An Empirical Model of Illegal Work

Frank McIntyre*  
Brigham Young University†  
20th April 2008

Abstract

This paper develops a model of workers choosing legal or illegal employment, where legality is defined as abiding by the minimum wage and participating in a set of payroll taxes and mandated non-wage benefits. Alternatively, agents may opt to be legally or illegally self-employed. Enforcement of a minimum wage law creates a kink in the trade-off between wages and benefits, causing illegal workers to clump at the minimum wage. Using insights from the model one can directly estimate the costs of evading the law either with maximum likelihood or quantile regression. In either case, identifying information comes from the wage distribution and the plentiful cross-sectional information rather than time-series variation. Estimating the model on Brazilian data suggests that evasion costs are minimal, in which case the minimum wage law and mandated non-wage benefits are likely to be irrelevant to market outcomes.

JEL codes O15, O17, J32, J88, K42.

Keywords Minimum wage, Illegal work, Non-wage benefits, Tax evasion.

---

*Brigham Young University. Dept. of Economics. 130 FOB. Provo, UT, 84602. mcintyre@byu.edu.
†Many thanks to Tom McCurdy, John Pencavel, David McKenzie, Lars Leffgen, David Sims, Sara Lemos, Daniel Hamermesh and many seminar participants for their advice and ideas.
1 Introduction

Labor market regulation is ubiquitous, though its enforcement is not. Many countries legislate a worker's paradise of laws on benefits and wages, encouraging the growth of large, unregulated, shadow economies. It is an open question just exactly how damaging it is to workers to be in the unregulated or illegal sector of the market; if legal enforcement is lax enough, illegality may be extremely low cost.

This paper uses a large dataset from Brazil to measure the costs to workers of being illegal. The estimation accounts for the selection of workers into and out of the legal work force and the endogenous choice of worker benefits and self employment. It allows for firm and worker heterogeneity in evasion costs and variation in worker preferences and productivity.

Methodologically, this paper presents a way to recover estimates of the effect of minimum wage laws that relies principally on cross-sectional, not time-series, information. Many developing countries have large cross-sectional surveys that do not extend back far in time. The time series is often so short as to require unpalatable assumptions in order to identify minimum wage effects. This paper presents an alternative method for estimating the costs of labor market regulations such as the minimum wage, given many observations but not much detail and a relatively short time series.

The model allows workers to trade mandated benefits for wages, providing an alternative to unemployment for low wage workers. This trade-off creates a mass point of illegal workers at the minimum wage which is in line with empirical observations, but often missing from theoretical models. Even in developed countries, minimum wage distortions may not occur in employment, but in non-wage benefits. The model and estimation suggest that the one piece of evidence for labor regulation potency—the large spike of workers at the minimum wage, actually represents a substitution out of non-wage benefits into wages, rather than actual compensation increases. The model further suggests an intuitively plausible identification strategy comparing the returns to productivity above and below the minimum wage as a way to discern the costs of evading minimum wage laws.

Using either maximum likelihood estimation or quantile regression, I find that evading regulations is extremely low cost. There are no fixed costs of illegality and the marginal costs are small. The results are robust to a variety of different assumptions about unobserved distributions. The maximum likelihood an quantile regression results are consistent with a model where illegal workers sort into firms with low costs of evasion and so the market price of evasion is very low. To the extent that the two estimation approaches recover the underlying parameters, they indicate that labor regulation is neither very distortionary, nor very relevant to Brazilian workers. Using the same methodology, future work could easily estimate the same parameters for other developing countries without requiring specialized datasets or unique policy interventions.

McIntyre (2004) provides a more thorough review of the related research. In summary,
work on minimum wages in Brazil and other developing countries is limited, see Fajnzylber (2001) and Lemos (2002) for evidence from Brazil suggesting small employment and wage effects, Bell (1997) suggesting some employment effects in Colombia but none in Mexico, Terrell & Gindling (2005) showing wage effects in Costa Rica, Jones (1997) showing movement from the formal to informal sector in Ghana, Suryahadi et al. (2003) for a survey of Indonesia finding rising wages and falling employment, and Strobl & Walsh (2003) finding that the minimum wage had poor compliance rates in Trinidad with some evidence of wage and employment effects. In Brazil workers are found to move easily between formal and informal markets (Sedlacek & Paes de Barros (1990)), suggesting that segmented market models may not be as useful as competitive labor market models.

Arrowsmith et al. (2003) provide anecdotal evidence from interviews of 55 small British firms faced with a national minimum wage. Though they do not provide formal estimates of minimum wage effects, they conclude that partially or fully evading the law was an important response to the new regulation. More recently, Tonin (2008) estimates a model on Hungary data that plausibly links informality to spikes at the minimum wage. While the exact mechanism differs from the one used here, relying upon income underreporting, Tonin also uses a cross-country dataset to show that informality is strongly correlated to larger minimum wage spikes—a result very much in line with the mechanism presented here.

Research on the U.S. minimum wage, ably reviewed in Brown (1999), does not fully address the central question here of compliance. The U.S. literature finds little evidence that non-wage benefits are affected by minimum wages, but much of the research concentrates on on-the-job training (see Acemoglu & Pischke (1999), Fairris & Pedace (2004), and Neumark & Wascher (2001)), which has different properties than mandated benefits like social security. Payroll tax evasion’s relation to the minimum wage has received little attention, and understandably so given the difficulty of getting the data.

Shadow economies have an extensive literature reviewed in Schneider & Enste (2000), but there is little empirical evidence on how labor market regulation specifically affects informality. Using cross-country evidence, Loayza (1997) finds that an index of labor regulation is correlated with larger informal sectors. Saavedra & Chong (1999) estimate the costs of informality in Peru allowing for the endogeneity of informality. They find that the coefficient on education in a wage equation is higher among formal workers. Unfortunately, their definition of informality does not include minimum wage violations, making direct comparisons difficult, as accounting for the minimum wage in wage equations represents a major challenge.

The past literature tends to be reasonably but loosely tied to an underlying theoretical model and often relies on time-series and/or cross-state variation for identification (though Flinn (2002) provides an interesting exception). None of this literature brings together evasion costs, non-wage benefit regulations and minimum wage regulations in a formal economic model with directly estimable parameters. Further, the model identification presented here
comes from the wage distribution rather than time series data. While the empirical results are consistent with past work, the model allows one to interpret findings more easily. It suggests that minimum wage laws have little effect on employment because of the low evasion costs, while the estimated wage gains from minimum wage increases may well be due to workers substituting out of non-wage benefits, rather than receiving higher total compensation.

Section 2 describes the institutional setting for the estimation. Section 3 presents the model. Section 4 describes the data. Sections 5 and 6 describe the specification and results of the maximum likelihood estimation. Section 7 shows that one reaches the same conclusions using alternative estimates based on quantile regression. Section 8 concludes.

2 Institutional Setting

As shown in Table 1, Brazil has several large, mandated non-wage benefits, some of which the worker may fully value and others of which are only partially valued compared to their cost.\footnote{See Amadeo and Camargo (1997) for a careful summary. Note that in the model presented here, benefits are treated as a continuous variable ranging from full to nothing. This is a straightforward simplification of the process in which an agent values some benefits as much as cash, others partially, and others not at all. The model treats these as a composite “benefits” good, where the specific benefits and their amounts are not modeled. This captures the agent’s desire to take some amount of compensation in benefits, but often less than the legal amount, which is the basis for regulation evasion.} The taxes and benefits are all paid in proportion to the worker’s wage. When aggregated, they approximately double the cost of labor employment.\footnote{Some benefits have a cost that is multiplied by the wage plus other benefits. This compounding is what brings the total costs from 176% of the wage to double the wage. Brazil’s payroll taxes are comparable to those in many other countries. See Amadeo and Camargo (1997) for a summary across Latin American countries.} Analyses based solely on the observed wage could be very misleading if these benefits are only received by some workers. With imperfect enforcement workers will have an incentive to move partially valued benefits into fully valued wages so as to equate the marginal benefits of types of compensation.

The mandated annual bonus and individual worker unemployment fund are both redeemable as cash at some point in the future, thus their value depends on the agent’s time-preference and credit constraints. On the other hand, there are many taxes that fund government programs such as social security, accident insurance, or worker education programs. The benefits of these taxes are only tenuously linked to the individual paying them and so some workers may be very willing to ditch these benefits in exchange for a higher wage. Their willingness to pay for such benefits may reflect a preference for legality or honesty itself, rather than any clear cash benefit realized by the agent.

Figure 1 graphs the hourly minimum wage in constant 1994 Reals during the sample period of 1981-2002.\footnote{The minimum wage is deflated using the most widely used consumer price index in Brazil, the IPCA. There is one minimum wage for the whole country set annually in May but adjusted more frequently in times of high inflation. In the first three years of the sample the minimum wage for the South was slightly below the countrywide minimum.} These legal minima are high, especially compared to wages in the
poor North, as such there is a large and obvious incentive for workers and firms to evade the minimum wage law if that law precludes a labor contract.

Given the incentives, it seems certain that labor laws would be widely ignored if evasion were costless. There is though, evidence that the government goes to some lengths to enforce these laws. The Brazilian labor ministry employs thousands of inspectors charged with ensuring compliance with all labor laws, from payroll to health codes (Ministério do Trabalho, 1979). 3,285 inspectors were employed in 1990 but only 1,960 in 1995. Even though the number of inspectors stabilized in the late 90’s at about 2,400, the number of businesses and employees inspected currently fluctuates 20-30% from year to year. The ministry records indicate that between 15 and 20 million workers’ businesses are inspected in any given year (Ministerio do Trabalho e Emprego, (2000)).

Recent inspection effort, however, has been focused on stamping out slave and child labor, as opposed to minimum wage or payroll violations. Anecdotal evidence suggests that actual enforcement of these laws is often due to wronged workers complaining to the Brazilian labor courts. The labor ministry reports that non-wage benefit and minimum wage violations are prosecuted in roughly the same amounts; in 2001, the inspectors found 11,970 businesses violating the minimum wage law, 14,726 having unregistered workers, and 16,030 failing to pay FGTS, the unemployment fund benefit (Ministerio do Trabalho e Emprego, (2002)). These numbers suggest that non-wage benefit and minimum wage enforcement efforts are of comparable magnitude.

Firms in violation of these laws can be assessed up to 24 months of back-paid wages and benefits and a fine that can be 3 to 120 times the value of the monthly minimum wage. Labor violations have a statute of limitations of two years (Consolidação das Leis Trabalhistas, 1943; Brazil Legal Code 1989, 1999). Of course, what the law says and what actually occurs may be two different things. Setting aside the costs of bribery, one would like detailed information about fines actually paid and their frequency. Even then, many of the costs of evasion may actually revolve around unobserved bribery of officials. Although this information is not readily available, the estimation method here does not rely upon explicitly observing the costs of evasion, but infers those costs from the wages of workers.

3 Model

This section presents an empirically estimable model of a labor market with both legal and illegal workers. The model has two kinds of actors: individuals and firms. Firms compensate workers with a package of wages $w_i$ and a benefits rate $\tau_i$, with total compensation $w_i \tau_i$. Both benefits and wages have legal minimum levels ($B$ and $M$ respectively). Firms that pay less than the legal requirements incur a worker-specific evasion cost, $\Delta_i$, that, in equilibrium, firms pass on to the worker. Firms present each worker with a wage schedule $w_i(\tau_i)$ that higher than the North, which is accounted for in the data used. This paper uses a single, national, deflator. The estimation is robust to fixed regional price differences as controls for region are included.
maps out a budget constraint of wage/benefit packages across which the firm is indifferent. Agents optimize across wages and benefits given a productivity level, $t_i$ and a preference for non-wage benefits, $\theta_i$. The first section presents the model in its simplest form. Section 3.2 determines the equilibrium. Section 3.3 discusses the empirical implications and Section 3.4 considers important extensions.

3.1 Actors and Institutions

This section discusses the choices faced by firms, the utility maximization problem solved by individuals, and the specification of the evasion cost function.

Firms

Firms each have access to the same production technology to produce a homogeneous consumption good, taking labor as the only input. The cost of a worker depends on his desired mix of benefits and wages, as well as the costs of evading the law for informal workers. Thus firms maximize:

$$\Pi = y(T) - \sum_{i=1}^{N} \pi_i t_i \tau_i \Delta_i$$

(1)

where $y(T)$ is the production function, $T = \sum_{i=1}^{N} t_i$, $N$ is the number of employees at the firm, $t_i$ is the productivity of agent $i$, and $\pi_i$ is the "piece rate" price of a unit of productivity from worker $i$. $\tau_i$ is a multiplier on the wage covering the costs of non-wage benefits which, to be legal, are required to be at the level $B$. $\Delta_i$ captures the costs paid for evading the law for worker $i$ and is weakly decreasing in both $w_i < M$ and $\tau_i < B$. $\Delta_i = 1$ for all formal agents, who are those with $w_i \geq M$ and $\tau_i = B$. For illegal workers, $\Delta_i > 1$ and is discussed more below.

Individuals

All agents have utility defined over consumption, given an individual-specific preference for benefits: Consumption is produced from wages and benefits:

$$c_i = \psi(w_i, \tau_i \theta_i)$$

(2)

The function $\psi$ defines how individuals combine wages and non-wage benefits to produce consumable goods, given the preference parameter $\theta_i$. The function is weakly increasing and concave in both its arguments. A special case of benefit preferences would be where all workers value the benefits at their cost to the firm, in which case $\psi(\cdot) = w_i \tau_i$. Individuals maximize by choosing a wage/benefit combination according to the competitive equilibrium wage schedule $w_i(\tau_i)$ presented them by the firm.
Evasion Costs

The key to understanding the effects of regulation under incomplete enforcement is to understand the costs associated with evading the law. Define \( D = \ln(M) - \ln(\min(w, M)) + \ln(B) - \ln(\min(\tau, B)) \) as the monetary log distance between the agent’s wage/benefit package and the legal requirements. This distance is zero for all legal workers and rises as the agents move into either form of illegality. Specify evasion costs as

\[
\ln \Delta_i = \delta^0 \cdot I(D > 0) + \delta^1 \cdot D
\]

Where \( I(\cdot) \) is an indicator function, \( \delta^0 \geq 0 \) is the fixed cost of becoming informal, while \( \delta^1 \in [0, 1] \) captures the variable costs associated with increased illegality. The functional form treats evasion symmetrically in wages and benefits. Minimum wage laws and non-wage benefit violations are enforced and punished by the same agency and in the same courts, with similar fines imposed for either type of violation. Thus imposing a symmetric cost structure is a reasonable restriction, although clearly a restriction.\(^4\)

Although the model is competitive, allowing for a fixed cost of evasion does create a type of segmentation in the market. Workers barely productive enough to be legal can receive much better compensation than those only slightly less productive. Thus large fixed costs imply that market regulations impose a sharply higher cost on the illegal low-skilled workers in a fashion reminiscent of labor market segmentation models.

3.2 Equilibrium

Equilibrium requires that each worker receives a competitive wage schedule and picks his optimal point on that schedule. Also, each worker faces a price \( \pi_i \) for his labor that leaves firms indifferent across workers. Workers then receive a wage \( w_i = \pi_i t_i \). Legal worker benefits are at \( B \) and there is no evasion cost \( (\Delta_i = 1) \), so the legal or formal sector price is constant. Call this price \( \pi_F \), the piece-rate price for formal sector work.

Firms face an extra evasion cost per illegal worker, thus in equilibrium an illegal worker faces a price that makes him competitive with formal workers. Given a formal sector price of \( \pi_F \), illegal workers receive:

\[
\pi_i(\pi_F, \Delta_i, B \cdot \tau_i) = \pi_F \cdot \frac{B}{\tau_i} \cdot \Delta_i^{-1}.
\]

This price function has the firms exactly passing on their evasion costs and the value of

\[^4\text{If there are important variations in the evasion costs across types of illegality, the estimation will attempt to capture a common cost parameter that maximizes the observed likelihood. The parameter will not be correct but may be close enough to remain useful.}\]
unreceived benefits to workers. Observed wages,
\[ w_i = \pi_F \cdot t_i \cdot \frac{B}{\tau_i} \cdot \Delta_i^{-1}, \]  
(5)
are a combination of the evasion cost, productivity, and the extent (if any) to which the worker substitutes wages for benefits. Since \( \Delta_i \) itself depends on the wage, \( w_i \) is only implicitly defined by (5). Substitute the definition of \( \Delta \) into the wage schedule (5) to get the following wage function:
\[ \ln(w) = \ln(\pi_F t) - \delta^0 I(D > 0) + (1 - \delta^1) \cdot \ln \left( \frac{B}{\tau} \right) \quad w \geq M \]
\[ = \frac{\ln(\pi_F t) - \delta^0 - \delta^1 \ln M}{1 - \delta^1} + \ln \left( \frac{B}{\tau} \right) \quad w < M \]  
(6)
where the \( i \) subscripts are suppressed here and in the remainder of the section. The formal and informal markets function as one labor market with one price constant \( \pi_F \) that is adjusted when hiring an informal worker. The adjustment, for benefits not received and evasion costs paid, affects the wages observed in the data.

A worker’s optimization over wages and benefits is fairly standard, but requires some care owing to distortions from the evasion cost. Each agent is faced with a wage/benefits schedule \( w(\tau) \) from which he must pick his optimal point. To provide a tractable solution to the agent’s optimization problem, define the utility function \( \psi \) as:
\[ \ln \psi(w, \tau) = (1 + \theta) \cdot \ln \tau - \frac{1}{2} (\ln \tau)^2 + \ln w \]  
(7)
where \( \theta \in [0, \ln B] \). Over the available range of \( \theta \) and \( \tau \), this log-quadratic form is concave and increasing, which meets the requirements of the model. Further, the simple quadratic form implies the following first order tangency conditions:
\[ \ln \hat{\tau} = \begin{cases} \theta & w < M \\ \theta + \delta^1 & w > M, \tau < B \end{cases} \]  
(8)
Thus \( \theta \) is the log benefits level the agent would choose in a market with no evasion costs. Note that once benefits are full, agents can no longer add benefits and so they remain at the corner. Equation (8) only describes the tangency condition, \( \hat{\tau} \). To describe the equilibrium, next consider those whose optimal choice is at the minimum wage kink or at a corner.

As is obvious from Equation (5), agents face a kink at the minimum wage in the trade-off between benefits and wages, since agents stop being illegal in wages. The kink, illustrated in Figure 2 comes from changing the trade-off between benefits and wages. This creates a group of workers who stay at the minimum wage as their maximum utility point. Note that these workers dumped at the minimum wage are all illegal workers, in that they only clump
at the minimum wage because they are not getting full benefits.\(^5\)

For agents who may fall below the minimum wage, Define:

\[
\bar{\theta}(t) = \frac{\ln(\pi_{F}t) - \delta^{0} - \ln M}{1 - \delta^{1}} + \ln(B).
\]  

(9)

Then agents with \(\theta\) above \(\bar{\theta}(t)\) choose wages below \(M\) and \(\theta < \bar{\theta}(t) - \delta^{1}\) choose wages above \(M\) (once again, at a tangency). All those between these two levels, with \(\theta \in [\bar{\theta}(t) - \delta, \bar{\theta}(t)]\), remain at the minimum wage.

Workers capable of full legality compare the utility gained from the above choice with the gain from avoiding illegality altogether. Staying legal allows them to evade the fixed cost, \(\delta^{0}\). The utility comparison results in agents choosing full benefits if:

\[
\theta > \ln B - \delta^{1} - \sqrt{2 \cdot \delta^{0}}
\]

(10)

although this only applies to agents productive enough to be able to receive at least the minimum wage at full benefits.

Agents working at less than the minimum wage move freely between benefits and wages at a one-to-one tradeoff. This is because they are simply substituting between two kinds of illegality—as their benefits fall further below the mandated level, their wages rise closer to the minimum wage. Those working above the minimum wage who choose to forego benefits face a degraded price ratio of a \(1 - \delta^{1}\) percentage wage increase for each percentage point reduction in benefits. Dropping benefits raises marginal evasion costs by \(\delta^{1}\), but the resulting wage increase does not decrease evasion costs if the agent is already making a legal wage.

Given the above restrictions, each worker falls into one of five regions in \(t, \theta\) space. Let \(\Omega_{j}\) be the \(j\)th region where \(j \in \{1, 2, 3, 4, 5\}\). The five regions are:

\[
\begin{align*}
\Omega_{1} &= \{t, \theta \mid t \geq \frac{M}{\pi_{F}}, \theta \geq \ln B - \delta^{1} - \sqrt{2 \cdot \delta^{0}}\} \\
\Omega_{2} &= \{t, \theta \mid t \geq \frac{M}{\pi_{F}}, \theta < \ln B - \delta^{1} - \sqrt{2 \cdot \delta^{0}}\} \\
\Omega_{3} &= \{t, \theta \mid t < \frac{M}{\pi_{F}}, \theta \geq \bar{\theta}(t)\} \\
\Omega_{4} &= \{t, \theta \mid t < \frac{M}{\pi_{F}}, \theta \in [\bar{\theta}(t) - \delta^{1}, \bar{\theta}(t)]\} \\
\Omega_{5} &= \{t, \theta \mid t < \frac{M}{\pi_{F}}, \theta < \bar{\theta}(t) - \delta^{1}\}
\end{align*}
\]

(11)

which are summarized in Figure 3. Note that the condition \(w(B) \geq M\) is equivalent to

\(^5\)One can apply a modified version of this result to minimum wages in high enforcement countries. The benefits to the job might be defined not as legally required ones, but simply niceties of employment, such as those considered in Simon & Kaestner (2003). Regardless, if there are benefits workers can forego to make themselves worth the minimum wage, some may wish to do this instead of becoming unemployed. Since they value these other benefits as well as wages, once they hit the minimum wage they stop the tradeoff. Thus the minimum wage may distort non-wage benefits decisions more than labor supply decisions, even with perfect enforcement. If non-wage benefits are more elastic than employment, which seems reasonable, this may be an important place to look for distortions of the minimum wage in developed countries. Unfortunately, if informal workers are not plentifully available in survey data, there may be difficulty in identifying the parameters of interest.
\[ t \geq \frac{M}{\bar{\theta}} \text{ and } \bar{\theta} \text{ is defined above in Equation (9).} \]

The first region, \( \Omega_1 \), is the only set of legal workers. \( \Omega_2 \) are those workers able to be legal, but preferring illegality due to their low valuation of benefits. Those in \( \Omega_3 \) value benefits but are insufficiently productive to be legal. Given that they have paid the fixed costs of evasion these workers pick the exact level of benefits they find attractive, as they pay no additional evasion costs for moving to their optimal point. \( \Omega_4 \) consists of workers who forego enough benefits to reach the minimum wage, at which point they face the kink in the budget constraint illustrated in Figure 2. This makes further trade-offs undesirable and so agents in \( \Omega_4 \) lump at the minimum wage. The last region, \( \Omega_5 \), are workers who trade off enough benefits that their observed wage is above the minimum wage. Thus they are illegal due to lack of benefits and appear to be like agents in region \( \Omega_2 \). They differ in that they actually cannot be legal because adopting full benefits would lower their wage to below \( M \).

Given this model and a set of distributional assumptions, one can take an observed worker and map his wage and what one observes about his benefits choice (full or less than full) into a probability and estimate the parameters with maximum likelihood, which is done below.

### 3.3 Empirical Implications

One can see several avenues through which the evasion costs will distort the wage distribution (and so be identified). Equation (6) shows that the derivative of the wage with respect to covariates should be higher for those below the minimum wage, by a factor of \( \frac{1}{1-\theta} \). Next, the lump of informal workers at the minimum wage is a function of the slope parameter \( \delta^1 \). The gap in benefits receipt between those above and below the minimum wage is driven by the evasion cost function. In particular, if all workers pay a fixed cost, no workers with full benefits should be located in the space just below the minimum wage. The sharp dropoff in agents directly below the minimum wage can be attributed to both the fixed costs of the law and workers ditching benefits to get higher wages.

To help the reader visualize the effects of evasion on wages, Figure 4 graphs simulated wage distributions around \( M \) for workers with full benefits and those with less than full benefits. The top panel has positive fixed costs of evasion, \( (\delta^0 = 0.2) \), the bottom has positive marginal costs of evasion \( (\delta^1 = 0.2) \). As can be seen, fixed costs lead to sharp drops below the minimum wage among workers with full benefits while high marginal costs create lumps of workers at the minimum wage.

\[ w(B) \equiv \pi_t \geq M \Rightarrow \pi_t(B) = \pi_F \Rightarrow t > \frac{M}{\pi_F} \cdot w(B) \equiv \pi_t < M \Rightarrow \pi_t(B) < \pi_F \Rightarrow t < \frac{M}{\pi_F} \]

\(^6\)Of course, in the empirical estimation, no such empty spots are observed in the wage distribution, nor is maximum likelihood well suited to estimating such boundaries. The estimation deals with this by allowing the possibility that only some agents face a fixed cost, in which case one is looking for a dip, rather than a vacant spot.
3.4 Variations on the Firm

The above model does not allow for firms to differ or be imperfectly competitive. This section considers three extensions to the firm side of the model and how those extensions impact the estimation.

Adding Self Employment

Many firms get around the minimum wage law simply by not hiring any workers beyond the first—and thus fitting under self employment laws. If this option is everywhere as good as employed work, one would expect it to make minimum wage laws irrelevant, as all the displaced workers simply resort into self employed occupations.

The above model requires only slight modification to allow for self employment. The firms modeled above may have only one worker and so can readily be adapted to self employment. Agents must now have a choice to work as self employed workers, requiring that they have preferences between the two types of work. Lastly, one must specify what laws apply to the self employed.

For each worker, add an additive utility preference parameter for self employment, so that:

\[ u_i = \psi_i - S_i \gamma_i \]

where \( S_i = 1 \) if the agent is self employed and \( \gamma_i \) is the agent’s dislike for self employment, which may be positive or negative. Optimization for this simple form is straightforward. Calculate the maximum consumption \( \psi_i \) the agent can obtain in either state and then pick the greater \( u_i \).

Self employed workers do not have to obey the minimum wage laws, but do have to pay into the social security system. They do not have to pay some of the worker benefits, but do face other, less well-defined, costs of either licensing or operating their business illegally. Lacking better data on these costs, I use the same evasion cost and non-wage benefit costs used for employed workers. Of course, I do not impose minimum wage laws on the self-employed.

One may also wish to allow for workers to have differing levels of productivity in the self-employed vs. employed sectors. Preliminary, reduced-form, empirical results suggested that this was not a major source of difference between the types of employment. Wage regressions on self employed or employed workers reported similar coefficients and error variances. In fact, after segmenting by whether or not the agents received non-wage benefits, the two kinds of employment also had similar mean wages. This evidence does not prove that the two forms of employment are identical nor that particular agents may not be realizing a comparative advantage in one form of employment over the other, but it does suggest the differences may be small enough to be ignorable at this stage.
Heterogeneity in Evasion Costs

The above model does not allow firms to differ in their evasion costs. Suppose that each firm \( j \) has a unique evasion cost function, \( \Delta_j \). Formal workers will be indifferent to these differences, but in equilibrium, informal workers will move to the firms with low costs of evasion.

This is entirely analogous to the model of discrimination presented in Becker (1971). Given the similarities, the equilibrium solution is obvious as well. Evasion costs will be translated into a market price schedule for evasion, paid by all informal workers, \( \Delta_m \). This market price will equal the evasion costs of the marginal firm employing illegal workers.\(^8\)

This has two implications for the model specified above. First, I will not be estimating firm-specific evasion costs, but rather a market outcome which gives a price to evading the law. Since this market price is actually the quantity of interest, this is a useful feature of the model. Firm specific heterogeneity in evasion costs, whether it exists or not, does not affect the outcomes faced by workers.

The second implication is that, just as the presence of discriminatory firms does not imply that market outcomes are discriminatory, so the presence of firms with high evasion costs does not imply that regulations have an impact on the market. The equilibrium price for evading the law could well be zero, even though some firms, such as large corporations, will be systematically unable to evade the law.

Imperfect Competition

While I leave more complete modeling of imperfect competition to future work, some basic analysis suggests what the results would look like. A basic model of monopsony ends with firms paying workers less than their marginal product per unit of labor. Thus it implies a lower \( \pi_F \). As long as the labor market still has workers who are able to work for more than the minimum wage, competitive pressure will keep all the costs of evasion on the worker, thus the estimation of the model is unaffected, just with different prices for a unit of labor. Note that the standard result that a minimum wage can improve outcomes in an imperfectly competitive market may or may not hold. If it occurred it would be through a general equilibrium effect on prices, \( \pi_F \).

Possibly a more realistic model would be to allow for imperfectly competitive firms in either the legal or illegal sector separately. If the illegal sector is less competitive than the legal, it drives illegal wages down even further and, in the empirical estimation, this lower wage is exactly captured by the evasion cost parameters. Since these costs could plausibly be considered a cost of illegality, this seems reasonable.

Alternatively, if the legal firms have more market power than illegal ones, they will be

---

\(^8\)The fixed and variable cost components do complicate the analysis so that the market price schedule may not follow the loglinear form given here, even if it is loglinear for the underlying firms, in which case I will be estimating an approximation to the evasion price schedule.
anxious to drive down wages in the formal sector, but may be constrained by competition from the illegal firms. If illegal firms are able to compete away their workers, then the market is more competitive. If the illegal firms face evasion costs that must be passed on to workers, this provides room for the legal firms to extract rents from legal workers. In simple cases, the legal firms charge price rents equal to the going evasion cost incurred by going illegal.

In that case, estimation, which relies upon differences between legal and illegal workers, may estimate zero evasion costs because the costs are matched by monopsony rents in the legal sector. Though the resulting parameters would have a slightly different interpretation, they would still capture the marginal differences between being legal and illegal. Note that, in such a market, raising evasion costs could actually lower wages for all workers because it would lessen competition on legal firms and so lower their wage offers. I leave a more complete analysis for later work.

4 Data

The data are annual, cross-sectional, individual observations on wages, labor market participation, benefits, and family characteristics drawn from one of Brazil’s household surveys, the Pesquisa Nacional de Amostra de Domicílios (PNAD).\textsuperscript{9} The survey interviews approximately 300,000 individuals every year. The estimation uses surveys from 1981 to 2002 except 1991 and 2000, when there were national censuses, and 1994 when there were budgeting difficulties. Altogether, the dataset contains over five million observations. From this data, the estimation uses a random sample of 100,000 men.

Data are collected in October about experiences in September and include detailed labor force participation information, such as whether or not the individual participates in the social security system, possesses a legal work contract (which implies being registered with the government), and his hours worked and earnings for the month. The hourly wage is constructed from the hours and earnings data which is then deflated using the IPCA. Agents with a work contract or who pay social security are classified as taking full benefits.

The minimum wage is also deflated by the price index. Minimum wages are reported as a monthly salary which must be combined with the maximum hours one can work in order to get a minimum wage. Those working half-time are required to receive half the minimum wage. Thus it is truly a minimum wage, not just a minimum salary. The maximum hours worked changes over the sample period from 48 to 44 in 1988. For purposes of estimating

\textsuperscript{9}The PNAD micro data is available from the Brazilian census agency, the IBGE. More information is available at their web site. The PNAD is not the only household survey data available in Brazil. There is also the PME, a monthly CPS-style survey that rotates households in and out of the sample over one year. Unfortunately the PME is drawn exclusively from the largest metropolitan areas, and hence is not nationally representative. The rural workers missed by the PME are especially relevant to formal and informal work. These rural workers make decisions about whether to migrate to a neighboring city to look for a job in the formal labor market. They also migrate back to the rural areas when formal sector jobs are scarce. The PNAD includes these workers and so is preferable.
the spike at the minimum wage, minimum wage workers are those working at the minimum monthly salary if they report working 40, 44, or 48 hours, or if their wage is within 1 log point of one of the wages implied by these hours. The minimum wage and the IPCA price index come from the Brazilian Central Bank’s online database.

Full benefits are calculated based on the legal payroll taxes in Brazil (see section 2). The cost multiplier to the firm of these benefits is 0.7 log points, thus \( \ln B = 0.7 \).\(^{10}\) Wages are computed by dividing earnings for the month of September by average hours worked.

Given the variety of regulations, a worker may be illegally employed in a variety of ways. Here, a formal or legal worker is one paid at least the minimum wage, registered with the government, and for whom all payroll taxes are paid. Thus observed payment of social security taxes and work registration are proxies for payment of all payroll taxes and non-wage benefits. This assumes that social security taxes are the first benefits workers relinquish when abandoning non-wage benefits or that workers who are registered with the government receive the mandated benefits package, neither of which is likely to be a bad assumption.

Worker age is restricted to 30 to 45 years old. This narrow range of men was used so as to ensure that labor force participation is not a relevant concern. In this sample, labor force participation rates are about 90%. Years of schooling is aggregated for higher levels. Those with 9-11 years of schooling are assigned a schooling level of 10. Those with more than 11 are assigned a value of 14. Years are normalized with the first year, 1981, set to 0.

**Illegal Economy Observed**

The formal or legal sector is defined as those that are in compliance with all the labor laws, whether employed or self employed. The informal or illegal sector are workers in violation of one or more laws. Measurement of illegal activity is fraught with difficulty. Agents working illegally are often less forthcoming about their status. Labor laws in a country like Brazil can provide useful data on these workers for several reasons. First, labor laws inflict punishments on firms, not workers; so the worker has far less concern about how information from a survey might be used against them, because it is the employer who faces penalties.\(^{11}\) Second, although some attempts are made to enforce labor laws, surveys report widespread violations, thus enforcement is low enough to make reporting feasible even though the law may still have an effect on the economy. Third, the labor market is the subject of regular

\(^{10}\)Due to the 1988 constitution, payroll benefits actually differ before and after 1988. Thus up to 1988 \( \ln B = \ln(1.84) \), and afterwards \( \ln B = \ln(2.02) \). This regime shift may have other effects on the administration of benefits. To account for this, benefit preferences include an indicator variable that equals one for all years after the 1988 Constitutional change.

\(^{11}\)This is not to say that the worker might not have some incentive to lie if the worker believed that the information a) would be used against the employer and b) the worker would lose their job on account of this. I am unaware of any evidence that the ministry overseeing labor regulations attempts to coerce this type of information from the Census Bureau. This would involve a remarkable amount of coordination across government bureaucracies.
surveys across many workers, admitting a large sample of data with which to approach the question.

Table 2 summarizes market illegality across the North and South of Brazil in 2002. The North is a poor but well populated region where the median wage, R$0.70, is half the South's median wage, R$1.40. One measure of informality is the number of workers below the minimum wage. The federal minimum wage hits much higher up in the North's wage distribution. In the North, 41% are at or below the minimum wage while only 11% are at or below the minimum in the South.

Employees are legally required to pay social security taxes and be registered with a signed work contract. Payment of these taxes is mandatory but only one third of employees report paying in the North, while about two thirds pay in the South. Unregistered work exhibits a similar pattern. Self employment is also common, accounting for 36% of employment in the North and 26% in the South.

Figure 5 plots the wage distribution for men in 2002, separated by whether the workers are self employed and whether or not they pay social security taxes. As can be seen, self employed workers look quite similar to their employed counterparts, once one controls for payroll taxes. The major difference is the spike at the minimum wage clearly observed among employed, illegal workers. Thus for this group, spikes at the minimum wage are more clearly a phenomenon among informal workers than formal.

5 Maximum Likelihood Specification

Given the above model, estimation requires specifying the observed variables and unobserved distributions to allow the model to assign a probability to any observed outcome. One can then use maximum likelihood for the estimation.

Productivity and latent preferences for benefits are allowed to be correlated across workers. Write these terms as:

$$\theta_i^* = W_i \theta + \eta_i$$

$$\ln t_i = X_i \beta + \varepsilon_i$$

where $X_i$ and $W_i$ are vectors of productivity and preference shifters. $\eta_i$ is normally distributed with mean zero and a standard deviation $\sigma_\eta$. $\varepsilon_i$ is allowed to be a mixture of two normals with separate standard deviations and means, although, as a normalization, at least one of the means is required to be zero. The two error terms are bivariate (mixed)

---

12 Throughout the paper, "North" refers to all the states in the North and Northeast census regions. "South" refers to all the states in the South and Southeast census regions. The comparatively small Center-West region is excluded from the analysis. The data and sample restrictions are the same as described in Section 4.
normal, with correlation \( \rho_{u,u} \). Unique elements of \( W \), along with the distributional assumptions, serve to identify the unobserved correlation. Since I am not interested in identifying structural parameters in the benefits equation, I use this reduced form to recover what I need.

The latent benefits parameter \( \theta^*_i \) is censored at the top and bottom such that it falls within the range from 0 to \( \ln B \) but generating mass points at 0 and \( \ln B \) for those latent values that would rise above and below these points.

\[
\theta_i = \begin{cases} 
0 & \text{if } \theta^*_i < 0 \\
\ln B & \text{if } \theta^*_i > \ln B 
\end{cases}
\]

This allows mass points of agents to value non-wage benefits fully or not at all.

The evasion parameter \( \sigma^1 \) is estimated as a constant term restricted to lie between 0 and 1. The fixed cost of evasion, \( \sigma^0 \) is slightly trickier. The presence of a fixed cost for crossing to below the minimum wage creates a set of wages wherein no worker with full benefits may exist, because the fixed cost pushes them below. First, this means that maximum likelihood is trying to estimate boundaries, which is problematic. It also is obvious from a casual inspection of the data that no such hole exists among those taking full benefits. As such, one can reject out of hand a fixed cost for such workers.

Instead, I estimate both a probability of an agent facing a fixed cost and a constant, non-negative value for that cost. This probabilistic approach is not obviously rejected by visually inspecting the data and so may bear fruit, while also avoiding estimation problems.

To allow for self employment, add an equation for \( \gamma_i \), the preference for employment over self employment:

\[
\gamma_i = Z_i \gamma + \phi_i
\]

where the error term is trivariate normally distributed with the other two error terms and \( Z_i \) is a vector of covariates.

Let \( \pi^k_F \) be the formal sector price in market \( k \), where each agent is considered to be in one of \( K \) markets. Then one can recover formal sector prices, \( \pi^1_F, \ldots, \pi^K_F \) with dummy variables that divide the data into markets. In this paper, I equate markets with years, and so estimate market prices with year dummies.

The covariates in the productivity equation \( X_i \) include polynomials in age, education, and region. The benefit and self employment preference covariates, \( W_i \) and \( Z_i \), include \( X_i \) plus family demographics, and a year trend. Due to the large changes surrounding the 1988 Constitution, \( W_i \) also contains a dummy variable for post-1988.
Identification

Given the above model and a draw from the distributions of productivity, preferences and evasion costs, one can determine the probability that a worker adopts any observed work, wage, and benefits choice. Most of the model can be identified strictly off of the joint normality assumptions used, but because this is not very compelling on its own, a variety of robustness checks are included to determine the extent to which evasion cost estimates are driven by distributional assumptions.

The benefits equation is more difficult to pin down. First note that one can easily identify \( \frac{\theta}{\sigma_0} \) as in any standard probit model. But to identify the two parameters separately requires a shifter of known magnitude in the cutoff between full and partial benefits. Evasion costs create this shifter, causing agents above and below the minimum wage to have different cutoffs. Further identification comes from comparing wages of those who take full benefits and those who don’t. If covariates affect wages differently between the two regimes, this identifies the effect of different preferences for benefits.

Finally, the evasion parameters each have multiple sources of identification. \( \delta^1 \) affects the covariates of all illegal workers, it affects differential benefit take-up rates between those above and below the minimum wage, and it is the cause of a spike in workers at the minimum wage. \( \delta^0 \) is identified by wage drops among workers just below the minimum wage and by benefit take-up rates. As a side note, the self employment equation has a variance parameter which is identified solely in cases with positive evasion costs, as these create identifiable differences in wages between the employed and self-employed.

As this discussion should make clear, identifying the parameters does not require time-series variation. One route researchers take in dealing with short time series is to treat each state or region as a separate observation, creating a panel of state-years. While this can solve many problems, it has difficulty precisely identifying covariates that only vary at the national level, such as national (as opposed to local) labor laws, and is subject to concerns about measurement error in the price index.

Brazil’s inflation, which can run into thousands of percentage points a year, clearly causes concerns about accurate data. If measurement error from price indices is classical, it will attenuate employment effects, but will create a correlated bias when wages are regressed on the minimum wage. Since both regressor and regressand are affected by the same error, least squares regression will infer a positive relationship even if none exists. Thus measurement error would bias a model that relies upon time series variation towards finding that the minimum wage raises wages but has no employment effect.

The approach here allows identification by modeling how the minimum wage or benefit level affects an individual, utility-maximizing agent. Since the minimum wage will affect different agents very differently, this creates variation in the “treatment” which can be used for identification, even if the level of the minimum wage is held constant. Like the other approaches discussed, it depends on untestable assumptions; but when time-series are short
and suffer from intertemporal measurement error these assumptions may be a welcome alternative approach.

6 Maximum Likelihood Results

Table 3 reports estimates of the model for 100,000 men. The top panel of Table 3 reports on the evasion costs while the bottom three panels give the covariates from the productivity and preference equations along with their relevant distributional parameters. The coefficients in the latent index functions for benefits and self employment are normalized by the standard deviation of the equation error terms (thus making them comparable to probit estimates that use a similar normalization). One may be concerned that the reported standard errors may be relying upon the functional form assumptions used in the maximum likelihood routine. Bootstrapped standard errors were essentially the same as those reported here, suggesting that the analytic approximations are accurate.

The productivity distribution gives results that are to be expected; there are nonlinear productivity increases associated with age and education, and productivity is higher in the South. The mixture of normals shows obvious signs of non-normality with the second distribution being noticeably more dispersed, indicating kurtosis in the productivity distribution.

The benefits preference coefficients show reasonable values. The 1988 reform resulted in an increased valuation of benefits. Older and more educated agents value benefits more. Those in the South strongly prefer more benefits and the valuation of benefits has a secular decline over time. Family size and number of children both have small but negative effects on the benefits distribution, perhaps due to the safety net provided by the family and, in the case of children, old age care. Values for the benefits parameter lie between 0 and approximately 0.7. The mean value of theta is about 0.65 so almost half of agents value benefits at or above their cost.

Preferences for employment over self employment are not well explained by observables—the coefficients tend to be rather small with the exception of geography where southerners prefer employment over self employment.

The evasion cost parameters show that there is a small amount of distortion coming from legal enforcement. The marginal evasion tax $\delta^1$, which is a cost based on how far you are from legality, is .081. This means that agents pay evasion costs of about 8.1% of their distance from the legal requirements. Thus a worker below the minimum wage who had their productivity decline by 1% would be paid $\frac{1}{1-.081} = 1.09\%$ less.

The agent’s fixed cost of evasion has two parameters—first is the probability that the agent has a cost of evasion and the second is how large that cost is. The maximum likelihood estimate indicates that there is zero probability of a fixed evasion cost, leaving the second parameter unidentified. Using the likelihood function, I construct a confidence interval by fixing the evasion cost parameter $\delta^0$ to 0.01. Although not as general as some other testing
options (see Andrews (1994) for a discussion of inference at borders), it is both simple and revealing, as the assigned confidence interval is very narrow. The reported standard error is backed out from this confidence interval assuming normality.

Combined, these results suggest that evasion is actually not very costly in Brazil. The lack of a fixed cost of evasion suggests that the market is not segmented and that informal work gives returns much like those in the formal sector. Since these results are averages, it is still possible that some idiosyncratic individuals face high costs of evasion.

The estimates are consistent with a model wherein low-productivity workers are able to sort into employment with no or low costs of evasion. Agents generally do not fully value their benefits, but face small costs of evading the law, so market regulation has little effect.

**Model Fit**

To determine out-of-sample predictive power, draw a sample of workers not used for the estimation and compare their outcomes to those simulated from the estimated parameters. This section compares the wage distributions, benefit levels across wage levels, and a variety of aggregate and individual measures of fit.

Figure 6 simulates the wage distribution for 2002 across the 4 kinds of workers. For comparison, Figure 5 plotted the empirical wage distribution in 2002. The distributions share similar characteristics, so the model is capable of replicating the observed distribution, at least in its rough outline.

Figure 7 compares wages to the percentage receiving full benefits, in the simulation (the solid line), and empirically (the dashed line). The simulation stays fairly close to empirical reality, though it predicts too many full benefits holders in the far left tail of the wage distribution. It also shows a slightly shallower wage gradient for self-employed workers. Figure 8 plots self employment rates at different wages which the simulation appears able to match.

Table 4 compares the simulation with the empirical results on several aggregate and individual predictors. Summary statistics include percentages formal, self-employed, above or at the minimum wage, and several moments and quantiles of wages. In general, the simulated and empirical models are quite close. The simulation underestimates the number of informal and formal workers at the minimum wage, perhaps because some of this clumping is due to reporting error. Regardless, there is room for future improvement.

Note that the percentage at the minimum wage actually includes a small window around \( M \), accounting for the non-zero number of formal minimum wage workers in the simulation. Empirically, most workers at the minimum wage are informal, which fits the model. The simulation does quite well in matching higher moments of the wage distribution as well as the wage quantiles, where it is not far off for any value.

The last two rows of Table 4 move from the aggregate to the individual. They record how often the model's outcome for receiving benefits or being self employed accords with
what the agent actually chose. 62% of work decisions are simulated correctly and 61% of benefit decisions. Considering that one can be right at guessing a coin flip half of the time, 50% is a lower bound for confidence in the predictive power of any such model. Beyond that, the expected quality of the prediction depends on the variance of unobservables.

The next column over from these estimates compares two simulations to determine how often results coincide across simulations. These numbers are almost identical to the empirical results, with benefits decisions correct 62% of the time and self employment correct 61%. Thus the model predicts the empirical results as well as it predicts a sample of data drawn from the model. Although this can be true even if the model is false, it is reassuring that the data and the simulation conform closely in predicting individual, and not just aggregate, behavior.

In summary, the model fits the data well. Although there are points of departure, the simulated distributions appear to be much like the empirical distributions. The benefits and self employment distributions across wages match fairly well and one can predict empirical individual outcomes with the same accuracy as one can predict simulated outcomes. All of this suggests that the proposed model is a reasonable approximation to the observed data.

Robustness

There are a variety of concerns one might have, apart from model fit, about the identification strategy. First, the exclusion restrictions used to identify the productivity and preference equations are not very compelling, and so the model relies upon the joint normality of the data to identify, in particular, the correlation estimates. These estimates are not of much interest in themselves, but to the extent that they bias the evasion cost estimates the conclusions are suspect.

To examine this problem, I reestimate the model using a variety of different distributional and functional form assumptions. Regardless of the changing assumptions, the evasion costs are estimated as close to zero. Thus it appears that the evasion costs result is not driven by basic errors in the functional forms. For these additional specifications, I report solely on the evasion cost parameters.

Figure 9 plots what happens to the evasion cost parameter $\delta^1$ as the correlation between productivity and preference for benefits is moved up and down from its estimated value of 0.31 (highlighted with a line). This correlation has little effect on the parameter of interest, nor does it ever cause the fixed cost parameter to move off of its point estimate of 0.

Figures 10 and 11 plot changes in the two other correlation parameters—between the self employment preference and either productivity or preference for benefits. The productivity correlation has no effect on evasion cost estimates. The benefits correlation has a small but noticeable effect, but does not change the basic inference, nor do the different assumptions ever yield a positive fixed cost estimate. Thus the evasion cost parameters appear to be
reasonably robust to misspecifying the unobserved correlation across equations.

A more drastic approach is to drop all self employed workers and estimate the model without a self employment option. This should reveal to what extent estimates are dependent on correctly specifying the self employed choice. It appears that this has little effect on the estimates, which are reported in row 1 of Table 5.

Another concern is that agents may be reporting themselves at the minimum wage as a useful shorthand, even though their wage is substantially different. This is a particular example of the broader concern that measurement error may be driving, or obfuscating, the results. It can be quite difficult to model measurement error that allows for clumping at the minimum wage as well as sectoral choice of employment. As a check of the seriousness of the problem, I introduce measurement error to the data and re-estimate it, to see whether measurement error is a problem on the margin.

To do this, I assume that agents report correct wages three fourths of the time, and the rest of the time they either have classical error with a standard error of 20 log points, which is reported in row 2, or they round to the nearest multiple or half step of the minimum wage, representing something like the "lighthouse effect" documented previously (for example, Neri et. al (2001)). This clumping simulation is reported in row 3. Classical measurement error lowers the estimate of $\delta^1$ slightly but has no other effect. Clumping more observations at $M$ predictably causes a somewhat higher estimate of $\delta^1$, of 0.126, but once again the fixed cost estimate is zero. Thus measurement error does not appear to be driving our conclusions. Further, excessive and error-ridden clumping at the minimum wage appears to bias our evasion cost estimate up, making them still interpretable as upper bounds.

Allowing for the productivity distribution to be non-normal certainly relieves some concern about overly rigid functional forms. In row 4 I re-impose strict normality on the productivity distribution to see how sensitive results are to the distributional assumption. Once again, the evasion cost parameters are largely unaffected, suggesting that this restriction is unimportant in determining the results.

The base model uses a quadratic in education to capture education effects. Since I am interested in behavior in the bottom tail, small errors in this quadratic assumption could potentially stack up to give incorrect inference. I check for this by putting in dummy variables for each potential schooling level. As seen in row 5, this does not affect the results.

The current estimates rely upon knowing ex ante the size of legally required non-wage benefits; but this estimate may well be off. To check for sensitivity to error in $B$, I re-estimate the model twice, moving $B$ up or down by 30 log points from its actual value. These results are in rows 6 and 7. This does not affect the fixed costs, and marginal evasion costs are only barely affected.

As discussed in the data section, I round wages within 1% of $M$ to equal $M$. Since the clump at the minimum wage is source of identification for evasion costs, this decision to round may be very important. Some rounding is needed just to deal with small errors.
in wages. I consider more extravagant rounding by rounding everything within 3% of the relevant $M$. Indeed, this does cause marginal costs to rise to 0.16, the highest value reported in the table.

One point of concern is that the utility function specified in (7), although analytically tractable, is quite narrow; it only has one parameter. Since the lump at the minimum wage is determined by the kink point created by evasion costs, the evasion cost parameters may be sensitive to misspecifying the utility function— and in particular by not allowing the second derivative to be freely determined. I relax that assumption here by specifying a new utility function:

$$\ln \psi (w, \tau) = (1 + \theta_1) \cdot \ln \tau - \frac{\theta_2}{2} (\ln \tau)^2 + \ln w$$

(15)

where now the diminishing returns parameter $\theta_2$ is estimated instead of assumed to equal one. This function gives a first order condition of $\tau^* = \frac{\partial \psi}{\partial \tau} - \frac{\delta^1}{\theta_2}$. Furthermore, the size of the lump at the minimum wage is now a function of $\frac{\partial \psi}{\partial \tau}$. Obviously, this has a profound potential impact on the evasion cost estimates. As reported in row 9, while fixed costs remain at 0, marginal evasion costs are unidentified, although the point estimate is close to 0. As a side note, the estimate of $\theta_2$ is also very imprecise with a point estimate very close to zero, implying that preferences for non-wage benefits are almost linear. This suggests a world in which agents value benefits either fully (because they prefer being formal) or not at all. Of course, since the estimates are so imprecise, it is very difficult to say what the right second derivative term might be. To the extent that preferences are essentially linear through this range, one gets substantial lumping at the minimum wage even with very modest marginal evasion costs.

More generally, this simulation highlights that the lump of informal workers at the minimum wage, so often used as a sign of the minimum wage having a dramatic effect on wages, is uninterpretable unless one has some knowledge of the extent to which a preference for non-wage benefits has diminishing returns. Even very small evasion costs, and thus fairly irrelevant labor market regulations, can generate observable spikes at the minimum wage if the underlying tradeoff between legal benefits and wages is linear or close to being linear.

As noted above, one of the important mechanisms for identifying evasion costs is that one should observe higher wage gradients on covariates for workers below the minimum wage, as opposed to above it. This is because increasing productivity not only raises the expected piece rate, it also lowers evasion costs at a rate $\delta^1$. Given this fact, identification could be compromised if the coefficients varied across the wage distribution for other, unaccounted for, reasons.

As a robustness check, allow the productivity terms to vary with the error term. Recall the productivity equation, $\ln t_i = X_i \beta + \varepsilon_i$, and now specify

$$\beta = \beta_0 + \beta_1 \varepsilon_i$$

(16)
for all terms except the constant. In this way, $\beta$ can vary with the error term, albeit not as freely as would be ideal. Similarly, allow for the price vector to be a function of the error term:

$$\pi_{it} = \pi_{0t} + \pi_{1t} \varepsilon_i$$

(17)

so that I no longer impose the same market fluctuations in price across all ability levels.

Row 10 reports the evasion cost estimates. Although such interactions do exist, allowing for them does not substantially affect the evasion cost parameter estimates. Row 11 further allows for the existence of a two parameter utility function, $\{\theta_1, \theta_2\}$, discussed above in conjunction with the flexible parametrization of $\beta$. Once again, the marginal evasion cost parameter is unidentified but the fixed cost remains at 0—suggesting that the results about fixed costs are somewhat robust to either form of misspecification.

Despite a wide variety of checks, the same basic story emerges. Cross sectional wage distributions give no evidence of substantial costs to evading labor laws. As could be expected, this result is tempered by the fact that unobserved evasion costs are difficult to separate out from unobserved preferences. In this case, the data and model strongly reject pervasive fixed costs to working illegally, but are less conclusive on marginal costs owing to the difficult identification problems. Of course, marginal evasion costs alone, without some fixed cost component, are unlikely to generate large market distortions. For example, if marginal evasion costs were estimated at .2, agents paid 50 log points below the minimum wage pay only a 10% evasion tax. While this is not immaterial, it is still fairly small.

7 Quantile Regression

In addition to the above maximum likelihood estimation, a simplified version of the model yields empirically estimable evasion coefficients from the productivity equation. This section uses quantile regression to recover evasion costs under very different identifying assumptions.

Specification

If a worker working for less than the minimum wage increases his productivity, this should increase his wage both from the direct increase in productivity and, indirectly, from making him less illegal. Ignore the endogenous choice of nonwage benefits and self employment and the wage function becomes:

$$\ln(w) = \pi_t + X \beta + \varepsilon \frac{\pi_t + X \beta + \varepsilon - \delta_1 - \delta_2 \ln M}{(1-\delta)} \quad w \geq M$$

$$= \pi_t + X \beta + \varepsilon \quad w < M$$

(18)

$\beta$ tracks changes in productivity, while, for those below the minimum wage, $\frac{\beta}{1-\delta}$ tracks the joint effect of productivity and lowering marginal evasion costs. Were it not for the severe censoring problem at the minimum wage, one could recover evasion costs by OLS
regression on these two equations.\textsuperscript{13}

Consider a linear model with only one varying covariate, education. Call the estimated coefficient on education from below the minimum wage $b_L$ and from above $b_H$. Then $\delta^1 = 1 - \frac{b_H}{b_L}$. For example, if the return to a year of education for the high skilled were $.08$, but $.1$ for the low skilled, $\delta^1 = 1 - \frac{.08}{.1} = .2$. Note that, although this result falls out particularly nicely from this model, it is intuitively clear that, to the extent workers pay more evasion costs for being more illegal, the returns to productivity should be higher below the minimum wage compared to above. This procedure captures that logic.

Although OLS regression would be problematic, quantile regression is well suited to this sort of analysis. Let $u$ be the error distribution quantile and estimate the following at every integer value of $u$ from one to ninety-nine.

$$w_i = \pi_i(u) + S_i \ast b(u) + \text{error}_{i\delta}$$

where $S_i$ is years of schooling. To account for direct effects of the minimum wage and other price movements, the regression includes a full set of year fixed effects in each regression, but no other covariates. Although quantiles do not exactly conform to particular points in the wage distribution, there is sufficient variation in the error term that very low values for $u$ are completely drawn from agents below the minimum wage and very high values are drawn from agents above the minimum wage. Lastly, pick candidates for $b_L$ and $b_H$ from quantiles drawn from entirely above or below the minimum wage and compute $\delta^1$.

I concentrate exclusively on low skilled workers by restricting the sample to those with 4 or fewer years of schooling. The estimation separately considers two age groups: ages 20 to 25 and 30 to 35. The younger sample will have more workers below the minimum wage, but may have more issues with unemployment as only about 80% of men work. The older sample has higher employment which minimizes employment selection concerns, but fewer workers below the minimum wage. Also, it is a subset of the data used in the prior maximum likelihood estimation. Unreported maximum likelihood estimation confirms that estimated evasion costs for these workers are comparable to that found in the entire sample (about 8-9% marginal cost with no fixed costs).

\textbf{Results}

Figures 12 and 13 graph the quantile regression coefficients on education with the x-axis being the average distance from the minimum wage for workers at that percentile of the error distribution. The figures also graph the 95\% pointwise confidence intervals. Thus the

\textsuperscript{13}Unfortunately the minimum wage's direct effect in the numerator of the second equation will be absorbed in any regression that includes the set of year dummies used to estimate $\pi_i$ and the fixed evasion cost, $\delta^0$, will be inseparable from the regression constant term. If one had confidence in the policy exogeneity of the minimum wage as well as a way to statistically deal with the large group of agents moving directly in tandem with the minimum wage owing to the clumping, one could potentially read the marginal evasion cost directly from the coefficient on the minimum wage variable.
3% quantile for workers age 20 to 25 has a return to schooling slightly over 0.12; the return is flat until one approaches the minimum wage, at which point it falls steeply to under 0.1 before rising again in the top half of the wage distribution.

The lumping of workers at the minimum wage clearly has a dramatic effect on quantiles close to the minimum.\textsuperscript{14} Fortunately the estimator does not rely on those workers, but focuses on the top and bottom of the distribution to provide candidates for $b_L$ and $b_H$. I consider the 3rd, and 5th percentiles at the bottom and the 80th, and 90th percentiles at the top. These percentile choices are to make sure that the observations local to the quantile are drawn either entirely from above or below the minimum wage.\textsuperscript{15} Table 6 reports the implied estimates of $\beta$, where standard errors are bootstrapped.

Based on the point estimates, one would infer a small, positive marginal evasion cost for the older workers somewhere between zero and six percent, putting them close to the maximum likelihood results. For younger workers, there is no evidence of any evasion cost. In fact, the point estimates based on the 80th percentile are negative, though not significantly so. With one exception, none of the results are statistically different from zero at conventional significance levels. Thus they confirm the conclusion that workers pay little if any wage penalty for evading labor market regulations.

The limitations of this method are obvious. First, no accounting is made for endogenous choice of self employment or nonwage benefits. Those issues are better dealt with by estimating the full maximum likelihood model. Second, the estimation relies upon assuming a stable latent productivity gradient, $\beta$, for education. Of course, it could be that the productivity gradient ($\beta$) varies with the quantile, in which case we would not be uncovering evasion costs at all. On the positive side, the return does appear reasonably stable everywhere outside of the minimum wage zone, and the upper half of the distribution is particularly stable, thus the assumption of a constant underlying education coefficient may not be baseless. I also deal with this concern in the robustness checks to the maximum likelihood estimation by allowing all productivity covariates to have coefficients that vary with the error term. Adding that flexibility did not alter the substantive conclusions.

In summary, the quantile regression approach provides a check on the maximum likelihood estimates using alternative identifying assumptions. The two methods appear to confirm one another, which strengthens one's faith in the conclusions.

\textsuperscript{14} As noted in the Data section, the minimum wage is designated as a salary at a fixed number of hours. A common minor way to evade the law is to have workers work 44 or 48 hours instead of 40. Thus lumping in the data shows up about 1 logs below the statutory minimum wage.

\textsuperscript{15} For younger workers, the acceptable range would be about the bottom tenth of the distribution and the top quarter. For older workers, only the bottom 5 percentiles are clearly away from the minimum wage, while the top third is similarly insulated. The values used are drawn from the intersection of the two ranges.
8 Conclusions

Mandated non-wage benefits and the minimum wage law have little effect on the actual labor market. Evasion costs, while present, are sufficiently low that they provide little impediment to labor market operations. This may be due to the fact that workers preferring illegality can sort into firms that can easily evade the law, avoiding the firms that cannot.

This result holds up through a variety of robustness checks and two different estimation methods using cross-sectional, as opposed to time-series, variation to identify labor market distortions. Although time-series variation can be an excellent tool in many cases, when the series are short, the variation is not very informative. The approaches used here represent a useful alternative in cases where cross-sectional data is comparatively plentiful.

To the extent that these conclusions hold internationally, it appears that some labor market regulations may be largely irrelevant to the workings of the market-place. Enforcement may vary dramatically by country; drawing conclusions will require studying more places. Fortunately, the methodology used here is accessible to the very kinds of survey data most widely available.

The model presented is limited in a variety of ways that may be fruitful avenues of future work. To limit the complexity, the model was estimated on workers with high job market attachment and unemployment was ignored. The lack of employment effects simplifies the estimation, but this is unlikely to hold for female workers or children. Of course, if one thinks that evasion is likely to be even easier for these groups, the results here suggest that they too are largely unaffected by regulation.

The model, as presented, estimates static evasion costs, but both enforcement and the minimum wage level are policy tools which the federal government may use differently over time. Since the estimation does not rely on time series variation in the minimum wage, one needn't be concerned about the correlation between macroeconomic outcomes and the minimum wage, but one may wish to estimate how evasion costs vary over time and what determines the level of enforcement.

The maximum likelihood estimates rely upon a specific model of how the minimum wage affects the economy coupled with assumptions about unobserved distributions of structural error terms. The distributional assumptions could be relaxed or modified to provide a better fit to the data, although robustness checks and alternative estimates from quantile regression suggest that this is not likely to change the results.

This research does not attempt to model the dynamic process that causes workers to move between jobs in response to changes in the minimum wage. It may be that there are important lags in response to minimum wage changes. This model treats those lags as being shorter than one year (usually the time between the setting of the wage in May and the survey in September). Developing a model with some element of adjustment costs could prove fruitful.

Although work remains to be done, the model presented here provides useful ways to
estimate the effect of labor laws in the presence of incomplete enforcement.

References


Table 1: Labor Costs in Brazil

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Cost as a fraction of the wage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual Bonus</td>
<td>0.08</td>
</tr>
<tr>
<td>Personal Unemployment Fund (FGTS)</td>
<td>0.08</td>
</tr>
<tr>
<td>Other Direct Payments and Subsidies</td>
<td>0.22</td>
</tr>
<tr>
<td>Paid Leisure</td>
<td>0.12</td>
</tr>
<tr>
<td>Social Security, accident insurance, worker training</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Payments post-1988
Source: Table 7.1 in Amadeo and Camargo (1997)

Table 2: Brazilian Men in 2002, Age 30-45

<table>
<thead>
<tr>
<th>Statistic</th>
<th>North</th>
<th>South</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (millions)</td>
<td>5.6</td>
<td>11.5</td>
</tr>
<tr>
<td>Median Years of Schooling</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Median Age</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Median Wage (1994 Reais)</td>
<td>0.70</td>
<td>1.40</td>
</tr>
<tr>
<td>Fraction Working</td>
<td>0.86</td>
<td>0.88</td>
</tr>
<tr>
<td>Fraction Working at Minimum Wage</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td>Fraction Working Below Minimum Wage</td>
<td>0.30</td>
<td>0.07</td>
</tr>
<tr>
<td>Fraction Paying Social Security</td>
<td>0.38</td>
<td>0.65</td>
</tr>
<tr>
<td>Fraction Formally Registered</td>
<td>0.50</td>
<td>0.76</td>
</tr>
<tr>
<td>Fraction Self Employed</td>
<td>0.36</td>
<td>0.26</td>
</tr>
<tr>
<td>Evasion Costs</td>
<td>Model Fit</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>( P(\delta^0 &gt; 0) )</td>
<td>Log Likelihood: -219134</td>
<td></td>
</tr>
<tr>
<td>( \delta^1 )</td>
<td>Observations: 100000</td>
<td></td>
</tr>
<tr>
<td>0.0000</td>
<td>0.0008 [1]</td>
<td></td>
</tr>
<tr>
<td>0.0810</td>
<td>0.0025</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Productivity Equation</th>
<th>Benefits Equation</th>
<th>Self-Employment Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td>0.0772</td>
<td>0.0495</td>
<td>0.0170</td>
</tr>
<tr>
<td>0.0094</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Squared</td>
<td>Age Squared</td>
<td>Age Squared</td>
</tr>
<tr>
<td>-0.0008</td>
<td>-0.0005</td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Education</td>
<td>Education</td>
</tr>
<tr>
<td>0.1293</td>
<td>0.1661</td>
<td>0.0035</td>
</tr>
<tr>
<td>0.0020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education Squared</td>
<td>Education Squared</td>
<td>Education Squared</td>
</tr>
<tr>
<td>0.0025</td>
<td>-0.0041</td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>South</td>
<td>South</td>
</tr>
<tr>
<td>0.2975</td>
<td>0.4204</td>
<td>0.1785</td>
</tr>
<tr>
<td>0.0051</td>
<td>0.0090</td>
<td></td>
</tr>
<tr>
<td>Constant 1</td>
<td>Family Size</td>
<td>Family Size</td>
</tr>
<tr>
<td>-2.4227</td>
<td>-0.0155</td>
<td>-0.0061</td>
</tr>
<tr>
<td>0.1731</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td>Constant 2</td>
<td>Num. of Children</td>
<td>Num. of Children</td>
</tr>
<tr>
<td>-2.2635</td>
<td>-0.0152</td>
<td>-0.0073</td>
</tr>
<tr>
<td>0.1736</td>
<td>0.0035</td>
<td></td>
</tr>
<tr>
<td>( \sigma_1 )</td>
<td>Year</td>
<td>Year</td>
</tr>
<tr>
<td>0.6546</td>
<td>-0.0336</td>
<td>-0.0061</td>
</tr>
<tr>
<td>0.0055</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td>( \sigma_2 )</td>
<td>Post 89 Regime</td>
<td>Constant</td>
</tr>
<tr>
<td>1.0703</td>
<td>0.4775</td>
<td>1.1258</td>
</tr>
<tr>
<td>0.0159</td>
<td>0.0249</td>
<td></td>
</tr>
<tr>
<td>P(1st Normal)</td>
<td>Constant</td>
<td>Std. Dev. of Error</td>
</tr>
<tr>
<td>0.7777</td>
<td>1.4046</td>
<td>0.2636</td>
</tr>
<tr>
<td>0.0167</td>
<td>0.3360</td>
<td></td>
</tr>
<tr>
<td>Std. Dev. of Error</td>
<td>Corr(Prod, Self)</td>
<td>Corr(Benefits, Self)</td>
</tr>
<tr>
<td>0.4987</td>
<td>0.1040</td>
<td>0.1040</td>
</tr>
<tr>
<td>0.0079</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corr(Prod, Benefits)</td>
<td>Corr(Benefits, Self)</td>
<td></td>
</tr>
<tr>
<td>0.3078</td>
<td>0.6791</td>
<td></td>
</tr>
<tr>
<td>0.0059</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[1] As a probability, the value is constrained to be between 0 and 1, thus the standard errors are only relevant moving right from the point estimate. See the text for more details. Asterisks for significance have been repressed as the results are almost universally significantly different from zero.
Table 4: Empirical vs. Simulated Outcomes

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Empirical</th>
<th>Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction Full Benefits</td>
<td>0.627</td>
<td>0.624</td>
</tr>
<tr>
<td>Fraction Self Employed</td>
<td>0.286</td>
<td>0.290</td>
</tr>
<tr>
<td>Fraction Wage less than $M$</td>
<td>0.128</td>
<td>0.145</td>
</tr>
<tr>
<td>Fraction Wage = $M$, Partial Benefits</td>
<td>0.048</td>
<td>0.025</td>
</tr>
<tr>
<td>Fraction Wage = $M$, Full Benefits</td>
<td>0.029</td>
<td>0.003</td>
</tr>
<tr>
<td>Average Log Wage</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>Log Wage Standard Deviation</td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td>Log Wage Skewness</td>
<td>0.29</td>
<td>0.26</td>
</tr>
<tr>
<td>Log Wage Kurtosis</td>
<td>3.20</td>
<td>3.24</td>
</tr>
<tr>
<td>10th Percentile of Log Wages</td>
<td>-0.96</td>
<td>-1.02</td>
</tr>
<tr>
<td>50th Percentile of Log Wages</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>90th Percentile of Log Wages</td>
<td>1.62</td>
<td>1.59</td>
</tr>
<tr>
<td>Fraction Full Benefits Predicted Correctly</td>
<td>0.620</td>
<td>0.618</td>
</tr>
<tr>
<td>Fraction Self Employment Predicted Correctly</td>
<td>0.607</td>
<td>0.605</td>
</tr>
</tbody>
</table>

As in the estimation, both the simulated and empirical values for Fraction Wage = $M$ include all those within 1 log point of the minimum wage; see Section 4.
<table>
<thead>
<tr>
<th>Estimation Method</th>
<th>$P(\delta^0 &gt; 0)$</th>
<th>$\delta^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Remove Self Employed</td>
<td>0</td>
<td>0.098**</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>2. Adding Classical Measurement Error</td>
<td>0</td>
<td>0.076**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>3. Adding Clumping Measurement Error</td>
<td>0</td>
<td>0.126**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>4. Single Normal Distribution</td>
<td>0</td>
<td>0.088**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>5. Education Dummy Variables</td>
<td>0</td>
<td>0.100**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>6. Change $\ln B + .3$</td>
<td>0</td>
<td>0.096**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>7. Change $\ln B - .3$</td>
<td>0</td>
<td>0.081**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>8. Round to $M$ all wages within 3 log points</td>
<td>0</td>
<td>0.160**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.004</td>
</tr>
<tr>
<td>9. 2-Parameter Benefits Preference</td>
<td>0</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>[1]</td>
</tr>
<tr>
<td>10. Interact Productivity Error and Coefficients</td>
<td>0</td>
<td>0.073**</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>11. Interact Productivity Error and Coefficients + 2-Parameter Benefits Preference</td>
<td>0</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>[1]</td>
</tr>
</tbody>
</table>

** = significant at the 1% level. Row labels are explained in the Robustness section of the text. [1] Although the parameter has a point estimate, it is so poorly identified that no value between 0 and 1 can be rejected.
<table>
<thead>
<tr>
<th>$b_L$</th>
<th>$b_H = 80$th quantile</th>
<th>90th quantile</th>
<th>$80$th quantile</th>
<th>90th quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd Quantile</td>
<td>-0.036</td>
<td>0.027</td>
<td>0.057*</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.046)</td>
<td>(0.029)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>5th Quantile</td>
<td>-0.047</td>
<td>0.017</td>
<td>0.019</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.040)</td>
<td>(0.026)</td>
<td>(0.032)</td>
</tr>
</tbody>
</table>

* = significant at the 5% level.
Figure 1: Minimum Wage in Brazil, 1981-2002
Figure 2: Example Trade-off Between Wages and Benefits
Figure 3: Optimal Benefits in $\theta$ and $t$ Space
Figure 4: Simulated Wage Distributions With Positive Evasion Costs

**Fixed Cost of Evasion**

- Partial/No Benefits
- Full Benefits

**Marginal Cost of Evasion**

- Partial/No Benefits
- Full Benefits

Graphs by Benefits
Figure 5: Log Wage Distributions by Employment Type, 2002

- Employed—Evading Payroll Tax
- Employed—Paying Payroll Tax
- Self-Emp.—Evading Payroll Tax
- Self-Emp.—Paying Payroll Tax
Figure 6: Simulated Log Wages

- Employed—Evading Payroll Tax
- Employed—Paying Payroll Tax
- Self-Emp.—Evading Payroll Tax
- Self-Emp.—Paying Payroll Tax
Figure 7: Empirical and Simulated Full Benefit Rates

Employed

Self-Employed

Simulated

Empirical
Figure 8: Empirical and Simulated Self Employment Rates
Figure 9: Robustness to Changing Benefits Equation Correlation

![Diagram showing the correlation between Marginal Evasion Cost and Benefits/Productivity Imposed Correlation. The graph displays a scatter plot with points indicating the relationship between the two variables.]
Figure 10: Robustness to Changing Self-Employment Equation Correlation—Productivity

Productivity/Self-Employment Imposed Correlation vs. Marginal Evasion Cost

-3 -2 -1 0 .1 .2 .3 .4

Marginal Evasion Cost

-1 -0.8 -0.6 -0.4

0.4 0.6 0.8 1
Figure 11: Robustness to Changing Self-Employment Equation Correlation-Benefits
Figure 12: Quantile Regression Education Coefficients, 20-25
Figure 13: Quantile Regression Education Coefficients. 30-35

\[ \frac{\ln(w)}{\ln(M)} \]

Education coefficient

95% Conf. Interval