has a bigger effect on informality than the quality of benefits which is desirable because it has an impact both on firms and workers. Figures C.4 and C.5, show the effect of the quality and cost parameters of the enforcement and the benefits, negative in quality and positive in costs. As can be expected, quality parameters are more sensitive than cost ones, due to in the model costs enter only in firm’s budget constraint.

\[
\frac{\partial \theta^*_i}{\partial \phi_1} < 0 \quad \frac{\partial \theta^*_i}{\partial \phi_2} \ ? \quad \frac{\partial \theta^*_i}{\partial a} < 0 \quad \frac{\partial \theta^*_i}{\partial d} < 0
\] (27)

The comparative statics in the set of equations in 28 show that government enforcement $e^*$ is decreasing in the fines parameters ($\phi_1$), decreasing in the quality of benefits ($d$) and with an uncertain sign respect to the quality of government enforcement ($a$). The equilibrium behavior when these parameters change is also shown in Figures C.3 and C.4, in which the equilibrium is solved by fixing the other parameters. The effect of $a$ on $e^*$ is positive with this set of parameters, which was uncertain in the analytical analysis, and the effect of $b_1$ and $d$ are negative (but the intensity for both also depends on the level of $a$ as is shown in Figure C.3 and C.4). Not only that, the magnitude of change on $e$ is low, especially with respect to the pair $a$ and $d$.

\[
\frac{\partial \kappa^*}{\partial \phi_1} < 0 \quad \frac{\partial \kappa^*}{\partial a} \ ? \quad \frac{\partial \kappa^*}{\partial d} < 0
\] (28)

\[
\frac{\partial \kappa^*}{\partial \phi_1} > 0 \quad \frac{\partial \kappa^*}{\partial a} > 0 \quad \frac{\partial \kappa^*}{\partial d} > 0
\] (29)

The benefits $\kappa$ increases in fines ($\phi_1$), the quality parameter of the enforcement ($a$) and the benefits ($d$), as is shown in Figure C.4. More resources have a positive effect on optimal benefits offered by the government, as more institutional quality. Figures C.3 and C.5, show the effect on $\kappa$ of $d$ and $b_2$, which are positive and negative respectively.

In the set of Equations in 30 and 31 I present the effect of the production function parameters on the model outcomes. In the case of the level of substitutability, the signs depend on the terms $B_1$ or $B_2$. If these terms are higher than one, a higher level of substitutability ($\delta_j \to -1$) leads to a lower level of informality, but if $B_1$ or $B_2$ are between 0 and 1, higher substitutability ($\delta_j \to -1$) leads to higher informality. These terms would be lower than one depending on how low $\psi_1$ and $\psi_2$ are, because the second term is always positive, and in the

\textsuperscript{11}The analytical development is presented in Appendix C.
\textsuperscript{12}The graphical analysis is presented in Appendix C, where there is a numerical exercise fixing the parameters as Argentina, and then computing the solution on a grid of two parameters. The parameters for Argentina are shown in Table A.1.
third one, the percentage of effective fines should be greater than the percentage of taxes.

\[ \frac{\theta^*_L}{\partial \delta_2} > 0 \quad \frac{e^*}{\partial \delta_2} > 0 \quad \frac{\kappa^*}{\partial \delta_2} < 0 \]
if \( B_2 = \left[ \frac{\psi_2}{1 + K(\kappa)} \right] > 1 \) (30)

\[ \frac{\theta^*_L}{\partial \delta_2} < 0 \quad \frac{e^*}{\partial \delta_2} < 0 \quad \frac{\kappa^*}{\partial \delta_2} > 0 \quad \text{ if } 0 < B_2 < 1 \]

\[ \frac{\theta^*_L}{\partial \delta_1} > 0 \quad \frac{e^*}{\partial \delta_1} > 0 \quad \frac{\kappa^*}{\partial \delta_1} < 0 \]
if \( B_1 = \left[ \frac{\psi_1}{1 + K(\kappa)} \right] > 1 \) (31)

\[ \frac{\theta^*_L}{\partial \delta_1} < 0 \quad \frac{e^*}{\partial \delta_1} < 0 \quad \frac{\kappa^*}{\partial \delta_1} > 0 \quad \text{ if } 0 < B_1 < 1 \]

In the first two panels of Figure C.2, I observe how the share of informality changes with the level of substitutability between formal and informal workers in the case of the educated and non-educated workers (\( \delta_1 \) and \( \delta_2 \)) and the shares in the production function (\( \psi_1 \) and \( \psi_2 \)). If these shares are lower, \( B_j \) is lower than 1, then the sign of the comparative statics of the informality changes when \( \psi's \) are around 0.4.

4 Results

The empirical strategy is to estimate both the production function and the quality and the cost functions. However, Botero-García (2010) calibrates the elasticity of substitution of a CES production function with informal and formal workers, the level of substitutability for low educated worker is -0.5, and between educated and non-educated workers it is 3.3. Ulyssea (2010) also calibrates the parameters of the general level of substitutability between formal and informal production at -0.3. My empirical strategy is to estimate the parameters of the production function from the data. The production function is estimated by an approximation of a linear regression. The other parameters of the model are estimated with the Method of Moments, minimizing the square of the distance between each moment through the simulated model and data for each country. The moments that I use are, informality for low and high educated, the quality of government enforcement, the quality of benefits provided, public expediters participation in the total and a lump sum tax.

4.1 Estimation of the production function

In order to estimate the CES production function (and the nested production function) I use the linear Taylor-series approximation, which was first developed by Kmenta (1967).

\[ y = \gamma \left[ \psi_1 x_1^{-\alpha} + (1 - \psi_1) x_2^{-\alpha} \right] ^ {\frac{1}{\alpha}} \]

The second order Taylor Approximation at \( \rho = 0 \)

\[ y = \gamma x_1^{\psi_1} x_2^{(1-\psi_1)} \exp(-0.5 \alpha \psi_1 (1 - \psi_1) (\ln x_1 - \ln x_2)^2) \]

In the case of my paper the production function is defined as follows, \( y = \gamma l_H^{\psi_1} l_L^{\psi_2} \),

\[ l_H = \left[ \psi_1 (l_H^f)^{-\delta_1} + (1 - \psi_1) (l_H^l)^{-\delta_1} \right] ^ {\frac{1}{\alpha - 1}} \]

\[ l_H = \left[ \psi_1 (l_H^f)^{-\delta_1} + (1 - \psi_1) (l_H^l)^{-\delta_1} \right] ^ {\frac{1}{\alpha - 1}} \]

\[ \text{and} \]

\[ \text{Kmenta (1967) justifies this only by mathematical convenience and in order to estimate around the Cobb Douglass shape.} \]
\[ t_L = \left[ \psi_2(t_L^f)^{-\delta_2} + (1 - \psi_2)(t_L^i)^{-\delta_2} \right]^{-\frac{1}{\psi_2}}, \] using the same methodology to estimate it we obtain that:

\[
\ln y \approx \ln \gamma + \rho_1 v_1 \psi_1 \ln t_H^f \ln t_H^i - \frac{1}{2} \rho_1 v_1 \psi_1 (1 - \psi_1) \delta_1 (\ln t_H^f - \ln t_H^i)^2 \\
+ \rho_2 v_2 (1 - \psi_2) \ln t_L^f + \rho_2 v_2 (1 - \psi_2) \ln t_L^i - \frac{1}{2} \rho_2 v_2 \psi_2 (1 - \psi_2) \delta_2 (\ln t_L^f - \ln t_L^i)^2
\]

To estimate:

\[
\ln y = \beta_0 + \beta_1 \ln t_H^f + \beta_2 \ln t_H^i + \beta_3 (\ln t_H^f - \ln t_H^i)^2 + \beta_4 \ln t_L^f + \beta_5 \ln t_L^i + \beta_6 (\ln t_L^f - \ln t_L^i)^2 + \epsilon
\]

The main parameters of the production function are estimated:

\[
\psi_1 = \frac{\beta_1}{\beta_1 + \beta_2}, \quad \delta_1 = -\frac{2\beta_3}{\beta_2 \psi_1}, \quad \psi_2 = \frac{\beta_4 + \beta_5}{\beta_5 \psi_2}, \quad \delta_2 = -\frac{2\beta_6}{\beta_5 \psi_2}
\]

Table 4: Estimation of the production function

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>IV(1)</th>
<th>IV(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln t_H^f )</td>
<td>2.688***</td>
<td>3.227***</td>
<td>3.687***</td>
</tr>
<tr>
<td></td>
<td>(0.57)</td>
<td>(0.55)</td>
<td>(0.57)</td>
</tr>
<tr>
<td>( \ln t_H^i )</td>
<td>2.641***</td>
<td>3.493***</td>
<td>3.875***</td>
</tr>
<tr>
<td></td>
<td>(0.47)</td>
<td>(0.42)</td>
<td>(0.51)</td>
</tr>
<tr>
<td>( \ln t_L^f )</td>
<td>2.031***</td>
<td>2.180***</td>
<td>1.649***</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.45)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>( \ln t_L^i )</td>
<td>2.189***</td>
<td>2.312***</td>
<td>2.267***</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.48)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>( \ln t_H^f - \ln t_H^i)^2 )</td>
<td>0.346***</td>
<td>0.472***</td>
<td>0.615***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>( \ln t_L^f - \ln t_L^i)^2 )</td>
<td>0.329***</td>
<td>0.375***</td>
<td>0.354***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Constant</td>
<td>16.97***</td>
<td>17.67***</td>
<td>17.25***</td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(0.71)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>Year</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Sector</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.573</td>
<td>0.689</td>
<td>0.667</td>
</tr>
<tr>
<td>( N )</td>
<td>987</td>
<td>807</td>
<td>865</td>
</tr>
</tbody>
</table>

The results of the production function estimation appear in Table 4, the dependent variable is the logarithm of the aggregate value by sector and the independent variables are the logarithm of the informality shares. In Table 4, I estimate the parameters for all of the countries together. The level of substitutability \((\delta's)\) is slightly higher in the case of the non educated workers, but even for the educated it is relatively high. The shares of workers are lower than 0.5, so the terms \(B_1\) and \(B_2\) are closer to 1, if these terms are lower than 1, the informality decrease with respect to \(\delta_1\) and \(\delta_2\).

The second and third columns present the IV estimation instrumented by the lags. In the second column, the estimation is through the inclusion of the error term of the first steps. Note that the joint test of these error
terms is significative\textsuperscript{14}. In the third column, the estimation is instrumented by the predicted estimation from the first step. In the next section, I will use the estimation from the second column.

4.2 Estimation of the quality and cost parameters by country

In the estimation of the model’s parameters through the Methods of Moments, I estimate the parameter minimizing the distance between simulated outcomes of the model and the data for the quality of government enforcement \( q(e) \), the quality of the benefits that the formal workers receive from the government \( K(\kappa) \), the level of informality in both levels of education, the level of lump sum taxes, and the share of public expenditure.

<table>
<thead>
<tr>
<th>Country</th>
<th>Model Data</th>
<th>Model Data</th>
<th>Model Data</th>
<th>Model Data</th>
<th>Model Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>0.2616</td>
<td>0.338</td>
<td>0.2193</td>
<td>0.228</td>
<td>0.2343</td>
</tr>
<tr>
<td>Brazil</td>
<td>0.1244</td>
<td>0.079</td>
<td>0.0668</td>
<td>0.026</td>
<td>0.1151</td>
</tr>
<tr>
<td>Colombia</td>
<td>0.2002</td>
<td>0.232</td>
<td>0.2907</td>
<td>0.431</td>
<td>0.3165</td>
</tr>
<tr>
<td>Peru</td>
<td>0.1884</td>
<td>0.209</td>
<td>0.1458</td>
<td>0.161</td>
<td>0.1602</td>
</tr>
<tr>
<td>Uruguay</td>
<td>0.0001</td>
<td>0.000</td>
<td>-0.0003</td>
<td>0.000</td>
<td>-0.0002</td>
</tr>
</tbody>
</table>

As is discussed in the introduction there are no clear variables to match in the case of the quality functions, so I want to compare how the fit is in the different countries\textsuperscript{15}. In the case of the quality of government enforcement I match the indicator with compliance with the law, and the quality of the benefits is matched with the indicator taxes well spent. Informality shares are estimated with the National Household Surveys, the public expenditures are matched with the figure from the National Accounts, and the level of the lump sum tax is matched at zero\textsuperscript{16}.

Firstly, I present the result of the estimation for each country independently. The model fits the data quite well in general and in particular the ranking of the countries is respected, as is presented in Table 5. This latter point is somewhat important given the comparative objective of this paper, principally, if the comparison is between the countries with better performance, such as Brazil and Uruguay, with those with low performance as in the case of Peru. However, there are some features that I want to highlight: informality for the low educated workers is underestimated, and probably the main problem is the overestimation in the high educated ones. The quality of government enforcement and quality of benefits are well estimated with the exception of the government enforcement in Brazil. The share of public expenditure is well estimated, and the taxes, which is a residual variable, fit well at zero.

In Figures 4 and 5 the estimation of the parameters for all five countries can be observed. Given that the effect of the quality parameters and the enforcement go in the same direction, it is not possible to observe a clear ranking in them. Thus, I will compare two groups of parameters, one that resume the tradeoff between quality and cost of the enforcement, and the second one respect to the benefits. So, the term \( \left( \frac{\alpha(\phi_1 - \phi_2)}{4\kappa} \right) \) link enforcement with informality in Equation 24 is higher in countries with a lower level of informality, especially when the extreme countries (Uruguay or Brazil and Peru) are compared. Considering the countries with a low levels of informality as Uruguay and Brazil, the term \( \left( \frac{\alpha(\phi_1 - \phi_2)}{4\kappa} \right) \) are 6.5 and 5.3 respectively and in the Peruvian case is only 3.3.

The second term that I analyze is \( \left( \frac{d}{4\phi} \right) \), which link benefits and informality. In this case the ranking is not so clear as in the latter case. Again, the term of Peru is the lowest one (1.08), while the terms are quite similar for Argentina, Colombia and Uruguay (2.45, 2.14 and 2.22 respectively). Note as in the comparative statics section is analyzed, informality is more sensitive to enforcement parameters than the benefits ones.

\textsuperscript{14}F(4, 775) = 9.41 \textit{Prob} > F = 0.00

\textsuperscript{15}The variables to fit are taken by surveys which are published for the IADB

\textsuperscript{16}The wages are not considered as moments because there is no data about work hours in Colombia, and there is a lack of information to construct (comparatively) the formal wages including benefits for all the countries.
Figure 4: Estimation of all fines and quality parameters

Note: Production function’s parameters $\delta_1 = -0.563$, $\delta_2 = -0.669$, $\psi_1 = 0.480$, $\psi_2 = 0.485$, $\rho_1 \upsilon_1 = 6.72$ and $\rho_2 \upsilon_2 = 4.92$. The standard deviations are estimated by the gradient of the moments vector.

Figure 5: Estimation of cost parameters

Note: Production function’s parameters $\delta_1 = -0.563$, $\delta_2 = -0.669$, $\psi_1 = 0.480$, $\psi_2 = 0.485$, $\rho_1 \upsilon_1 = 6.72$ and $\rho_2 \upsilon_2 = 4.92$. The standard deviations are estimated by the gradient of the moments vector.
In order to disentangle how each parameter can differentiate the optimal level of the quality functions and the level of the informality for these five countries, I will estimate the parameters for all countries together, leaving only one parameter being different country by country. This exercise allow me to analyze how powerful the model is creating heterogeneity, and which is the role in the heterogeneity of each parameter.

First of all, in Figures 6 - 8 show the estimation and in Tables A.2 - A.6 the moment matching when the fines \( \phi_1 \), the market imperfection \( \phi_2 \), quality of the government enforcement \( a \), the quality of the benefits \( d \), and the cost parameters \( b_1 \) and \( b_2 \) are different country by country using the levels of informality, the quality of enforcement and benefits, the lump sum tax and public expenditure as moments. Obviously, the model fits worse than when the countries all have different parameters, but in general the model fits pretty well and the ranking of countries that the data shows are respected.

Table A.2 shows the model matching when the fines \( \phi_1 \) are different country by country, informality among high educated is the moment which has most problems to fit with a clear overestimation (about double of the data), but this moment has also problems when I fit the general model. Regarding the fit with the other moments, there is no clear tendency (neither underestimation nor overestimation for all countries), but is Colombia the country with more fit problems.

In the first graph of Figure 6, I observe the performance (point estimation and confidence interval) of \( \phi_1 \), which is the amount of fines that the firm would have to pay to have workers in informality, leaving the other parameters constant country by country. There is a clear and direct relation between the level of fines and level of informality, in particular it is higher in Brazil and Uruguay, 1.4 and 0.88, than in Argentina and Peru, 0.2 and 0.025 respectively. In the case of Colombia, it has a relatively high level of fines that is reflected in much lower levels of informality in low educated workers, which are not documented in the data (see Table A.2). The ability of \( \phi_1 \) to capture heterogeneity in the model is a feature which is in line with the empirical evidence of the literature.

Table A.3 shows the moment matching when the market imperfection \( \phi_2 \) is different country by country, the high informality is again overestimated, and the low informality is well estimated for those countries with low levels (Brazil and Uruguay) but there is an underestimation for the other ones.

In the second graph of Figure 6, there is the parameter estimation when \( \phi_2 \) (market imperfection) is different country by country. Countries with high levels of informality, such as Argentina and Peru present high values of \( \phi_2 \) and Brazil and Uruguay have lower values. In this case, Colombia is again the country in which the order does not fit, and it is reflected in the underestimation for almost ten points in the informality of low educated workers (see Table A.3). The market imperfection parameter is not able to capture the heterogeneity as well as the fines parameter was able to.

In the case of the parameters of the quality functions (enforcement and benefits). The moment matching of these estimation are presented in Tables A.4 and A.5. The matching which corresponds to the \( a \) estimation (Table A.4) shows the overestimation of the high educated informality for all countries, and the underestimation of the low educated informality for Brazil, Peru and Uruguay, but a good fit for Argentina and Colombia. The matching of \( d \) estimation (Table A.5) is similar to the previous one. For these two parameters, Colombia has not present problems as in the previous ones and follow the general trend.

In Figure 7, I present the estimation when the parameters of the quality functions change. Both quality parameters allow the model to capture the heterogeneity in the moments.

The upper graph of Figure 7 shows the estimation of the quality of government enforcement, and the right panel the quality of the benefits. The estimations of the quality parameters show the same order as the informality and not the order of the quality moments among countries. These estimations show that better quality parameters means lower informality.

Finally, Figure 8 shows the estimation of the cost function, and Table A.6 is the moment matching of this estimation. The worst fit in this case is Argentina and Uruguay that underestimate the low educated informality (simulated moments are the half of the data). However, cost function is able to reproduce the heterogeneity,
Figure 6: Estimation of fines ($\phi_1$) and market imperfection ($\phi_2$) parameters

Note: Production function’s parameters $\delta_1 = -0.563$, $\delta_2 = -0.669$, $\psi_1 = 0.480$, $\psi_2 = 0.485$, $\rho_{1v1} = 6.72$ and $\rho_{2v2} = 4.92$. The standard deviations are estimated by the gradient of the moments vector.
Figure 7: Estimation of quality function parameters

Note: Production function’s parameters $\delta_1 = -0.563$, $\delta_2 = -0.669$, $\psi_1 = 0.480$, $\psi_2 = 0.485$, $\rho_1 \psi_1 = 6.72$ and $\rho_2 \psi_2 = 4.92$. The standard deviations are estimated by the gradient of the moments vector.
Figure 8: Estimation of cost function parameters

Note: Production function's parameters $\delta_1 = -0.563$, $\delta_2 = -0.669$, $\psi_1 = 0.480$, $\psi_2 = 0.485$, $\rho_1 \psi_1 = 6.72$ and $\rho_2 \psi_2 = 4.92$. The standard deviations are estimated by the gradient of the moments vector.
the model matches the order well but it has some problems in the case of Argentina. The estimation shows
that lower cost function parameters lead to low levels of informality.

Specifically through the estimation, the model is able to capture the heterogeneity of the five Latin American
countries using either all parameters or each one individually. Fines and quality parameters in enforcement and
benefits have enough capacity to replicate the heterogeneity. Market imperfection and cost parameters works
well but with less heterogeneity.

## 5 Conclusion

In this paper, I develop a model where the government determine the level of enforcement and the provision of
benefits, maximizing the utility of the households. In this framework, informality is allowed and optimal even
if its destruction is costless. I focus on the intensive margin of informality where firms and household decide
how to split the work hours.

The model captures the quality of government enforcement and benefits, and the informality for two levels of
education in five different Latin American countries which have a great heterogeneity among them. The model
allow us to analyze the effectiveness of some polices that governments can carry out to mitigate informality.

The first result is the estimation of a production function where formal and informal workers coexist when
there are two levels of education. I estimate the production function using data from all countries, and the
level of substitutability of the low educated workers is higher than the level of the high educated workers. This
estimation is in line with other papers that calibrate the substitutability rates.

The second result is the estimation of the parameters in the quality of government enforcement, the quality
of the benefits that the workers receive in formal employment, the fines, the market imperfection and the cost
function by the Method of Moments. Through this estimation, the model can capture the ranking of countries
as well as the informality for different educative levels. However, the model has some difficulties in capturing the
level of informality for high educative levels. The model is robust estimating the ranking when I make estimates
using the same parameters for all the countries and leaving only one different country by country. This is the
third main result, all parameters allow the heterogeneity with the only exception of the market imperfection
parameter $\phi_2$. The fines ($\phi_1$), quality parameters ($a$ and $d$) and cost parameters ($b_1$ and $b_2$) allow the model to
generate heterogeneity, and the moments are matched quite well.

Finally, the model through the idea that the government can choose level of endogenous enforcement captures
the main features of these economies and produces a good estimation of the parameters describing the countries’
heterogeneity.
References


Botero-García, J. (2010). El impacto de las restricciones a las exportaciones colombianas a Venezuela. DOCUMENTOS DE TRABAJO CIEF 010614, UNIVERSIDAD EAFIT.


A Appendix: Tables and figures.

Table A.1 is the estimation of the parameters for all five countries.

Table A.1: Estimation of all quality and cost parameters

<table>
<thead>
<tr>
<th></th>
<th>Argentina</th>
<th>Brazil</th>
<th>Colombia</th>
<th>Peru</th>
<th>Uruguay</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>0.8951</td>
<td>0.9180</td>
<td>1.5336</td>
<td>0.6329</td>
<td>1.9675</td>
</tr>
<tr>
<td>(d)</td>
<td>0.8426</td>
<td>0.6521</td>
<td>0.8087</td>
<td>0.6247</td>
<td>1.6775</td>
</tr>
<tr>
<td>(b_1)</td>
<td>0.0944</td>
<td>0.0645</td>
<td>0.1138</td>
<td>0.0001</td>
<td>0.0209</td>
</tr>
<tr>
<td>(b_2)</td>
<td>0.0861</td>
<td>0.1159</td>
<td>0.0943</td>
<td>0.1442</td>
<td>0.1888</td>
</tr>
<tr>
<td>(b_3)</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>(\phi_1)</td>
<td>1.1710</td>
<td>1.6423</td>
<td>1.2326</td>
<td>0.1319</td>
<td>0.3776</td>
</tr>
<tr>
<td>(\phi_2)</td>
<td>0.1878</td>
<td>0.1547</td>
<td>0.0000</td>
<td>0.1298</td>
<td>0.1016</td>
</tr>
</tbody>
</table>

Production function’s parameters \(\delta_1 = -0.563, \delta_2 = -0.669, \psi_1 = 0.480, \psi_2 = 0.485, \rho_1 \psi_1 = 6.72\) and \(\rho_2 \psi_2 = 4.92\). The standard deviations are estimated by the gradient of the moments vector.

The Tables A.2 - A.5 show the model matching when the estimation is made leaving only one parameter \((\phi_1, \phi_2, a, d, b_1\) and \(b_2)\) different among countries.

Table A.2: Moment matching leaving free the fine parameter \(\phi_1\).

<table>
<thead>
<tr>
<th></th>
<th>Argentina Model Data</th>
<th>Brazil Model Data</th>
<th>Colombia Model Data</th>
<th>Peru Model Data</th>
<th>Uruguay Model Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_i^L)</td>
<td>0.3731 0.338</td>
<td>0.1817 0.228</td>
<td>0.2363 0.349</td>
<td>0.4040 0.485</td>
<td>0.2436 0.240</td>
</tr>
<tr>
<td>(\theta_i^H)</td>
<td>0.1609 0.079</td>
<td>0.0964 0.026</td>
<td>0.1158 0.057</td>
<td>0.1710 0.117</td>
<td>0.1182 0.024</td>
</tr>
<tr>
<td>(q(e^*))</td>
<td>0.3075 0.232</td>
<td>0.4041 0.431</td>
<td>0.4069 0.317</td>
<td>0.1589 0.162</td>
<td>0.4061 0.508</td>
</tr>
<tr>
<td>(K(\kappa^*))</td>
<td>0.1582 0.209</td>
<td>0.1613 0.161</td>
<td>0.1639 0.167</td>
<td>0.1546 0.101</td>
<td>0.1641 0.358</td>
</tr>
<tr>
<td>(T)</td>
<td>0.0003 0.000</td>
<td>-0.0002 0.000</td>
<td>-0.0002 0.000</td>
<td>0.0005 0.000</td>
<td>-0.0002 0.000</td>
</tr>
<tr>
<td>(SCP)</td>
<td>0.1338 0.162</td>
<td>0.1948 0.246</td>
<td>0.1718 0.213</td>
<td>0.1267 0.110</td>
<td>0.1692 0.151</td>
</tr>
</tbody>
</table>

Table A.3: Moment matching leaving the market imperfection \(\phi_2\) free.

<table>
<thead>
<tr>
<th></th>
<th>Argentina Model Data</th>
<th>Brazil Model Data</th>
<th>Colombia Model Data</th>
<th>Peru Model Data</th>
<th>Uruguay Model Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_i^L)</td>
<td>0.2784 0.338</td>
<td>0.2214 0.228</td>
<td>0.2599 0.349</td>
<td>0.2822 0.485</td>
<td>0.2176 0.240</td>
</tr>
<tr>
<td>(\theta_i^H)</td>
<td>0.1300 0.079</td>
<td>0.1106 0.026</td>
<td>0.1238 0.057</td>
<td>0.1312 0.117</td>
<td>0.1092 0.024</td>
</tr>
<tr>
<td>(q(e^*))</td>
<td>0.1532 0.232</td>
<td>0.4521 0.431</td>
<td>0.3964 0.317</td>
<td>0.1488 0.162</td>
<td>0.4436 0.508</td>
</tr>
<tr>
<td>(K(\kappa^*))</td>
<td>0.1305 0.209</td>
<td>0.1302 0.161</td>
<td>0.1310 0.167</td>
<td>0.1305 0.101</td>
<td>0.1301 0.358</td>
</tr>
<tr>
<td>(T)</td>
<td>-0.0000 0.000</td>
<td>-0.0002 0.000</td>
<td>-0.0004 0.000</td>
<td>-0.0000 0.000</td>
<td>-0.0002 0.000</td>
</tr>
<tr>
<td>(SCP)</td>
<td>0.1419 0.162</td>
<td>0.1737 0.246</td>
<td>0.1505 0.213</td>
<td>0.1410 0.110</td>
<td>0.1741 0.151</td>
</tr>
</tbody>
</table>
Table A.4: Moment matching leaving the quality of the government enforcement parameter \( a \) free.

<table>
<thead>
<tr>
<th></th>
<th>Argentina Model Data</th>
<th>Brasil Model Data</th>
<th>Colombia Model Data</th>
<th>Peru Model Data</th>
<th>Uruguay Model Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta^*_L )</td>
<td>0.2983 0.318</td>
<td>0.1842 0.228</td>
<td>0.3173 0.349</td>
<td>0.3946 0.485</td>
<td>0.1674 0.240</td>
</tr>
<tr>
<td>( \theta^*_H )</td>
<td>0.1366 0.079</td>
<td>0.0973 0.026</td>
<td>0.1428 0.057</td>
<td>0.1679 0.117</td>
<td>0.0910 0.024</td>
</tr>
<tr>
<td>( q(e^*) )</td>
<td>0.1952 0.232</td>
<td>0.4871 0.431</td>
<td>0.1578 0.317</td>
<td>0.0227 0.162</td>
<td>0.5467 0.508</td>
</tr>
<tr>
<td>( K(\kappa^*) )</td>
<td>0.1578 0.209</td>
<td>0.1559 0.161</td>
<td>0.1570 0.167</td>
<td>0.1504 0.101</td>
<td>0.1545 0.358</td>
</tr>
<tr>
<td>( T )</td>
<td>-0.0001 0.000</td>
<td>-0.0002 0.000</td>
<td>-0.0000 0.000</td>
<td>0.0004 0.000</td>
<td>-0.0001 0.000</td>
</tr>
<tr>
<td>SCP</td>
<td>0.1524 0.162</td>
<td>0.1931 0.246</td>
<td>0.1474 0.213</td>
<td>0.1291 0.110</td>
<td>0.2020 0.151</td>
</tr>
</tbody>
</table>

Table A.5: Moment matching leaving the quality of the benefits \( d \) free.

<table>
<thead>
<tr>
<th></th>
<th>Argentina Model Data</th>
<th>Brasil Model Data</th>
<th>Colombia Model Data</th>
<th>Peru Model Data</th>
<th>Uruguay Model Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta^*_L )</td>
<td>0.2759 0.348</td>
<td>0.2481 0.228</td>
<td>0.3284 0.349</td>
<td>0.3630 0.485</td>
<td>0.2021 0.240</td>
</tr>
<tr>
<td>( \theta^*_H )</td>
<td>0.1291 0.079</td>
<td>0.1198 0.026</td>
<td>0.1464 0.057</td>
<td>0.1577 0.117</td>
<td>0.1638 0.024</td>
</tr>
<tr>
<td>( q(e^*) )</td>
<td>0.3020 0.232</td>
<td>0.2926 0.431</td>
<td>0.3195 0.317</td>
<td>0.3278 0.162</td>
<td>0.2721 0.508</td>
</tr>
<tr>
<td>( K(\kappa^*) )</td>
<td>0.2079 0.209</td>
<td>0.2861 0.161</td>
<td>0.0803 0.167</td>
<td>0.0066 0.101</td>
<td>0.4398 0.358</td>
</tr>
<tr>
<td>( T )</td>
<td>-0.0000 0.000</td>
<td>0.0001 0.000</td>
<td>-0.0002 0.000</td>
<td>-0.0003 0.000</td>
<td>-0.0004 0.000</td>
</tr>
<tr>
<td>SCP</td>
<td>0.1544 0.162</td>
<td>0.1621 0.246</td>
<td>0.1444 0.213</td>
<td>0.1408 0.110</td>
<td>0.1789 0.151</td>
</tr>
</tbody>
</table>

B  Appendix: Household problem

The complete F.O.C. of the Representative household problem are:

\[
C \frac{\partial L}{\partial C} = 0 : \frac{\partial U(C)}{\partial C} + \lambda^H_H = 0 \quad (B.1)
\]

\[
\theta^*_L \frac{\partial L}{\partial \theta^*_L} = 0 : \theta^*_L \left( -\lambda^H_H \left( \omega^*_L \left( 1 - \phi_2 q(e) \right) \right) + \lambda^H_H \right) = 0 \quad (B.2)
\]

\[
\theta^*_H \frac{\partial L}{\partial \theta^*_H} = 0 : \theta^*_H \left( -\lambda^H_H \left( \omega^*_H \left( 1 + K(\kappa) \right) \right) + \lambda^H_H \right) = 0 \quad (B.3)
\]

\[
\theta^*_L \frac{\partial L}{\partial \theta^*_H} = 0 : \theta^*_L \left( -\lambda^H_H \left( \omega^*_L \left( 1 - \phi_2 q(e) \right) \right) + \lambda^H_H \right) = 0 \quad (B.4)
\]

\[
\lambda^H_H \frac{\partial L}{\partial \lambda^H_H} = 0 : \lambda^H_H \left( C - \sum_s \left[ \omega^*_s \theta^*_L \left( 1 - \phi_2 q(e) \right) - \omega^*_s \theta^*_H \left( 1 + K(\kappa) \right) \right] + T - \Pi \right) = 0 \quad (B.5)
\]

\[
\lambda^H_L \frac{\partial L}{\partial \lambda^H_L} = 0 : \lambda^H_L \left( -x + \theta^*_L + \theta^*_H \right) = 0 \quad (B.7)
\]

\[
\lambda^H_H \frac{\partial L}{\partial \lambda^H_H} = 0 : \lambda^H_H \left( x - 1 + \theta^*_H + \theta^*_H \right) = 0 \quad (B.8)
\]
C Appendix: Comparative Statics of the equilibrium variables

In this section I show the comparative statics of the informality shares, the level of government enforcement and the benefits ($\theta^i_L$, $\theta^i_H$, $\epsilon$ and $\kappa$) and the main exogenous variables ($\phi_1$, $\phi_2$, $\delta$, $\psi$, $a$ and $d$).

Some terms for $j = \{1, 2\}$:

\[
A_j = \frac{1}{\delta_j + 1} \left[ \psi_j \frac{1 + E(\kappa)}{1 - \psi_j 1 - \phi_2 q(\epsilon) (1 + \tau)} \right]^{\delta_j + 2} \psi_j \frac{1}{1 - \psi_j (1 - \phi_2 q(\epsilon)) (1 + \tau)} > 0
\]

\[
E_1 = (\omega^i_L \theta^i_L + \omega^i_H \theta^i_H) (\phi_1 - \phi_2) \frac{\partial^2 q}{\partial \kappa^2} - \frac{\partial^2 B}{\partial \kappa^2} < 0
\]

\[
K_1 = (\omega^i_L (1 - \theta^i_L) + \omega^i_H (1 - \theta^i_H)) \frac{\partial^2 K}{\partial \kappa^2} - \frac{\partial^2 B}{\partial \kappa^2} < 0
\]

\[
B_j = \left[ \psi_j \frac{1 + K(\kappa)}{1 - \psi_j 1 - \phi_2 q(\epsilon) (1 + \tau)} \right] > 0
\]

\[
C_j = \frac{1}{\delta_j + 1} \left[ \psi_j \frac{1 + K(\kappa)}{1 - \psi_j 1 - \phi_2 q(\epsilon) (1 + \tau)} \right]^{\delta_j + 2} \psi_j \frac{1 + K(\kappa)}{1 - \psi_j (1 - \phi_2 q(\epsilon)) (1 + \tau)} > 0
\]

Denominator $D$:

\[
D = \frac{\sqrt{x}}{(x - \theta^i_L)^2 \left[ 1 - x \frac{(1 - x - \theta^i_H)^2}{(x - \theta^i_L)^2} \right] E_1 K_1 + \left[ \omega^i_L \theta^i_L + \omega^i_H \theta^i_H \right] \frac{\partial K}{\partial \kappa} E_1 - \left( \omega^i_L + \omega^i_H \right) \frac{\partial^2 K}{\partial \kappa^2} K_1}{(1 - \phi_2 q(\epsilon)) (1 + \tau)} \geq 0
\]

In the interval of interest of the parameters the simulation shows that the positive terms are higher than the negative one, then:

\[
D > 0
\]

C.1 Comparative statics of $\theta^i_L$ and $\theta^i_H$

In this section I show the comparative statics of $\theta^i_L$ and $\theta^i_H$ with respect to the main exogenous parameters of the model. The change of $\theta^i_L$ and $\theta^i_H$ with respect to $\phi_1$ is negative, the amount of fines impact negatively on the equilibrium informal shares:

Numerator $N_{11}$:

\[
\frac{\partial \theta^i_L}{\partial \phi_1} = \frac{N_{11}}{D} < 0
\]

\[
N_{11} = A_2 \left( 1 + K(\kappa) \right) \left[ 1 - x \frac{(1 - x - \theta^i_H)^2}{(x - \theta^i_L)^2} \right] K_1 \left[ \frac{(1 + K(\kappa))}{1 - \phi_2 q(\epsilon)} \left( \omega^i_L \theta^i_L + \omega^i_L \theta^i_H \right) \frac{\partial^2 q}{\partial \kappa^2} \right] q(\epsilon) E_1 < 0
\]

Numerator $N_{21}$:

\[
\frac{\partial \theta^i_H}{\partial \phi_1} = \frac{N_{21}}{D} < 0
\]

\[
N_{21} = A_1 \left( 1 + K(\kappa) \right) \frac{x}{(x - \theta^i_L)^2} K_1 \left[ \frac{(1 + K(\kappa))}{1 - \phi_2 q(\epsilon)} \left( \omega^i_L \theta^i_L + \omega^i_L \theta^i_H \right) \frac{\partial^2 q}{\partial \kappa^2} \right] q(\epsilon) E_1 < 0
\]
The impact is higher in the lower (higher) skilled informal worker if:

\[
\frac{\partial \theta^i_L}{\partial \phi_1} \geq \frac{\partial \theta^i_H}{\partial \phi_1} \quad \text{if} \quad \frac{A_2(1 - x)}{(1 - x - \theta^i_H)^2} \geq \frac{A_1 x}{(x - \theta^i_L)^2}
\]

In the case of the comparative statics of \(\theta^i_L\) and \(\theta^i_H\), the sign is uncertain.

**Numerator \(N_{12}\):**

\[
N_{12} = -A_2 \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \frac{1 - x}{(1 - x - \theta^i_L)^2} K_1 \left[ \left(1 + \phi_1 q(e)\right) q(e) E_1 + (\phi_1 + \phi_2) (\omega^l \theta^i_L + \omega^h \theta^i_H) \frac{\partial \eta}{\partial e} \right] \quad \text{?? (C.6)}
\]

**Numerator \(N_{22}\):**

\[
N_{22} = -A_1 \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \frac{x}{(1 - x - \theta^i_H)^2} K_1 \left[ \left(1 + \phi_1 q(e)\right) q(e) E_1 + (\phi_1 + \phi_2) (\omega^l \theta^i_L + \omega^h \theta^i_H) \frac{\partial \eta}{\partial e} \right] \quad \text{?? (C.7)}
\]

Analyzing the comparative statics of \(\theta^i_L\) with respect to \(\delta_2\), the sign depends on \(B_1\). If \(B_1\) is higher than 1, the effect of a higher level of substitutability impacts positively on the informality. In the case of \(\theta^i_H\), the sign is uncertain.

**Numerator \(N_{13}\):**

\[
N_{13} = \ln \frac{B_2 B_2^{2/3}}{B_1^{2/3}} \frac{1}{(2_2 + 1)^2} \times \left[ \frac{1 - x}{(1 - x - \theta^i_H)^2} E_1 K_1 - A_1 \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left(\phi^2 - \phi_2^2\right) \frac{\partial \eta}{\partial e} K_1 + A_1 (1 + \phi_1 q(e)) \frac{\partial K}{\partial e} \omega^l \theta^i L E_1 \right] \quad \text{?? (C.9)}
\]

\[
> 0 \quad \text{(by simulation)}
\]

if \(B_2 > 1\)

**Numerator \(N_{23}\):**

\[
N_{23} = A_1 \ln \frac{B_2 B_2^{2/3}}{B_1^{2/3}} \frac{1}{(2_2 + 1)^2} \left[ \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left(\phi^2 - \phi_2^2\right) \frac{\partial \eta}{\partial e} K_1 - (1 + \phi_1 q(e)) \omega^l \theta^i H E_1 \right] \quad \text{?? (C.10)}
\]

When \(\psi_2\) changes, impacts negatively on the low skilled informality and has an uncertain effect on the higher ones:

**Numerator \(N_{14}\):**

\[
N_{14} = - \frac{C_2}{(1 - \psi_2)^2} \times \left[ \frac{1 - x}{(1 - x - \theta^i_H)^2} E_1 K_1 - A_1 \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left(\phi^2 - \phi_2^2\right) \frac{\partial \eta}{\partial e} K_1 + A_1 (1 + \phi_1 q(e)) \frac{\partial K}{\partial e} \omega^l \theta^i L E_1 \right] < 0
\]

**Numerator \(N_{24}\):**

\[
N_{24} = - \frac{C_2}{(1 - \psi_2)^2} \times \left[ \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left(\phi^2 - \phi_2^2\right) \frac{\partial \eta}{\partial e} K_1 - (1 + \phi_1 q(e)) \omega^l \theta^i H E_1 \right] < 0
\]

The sign of \(\theta^i_H\) when \(\delta_2\) changes depends on \(B_2\). If \(B_2\) is higher than 1, the effect of a higher level of substitutability impacts positively on the informality. In the case of \(\theta^i_L\), the sign is uncertain.
Numerator $N_{15}$:
\[
N_{15} = \ln B_1 B_2^{-1} \left(\frac{1}{(\delta_1 + 1)^2}\right) \left[A_2 \frac{1 + K(\kappa)}{1 - \phi q(\epsilon)} (\phi_2 - \phi_2^2) \frac{\partial \theta}{\partial \epsilon} \right]^2 \omega^L K_1 - A_2 (1 + \phi q(\epsilon)) \frac{\partial K}{\partial \epsilon} E_1 \right] \right] \frac{\partial K}{\partial \epsilon} \omega^L E_1 \right] \right] \frac{\partial K}{\partial \epsilon} \omega^L E_1 \right]}
\]

Numerator $N_{25}$:
\[
N_{25} = \ln B_1 B_2^{-1} \left(\frac{1}{(\delta_1 + 1)^2}\right) \left[A_2 \frac{1 + K(\kappa)}{1 - \phi q(\epsilon)} (\phi_2 - \phi_2^2) \frac{\partial \theta}{\partial \epsilon} \right]^2 \omega^L K_1 - A_2 (1 + \phi q(\epsilon)) \frac{\partial K}{\partial \epsilon} E_1 \right] \right] \frac{\partial K}{\partial \epsilon} \omega^L E_1 \right]}
\]

The change of $\psi$ has an uncertain effect on $\theta_L^i$, and a negative effect on $\theta_H^i$:

Numerator $N_{16}$:
\[
N_{16} = -\frac{C_1}{(1 - \psi_1)^2} \left[A_2 \frac{1 + K(\kappa)}{1 - \phi q(\epsilon)} (\phi_2 - \phi_2^2) \frac{\partial \theta}{\partial \epsilon} \right]^2 \omega^L K_1 - A_2 (1 + \phi q(\epsilon)) \frac{\partial K}{\partial \epsilon} E_1 \right] \right] \frac{\partial K}{\partial \epsilon} \omega^L E_1 \right]}
\]

The change of the quality parameters $a$ and $d$ on $\theta_L^i$ and $\theta_H^i$ is negative. If the quality parameters are higher the informal shares are reduced:

Numerator $N_{17}$:
\[
N_{17} = A_2 \frac{1 + K(\kappa)}{1 - \phi q(\epsilon)} (1 - \theta_H^i) \left(\phi_1 + \phi_2\right) \left[\phi_1 + \phi_2\right] (\omega^L \theta_L^i + \omega^H \theta_H^i) \frac{\partial^2 q}{\partial \epsilon \partial \epsilon} K_1 < 0 \right] \]

Numerator $N_{27}$:
\[
N_{27} = A_1 \frac{1 + K(\kappa)}{1 - \phi q(\epsilon)} \left(1 - \theta_H^i\right) \left(\phi_1 + \phi_2\right) K_1 \left[(\omega^L \theta_L^i + \omega^H \theta_H^i) \frac{\partial^2 q}{\partial \epsilon \partial \epsilon} - \frac{\partial^2 q}{\partial \epsilon \partial \epsilon} E_1 \right] < 0 \right] \]

Numerator $N_{28}$:
\[
N_{28} = A_1 \frac{1 + K(\kappa)}{1 - \phi q(\epsilon)} \left(1 - \theta_H^i\right) \left(\phi_1 + \phi_2\right) \left[-\frac{\partial K}{\partial \epsilon} E_1 K_1 + \left(\omega^L (x - \theta_L^i) + \omega^H (1 - x - \theta_H^i)\right) \frac{\partial^2 K}{\partial \epsilon \partial \epsilon} E_1 \right] < 0 \right] \]
The impact of the taxes $\tau$ on $\theta_L$ and $\theta_H$ is positive. If there are higher taxes (contributions in the model) the informality is higher:

Numerator $N_{19}$:

$$\frac{\partial \theta_L}{\partial \tau} = \frac{N_{19}}{D} > 0 \quad (C.34)$$

$$N_{19} = A_2 \frac{1 - x}{(1 - x - \theta_H)^2} \frac{1 + K(\kappa)}{1 + \tau} (1 + \phi_1 q(e)) E_1 K_1 > 0 \quad (C.35)$$

Numerator $N_{29}$:

$$\frac{\partial \theta_H}{\partial \tau} = \frac{N_{29}}{D} > 0 \quad (C.36)$$

$$N_{29} = A_1 \frac{x}{(x - \theta_L)^2} \frac{1 + K(\kappa)}{1 + \tau} (1 + \phi_1 q(e)) E_1 K_1 > 0 \quad (C.37)$$

The effect of the cost parameters ($b_1$ and $b_2$) on the informality is positive in all cases. If the enforcement and the benefits are more expensive, the informality goes up.

Numerator $N_{110}$:

$$\frac{\partial \theta_L}{\partial b_1} = \frac{N_{110}}{D} > 0 \quad (C.38)$$

$$N_{110} = - \frac{A_2 (1 + K(\kappa)) (\phi_1 + \phi_2)}{1 - \phi_2 q(e)} \frac{1 - x}{(1 - x - \theta_H)^2} \frac{\partial^2 B}{\partial \theta_L} \frac{\partial q}{\partial b_1} E_1 K_1 < 0 \quad (C.39)$$

Numerator $N_{210}$:

$$\frac{\partial \theta_H}{\partial b_1} = \frac{N_{210}}{D} > 0 \quad (C.40)$$

$$N_{210} = - \frac{A_1 (1 + K(\kappa)) (\phi_1 + \phi_2)}{1 - \phi_2 q(e)} \frac{x}{(x - \theta_L)^2} \frac{\partial^2 B}{\partial \theta_H} \frac{\partial q}{\partial b_1} E_1 K_1 < 0 \quad (C.41)$$

Numerator $N_{111}$:

$$\frac{\partial \theta_L}{\partial b_2} = \frac{N_{111}}{D} > 0 \quad (C.42)$$

$$N_{111} = - A_2 \frac{1 - x}{(1 - x - \theta_H)^2} \frac{\partial^2 B}{\partial b_2} \frac{\partial K}{\partial \theta_L} E_1 < 0 \quad (C.43)$$

Numerator $N_{211}$:

$$\frac{\partial \theta_H}{\partial b_2} = \frac{N_{211}}{D} > 0 \quad (C.44)$$

$$N_{211} = - A_1 \frac{x}{(x - \theta_L)^2} \frac{\partial^2 B}{\partial b_2} \frac{\partial K}{\partial \theta_H} E_1 < 0 \quad (C.45)$$

### C.2 Comparative statics of $e^*$

The effect of $\phi_1$ on the equilibrium enforcement is undetermined, but in the simulations there are a positive effect in the lower values of $\phi_1$ and then a negative one:

Numerator $N_{31}$:

$$\frac{\partial e}{\partial \phi_1} = \frac{N_{31}}{D} n.d. \quad (C.46)$$

$$N_{31} = A_2 (1 + K(\kappa)) q(e) \frac{1 - x}{(1 - x - \theta_H)^2} (\phi_1 - \phi_2) \frac{\partial q}{\partial \theta_L} + K_1$$

$$- (1 - \phi_2 q(e)) (\omega_L \theta_L + \omega_H \theta_H) \frac{\partial q}{\partial \theta_H} \frac{\partial K}{\partial \kappa} \left( \frac{1 - x}{1 - x - \theta_H} A_2 \frac{\partial K}{\partial \kappa} \omega_L K_1 \right)$$

$$+ \frac{x}{(x - \theta_L)^2} \left( - \frac{1 - x}{(1 - x - \theta_H)^2} (\omega_L \theta_L + \omega_H \theta_H) \frac{\partial q}{\partial b_1} K_1 \right)$$

$$+ A_1 (1 + K(\kappa)) q(e) \phi_1 (\phi_1 - \phi_2) \frac{\partial q}{\partial \theta_L} K_1 - A_1 (1 + \phi_1 q(e)) \left( \frac{\partial K}{\partial \kappa} \right)^2 \omega_H (\omega_L \theta_L + \omega_H \theta_H) \frac{\partial q}{\partial \theta_H} n.d. \quad (C.47)$$

When $\phi_2$ is higher the sign of the effect on $e^*$ is uncertain.
Numerator $N_{32}$:
\[
\frac{\partial e}{\partial \phi_2} = \frac{N_{32}}{D}. \tag{C.48}
\]
\[
N_{32} = (1 + \phi_1 q(e)) \left( \frac{\partial K(e)}{\partial \phi_2} \right)^2 \left( \frac{\partial q}{\partial e} \right) \left( \omega_i L_t + \omega_i \theta_i \right) \left( \frac{A_1 \omega_i x}{(x - \theta_i^e)^2} + A_2 \omega_i (1 - x) \right)
\]
\[
\frac{1 + K(e)}{1 - \phi_2 q(e)}(\phi_1 - \phi_2) q(e) \left( \frac{\partial q}{\partial e} \right) K_1 \left( \frac{A_1 \omega_i x}{(x - \theta_i^e)^2} + A_2 \omega_i (1 - x) \right)
\]
\[
\frac{1 - x}{(1 - x - \theta_i^e)^2} \left( \omega_i L_t + \omega_i \theta_i \right) \left( \frac{\partial q}{\partial e} \right) K_1 n.d
\]

The comparative statics of $e^*$ with respect to $\delta_2$ depend on $B_2$. The sign goes in the same direction as the sign of the change in informality.

Numerator $N_{33}$:
\[
\frac{\partial e}{\partial \delta_2} = \frac{N_{33}}{D} > 0 \quad \text{if} \quad B_2 > 1 \tag{C.50}
\]
\[
\frac{\partial e}{\partial \delta_2} = \frac{N_{33}}{D} < 0 \quad \text{if} \quad B_2 < 1
\]
\[
N_{33} = -\ln B_2 B_2^{2+\frac{1}{\delta_2}} \left( \frac{1 - x}{(1 - x - \theta_i^e)^2} \right) \left( \omega_i L_t (\phi_1 - \phi_2) \frac{\partial q}{\partial e} K_1 \right) > 0 \quad \text{if} \quad B_2 > 1
\]
\[
N_{33} = -\ln B_2 B_2^{2+\frac{1}{\delta_2}} \left( \frac{1 - x}{(1 - x - \theta_i^e)^2} \right) \left( \omega_i L_t (\phi_1 - \phi_2) \frac{\partial q}{\partial e} K_1 \right) < 0 \quad \text{if} \quad B_2 < 1
\]

Comparative statics of $e^*$ respect to $\psi_2$:

Numerator $N_{34}$:
\[
\frac{\partial e}{\partial \psi_2} = \frac{N_{34}}{D} < 0 \tag{C.52}
\]
\[
N_{34} = \frac{C_2}{(1 - \psi_2^2) (1 - x - \theta_i^e)^2} \left( \omega_i L_t (\phi_1 - \phi_2) \frac{\partial q}{\partial e} K_1 \right) < 0 \tag{C.53}
\]

The comparative statics of $e^*$ with respect to $\delta_1$ depend on $B_1$. The sign goes in the same direction as the sign of the change in informality.

Numerator $N_{35}$:
\[
\frac{\partial e}{\partial \delta_1} = \frac{N_{35}}{D} > 0 \quad \text{if} \quad B_1 > 1 \tag{C.54}
\]
\[
\frac{\partial e}{\partial \delta_1} = \frac{N_{35}}{D} < 0 \quad \text{if} \quad B_1 < 1
\]
\[
N_{35} = \ln B_1 B_1^{2+\frac{1}{\phi_1}} \left( \frac{x}{(1 - \theta_i^e)^2} \right) \left( \omega_i L_t (\phi_1 - \phi_2) \frac{\partial q}{\partial e} K_1 \right) > 0 \quad \text{if} \quad B_1 > 1
\]
\[
N_{35} = \ln B_1 B_1^{2+\frac{1}{\phi_1}} \left( \frac{x}{(1 - \theta_i^e)^2} \right) \left( \omega_i L_t (\phi_1 - \phi_2) \frac{\partial q}{\partial e} K_1 \right) < 0 \quad \text{if} \quad B_1 < 1 \tag{C.55}
\]

As in the case of $\psi_2$, the effect on $e^*$ of a positive change in $\psi_1$ is negative, when the informal share has lower weight in the production function, the equilibrium share is lower and also the level of enforcement.

Numerator $N_{36}$:
\[
\frac{\partial e}{\partial \psi_2} = \frac{N_{36}}{D} < 0 \tag{C.56}
\]
\[
N_{36} = \frac{C_1}{(1 - \psi_1)^2} \left( \frac{x}{(x - \theta_i^e)^2} \right) \left( \omega_i L_t (\phi_1 - \phi_2) \frac{\partial q}{\partial e} K_1 \right) < 0 \tag{C.57}
\]

The effect of the quality parameters is uncertain in the case of $a$, which is one in the enforcement function, and negative in the one of the benefits function ($d$).
C.3 Comparative statics of $\kappa$:

The effect of the taxes on the equilibrium level of benefits is positive, more fines produce more resources for the formal workers.

Numerator $N_{37}$:

$$\frac{\partial e}{\partial a} = \frac{N_{37}}{D}$$

$N_{37} = -(1 + \phi_1 q(e)) \left( \omega_L^{'} \theta_L^{'} + \omega_H^{'} \theta_H^{'} \right) \left( \phi_1 - \phi_2 \right) \frac{\partial^2 q}{\partial \mathcal{E} \partial a} \left[ \frac{\partial K}{\partial e} \right]^2 \left( \frac{A_1 \omega_H^{'} x}{(x - \theta_H^{'})^2} + \frac{A_2 \omega_H^{'} (1 - x)}{(1 - x - \theta_H^{'})^2} \right)$

$$+ \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left( \phi_1 - \phi_2 \right) \frac{\partial K}{\partial e} \left( \omega_L^{'} \theta_L^{'} + \omega_H^{'} \theta_H^{'} \right) \left( \phi_1 - \phi_2 \right) \frac{\partial^2 q}{\partial \mathcal{E} \partial a} K_1$$

Numerator $N_{38}$:

$$\frac{\partial e}{\partial d} = \frac{N_{38}}{D}$$

$N_{38} = \left[ (1 + \phi_1 q(e)) \left( \phi_1 - \phi_2 \right) \frac{\partial K}{\partial e} \left( \omega_L^{'} \theta_L^{'} + \omega_H^{'} \theta_H^{'} \right) \left( \phi_1 - \phi_2 \right) \frac{\partial^2 q}{\partial \mathcal{E} \partial a} \frac{\partial K}{\partial e} \right] < 0$

Comparative statics of $e$ respect to $b_1$ and $b_2$:

Numerator $N_{39}$:

$$\frac{\partial e}{\partial r} = \frac{N_{39}}{D}$$

$N_{39} = \frac{1 + K(\kappa)}{1 + r} \left( 1 + \phi_1 q(e) \right) \left( \phi_1 - \phi_2 \right) \frac{\partial K}{\partial e} K_1 \left[ \frac{A_1 x \omega_H^{'} x}{(x - \theta_H^{'})^2} + \frac{A_2 (1 - x) \omega_H^{'} x}{(1 - x - \theta_H^{'})^2} \right] > 0$

C.3 Comparative statics of $\kappa^*$

The effect of $\phi_1$ on the equilibrium level of benefits is positive, more fines produce more resources for the formal workers:

Numerator $N_{41}$:

$$\frac{\partial \kappa}{\partial \phi_1} = \frac{N_{41}}{D}$$

$N_{41} = \left[ A_1 \omega_H^{'} x \frac{x}{(x - \theta_H^{'})^2} + \frac{A_2 \omega_H^{'} (1 - x)}{(1 - x - \theta_H^{'})^2} \right]$ \times \left[ \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left( \phi_1 + \phi_2 \right) \left( \omega_L^{'} \theta_L^{'} + \omega_H^{'} \theta_H^{'} \right) \left( \phi_1 - \phi_2 \right) \frac{\partial^2 q}{\partial \mathcal{E} \partial a} \frac{\partial K}{\partial e} - \left( 1 + K(\kappa) \right) q(e) \frac{\partial K}{\partial e} E_1 \right] > 0$

Comparative statics of $e$ respect to $\phi_2$:

Numerator $N_{42}$:

$$\frac{\partial e}{\partial \phi_2} = \frac{N_{42}}{D}$$

$N_{42} = \left[ 1 + K(\kappa) \right] \left( 1 + \phi_1 q(e) \right) q(e) \left( \phi_1 + \phi_2 \right) \left( \omega_L^{'} \theta_L^{'} + \omega_H^{'} \theta_H^{'} \right) \left( \phi_1 - \phi_2 \right) \frac{\partial^2 q}{\partial \mathcal{E} \partial a} E_1 \left[ \frac{A_1 \omega_H^{'} x}{(x - \theta_H^{'})^2} + \frac{A_2 \omega_H^{'} (1 - x)}{(1 - x - \theta_H^{'})^2} \right] < 0$
The effect of $\delta_2$ in the $\kappa^*$ depend on $B_2$ as in the other variables, but the sign is the opposite than in the case of the informal shares and the enforcement.

Numerator $N_{43}$:

$$\frac{\partial\kappa}{\partial\delta_2} = \frac{N_{43}}{D} < 0 \quad \text{if} \quad B_2 > 1$$

$$\frac{\partial\kappa}{\partial\delta_2} = \frac{N_{43}}{D} > 0 \quad \text{if} \quad B_2 < 1$$  \hfill (C.71)

$$N_{43} = \frac{\ln B_2 B_1^{1/(1+\omega)}}{(\delta_2 + 1)^2} \frac{(1 - x)}{(1 - \omega L)(\delta_2 + 1)^2} \frac{\partial K}{\partial\kappa} E_1 < 0 \quad \text{if} \quad B_2 > 1$$

$$N_{43} = \frac{\ln B_2 B_1^{1/(1+\omega)}}{(\delta_2 + 1)^2} \frac{(1 - x)}{(1 - \omega L)(\delta_2 + 1)^2} \frac{\partial K}{\partial\kappa} E_1 > 0 \quad \text{if} \quad B_2 < 1$$  \hfill (C.72)

The effect of $\psi_1$ and $\psi_2$ is positive on $\kappa$. When the weight of the formal workers in the production function is higher the level of benefits in equilibrium goes in the same direction:

Numerator $N_{44}$:

$$\frac{\partial K}{\partial\psi_2} = \frac{N_{44}}{D} > 0$$  \hfill (C.73)

$$N_{44} = -C_2 \frac{(1 - x)}{(1 - \psi_2)^2} \frac{(1 - x - \theta y)^2}{(1 - x - \theta_L)^2} \frac{\partial K}{\partial\kappa} E_1 > 0$$  \hfill (C.74)

The effect of $\delta_1$ is similar as in the case of $\delta_2$:

Numerator $N_{45}$:

$$\frac{\partial\kappa}{\partial\delta_1} = \frac{N_{45}}{D} < 0 \quad \text{if} \quad B_1 > 1$$

$$\frac{\partial\kappa}{\partial\delta_1} = \frac{N_{45}}{D} > 0 \quad \text{if} \quad B_1 < 1$$  \hfill (C.75)

$$N_{45} = \frac{\ln B_1 B_1^{1/(1+\omega)}}{(\delta_1 + 1)^2} \frac{(1 - x)}{(1 - \omega L)(\delta_1 + 1)^2} \frac{\partial K}{\partial\kappa} E_1 < 0 \quad \text{if} \quad B_2 > 1$$

$$N_{45} = \frac{\ln B_1 B_1^{1/(1+\omega)}}{(\delta_1 + 1)^2} \frac{(1 - x)}{(1 - \omega L)(\delta_1 + 1)^2} \frac{\partial K}{\partial\kappa} E_1 > 0 \quad \text{if} \quad B_2 < 1$$  \hfill (C.76)

Comparative statics of $\kappa$ respect to $\psi_1$:

Numerator $N_{46}$:

$$\frac{\partial K}{\partial\psi_1} = \frac{N_{46}}{D} > 0$$  \hfill (C.77)

$$N_{46} = -C_1 \frac{(1 - x)}{(1 - \psi_1)^2} \frac{(1 - x - \theta y)^2}{(1 - x - \theta_L)^2} \frac{\partial K}{\partial\kappa} E_1 > 0$$  \hfill (C.78)

The effect of the quality parameters as is expected is positive in the level of benefits in equilibrium:

Numerator $N_{47}$:

$$\frac{\partial\kappa}{\partial q} = \frac{N_{47}}{D} > 0$$  \hfill (C.79)

$$N_{47} = \frac{A_1 \omega L x}{(1 - \theta_L)^2} \frac{A_2 \omega L (1 - x)}{(1 - \theta_H)^2} \frac{1 + K(\kappa)}{\partial K} \frac{\partial K}{\partial\kappa}$$

$$\left(\omega_L \theta_L + \omega_H \theta_H\right) (\phi_2^2 \phi_2) \frac{\partial q}{\partial a} \frac{\partial q}{\partial E_1} > 0$$  \hfill (C.80)

Numerator $N_{48}$:

$$\frac{\partial\kappa}{\partial d} = \frac{N_{48}}{D} > 0$$  \hfill (C.81)
\[ N_{49} = \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left( \omega_L^2 \theta_L + \omega_H^2 \theta_H \right) \frac{\partial^2 K}{\partial e \partial d} \left( \phi_1^2 - \phi_2^2 \right) \left( \frac{A_1 \omega_H x}{(x - \theta_H^e)^2} + \frac{A_2 \omega_H^2 (1 - x)}{(1 - x - \theta_H)^2} \right) \]

\[ - \left( 1 + \phi_1 q(e) \right) \left( \frac{\partial K}{\partial \kappa} \frac{\partial^2 K}{\partial e \partial d} \right) E_1 \left( \frac{A_1 \omega_H x}{(x - \theta_H^e)^2} + \frac{A_2 \omega_H^2 (1 - x)}{(1 - x - \theta_H)^2} \right) \]

\[ - \frac{x}{(x - \theta_L^e)^2} \left( \frac{1 - x}{1 - x - \theta_L^e} \right)^2 \left( \omega_L^2 \theta_L + \omega_H^2 \theta_H \right) \frac{\partial^2 K}{\partial e \partial d} E_1 > 0 \quad (C.82) \]

The effect of taxes (\( \tau \)) is negative on the benefits, there are fewer formal workers and the government in equilibrium spends more money on the enforcement task:

Numerator \( N_{49} \):

\[ \frac{\partial \kappa}{\partial \tau} = \frac{N_{49}}{D} < 0 \quad (C.83) \]

\[ N_{49} = \frac{1 + K(\kappa)}{1 + \tau} \left( 1 + \phi_1 q(e) \right) \left( \frac{A_1 \omega_H x}{(x - \theta_H^e)^2} + \frac{A_2 \omega_H^2 (1 - x)}{(1 - x - \theta_H)^2} \right) < 0 \quad (C.84) \]

Comparative statics of \( \kappa \) respect to \( b_1 \):

Numerator \( N_{410} \):

\[ \frac{\partial \kappa}{\partial b_1} = \frac{N_{410}}{D} < 0 \quad (C.85) \]

\[ N_{410} = \frac{1 + K(\kappa)}{1 - \phi_2 q(e)} \left( \phi_1 + \phi_2 q(e) \right) \frac{\partial^2 K}{\partial q \partial e} \left( \frac{A_1 \omega_H x}{(x - \theta_H^e)^2} + \frac{A_2 \omega_H^2 (1 - x)}{(1 - x - \theta_H)^2} \right) < 0 \quad (C.86) \]

### C.4 Comparative statics with respect to a couple of parameters

Figure C.1: Changes in the equilibrium with the fines (\( \phi_1 \)) and the informal wage losses (\( \phi_2 \))

(a) \( \theta_H^{\ast} \)  

(b) \( \theta_L^{\ast} \)
Figure C.2: Changes in the equilibrium with the level of substitutability and the shares ($\delta_j$ and $\psi_j$).

(a) $\theta_H^* (\delta_1, \psi_1)$
(b) $\theta_L^* (\delta_2, \psi_2)$

(c) $q(e) (\delta_1, \psi_1)$
(d) $K(\kappa) (\delta_1, \psi_1)$

(e) $q(e) (\delta_2, \psi_2)$
(f) $K(\kappa) (\delta_2, \psi_2)$
Figure C.3: Changes in the equilibrium with the quality parameters ($a$ and $d$)

(a) $\theta_{H}^{\ast}$

(b) $\theta_{L}^{\ast}$

(c) $e^{\ast}$

(d) $\kappa^{\ast}$
Figure C.4: Changes in the equilibrium with quality and cost of the enforcement (a) and (b1)

(a) $\theta_{ij}^*$

(b) $\theta_{L}^*$

(c) Enforcement
Figure C.5: Changes in the equilibrium with quality and cost of the benefits (d) and (b_2)

(a) $\theta^*_H$

(b) $\theta^*_L$

(c) Benefits