How does informal employment affect the design of unemployment insurance and employment protection? To study this question, we extend the simple model of unemployment insurance developed by Blanchard and Tirole (2008) to add informality. In particular, we consider an extended model in which individuals can both work in the informal sector (the shadow economy), where employment is not observed by the government, and simultaneously claim unemployment insurance payments.

It is important to account for the possibility of informality because in many economies a large share of total economic activity takes place in the informal sector. According to the work of Schneider and Enste (2000) and Schneider, Buehn, and Montenegro (2010), the size of the informal economy is nonnegligible for developed countries and very large for developing ones (Table 1). For example, the U.S. shadow economy during the 2000s is about 9 percent of gross domestic product (GDP). The size of the informal sector is much larger for less-developed countries. For instance, during the 2000s informality as a share of GDP was about 40 percent in Brazil.

The authors use a simple model to study the optimal design of unemployment insurance and employment protection. Workers are risk averse and face the possibility of unemployment. Firms are risk neutral and face random shocks to productivity. Workers can participate in a shadow economy, or informal sector. The model yields several lessons. First, countries should encourage formal employment to address the issue of informal employment. In extreme cases, such encouragement translates into high severance payments and negative payroll taxes. Along these same lines, unemployment payments cannot be too large. In fact, when the risk of informality is extreme, the authors find that unemployment benefits should be negative, which is (in effect) a positive tax on the lack of formal employment. (JEL D82, H55, I38, J65)

How does the risk of informality (hereafter, informality) affect the optimal design of unemployment insurance and employment protection? To study this question, we extend the simple model of unemployment insurance developed by Blanchard and Tirole (2008) to add informality. In particular, we consider an extended model in which individuals can both work in the informal sector (the shadow economy), where employment is not observed by the government, and simultaneously claim unemployment insurance payments.

It is important to account for the possibility of informality because in many economies a large share of total economic activity takes place in the informal sector. According to the work of Schneider and Enste (2000) and Schneider, Buehn, and Montenegro (2010), the size of the informal economy is nonnegligible for developed countries and very large for developing ones (Table 1). For example, the U.S. shadow economy during the 2000s is about 9 percent of gross domestic product (GDP). The size of the informal sector is much larger for less-developed countries. For instance, during the 2000s informality as a share of GDP was about 40 percent in Brazil.
Studying informality and unemployment insurance is relevant because “unreported wages” are the most common form of informality (Schneider and Enste, 2000). If employment in the informal sector is unreported, individuals can receive labor earnings and collect unemployment benefits at the same time. In fact, according to Fuller, Ravikumar, and Zhang (2013), unemployment insurance overpayments in the United States amounted to $3.3 billion in 2011. About 65 percent of such overpayment cases were related to individuals collecting unemployment insurance despite being employed.1

The unemployment insurance model presented here is extremely simple. In particular, the model is static, since all decisions are made in a single period. Workers are risk averse, so they need insurance, but firms are risk neutral. Unemployment exists because productivity is random and firms are better off firing low-productivity workers. Informality is modeled as in Álvarez-Parra and Sánchez (2009). Individuals have the option of working in an informal sector at a low wage. However, unlike Álvarez-Parra and Sánchez (2009), here we assume that the optimal policy must prevent informality. Thus, we assume the costs of informality (not modeled) are high such that in equilibrium zero informality is optimal; we discuss this point in more detail later.

We use the model to study the optimal design of unemployment insurance and employment protection under the risk of informality. We focus on the implications of the risk of

### Table 1

<table>
<thead>
<tr>
<th>Country</th>
<th>Percent of GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>8.6</td>
</tr>
<tr>
<td>Japan</td>
<td>10.8</td>
</tr>
<tr>
<td>New Zealand</td>
<td>12.4</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>12.5</td>
</tr>
<tr>
<td>Australia</td>
<td>13.8</td>
</tr>
<tr>
<td>France</td>
<td>14.9</td>
</tr>
<tr>
<td>Canada</td>
<td>15.6</td>
</tr>
<tr>
<td>Germany</td>
<td>15.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>18.8</td>
</tr>
<tr>
<td>Portugal</td>
<td>22.6</td>
</tr>
<tr>
<td>Spain</td>
<td>22.8</td>
</tr>
<tr>
<td>Italy</td>
<td>27.0</td>
</tr>
<tr>
<td>Ecuador</td>
<td>33.4</td>
</tr>
<tr>
<td>Brazil</td>
<td>39.8</td>
</tr>
<tr>
<td>Zambia</td>
<td>49.9</td>
</tr>
<tr>
<td>Bolivia</td>
<td>66.4</td>
</tr>
</tbody>
</table>

SOURCE: Schneider, Buehn, and Montenegro (2010).
informality on the tax structure that implements the optimally designed unemployment insurance and employment protection. Despite its simplicity, the framework provides some sharp insights. First, in the economy with the risk of informality, low-productivity workers who would be unemployed in the formal economy without this risk are encouraged to work in the formal sector. Along these same lines, when the risk of informality is extreme—in the sense that the wage paid in the informal sector is high—the government may want to establish high severance payments. Similarly, in such situations, the government must set negative payroll taxes to make employment in the formal sector attractive. Another lesson from our framework is that unemployment payments cannot be too generous in economies with labor informality because unemployed workers could potentially claim unemployment benefits in the formal sector and work in the informal sector. In fact, if the risk of informality is extreme, unemployment benefits must be negative, which can be interpreted as a tax on informality.

The literature studying the optimal design of unemployment insurance in dynamic models is large and began with the seminal work of Hopenhayn and Nicolini (1997). The problem is modeled as a repeated principal-agent problem involving a risk-averse agent (the worker) and a risk-neutral principal (the government) that cannot monitor the agent’s search effort. These authors emphasize that unemployment payments should depend on the entire employment history of individuals. Their main finding is that unemployment payments should decrease during the unemployment spell and post-reemployment wage taxes should increase with the duration of unemployment.

Hopenhayn and Nicolini’s (1997) study, however, assumes that all employment takes place in the formal sector (employment is observable). In a related paper, Álvarez-Parra and Sánchez (2009) incorporate an informal sector into Hopenhayn and Nicolini’s (1997) framework. The key finding of Álvarez-Parra and Sánchez is that unemployment must decline abruptly to zero after several months. This design provides incentives for workers to concentrate on their job search during the first month of unemployment (when payments are being received). If individuals do not find a job during this period, they will spend most of their time working in the informal sector.

Our analysis is organized as follows. The following section presents the benchmark economy and characterizes the optimal design of unemployment insurance and employment protection and its implementation. In a separate section, we introduce the risk of informality into the model and study its impact on the optimal design and implementation. We then present the results of solving the model numerically. This section is the key to understanding the role of key parameters. The final section presents conclusions and policy implications.

**THE BENCHMARK ECONOMY**

The benchmark economy is used to illustrate the outcome with no informal sector or private information. The economy is composed of a continuum of mass 1 of ex ante identical workers. Workers are risk averse, with utility function $U$. Individuals who do not have a job can produce $b$ units of the consumption good at home.
On the other side, there is a continuum of mass 1 of ex ante identical entrepreneurs. Entrepreneurs are risk neutral, and each entrepreneur can start and run a firm. There is a fixed cost $I$ of creating a firm, which is the same for all entrepreneurs. If a firm is created, a worker is hired, and the productivity of the match is then revealed. Productivity is given by $y$ from the cumulative distribution function $G(y)$ with density $g(y) > 0$ on $[0,1]$. It is useful to define $Y(x) = \int_0^1 y dG(y)$ as the output produced in the formal sector. Note that $Y'(x) < 0$, and we assume that $Y''(x) < 0$.

The firm, but not the worker (or, for that matter, third parties such as an insurance company or the state) observes $y$. Then the firm must take a decision: It can either retain the worker and produce or it can lay off the worker, who then becomes unemployed. Realizations of productivity matches are i.i.d across firms, and there is no aggregate risk.

Finally, there is a government that can levy taxes and provide unemployment insurance. This implies the following sequence of events:

- firms are created;
- workers and firms are matched;
- shocks are observed;
- production and employment decisions are taken; and
- taxes are collected and transfers are distributed.

### Optimal Allocation under Full Information

The full-information allocation can be characterized as defining a threshold level of productivity below which workers are laid off. Call this threshold $\bar{y}$. Let $w$ be the payment to workers who remain employed and $s$ be the payment to workers who are laid off and become unemployed.

The optimal allocation maximizes the expected utility of the representative agent subject to the economy's resource constraint; that is,

$$\max_{(w,s,\bar{y})} G(\bar{y}) U(b+s) + (1-G(\bar{y})) U(w),$$

subject to

$$\int y G'(y) dy = G(\bar{y}) s + (1-G(\bar{y})) w + I.$$ 

The optimal design, $(w^*, s^*, \bar{y}^*)$, dictates that consumption by employed and unemployed workers is the same; that is, $b + s^* = w^*$. This implies that the government provides full insurance to the representative risk-averse worker. The optimal design also dictates that the threshold of productivity, $\bar{y}$, is such that all workers with a productivity match higher than what they can produce at home, $y \geq b$, will work. All other workers, on the other hand, are laid off and become unemployed. This translates into $\bar{y}^* = b$.

We can use the resource constraint to write

$$w^* = Y(b) + b G(b) + I,$$

and so it defines the consumption (or net wage) of the employed worker, $w^*$.
Implementation of the Full-Information Optimal Allocation

The implementation can be analyzed in three stages. At stage I, the government announces the payroll tax paid by the firm for each employed worker, \(\tau\); the severance tax paid by the firm for each laid-off worker, \(f\); and the unemployment benefits provided directly by the government, \(s\). At stage II, the firm and the worker are matched and sign a contract that defines the noncontingent wage if the worker is not fired, \(w\), and the productivity cutoff level, \(\bar{y}\). At stage III, the firm (and not the worker) observes the realization of the productivity match, \(y\), and decides to either retain the worker and pay wage \((w + \tau)\) or fire the worker and pay layoff costs, \(f\).

As usual, it is easier to solve the problem by working backwards. Given \((w', s', \bar{y}')\), the government tries to design taxes \((\tau, f)\) such that the optimal allocation studied earlier can be decentralized. Note that stages II and III define the optimal behavior of the representative firm.

Stage III: Productivity Cutoff Level. For the marginal worker with productivity \(\bar{y}'\), a firm must be indifferent between retaining him and firing him; that is,

\[
\bar{y}' - (w' + \tau) = -f.
\]

Stage II: Free Entry Condition. Ex ante, the representative firm is indifferent between starting the firm at cost \(I\) and not starting the firm. As the entrepreneurs operating the firms are risk neutral, the corresponding ex ante zero-profit condition reads

\[
I = \int_{y'}^1 yG'(y)dy - G(\bar{y}')f - (1 - G(\bar{y}'))(w' + \tau).
\]

Stage I: The Government's Budget Constraint. The government sets taxes such that its budget constraint is satisfied; that is,

\[
(1 - G(\bar{y}'))\tau = G(\bar{y}')(f + s').
\]

Note that the government's budget constraint coupled with the feasibility condition implies that the free entry condition (stage II) is satisfied. Hence, the conditions in stages I and III define the optimal \((\tau', f')\) given \((w', s', \bar{y}')\).

Importantly, as optimality dictates that (i) workers with productivity above the level of home production will produce \(\bar{y}' = b\) and (ii) the consumption of workers and unemployed individuals is the same, achieving full insurance—that is, \(w' = b + s'\). Moreover, the conditions in stages I and III imply that \(f' = s'\) and \(\tau' = 0\). That is, the contribution rate—the ratio of layoff taxes to unemployment benefits \((f'/s')\)—is 1. Payroll taxes are not used to make the firm fully internalize the cost of firing a worker.\(^4\)

THE ECONOMY WITH RISK OF INFORMALITY

Consider an extended version of the original setting that includes an informal labor market. That is, once a worker is matched with a firm,\(^5\) the worker can choose to move to work in the informal sector before the productivity draw has been revealed.
The employment status of an informal worker is unobservable and, in particular, cannot be distinguished a priori from unemployment. This means there is no technology available to detect employment in the informal sector and, consequently, individuals employed in the informal sector can pretend they are unemployed and collect unemployment benefits.

Finally, if the worker is fired, he cannot move to the informal sector immediately. The wage in the informal sector is denoted by $w_I$. Importantly, we assume that $w_I > b$.

Note that in our framework it is not immediately clear that eliminating informality is necessarily optimal. Actually, informality is included in the optimal allocation studied by Álvarez-Parra and Sánchez (2009). Here, we assume it is desirable not to have workers in the informal sector for reasons beyond the scope of this article (crime, terrorism, and so on). In particular, the assumption is that informality generates a social cost that it is always larger than any benefit it may generate. This assumption is formalized with a new constraint. Given $(w, s, \bar{y})$, this ex ante incentive compatibility constraint is

$$\left(1 - G(\bar{y})\right)U'(w) + G(\bar{y})U'(b + s) \geq U'(w_I + s),$$

which be rewritten as

$$\left(1 - G(\bar{y})\right)(U'(w) - U'(b + s)) \geq U'(w_I + s) - U'(b + s).$$

**Optimal Allocation with Risk of Informality**

Now we proceed to characterize the constrained optimal allocation in this setting. Notice that under these assumptions, incentives are provided so that no worker decides to work in the informal sector; thus, the optimal design must solve

$$\max_{(w, s, \bar{y})} G(\bar{y})U'(b + s) + \left(1 - G(\bar{y})\right)U'(w),$$

subject to

$$\int_{\bar{y}}^1 yG'(y)dy = G(\bar{y})s + \left(1 - G(\bar{y})\right)w + I,$$

$$\left(1 - G(\bar{y})\right)(U'(w) - U'(b + s)) \geq U'(w_I + s) - U'(b + s).$$

Let $\lambda$ and $\chi$ denote the Lagrange multipliers corresponding to the feasibility and the incentive compatibility constraints, respectively.

From the first-order conditions (FOCs) we obtain that

$$U'(b + s') = U'(w^*) + \frac{\chi}{\left(1 + \chi\right)G'(\bar{y})} U'(w_I + s').$$

Therefore, as $\lambda^* > 0$, this implies there is no full insurance because $U$ is concave; that is, $w^* > b + s'$. The risk of labor market informality limits the provision of risk sharing as the government must limit the amount of unemployment benefits to reduce the value of informality.
On the other hand, observe that the FOCs also imply that
\[ \bar{y}' = b - \left[ \left( (b + s') - w' \right) - \left( U'(w') \right)(U'(b + s') - U'(w')) \right]. \]

Therefore, since \( U \) is strictly concave and \( w' > b + s' \), it follows that the term in brackets is positive and so \( \bar{y}' < b \); that is, there is more employment in the allocation with risk of informality than in the allocation without risk of informality.

The intuition of this result can be grasped as follows. Under the risk of informality, the government wants not only to decrease the value of an informal job but also to increase the value of a formal job. Therefore, reducing \( \bar{y}' \) increases the likelihood that a worker will keep his formal job and be paid a salary \( w' > b + s' \). This is the main effect of the risk of informality on the optimal design of unemployment insurance and employment protection: Less-productive formal sector matches are preserved to provide incentives and thus make the informal sector less attractive.

**Implementation of the Informality Risk Optimal Allocation**

Given \((w', s', \bar{y}')\), the government tries to design taxes \((\tau, f)\) such that the (constrained) optimal allocation studied earlier can be decentralized. Again, we solve the problem backwards. Recall that stages II and III define the optimal behavior of the representative firm.

**Stage III: Productivity Cutoff Level.** For the marginal worker with productivity \( \bar{y}' \), a firm must be indifferent between retaining him and firing him and so the government must set \((\tau, f)\) such that
\[ \bar{y}' = (w' + \tau) = -f. \]

**Stage II: Free Entry Condition.** Ex ante, the representative firm must be indifferent between starting the firm at cost \( I \) and not starting the firm in an equilibrium where no worker moves to the informal sector because of the incentive compatibility constraint. So,
\[ I = \int_{\bar{y}'}^1 yG'(y)dy - G(\bar{y}')f - \left( 1 - G(\bar{y}') \right)(w' + \tau). \]

**Stage I: The Government’s Budget Constraint.** Taxes must satisfy the government’s budget constraint such that
\[ (1 - G(\bar{y}'))\tau = G(\bar{y}')(w' + s'). \]

As before, the government’s budget constraint coupled with the feasibility condition implies that the free entry condition is satisfied and so equations in stages I and III determine the optimal taxes \((\tau', f')\). The combination of taxes that implements the optimally designed system is \( \tau' = G'(\bar{y}')/(s' + \bar{y}' - w') \) and \( f' = s' + \tau' - \bar{y}' \).

The risk of informality implies that formal jobs should not be taxed but rather should be subsidized. To see this, note that if \( \tau' \geq 0 \), then the government’s budget constraint implies that \( f' \leq s' \) and so \( \tau' - f' \geq s' \). The condition derived above implies that
and therefore $b + s' \geq w'$. But this contradicts the fact that full insurance is not attained as long as $\lambda' > 0$. Consequently, this implies that $\tau' < 0$. Importantly, it follows from the government’s budget constraint that $f' \geq s'$. Thus, to provide incentives for firms to create more formal jobs, the government makes the contribution ratio larger than 1. That is, to finance subsidies, firms must pay more to fire a worker than what that fired worker will receive from unemployment.

**NUMERICAL EXERCISES**

In this section, we perform numerical exercises that add to the understanding of the implications of informality for the design of optimal policies. Several assumptions must be made about functional forms and parameter values as follows: The utility function is logarithmic, and the distribution of the productivity, $y$, is $N(1, \sigma^2)$ truncated between 0 and 1.

We solve for two economies. The first is the benchmark economy, referred to as *No informality*, which is basically the model developed by Blanchard and Tirole (2008). The second is the benchmark economy extended to allow for an informal sector as described in the previous section; we refer to this case as *Risk of informality*.

All figures display the optimal policy for different values of the informal wage-to-home production ratio, $w_I/b$. Recall that working in the informal sector requires workers to relinquish the opportunity to work in the formal sector. Thus, if $w_I/b < 1$, there is actually no risk of informality because its return is too low. Thus, we consider $w_I/b \in [1,2]$.

The *Risk of informality* case has significant effects on the optimal design of unemployment insurance and employment protection. In particular, it makes it desirable to design the tax structure such that it provides incentives for firms to hire (and not fire) less-productive workers. This tax structure makes the informal sector less attractive at the cost of not firing less-productive workers who would be unemployed in the absence of the risk of entering the informal market.

First, we show how the proportion of people unemployed changes as the returns to informality increase (Figure 1). Recall that in the *No informality* case, the employment threshold, $\bar{y}'$, is equal to $b$, as shown by the red line in the figure. For the *Risk of informality* case, one key result is that the risk of informality implies that the employment threshold must be lower and, as a consequence, unemployment must be lower. In particular, if the value of staying at home, $b$, is half of the informal sector wage, $w_I$ the threshold of productivity for employment implies that unemployment must be about 24 percent with no informality and about 10 percent with the informality risk. Thus, one lesson from our framework is that formal employment must be heavily subsidized in countries with high informality. This requirement is a result of the constraint in the design of the optimal policy that prevents workers from joining the informal sector. In the *Risk of informality* model, low-productivity workers are encouraged to work in the formal sector, whereas in the *No informality* model, they would be unemployed.
But how is a higher level of employment in the formal sector achieved? Figure 2, which exhibits the behavior of payroll taxes, shows the subsidy to employment. Taxes are always in the benchmark economy with no informality. The risk of informality requires subsidies to provide incentives to firms so they destroy relatively fewer matches with low-productivity workers. For high levels of \( \frac{w_I}{b} \), making employment in the formal sector attractive requires the government to subsidize formal jobs so that effective matches are more likely, thereby increasing the utility of these formal jobs relative to the informal ones. However, recall that the larger \( \frac{w_I}{b} \), the lower the unemployment benefits as informality needs to be made less attractive and, consequently, fewer and fewer subsidies are needed as incentives because unemployment payments are used more intensively to provide incentives.

Figure 3 shows the value of unemployment payments. Again, for values of \( b = w_I \), the two cases coincide. In that case, the unemployment payment is such that consumption by unemployed and employed workers is the same. In contrast, for values of \( \frac{w_I}{b} > 1 \), informality implies that unemployment payments must decrease as the return to informality increases. This happens because unemployment payments are received with probability 1 by workers who are tempted to join the informal sector, while they are received with probability less than 1—actually \( G(\bar{y}) \)—by those who decide not to join the informal sector. Thus, \( s \) must decrease to guarantee that workers do not participate in the informal sector. This is the second lesson from our framework: To make the informal sector less attractive, unemployment payments cannot be as large in economies with labor informality.
Figure 2
Payroll Taxes

Figure 3
Unemployment Insurance Payments
Figure 4 illustrates that severance payments are increasing with respect to the relative returns to informality $w_I/b$. Recall that these payments are used to implement the desired level of employment. Thus, since unemployment must decrease with $w_I/b$ (see Figure 1), it must be the case that severance payments increase. This is another lesson from our framework: When there is the risk of informality, severance payments must be used to prevent unemployment, but the payments are not transferred to unemployed individuals; rather, they are used to finance the negative payroll tax.

In addition, note that for very high values of $w_I/b$, the value of $s$ can indeed be negative. This is a key lesson from our framework: Unemployment benefits must be negative in situations of extreme informality. This requirement can be interpreted as an indirect tax on informality. Allowing negative unemployment payments is in contrast to the assumption in Álvarez-Parra and Sánchez (2009). In their case, unemployment payments are bounded below by zero. However, we can reconcile our implications with theirs. Since they studied a dynamic setting, the comparison must be the expected discounted sum of unemployment payments in their case and unemployment benefits in our static setting. They find that the duration of payments is shortened by the presence of informality. Thus, the expected discounted sum of benefits is reduced in their framework, as it is here in ours.

Figure 5 displays total consumption by the unemployed. For values of $w_I/b > 1$, the consumption by unemployed workers is lower with informality. Actually, the higher the relative return to informality, the lower is consumption by the unemployed. This is a direct conse-
Figure 5
Consumption by Unemployed Workers

Figure 6
Consumption by Employed Workers
sequence of the behavior of unemployment payments. Recall that such payments decrease to discourage workers from participating in the informal sector.

The consumption by workers, which coincides with their wage $w$, is shown in Figure 6. Two forces move $w$ for values of $w/b > 1$. First, as $w/b$ increases, lower-productivity workers become employed and naturally the wage decreases—this is the only force that could move the wage in the No informality model. Second, since $w$ is received only by workers who do not participate in the informal sector, providing incentives to keep workers out of the informal sector is helpful. This explains why consumption by employed workers is higher than in the No informality model. This is a further lesson from our framework: Working in the formal sector must be subsidized in economies with informality.

Finally, Figure 7 shows the cost of informality. This cost is computed in terms of the percent of consumption that workers would be willing to give up ex ante to eliminate the possibility of informality. Thus, this cost captures how costly it is to prevent such individuals from entering the informal sector. This cost is positive because informality is unobservable; therefore, the government must distort the optimal allocation to deal with the risk of informality. In addition, the cost is increasing in the relative returns to informality because as $w/b$ increases, it is increasingly difficult to prevent the existence of informality.

**Figure 7**
Welfare Cost of Informality

![Graph showing the cost of informality](image_url)
CONCLUSION

This article studies optimal unemployment insurance and employment protection in the presence of the risk of informality. The analysis is performed in a simple static model along the lines of Blanchard and Tirole (2008) that is modified to incorporate the risk of informality.

We find that employment in the formal sector must be encouraged when the risk of informality exists. In particular, if the wage that individuals can obtain in the informal sector is quite high (extreme risk of informality), the government should set high severance payments and negative payroll taxes. For similar reasons, unemployment benefits cannot be as large in economies with the risk of informality as in those with no such risk. In fact, if the risk of informality is extreme, informality is taxed by setting negative unemployment benefits.

NOTES

1 This number is likely to be downward biased because it is based on random audits that most likely include only unemployment insurance recipients with jobs in the formal sector.

2 We assume that the cost is small enough, such that the return of creating a firm relative to what agents can produce at home implies that having a firm in the formal sector is optimal.

3 Note here that the “planner” in this economy is directly choosing (i) who works and who does not and (ii) the consumption by every agent, as usual.

4 Here, the term “internalization” is used in the sense of Pigou (1920).

5 This can be interpreted broadly, either because the match between the worker and the firm is new or because the worker was previously working for that firm, but his productivity follows an i.i.d. process with draws every period.

REFERENCES


