Estimating the Effects of Minimum Wage in a Developing Country: A Density Discontinuity Design Approach

Hugo Jales*

October 31, 2013

Abstract

This paper proposes a new framework to empirically assess the effects of Minimum Wage in a developing country. This approach allows to jointly estimate the effects of minimum wage on unemployment, average wages, sector mobility, wage inequality, the size of the informal sector and on labor tax revenues. The paper shows that under reasonable assumptions the parameters that governs how minimum wage affects the economy and the **joint** distribution of latent sector and wages can be identified using only cross-sectional data on sector and wages. The identification strategy builds on the "Density Discontinuity Design" approach by Doyle (2008), but nests it with a parametric specification for the conditional distribution of sector given the wages. The idea that drives the identification is that the discontinuity of the wage distribution around the minimum wage identify the size of the non-compliance with the law, whereas the shape of the relationship between **latent** sector and wages can be recovered using the information of the conditional probability of sector given the wages for values above the minimum. I apply the method in the "PNAD", a nationwide representative Brazilian cross-sectional dataset from years 2001 to 2009. I show on the application that the assumptions used are not violated in the context of the Brazilian labor market. The results show that the probability of migration between sectors is very small, around 10%, but the relative size of the informal sector in the economy is still increased due to high unemployment effects - of around 60% - on the formal sector of the economy. In addition, minimum wage legislation strongly affects wage inequality, reducing up to 20% the standard deviation of log-wages, and reduces revenues from labor taxes up to 15% .

Keywords: Minimum Wage, Informality, Unemployment, Density Discontinuity Design, Wage Inequality

JEL Codes: J60 , J31 , J30

*PhD. Candidate at University of British Columbia. Email: hugo.jales@alumni.ubc.ca. I am grateful to Thomas Lemieux for the advise during this project. I am also grateful to the participants at the CEA conference and the participants at EESP-FGV, UFMG, UBC and SFU academic seminars for helpful suggestions that improved the quality of this paper. All remaining errors are my own.

1 Introduction

Despite its widespread use, there is still controversy regarding the economic effects of minimum wages. In a simple one-sector competitive markets model economic theory predicts that there will be some unemployment effects as long as the minimum wage is higher than the market clearing wage. If there is some market power from the employer, then the introduction of minimum wage can lead to both employment and wage increases. In an economy with a large informal sector, where some employers do not comply with the minimum wage legislation, minimum wage might not generate unemployment effects even in the absence of market power from the employer. This will hold as long as the workers can freely migrate from one sector to the other and the informal sector is large enough to accommodate this movement.

Since the theory can easily accommodate such opposite predictions about its impact, the task of understanding the effects of minimum wage becomes mostly empirical.

This paper develops a two-sector model to assess the impacts of minimum wage on (a) unemployment, (b) average wages, (c) wage inequality, (d) sector mobility (e) size of the informal sector and (e) labor tax revenues. It shows the conditions for identifying model parameters and latent joint distribution of sector and wages distribution, that is the distribution that would prevail in the absence of the policy. The identification strategy relies on the discontinuity of wage density at the minimum wage and the differences on the response to the minimum between formal and informal sectors.

The contributions of this paper to the literature are the following: (i) It documents key empirical facts about the relationship between formal and informal wage densities that were overlooked in previous research, namely the similarity between them conditioning on values above the minimum wage.(ii) It provides a novel identification strategy that combines a non-parametric Density Discontinuity Design with a parametric model for the conditional distribution of sector given the wage. In particular, it shows that under reasonable conditions the parameters that describes the effects of minimum wage and the underlying latent joint distribution of sector and wages are identified using only cross-sectional data. (iv) It estimates a sector mobility parameter, that is the probability of a worker in the formal sector to move to the informal as a response to the minimum wage. (iii) It shows how to test some of the assumptions invoked to identify the parameters of the model and it shows that those assumptions are valid on the empirical application. (iv) It jointly estimates the effects of Minimum Wage in the distribution of sector and wages and further estimate its impact on labor tax revenues. To the best of our knowledge, this is the first paper that attempts to identify both the latent share of the formal sector and the effects of the minimum wage on labor taxes revenues.

The model is estimated using the years of 2001 to 2009 of the PNAD dataset which are repeated cross sections of an annual household survey representative of the Brazilian population. The main results are: The probability of mobility between sectors is very small, of around 10%. Despite this fact, the size of the informal sector in the economy is still inflated given the high unemployment effects on the formal sector of the economy. Unemployment effects of minimum wage are, as expected, highly correlated with the real value of the minimum wage. Also, minimum wage strongly affects average wages (promoting an increase of around 20%) , wage inequality (-16% effect on standard deviation of log wages and -1% on the Gini Index), and labor tax revenues (-10%).

2 Related Literature

It is usually hard to estimate policy effects when there is no policy variation. In a simple regression framework, absence of policy changes is equivalent to failure of the "Rank Condition", which is necessary to identify the parameters of interest. In a randomized controlled trial, the benchmark case for the policy evaluation problem, absence of policy variation is equivalent to a dataset consisting only of treated individuals, a dataset without a control group.

The task of estimating the effects of minimum wage on labor market outcomes is quite similar to these problems. This is the case because the institution was created several decades ago and in no dataset will we be able to observe individuals that were not subjected to the policy. Of course, it is possible to estimate the effects of **changes** in the level of minimum wage on labor market outcomes, given that it presents some variation. [1] Minimum wages in Brazil have been set at the federal level since the 1980s, which does not allow researchers to easily separate time effects from minimum wage effects. In other words, although there is some variation in minimum wage that can be explored to identify its effects, it is not of a great quality, like the one observed in North America, where minimum wage varies across time and states/provinces. This type of variation allows for more flexible econometric techniques, such as differences-in-differences, which are robust to a broader range of sources of unobserved heterogeneity. Another feature of minimum wage changes in Brazil is that they were recently (since 2005) linked to inflation and GDP growth, which poses more challenges to the use of time series variation to estimate its effects. In this scenario, it is even harder to disentangle the effects of minimum wages from other sources of changes in the wage distribution that are only due to increases in economic activity.

Luckily, the economic theory can sometimes help the identification. By imposing natural conditions that arise from microeconomic theory of markets in the presence of minimum wage regulation, one can come up with a framework that allows the identification of minimum wage effects based on a single cross-section, that is, using data where all the individuals are faced with the same level of the policy. The approach taken here is an extension of that developed by Doyle (2007), which follows the influential work by Meyer and Wise (1983). This paper extends their model to a two-sector model with sector mobility. The extension allows for estimation of the effects of minimum wages on size of the informal sector and significantly clarifies the conditions necessary for the density below the minimum wage to be informative of latent wage distribution.

3 Model

In an early attempt to estimate the effects of the minimum wage, Meyer and Wise (1983) explored the distortion introduced in the wage density. First, a parametric model for the latent wage density is specified. Then, the parameters of this model are estimated taking into account the fact that the

[1] This approach has been taken by several researchers. The most recent example of this strategy for the Brazilian economy is Lemos (2009).

observed density is a truncated version of the latent density as a result of the minimum wage. Using these parameters, by comparing the latent and observed wage distribution the impacts of minimum wage on unemployment and average wages are estimated. Their results imply large unemployment effects from minimum wage legislation among young workers, in the magnitude of 30% to 50% (compared to the scenario under no minimum wage).

Later on, several papers tried to estimate the effects of minimum wage using state, time and even state borders over time variation (see, for example, Card and Krueger (1994). This approach, more in line with the usual tools of policy evaluation, led researchers to conclude that the sizable unemployment effects from minimum wage were due to restrictive parametric assumptions that were assumed by Meyer and Wise (1983) (Dickens, Machin and Manning, 1994). To evaluate this claim, Doyle (2007) developed a non-parametric version of Meyer and Wise's technique that also explores the discontinuity on the density generated by the minimum wage but is based only on a continuity assumption of the latent wage distribution. His results showed that sizable unemployment effects are estimated even without imposing parametric assumptions on the wage density.

This technique is especially relevant for Brazil, given that, as pointed out before, the country lacks some important regional variation in the minimum wage that is necessary to use popular policy evaluation tools, such as differences-in-differences. However, given that informality (and non-compliance) is sizable in the country, it is important to be able to accommodate movements between sectors in the model to get a good description of the Brazilian environment.

In the model, a worker i is characterized by a pair of wage (w_i) and a sector (s_i), which I will denote by one if it is the formal sector and zero otherwise. Compliance with the minimum wage legislation is perfect in the formal sector, but not in the informal sector. In addition, for each worker define a pair (w_i^*, s_i^*) denoting the counterfactual - or latent - wage and sector under the absence of the minimum wage [2]. Finally, define F the distribution of w^* and $B(\rho(w^*))$ the (Bernoulli) distribution of s^*. Define the proportion of workers in the formal sector by in the presence of minimum wage by s. Define also H the distribution of observed wages. I will assume that the **latent** wage and sector distribution have the following characteristics:

Assumption 1. *Latent wage and sector distributions:*

$$
\begin{gathered}
w^* \sim F \\
s^*(w^*) \sim B(\rho(w^*)) \\
\rho(w^*) \equiv Pr(s^* = 1 | w^*) = m(w^*; \beta)
\end{gathered}
$$

Where m: $R^{k+1} \to [0,1]$ is a function known up to the parameters β. A couple of things should be noted here. First of all, I do not restrict the distribution of latent wages, F, to belong to any parametric family. In this sense, the approach here is completely flexible. Second, the assumption above means state that the joint distribution of **latent** wage and sectors can be described by a parametric model through the function m() that governs the conditional distribution of sector given the wage.

[2] Note that this is a non standard policy evaluation problem where all individuals are treated. This forces the use of a model to identify the effects of the policy since the common support assumption fails to hold for everyone in the data. Moreover, the distinction between Average Treatment Effects and Treatment Effects on the Treated becomes irrelevant.

Now, given the usual results in microeconomic theory, we know that the workers in sectors operating in competitive markets whose wages would be below the minimum might become unemployed following introduction of the minimum wage. If there is some bargaining involved in the wage determination or if there is market power from the employers, some workers will bump at the minimum as a result of the legislation. Finally, since compliance with the minimum is imperfect in some markets, workers might migrate from the formal to the informal sector to avoid unemployment. In terms of the model, this leads to the following assumptions (Doyle, 2007):

Assumption 2. *Minimum wage effects:* [3]

For wages below the minimum wage:

If $s^* = 0$, then:

With probability P_1^i the wages continue to be observed. With the complementary probability P_2^i the worker earns the minimum wage [4].

If $s^* = 1$ then:

With probability P_1^f the wage continues to be observed, meaning that the worker successfully transits from the formal to the informal sector [5]. In this case, the observed sector will be $s = 0$, being different from the latent sector. With probability P_2^f the worker earns the minimum wage. With the complementary probability ($U^f = 1 - P_1^f - P_2^f$), the worker becomes unemployed.[6]

Assumption 3. *Known spillover effects:*

$$w^* > Mw \rightarrow \begin{bmatrix} w \\ s \end{bmatrix} = \begin{bmatrix} \Lambda(w^*) \\ s^* \end{bmatrix}$$

Where Λ: [Mw,∞] $\rightarrow$ [Mw,∞] is a known invertible spillover function with inverse given by $\lambda(\Lambda(w)) = w$. One important special case is the absence of spillovers, where $\Lambda()$ is given by the identity function. [7]

Assumption 4. *Continuity of the latent wage distribution:*

The latent wage distribution and its first derivative is assumed to be continuous everywhere. In particular:

$$lim_{\epsilon \rightarrow 0^+} f(Mw - \epsilon) = lim_{\epsilon \rightarrow 0^+} f(Mw + \epsilon)$$

[3] This assumption is implied by a variety of primitive assumptions on the technology, the distribution of workers heterogeneity and wages determination mechanism.

[4] The first reason for allowing workers in the informal sector to bump in the minimum wage is for the model to account for the empirical fact that they actually do so. The economic logic behind this regularity is under debate. One hypothesis is that the minimum wage acts as a signal to the agents of a fair price for unskilled labor, which might affect the way workers in the informal sector bargain with their employers. This feature is closely related to the "self-enforcing" nature of minimum wages.

[5] The assumption that the wage is exactly simplifies the exposition. The same results holds when this assumption is replaced to the worker drawing a wage from the below the minimum wage conditional distribution of wages. This modification does not change the results of the model.

[6] To ease the exposition I assumed that P_2^f and U_f do not vary as a function of the latent wage. In this case that they vary over the latent wages, the parameter recovered by assuming that they are constants is the expectation of the distribution of P_2^f and U^f over the distribution of wages below the minimum. Importantly, this result holds only as long as P_1^f remains constant as a function of the wage.

[7] The model can easily incorporate stochastic spillovers as well.

$$lim_{\epsilon\to 0^+} f'(Mw-\epsilon) = lim_{\epsilon\to 0^+} f'(Mw+\epsilon)$$

As discussed in Doyle (2007), this third assumption exploits the fact that the distribution of worker productivity is likely to be smooth [8], but the observed density of wages has a jump around the minimum wage. This jump might give exactly the information necessary to trace back the effects of the policy on the outcomes of interest.

The goal is to recover the unknown parameters $\theta \equiv (P_1^f, P_1^i, P_2^f, P_2^i, U_f)'$, and the function $\rho(w^*)$. Using estimates of these parameters one can recover the underlying density of wages F, that is, the density that would prevail in the absence of minimum wage. By comparing these two distributions one can evaluate how much the minimum wage affected labor market outcomes, such as wage inequality, unemployment and so on. Notice that by defining the latent sector and the sector-specific parameters a broader range of implications of minimum wage becomes assessable, such as changes in tax revenues and movements between sectors.

It is helpful to understand the implications of the model using limiting cases for the parameter values. For example, if P_1^f tends to zero, there is no mobility between sectors. In this case, unemployment size will be given by $F(Mw|s^*=1)(1-P_2^f)$, which simply means that the unemployment will be higher the smaller the probability of workers to bump at the minimum wage, the higher the mass of workers for which the minimum wage 'bites', and the bigger the formal sector size. On the other extreme, when P_1^f tends to one, all workers in the formal sector manage to find a job with the same wage in the informal sector, which also implies no unemployment effects from the minimum wage. Effects of minimum wage on average wages are maximized in the limiting case where P_2^f and P_2^i tend to one. In terms of market structures that could generate these values, P_1^f tends to one if the economy can be described by a simple two-sector model with imperfect compliance with the minimum wage and costless sector mobility. P_2^f tends to be higher if the economy is mostly consisted of employers with monopsonistic power in the labor market, and U^f tends to be higher if the labor market operates close to perfect competition and mobility to the informal sector is limited.

4 Identification

It is not possible to directly use the techniques developed in Doyle (2007) in each sector separately, since I introduced movements between them. Also, there is no guarantee that the unconditional probability of non compliance will not vary over the wage distribution, even when the sector specific probabilities (P_1^f and P_1^i) do not vary with respect to the wage. So, to solve the model, a different approach must be used. Below, I state the main identification results of this paper, which concerns the identification of (a) the latent **joint** distribution of sector and wages, that is, the distribution that would prevail in the absence of minimum wage; (b) the vector of parameters θ which governs how the minimum wage affects the economy, and (c) the effects of Minimum Wage on functionals

[8] This condition is implied by most of the standard models of wage determination, like Mincer equations with normally distributed errors. It is also satisfied when latent wages are lognormally distributed, which is also often assumed in the literature.

of the distribution of sector and wages (d) the effects of Minimum Wage on Labor tax revenues. On the following exposition, assume that the econometrician observes a random sample of the pair (w_i, s_i) of size N from a population of interest.

Lemma 4.1 (***Identification of sector specific parameters***). Under assumptions 1,2,3 and 4 and some technical conditions, θ is identified. Proof: See Appendix 1.

Lemma 4.2 (***Identification of latent distributions***). Under assumptions 1,2,3 and 4 and some technical conditions, the latent **joint** distribution of sector and wages is identified. Proof: See Appendix 1.

Finally, define $TE_\nu \equiv \nu(\mathbf{F_{w^*}}) - \nu(\mathbf{H_w})$ and $STE^j_\nu \equiv \nu(\mathbf{F_{w^*|s=j}}) - \nu(\mathbf{H_{w|s=j}})$ for j $\in 0, 1$, where $\nu()$ is a functional of a distribution function, such as the expectation, the standard deviation or Gini coefficient. These are respectively the "Functional Treatment Effect" and the "Sector Specific Treatment Effects" of minimum wage on the distribution of wages. Similarly, define $ATE_{s,s^*} \equiv Pr(s^*_i = 1) - Pr(s_i = 1)$, the effect of minimum wage on the share of the formal sector of the economy.

Corollary 4.3 (***Identification of Minimum Wage treatment effects***). Under assumptions 1,2, 3 and 4 and some technical conditions, TE_ν and STE^j_ν and ATE_{s,s^*} of minimum wage are identified. Proof: See Appendix 1.

Corollary 4.4 (***Identification of Minimum Wage Effects on Labor Tax Revenues***). Under assumptions 1,2, 3 and 4 and some technical conditions, the effects of minimum wage on labor tax revenues are identified. Identification of the effects of minimum wage on labor tax revenues hold as long as it can be written as a functional of the latent and observed wage distributions and the model parameters θ. See the section on taxes for further discussion of this issue.

The intuition for the identification can be summarized by a simple three-step procedure. First, the latent share of the formal sector is identified by the share above the minimum. Next, a fake dataset is generated. In this dataset, the spike at the minimum is removed by substituting the observations truncated at the minimum by random draws of the wage distribution conditional on being below the minimum. [9] After removing the truncation, more observations are then added until the distribution shows no gap at the minimum wage. The model parameters are then identified by computation of simple proportions based on this generated dataset and on the observed one. Figure 1 illustrates such process using a uniform distribution with a minimum wage set at .5. Of course, if the proportion of workers in each sector changes with respect to the wages, the process of adding observations need to be adjusted accordingly, but the intuition remains the same.

5 Estimation

A crucial step for obtaining estimates of the objects of interest such as the model parameters and the counterfactual distributions involves the estimation of a ratio of one sided limits of the density at the minimum wage.

[9] This can be done by simply reweighting the data, given zero weight to the observations at the minimum, and reweighting the observations below the minimum by the ratio $Pr(w_i = Mw)/Pr(w_i < Mw)$.

The estimation of these quantities can be performed by non-parametric methods. Notice that since the density is discontinuous around the minimum wage, only observations below the minimum are informative of the density $h(Mw-\epsilon_n)$ (and similarly for the density above the minimum). This implies that the estimators of these quantities will behave like if the minimum wage was a boundary point of the density, which has implications in terms of bias and variance.

Therefore, it is advisable to use methods such that the performance of the density estimator is satisfactory on points that are close to the support boundaries. I used local linear density estimators, which have the same order of bias on the boundary as in interior points of the distribution. This estimation strategy here will closely follow Mccrary(2008) in the context of testing for manipulation of the running variable in RD designs.

A standard approach to non-parametrically estimate densities at boundary points is to use a local linear density estimator. This estimator builds on the idea of local linear conditional mean estimators. It starts by dividing the support of the density in a set of bins. After, a "response variable" is defined as the bin counts of these disjoint intervals. After this process, one is left with a vector containing the "independent variable" which are the bin centers and a correspondent "dependent variable", the bin counts. Finally, standard local polynomial smoothing estimates are applied to this constructed variables.

Define $g(w_i)$ as the discretized version of the wage support for a binsize equal to b.

$$\text{g}(\text{w}_i) = \begin{cases} \lfloor \frac{w_i - Mw}{b} \rfloor b + \frac{b}{2} + Mw \text{ if } w_i \neq Mw \\ Mw \text{ if } w = Mw \end{cases}$$

Where $\lfloor a \rfloor$ is the greatest integer in a. Clearly, it holds that $g(w_i) \in \chi \equiv \{..., Mw - 5\frac{b}{2}, Mw - 3\frac{b}{2}, Mw - \frac{b}{2}, Mw, Mw + \frac{b}{2}, Mw + 3\frac{b}{2}, Mw + 5\frac{b}{2}, ...\}$. I will call the jth element of this set X_j. [10].

Define the normalized cellsize for the jth bin, $Y_j = \frac{1}{Nb}\sum_{i=1}^{N} \mathbb{I}\{g(w_i) = X_j\}$

Let $K(.)$ by a symmetric kernel function satisfying the usual properties. Then, the local linear estimator of the density and its derivative are defined, for $w \neq Mw$ as:

$$\begin{bmatrix} \hat{h}(w) \\ \hat{h}'(w) \end{bmatrix} = \text{argmin}_{(a,b)'} \sum_i^J (Y_j - a - b(X_j - w))^2 K(\tfrac{w - w_i}{h})(\mathbb{I}\{X_j > Mw\}\mathbb{I}\{w > Mw\} + \mathbb{I}\{X_j < Mw\}\mathbb{I}\{w < Mw\})$$

For a bandwidth h satisfying the conditions $nh \to \infty$ and $(nh)^{1/2}h^2 \to 0$, it can be shown that as $n \to \infty$:

$$(Nh)^{1/2}(\widehat{h(w)} - h(w)) \to N\left(0, h(w)\int K^2(u)du\right).$$

[10]As discussed in Mccrary(2008) the endpointsX_1 and X_j may always be chosen arbitrarily small (large) so that all points in the support of the distribution of wages are inside one of the bins.

And:

$$(Nh)^{3/2}(\widehat{h'(w)} - h'(w)) \to N\left(0, h'(w)\int K'(u)^2 du\right).$$

I approach to estimation that this paper will follow is the analogy principle, replacing population objects by their respective sample counterparts whenever feasible. So, the estimator of $P_1(Mw)$ will be given by:

$$\hat{P}_1(Mw) = \frac{\hat{h}(Mw^-)}{\hat{h}(Mw^+)}$$

Where $\hat{h}(Mw^-)$ is the estimator of the density just below the minimum wage value using the local linear density estimator.

Then, using the delta method, it can be shown that:

$$(Nh)^{1/2}(\hat{P}_1(Mw) - P_1(Mw)) \to N(0, \Sigma)$$

$$\hat{P'}_1(Mw) = \left(\frac{\hat{h}'(Mw - \epsilon_n)}{\hat{h}'(Mw + \epsilon_n)} - \frac{\hat{h}(Mw - \epsilon_n)}{\hat{h}(Mw + \epsilon_n)}\right) \cdot \frac{h'(Mw + \epsilon)}{h(Mw + \epsilon)}$$

Then, again using the delta method and for a particular choice of bandwidth:

$$(Nh)^{3/2}(\widehat{P}_1'(Mw) - P_1'(Mw) \to N(0, V)$$

To complete the process of recovering the structural parameters θ one need estimates of $\rho(Mw)$ and $\rho'(Mw)$. Given Assumption 1, these estimators can be defined as:

$$\rho(Mw) = m(Mw, \widehat{\beta})$$

$$\rho'(Mw) = m'(Mw, \widehat{\beta})$$

Where:

$$\widehat{\beta} = argmin_b \sum_{i=1}^{N} (s_i - m(w_i; b))^2 \mathbb{I}\{w_i > Mw\}$$

Finally, using the estimate $\hat{\rho}(Mw)$ of the latent share of the formal sector, we can define the plug-in estimator for the parameters P_1^f and P_1^i:

$$\hat{P}_1^i = \hat{P}_1(Mw) - \frac{\hat{\rho}(Mw)}{\hat{\rho}'(Mw)} \cdot \hat{P}_1'(Mw)$$

$$\hat{P}_1^f = [\hat{P}_1(Mw) - (1 - \hat{\rho}(Mw)) \cdot \hat{P}_1^i] \cdot \hat{\rho}(Mw)^{-1}$$

These estimators are also consistent and asymptotically normally distributed for a suitable choice of bandwidth.

$$\hat{D} = [\sum_{i}^{J} \frac{Y_i b}{\hat{P}_1(X_i)} \mathbb{I}\{X_i < Mw\} + 1 - \widehat{H}(Mw)]^{-1}$$

Then:

$$\hat{f}(w) = \begin{cases} \frac{\hat{h}(w)\hat{D}}{\hat{P}_1(w)} \text{ if w} < \text{Mw} \\ \hat{h}(w)\hat{D} \text{ if w} \geq \text{Mw} \end{cases} \tag{1}$$

Finally, defining

$$\widehat{\omega}(w_i) = \begin{cases} \frac{\hat{D}}{\hat{P}_1(w_i)} \text{ if } w_i < \text{Mw} \\ \hat{D} \text{ if } w_i \geq \text{Mw} \end{cases}$$

Thus it becomes clear that we can reweight the observations in the observed data to obtain the latent distribution. This eases the computation of the treatment effects.

$$\widehat{TE}_\nu = \nu(\hat{H}) - \nu(\hat{F})$$

$$\widehat{STE}_\nu^j = \nu(\hat{H}_{|s=j}) - \nu(\hat{F}_{s=j})$$

$$\widehat{ATE}_{s,s^*} = \widehat{Pr}(s=1) - \widehat{Pr}(s^*=1) = N^{-1}\sum_{i}^{N} s_i - \sum_{i=1}^{J} \hat{\rho}(X_i)\hat{f}(X_i)b$$

Consistency of the estimators of θ, β and consequently $\rho(w^*)$ and $f(w^*)$ follows directly from the identification equations and the consistency of the estimators of $h(w)$ and $h'(w)$. Closed formulas for the asymptotic variances can be easily derived, but I will rely on re-sampling methods to estimate them on the empirical application.

6 Testing

Once one successfully estimated the model parameters several Hypothesis regarding the behavior of the economy might be tested. Null hypothesis of special interest are zero restriction on the model parameters such as P_1^f and U^f. As discussed before, if $P_2^f = 0$ the economy can be described by a two-sector model without sector mobility. Other interesting condition to test is $P_1^f = 1$ which tests if the economy can be described by a two sector model with costless sector mobility.

Perhaps even more importantly, some of the model's maintained assumptions are partially testable. Assumption 4, continuity of the latent wage distribution, can be verified by visual inspection of the histogram and the kernel density estimates using different values for the bandwidth. Formally, this condition can be tested by checking if there is statistically significant differences between the left and right limits of density estimates at wage points different than the minimum wage. Assumption 1 can be tested by comparing the fit of the parametric model with nonparametric smoothing estimates. Define the statistic:

$$I \equiv \int_{Mw}^{\infty} (m(u;\beta) - \rho(u))^2 h(u) du$$

Where $\beta = argmin_b E[(s_i - m(w_i;b))^2 \mathbb{1}\{w_i > Mw\}]$. Then, correctness of the specification of the model for $\rho(w^*)$ implies that $I = 0$.

$$\hat{I}_n \equiv \sum_{i=1}^{J} (\hat{m}(X_i;\hat{\beta}) - \tilde{\rho}(X_j))^2 \hat{h}(u) b \mathbb{1}\{X_j > Mw\}$$

Where $\tilde{\rho}(u) \equiv \frac{\sum_i^N s_i K(\frac{w_i - u}{\sigma})}{\sum_i^N K(\frac{w_i - u}{\sigma})}$ is a nonparametric estimator of the conditional mean of the sector given the wage. Comparing the fit of the parametric model with the one from the nonparametric one can help to identify the proper functional form for the sector wage relationship. This is specially important given part of the identification relies on extrapolating this conditional mean function to values below the minimum wage.

The most surprising result is that even assumption 2 is also testable. The fact that it is used to define and identify the parameters of the model might give the impression that the resulting estimated latent density must "conform" with the estimated value of θ, in the same sense that the error terms in a linear regression are always by construction orthogonal to the vector of covariates. The key feature that allows one to test assumption 2 - against the alternative that the parameters of the model are not constant over the wages conditional on the latent sector - is that under the null the second derivative of the wage density should show a discontinuity similar to the one presented on the first derivative. By comparing those two one can assess if the parameters P_1^f and P_1^i do vary over the wages around the minimum.

To see why this is the case, first one need to look at the second derivative of the observed wage density:

$$h''(w) = \begin{cases} \frac{P_1''(w)f(w)}{D} + 2\frac{P_1'(w)f'(w)}{D} + \frac{P_1(w)f''(w)}{D} \text{ if w < Mw} \\ \frac{f''(w)}{D} \text{ if w > Mw} \end{cases} \tag{2}$$

Now, if the continuity assumption on the latent wage distribution is strengthened up to the second derivative, then we have that:

$$P_1'(Mw) = \frac{1}{2}\left[\frac{h''(Mw - \epsilon)}{h''(Mw + \epsilon`)} - \frac{h(Mw - \epsilon)}{h(Mw + \epsilon)}\right] \frac{h''(Mw + \epsilon)}{h'(Mw + \epsilon)}$$

The interesting feature of this strategy is that under the null that if the model parameters are indeed structural in the sense that the probabilities (P_1^f, P_1^i) do not vary across different wages, then $P_1'(Mw)$ estimated by the equation above should converge to exactly the same value as the one estimated using the baseline equation, from the first derivative.

So, by comparing estimates of $P_1'(Mw)$ obtained by the expression above with the ones based on the first derivative of the wage distribution one can formally test the null hypothesis that the model parameters are constant over the wages against the alternative that there is a linear relationship between the model parameters and the wages. Under the null hypothesis that the model parameters are constants over the wages, the two estimates should be similar. If the model parameters vary over the wages, then $P_1'(Mw)$ defined above will be different from the one using the first derivative. This difference allows us to test the properness of Assumption 2.

If one is willing to impose further smoothing conditions on the latent wage distribution, it is possible to identify the model imposing very flexible conditions on the relationship between the parameters and the wages. For example, if one believes that (P_1^f, P_i^i) is appropriately described by a quadratic (cubic) function, then one needs to go up to the third (fourth) derivative of the wage density to estimate the model parameters. And even in this scenario, the assumption that this functional form for the parameters is correctly specified can be tested against the alternative that a more flexible polynomial better describes the data. This can be done in the same fashion as the tested presented above, that is, by comparing the estimates with the ones obtained using further derivatives of the wage density.

The results from this section are summarized in the following lemmas, which concerns testing hypothesis regarding the model parameters and the model assumptions.

Lemma 6.1 (***Hypothesis Tests for*** θ). Let θ^k be the kth entry of the vector θ. Under Assumptions 1 to 4 and for a suitable choice of bandwidth, $\frac{\hat{\theta}^k - \hat{\theta}_0^k}{\widehat{SE}(\hat{\theta}^k)} \rightarrow N(0,1)$

Lemma 6.2 (***Testing Assumption 1***). Under standard technical conditions that allow non parametric estimation of the conditional mean of sector given wages, Assumption 1 is testable through the statistic I_n.

Lemma 6.3 (***Testing Assumption 2***). Under stronger smoothness conditions on the wage distribution that allows for non parametric estimation of the second derivative of the wage density and the assumption that $f''(Mw) \neq 0$, Assumption 2 can be tested against the alternative that $P_1(Mw)$ is a quadratic function of wages.

Lemma 6.4 (***Testing Assumption 4***).

$$T(w) = \frac{\frac{\hat{h}(w^-)}{\hat{h}(w^+)} - 1}{\widehat{SE}(\frac{\hat{h}(w^-)}{\hat{h}(w^+)})}$$

Then, for $w \neq Mw$ T(w) $\rightarrow N(0,1)$ if the latent wage distribution is continuous.

7 Role of Covariates and Unobserved Heterogeneity

Preliminary Version – [To be rewritten]

By exploring the different effects of minimum wage between sectors and the discontinuity of the density of wages around the minimum one can estimate how the economy responds to this policy. This approach has some similarities to the quasi-experimental Regression Discontinuity Designs.

Since one of the main advantages of the Regression Discontinuity Designs is to provide a way to avoid most of the endogeneity considerations on using observational data to infer causality, it is useful to discuss how much of these advantages are also present in this method.

Assume there is a variable X – say, for example, age – distributed in a bounded domain that is known to affect individual labor market conditions. One example is when workers with different values of X draw from different latent wage distributions. Another way that X can affect the worker's labor market conditions is through the model parameters. For example, after the introduction of minimum wage younger workers might be more likely to move into the informal sector than older workers, which in the model would be represented by a higher P_1^f. In these cases, is it necessary to estimate the model conditional on X for the inferences to be valid?

In the following discussion, I will always assume continuity of the x-specific latent wage distribution; absence of spillovers and a covariate specific version of assumption 2.

The sufficient conditions for the inferences based on the unconditional wage distribution to be valid in the presence of covariates are the following:

Case 1:

Assumption 5. *Latent wage and sector distributions in the presence of covariates:*

$$\begin{gathered} w_x^* \sim F_x \\ s_x^* \sim B(\rho(w_x^*)) \\ \rho(w_x^*) = \tilde{m}(w^*; \beta_x) \\ \rho(w^*) = m(w^*; \beta) \end{gathered}$$

Assumption 6. *Equality of Parameters*

$$\theta(x) = \theta \quad \forall x \in \chi$$

In this case, the role of X is to change the latent joint distribution of sector and wages. The first assumption defines the role of x by changing both the latent wage distribution and the model for the conditional mean of sector given the wage. The last line of this assumption ensures that a (perhaps different) model for the aggregate data still holds.[11]

In summary, when the effect of X occurs through changes in the latent joint distribution of sector and wages but not through differential responses to the minimum wage law then X can be safely ignored when making inferences with regard to the unconditional distribution.

Case 2:

Assumption 7. *Latent wage and sector distributions in the presence of covariates:*

[11] In general this model will be more complex than the covariate-specific one. A simple sufficient condition to guarantee that such a model will exist is when the third line of Assumption 5 is strengthened to $\rho(w_x^*) = m(w^*; \beta)$.

$$
\begin{gathered}
w_x^* \sim F \\
s_x^* \sim B(\rho(w_x^*)) \\
\rho(w_x^*) = m(w^*, \beta_x) \\
F_x(w^*) = F_{x'}(w^*) \quad \forall x, x' \in \chi
\end{gathered}
$$

Assumption 8. *Equality of a subvector of parameters*

$$
U_f(x) = U_f(x') \qquad \forall x, x' \in \chi
$$

It is clear that even after restricting the latent wage distribution to be the same for all values of X, inference based on the unconditional distribution ignoring the covariate will only be valid if unemployment effects are the same regardless of X. The intuition for this result is that in this case the relative size of each group in the observed data is not changed by the introduction of the minimum wage. The remaining parameters $(P_1^f, P_i^1, P_f^2, P_i^2, \rho)$ recovered from the aggregate data will be weighted averages of the covariate specific ones, with correct weights to reflect the share of each group of values of X in the population. These, of course, are much stronger conditions than those in Case 1, since the role of covariates are severely limited when they are only allowed to determine the wages through the differences in minimum wage effects.

When both the latent wage distribution and the parameters are allowed to vary over X the estimate of P_1 can be interpreted as a local effect, since it recovers the likelihood of non-compliance for those with latent wages around the minimum wage. Preliminary results from simulations showed that is necessary an unreasonably large degree of heterogeneity on both the latent distributions and the model parameters for the inference based on unconditional distribution show sizable distortions.

The relevance of these results is quite small if the wage determinants are observable, since under assumption 5 the model can be easily estimated conditional on these variables. If the estimation is performed conditioning on the covariates, then assumption 6 can be dropped, meaning that the model parameters can be different for different values of X. However, things are different when not all wage determinants are observable. Failure to observe wage determinants is a major source of bias in inferences based on regression models. In this design this is not the case, as long as the model parameters remain constant over the distribution of the variables that are ignored, which seems to be a much easier condition to satisfy than the zero correlation usually assumed in regression models. In this sense, this research design resembles most of the characteristics of Regression Discontinuity Designs, overcoming the difficulties to assess causal effects from observational data that are due to endogeneity considerations. The reason for that is that the identification does not rely on the variation of minimum wage to assess its impact. Instead, it relies on the sharp contrast between the effect of minimum wage across individuals whose wages would fall on each side of it. Thus, concerns with omitted variable biases should be much more limited.

8 Empirical Application - The Effects of Minimum Wage in Brazil

8.1 Data and Descriptive Statistics

On the empirical application I will strengthen assumptions 1 and 2 to the following:

Assumption 9. *Latent wage and sector distributions:*

$$\begin{gathered} w^* \sim F \\ s^* \sim B(\rho) \\ s^* \perp w^* \end{gathered}$$

Assumption 10. *No spillover:*

$$w^* > Mw \to \begin{bmatrix} w \\ s \end{bmatrix} = \begin{bmatrix} w^* \\ s^* \end{bmatrix}$$

This greatly simplifies the estimation, as it can be seen in the Appendix 2. Importantly, the orthogonality between latent sector and wages can be tested. Below I provide evidence that it is not violated on the context of the Brazilian labor market.

To evaluate the effects of minimum wage on labor market outcomes I used the years of 2009 to 2001 of the dataset known as PNAD. This data has been collected by the IBGE – which is a Portuguese acronym for "Brazilian Institute of Geography and Statistics" – since 1967 and contains characteristics of income, education, labor force participation, migration, health and other socioeconomic characteristics of the Brazilian population. Workers who do not report wages, those who work in the public sector (since there is no informal public sector) and workers who are older than 60 years old or younger than 18 years old were removed from the sample. The variable of interest – wage – is measured at the monthly level, which is the most natural unit within the institutional context of Brazil. The real wage value was computed using the "IPCA", which is the consumer price index used by the Central Bank in the "Inflation Target System". Importantly, the dataset includes information regarding the worker's labor contract status, which was used to define formality.

In Brazil, all workers carry an official document called "Carteira de Trabalho" (worker's card). This document is signed by the employers in the formal act of hiring. Lack of a formal signed labor contract means that the employer is not enforced to collect labor taxes neither to comply with minimum wage and other kinds of regulation. The Brazilian economy is known to be characterized by a large informal sector. Tables 1, 2 and 3 illustrate this fact and describe the main features of the data. All estimates are computed considering the weights in the survey design.

It is clear from Table 2 that workers in the informal sector earn on average 35% less than workers on the formal sector. In addition, in terms of the observable characteristics, workers in the informal sector are more likely to be man, nonwhite, less educated and young. Considering the likelihood of earning minimum and sub-minimum wages, Table 3 shows there is considerable variation on these probabilities across population subgroups. For example, white workers have a probability of earning minimum wage that is 40% smaller than that for nonwhite workers. Workers with less than 5 years of education have around 20% of probability of earning minimum wage, whereas such likelihood is only 5% for workers with more than 12 years of education. When different regions are compared, remarkable heterogeneity is found as well. Workers in the South Region have only 5% of probability of earning minimum wage. On the other extreme, workers in the Northeast have around 24% of probability of earning minimum wage. The same pattern appears when we look at the probability of earning sub-minimum wages.

The history of minimum wages in Brazil started in the Getulio Vargas government on May 1st, 1940. Initially, the minimum wage did vary across regions to accommodate differences in price levels across the country. Later, in 1984, regional minimum wages were unified into a single wage at national level. Importantly, the Constitution of 1988 prohibited use of the minimum wage as a reference for wage bargaining for other categories of workers and contracts. This aimed at reducing over-indexation of the economy, which was thought to be fueling inflation. The periodicity of changes in the minimum wage has been yearly since the economic stabilization in 1994. Graphs 2 and 3 show the evolution of minimum wages along with different statistics of wage distribution over the last decade.

By looking at Graphs 2 and 3, the challenge of relying on time series variation to identify the effects of minimum wages becomes clear, since there is almost as much evidence in favor of minimum wage effects on the 20th percentile as there is on the 80th percentile [12]. The correlation between minimum wage changes and changes on such high percentiles of wage distribution are probably a reflection of the pro-cyclical nature of changes on minimum wages. Given this, effects of minimum wages on other statistics of wage distribution – such as average wages or lower quantiles – that are based on time series variation should be cautiously interpreted as well.

9 Results

Figure 5 shows that, as a consequence of sizable unemployment effects, the observed density above the minimum wage is higher than the latent density. Furthermore, due to both truncation at the minimum and unemployment, the observed density below the minimum wage is greater than the latent density. Looking at the point estimates and standard errors [13] in Table 4, we see that unemployment effects are sizable, as a result of limited mobility across sectors. Interestingly, the results are comparable to the other uses of this approach. Doyle, for example, found that around 60% of young workers that would earn below the minimum became unemployed. The high unemployment effects in a country with a large informal sector are clearly due to two reasons: First, as stated above, there is very little evidence of mobility across sectors: My estimates are around 10%, with a maximum of 22%. Second, the small probability of truncation at the minimum wage gives evidence that small wages on the wage distribution are more likely due to low productivity

[12] The same feature was noticed by Lee(2008) when analyzing U.S. data.

[13] Standard errors were computed through non-parametric bootstrap. Bandwidth was selected using MSE minimizing bandwidth (eight times the Silverman's rule of thumb) in a Monte Carlo exercise imposing a log-normal distribution for wages. As a robustness test, I apply the automatic bandwidth method proposed by Mccrary(2008) and got similar results. A theoretical method for bandwidth selection in this setup is object of ongoing research.

than to low bargaining power.

Another interesting finding is the enormous mass of workers that would earn wages below the minimum wage in the absence of the minimum wage imposition. Estimates are around 34%, but a clear upward trend is visible, since the estimates go from the minimum value at 25% in 2001 to the maximum of 43% in 2007 and 2009.

In addition, the high correlation over time of unemployment estimates (U) and the proportion of workers that would earn wages below the minimum in the absence of the minimum wage imposition ($F(Mw)$) is remarkable. Using aggregate unemployment estimates, the correlation coefficient is .85; using unemployment in the formal sector (U_f), which also takes variations on the estimate of ρ into account, the correlation is .86. This finding suggests there is not much more room for effective increases in the real value of the minimum wage. The correlation coefficient between aggregate probability of truncation at the minimum wage and $F(Mw)$ is equal to -.71. And the probability of a worker bumping at the minimum is already small (around 27%). As the results above indicate, it is more likely to go down in the event of further increases. Putting together, it is safe to say that future increases in minimum wage will probably be more harmful in terms of unemployment and less fruitful in terms of shaping the wage distribution through the truncation effect.

On the other hand, Table 5 shows how strongly the minimum wage affects the shape of (log)wage distribution. Here I compute the effects of minimum wage on usual measures of wage inequality, such as standard deviation of log wages, gini and so on. Clearly, minimum wages have a large positive impact on average wages (conditional on employment). The maximum difference is .39 log points in 2007 and the minimum is .18 in 2002. Minimum wages also reduce wage inequality, as measured by differences in quantiles, standard deviation, or the gini coefficient. These estimates give a clear picture of the trade-off faced by policy makers when choosing the minimum wage level. On one hand, there is a gain in terms of reducing wage inequality and increasing average wages. On the other hand, workers tend to have more difficulty finding jobs.

Table 6 shows how heterogeneous the parameters are across population subgroups defined by observed covariates. Based on model estimates for the year of 2009, there are some differences in the estimates across gender, race and education. Interestingly, the null hypothesis of zero sector mobility for the groups of woman, black and individuals with less than 12 years of education cannot be rejected. This suggests that the very small but yet significant estimates of the unconditional model might be due to bias generated by ignoring the role of covariates; however, it can also be only due to the loss of power of using a smaller set of observations. On the other hand, even with changes in significance for some coefficients, the magnitude of differences between estimates across subgroups is still quite small. This finding suggests a limited role for “omitted variable biases” in the estimates of the unconditional version of the model.

9.1 Tax Revenues and Size of Informal Sector

A simple comparison of Tables 1 and 4 shows that the minimum wage compresses the share of the formal sector of the economy. This occurs through two different but related channels: First, minimum wage reduces the size of the formal sector as long as unemployment effects are greater than zero, as it was found in Brazil. Second, minimum wage increases the size of the informal sector, through sector movements that are driven by the introduction of the minimum wage. These later effects were shown to be relatively small in this application. Overall, the share of the formal sector in the Brazilian economy is reduced by around 10%.

For that reason, minimum wages end up indirectly affecting the government budget. Here I used the term indirectly because minimum wages already affect the government budget through the spending channel. This is due to the indexation of pensions to the minimum wage, which is usually the effect with which policy makers are concerned when discussing minimum wage increases.

But here I focus on the indirect channel, which is often ignored. Minimum wages affect the shape of wage distribution, the relative size of the formal sector and the likelihood of employment. Each of these has the potential of changing tax revenues. Therefore, the goal of this section is to get an estimate of these effects. I here considered the effects on the INSS tax revenues, which is the Brazilian labor tax. The INSS is the tax collected to fund the social insurance system in Brazil, and the rate is 20% for companies inserted in the regular system of taxation and 12% for small companies that opt for the "simplified" system.

To estimate the effects, I will rely on the following assumption:

Assumption 11. *No Tax Revenues in the Informal Sector*

Given this assumption, the effects of minimum wages on the revenues from a labor taxes of rate $\tau(w)$ are identified. By definition, the tax revenues will be:

$$T^1 = \sum_{i=1}^{N^1_{s=1}} \tau(w_i) w_i$$

$$T^0 = \sum_{i=1}^{N^0_{s=1}} \tau(w_i) w_i$$

Where T^j represents the tax revenues and $N^j_{s==1}$ is the size of the population employed in the formal sector under each scenario– $j = 1$ indexes under the minimum wage imposition and $j = 0$ in its absence –.[14] The object of interest is the ratio of these two quantities. After some algebra, it can be shown that:

$$R \equiv \frac{T^1}{T^0} = \frac{s}{\rho} \cdot (1 - F(mw)U) \cdot \frac{E(\tau(w)w|s = 1)}{E(\tau(w^*)w^*|s^* = 1)}$$

Where s is the observed share of the formal sector and ρ the latent one. This expression is further simplified in the Brazilian case, where labor taxes are a constant fraction of the wages. In this case, R is given by:

$$R \equiv \frac{T^1}{T^0} = \frac{s}{\rho} \cdot (1 - F(mw)U) \cdot \frac{E(w|s = 1)}{E(w^*|s^* = 1)}$$

Interestingly, the effects on tax revenues can be decomposed into three components: Compression of the formal sector, reduction in the workforce size through unemployment effects, and change

[14] Notice that I abused on notation and used N to refer here to the size of the population, not the size of the sample.

in expected wages in the formal sector. We already know that the minimum wage increases the expected wages compared to the latent wage distribution. The question is whether it increases the expected wages enough so that it compensates for the reduction of employment in the informal sector due to sector migration and unemployment. Notice this ratio also answers a related question: Is it the mass of wages, the sum of wages of all workers in the formal sector, higher under the minimum wage or in its absence? Since the tax rate τ is a constant function of the wages, the effects on tax revenues are proportional to the effects on the mass of wages.

Table 7 shows that the minimum wage clearly reduces the mass of wages in the formal sector, with a corresponding loss on labor tax revenues. This is due to the sizable unemployment effects and reduction in the formal sector size, which more than compensate for the increase in expected wages.

To give an idea of the magnitude of this difference, a different exercise will be performed. In such exercise I will first ask the reader to ignore the model for a moment and focus only on the expression for R. Its three components are potentially independent pieces. The first two pieces s/ρ and $1 - F(mw)U$ account for the differences in size of the population employed in the formal sector. The last piece accounts for differences in wages. Under the model assumptions, all these parameters can be estimated and, by doing so, R is also estimated. However, some readers might have different degrees of confidence in some of the identifying assumptions or guesses about the quantities present in the expression for R that are based on different assumptions. Most importantly, some of my estimates do not rely on all four assumptions. The estimates of s/ρ rely entirely on assumptions 1 and 3, for which the validity its hard to question in this application. Therefore, a related but different question could be: How much bias do I need to have on my estimates of the other parameters to get the wrong conclusion about the sign of minimum wage effects on tax revenues? The answer for this question is: a lot.

For example, fixing all other parameters, unemployment estimates in 2009 would need to be 39% percent smaller for the revenues under the minimum wage to be equal to the revenues under its absence. Similarly, expected wages in 2009 under the absence of minimum wages have to be at least 15% smaller than my estimates to achieve tax revenue equivalence between both scenarios. Anything greater than that would imply a smaller mass of wages in the presence of minimum wage. In summary, the mass of wage seems to be significantly compressed by the minimum wage legislation, going from small 2% estimates in 2001 to surprising 15% in 2009 [15].

10 Testing the Underlying Hypothesis and Robustness Checks

One of the advantages of the research design is that it is possible to indirectly test the most important model assumptions. Firstly, I will demonstrate that the density of wages in the informal sector in Brazil is informative about the shape of the overall latent wage distribution. For this to be true, the first model assumption must hold, i.e., the latent wage density should be the same in both sectors. This hypothesis is testable in two different ways. One is to look at the proportion

[15] Appendix I has other counter-factual exercises to evaluate the effects of minimum wages on tax revenues under a complete absence of unemployment effects. I show that the model can still be identified under this alternative hypothesis. In this extreme case, the minimum wage would increase tax revenues by 9% in 2001 and 3% in 2009.

of workers in each sector as a function of wages. If the assumption holds, this proportion should not vary together with the wage for wage values that are above the minimum. Of course, a naive regression of formality on wages should mechanically detect a negative relationship because no worker in the formal sector can earn below the minimum wage. However, after restricting our attention to wage values well above the minimum, the relationship should disappear. Another related way to do the same is by looking at the estimated wage densities, again restricting to wage values above the minimum. If the model is correct, differences in wage densities for values above the minimum between sectors are only due to rescaling and movements between sectors. Thus, by conditioning on values above the minimum, the effects of rescaling and sector movements should have no effect, and the densities should be approximately the same.

Graphs 6 and 7 give outstanding visual evidence of the adequacy of such assumption within the Brazilian context. The proportion of workers in the formal sector of the economy does not systematically vary with the wage, which gives confidence in the assumption that the underlying latent density should be the same between sectors. The plots of kernel density estimates across sectors point in the same direction: Workers in the formal and informal sectors apparently draw from the same distribution conditional on being above the minimum. This suggests that the differences between the overall distribution of wages occur as a result of the different ways the sectors respond to the minimum wage.

Table 8 shows the estimates of the elasticity of formality with respect to the wage by year, using different restrictions on the sample. The strong relationship between sector distribution and wages becomes much weaker after one conditions the regression to be above the minimum wage. Looking at the coefficient conditioning at higher values, the sign even changes to negative, which gives further evidence that sector distribution might be truly orthogonal to the wages at the latent wage distribution. As expected, several non different from zero estimates were found.

These results also allow us to reinterpret the observed difference in terms of demographic characteristics between sectors shown in Table 2. The higher proportion of nonwhite, less educated (and so on) is not due to structural differences between sectors beyond the way they respond to the minimum wage. In fact, it seems to be only a consequence of the fact that these workers have a higher probability of having a latent wage lower than the minimum, which makes it more likely for a worker in the informal sector to have these characteristics. This can be seen by looking at the differences in observable characteristics of the workers between sectors conditioning on values above the minimum wage. Table 9 shows a significant and sizable decrease on most of the differences between worker's characteristics across sectors after conditioning on wages above the minimum.

Other maintained assumption of the model is that the latent wage density is continuous around the minimum. If the wage density is continuous, then our estimates should not find any effect when the model for values different than the minimum is estimated. The table below reports the estimates of P_1 for several values, all of them different than the actual value of the minimum wage at the respective year. If the continuity assumption holds, the estimate of P_1 should not be statistically different from one.

As expected, the estimates fluctuate around one. This suggests that wage distribution presents no jumps for values other than the minimum wage. This increases the confidence that the continuity assumption holds for the latent wage density, and the spike observed in the data is only due to minimum wage legislation [16].

Finally, as discussed earlier, one key parameter of the model is defined as the ratio of the wage density above and below the minimum wage. On the baseline specification, the estimation was performed using local linear density estimators. This non-parametric method is advisable, since the order of bias at the domain boundary is the same as in interior points. As a robustness check, Table 11 shows the parameter estimates when P_1 is estimated using the approach proposed by Mccary (2008). It is clear from the comparison of both tables that although the point estimates are slightly different, qualitative implications are similar.

11 Conclusion

This papers explores the discontinuity on wage density generated by the otherwise apparently continuous wage distribution to assess the impacts of minimum wages on a broad range of labor market outcomes and policy relevant variables, such as size of the formal sector and labor tax revenues. The results show that minimum wage significantly alters the shape of wage distribution and reduces wage inequality. On the other hand, minimum wages come with a high cost of unemployment effects and reduction in the size of the formal sector of the economy. Together, these effects imply a reduction on tax revenues collected by the government to support the social welfare system.

The research design based on the sharp contrast of minimum wage effects between workers on each side of the minimum wage value allows for indirect tests of the underlying identification hypothesis of the model. The graphical and statistical evidence is in favor of the assumptions used and provides greater confidence in the inference based on the model estimates. Finally, the robustness check showed similar results when compared to the baseline estimator.

For future work, it could be it could be enlightening to further investigate presence of heterogeneity on the impacts of minimum wage across population sub-groups. Also, it could be helpful show the conditions under which is possible to jointly estimate spillovers together with the effects of minimum wage on the bottom part of the wage distribution. These extensions are object of ongoing research.

References

[1] Glauco Arbix. A queda recente da desigualdade no brasil. *Nueva Sociedad*, 2007.

[2] David Card and Alan B. Krueger. Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania. *American Economic Review*, 84:774–775, 1994.

[16] Notice that the standard errors are somewhat too small; therefore, we reject the null of no gap for most several years. But importantly, we tend to have estimates above one as much as we have below it, which means that we are as likely to find a jump on the latent density at the minimum wage on each direction, which is sufficient for my identification strategy.

[3] Ricardo Paes de Barros. A efetividade do salario minimo em comparacao a do programa bolsa familia como instrumento de reducao da pobreza e da desigualdade. In *Desigualdade de renda no Brasil : uma analise da queda recente*. IPEA, 2006.

[4] Ricardo Paes de Barros and Rosane Silva Pinto de Mendonca. Os determinantes da desigualdade no brasil. Julho 1995.

[5] Richard Dickens, Stephen Machin, and Alan Manning. The effects of minimum wages on employment: Theory and evidence from britain. *Journal of Labor Economics*, 17(1):1–22, January 1999.

[6] Francisco H.G. Ferreira. Os determinantes da desigualdade de renda no brasil: Luta de classes ou heterogeneidade educacional? Working Paper, 2000.

[7] Sergio Firpo, Nicole M. Fortin, and Thomas Lemieux. Occupational tasks and changes in the wage structure. Working Paper, September 2009.

[8] Sergio Firpo, Nicole M. Fortin, and Thomas Lemieux. Unconditional quantile regressions. *Econometrica*, 77:953–973, 2009.

[9] Sergio Firpo, Nicole M. Fortin, and Thomas Lemieux. Decomposition methods in economics. In *Handbook of Labor Economics*. Elsevier, 2011.

[10] Sergio Firpo and Mauricio Cortez Reis. O salario minimo e a queda recente da desigualdade no brasil. In *Desigualdade de renda no Brasil : uma analise da queda recente*. IPEA, 2006.

[11] A. Fishlow. Brazilian size distribution of income. pages 391–402. American Economic Association: Papers and Proceedings, 1972.

[12] Miguel Nathan Foguel and Joao Pedro Azevedo. Uma decomposicao da desigualdade de rendimentos do trabalho no brasil: 1995-2005. In *Desigualdade de renda no Brasil : uma analise da queda recente*. IPEA, 2006.

[13] Rodolfo Hoffmann. Transferencias de renda e a reducao da desigualdade no brasil e cinco regioes entre 1997 e 2004. *Economica*, v.8:55–81, 2006.

[14] Joseph J. Doyle Jr. Employment effects of a minimum wage: A density discontinuity design revisited. *Working Paper*, 2007.

[15] Chinhui Juhn, Kevin M. Murphy, and Brooks Pierce. Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101:410–442, 1993.

[16] C.G. Langoni. Distribuicao da renda e desenvolvimento economico do brasil. Rio de Janeiro: Expressï¿œo e Cultura, 1973.

[17] David S. Lee. Wage inequality in the united states during the 1980s: Rising dispersion or falling minimum wage? *Quarterly Journal of Economics*, 114(3):977–1023, 1999.

[18] Thomas Lemieux. Minimum wages and the joint distribution employment and wages. 2011.

[19] Sara Lemos. Minimum wage effects in a developing country. *Labour Economics,*, 16(2):224–237, 2009.

[20] Ana Flavia Machado, Ana Maria Hermeto de Oliveira, and Mariangela Antigo. Evolucao recente do diferencial de rendimentos entre setor formal e informal no brasil (1999 a 2005): evidï¿œencias a partir de regressoes quantilicas. In *Desigualdade de renda no Brasil : uma anaise da queda recente*. IPEA, 2006.

[21] Jose A .Machado and Jose Mata. Counterfactual decomposition of changes in wage distribution using quantile regression. *Journal of Applied Econometrics*, 20:445–465, 2005.

[22] Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698 – 714, 2008.

[23] Blaise Melly. Decomposition of differences in distribution using quantile regression. *Labour Economics*, 12(4):577–590, April 2005.

[24] Blaise Melly. Estimation of counterfactual distributions using quantile regression. 2006.

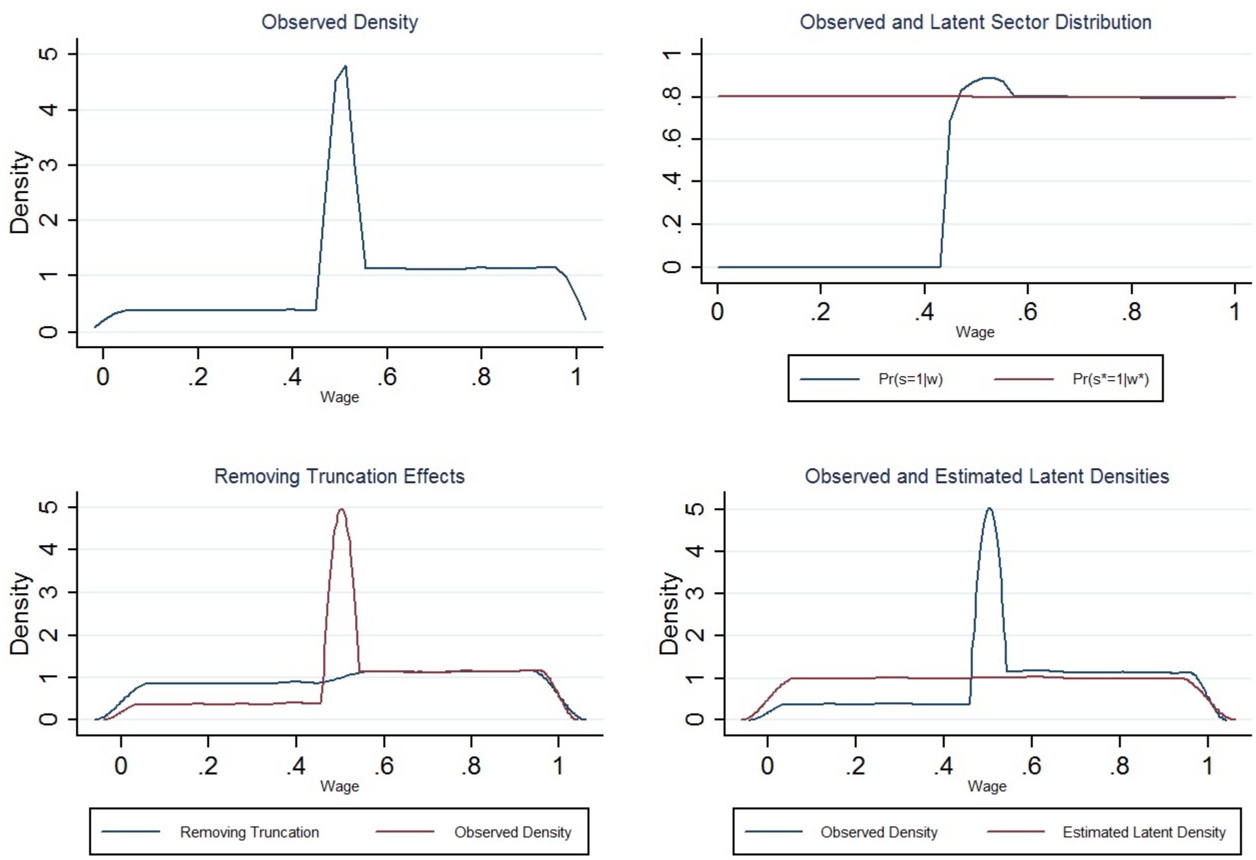
[25] R.H. Meyer and D.A. Wise. Discontinuous distributions and missing persons: The minimum wage and unemployed youth. *Econometrica*, 51, Issue 6:1677–1698, 1983.

[26] R.H. Meyer and D.A. Wise. The effects of the minimum wage on employment and earnings of youth. *Journal of Labor Economics*, 1:66–100, 1983.

[27] David Neumark, Wendy Cunningham, and Lucas Siga. The effects of the minimum wage in brazil on the distribution of family incomes: 1996-2001. Working Papper, 2004.

[28] Joao Saboia. O salario minimo e o seu potencial para a melhoria da distribuicao de renda no brasil. In *Desigualdade de renda no Brasil : uma analise da queda recente*. IPEA, 2006.

Figure 1: Example: Observed, Latent and Estimated Distributions



True and (estimated) parameters: P1f:.20 (.197) , P2f=.50 (.507), P1i=.90 (.902) , P2i=.10(.098) , Uf=.30(.302) , Fm=.5(.5) , rho=.80 (.800)

Figure 2: Nominal Wages and Minimum Wage Evolution

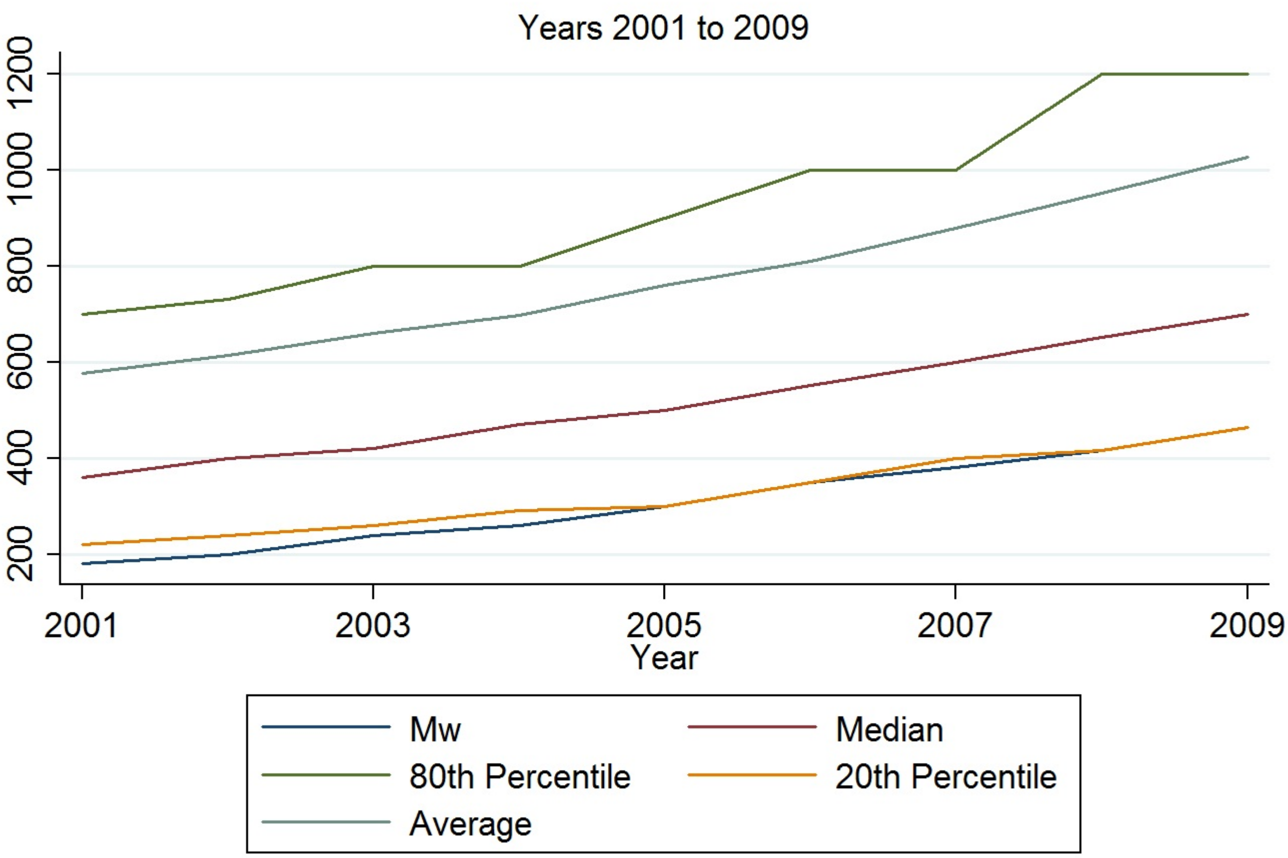
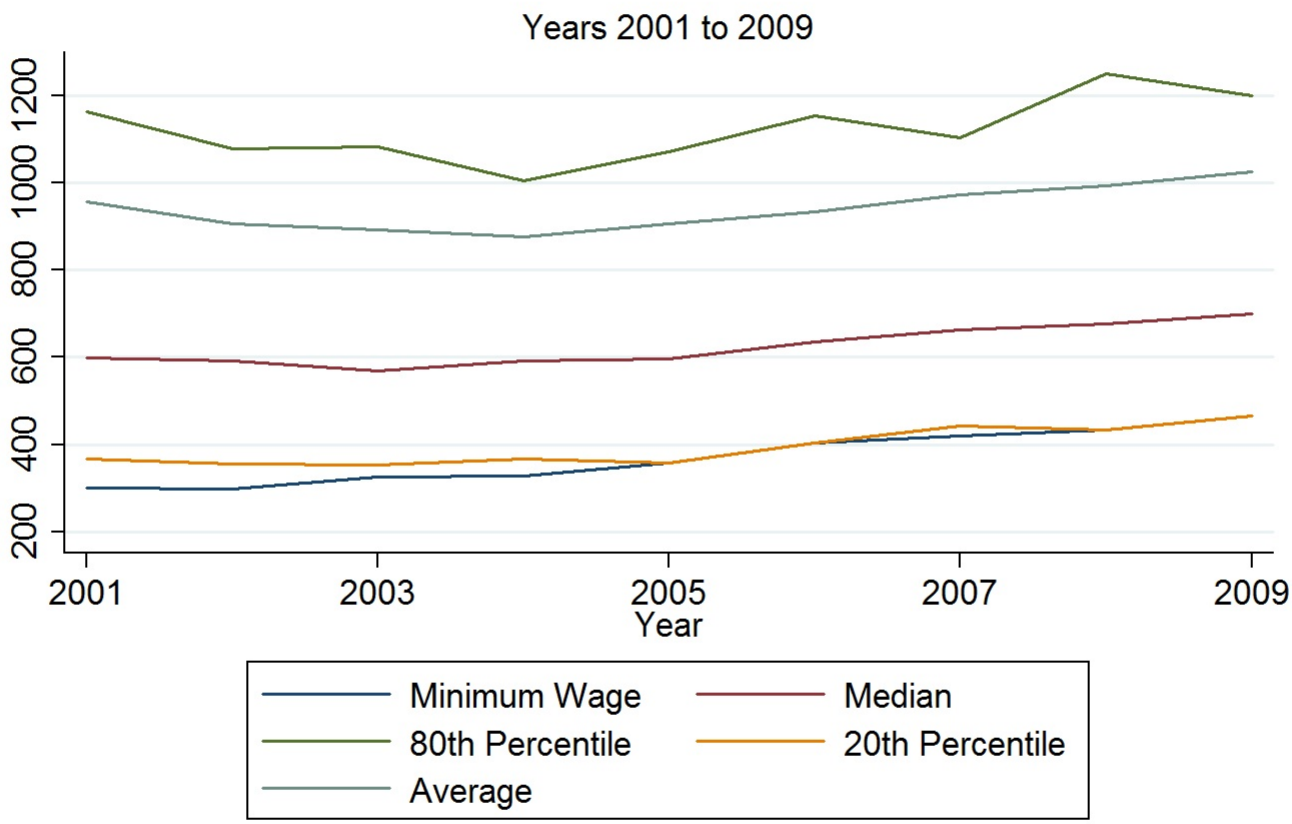


Figure 3: Real Wages and Minimum Wage Evolution



Note: Real Values computed using accumulated inflation as measured by the IPCA. Units in 2009 R$.

Figure 4: Wage Distribution

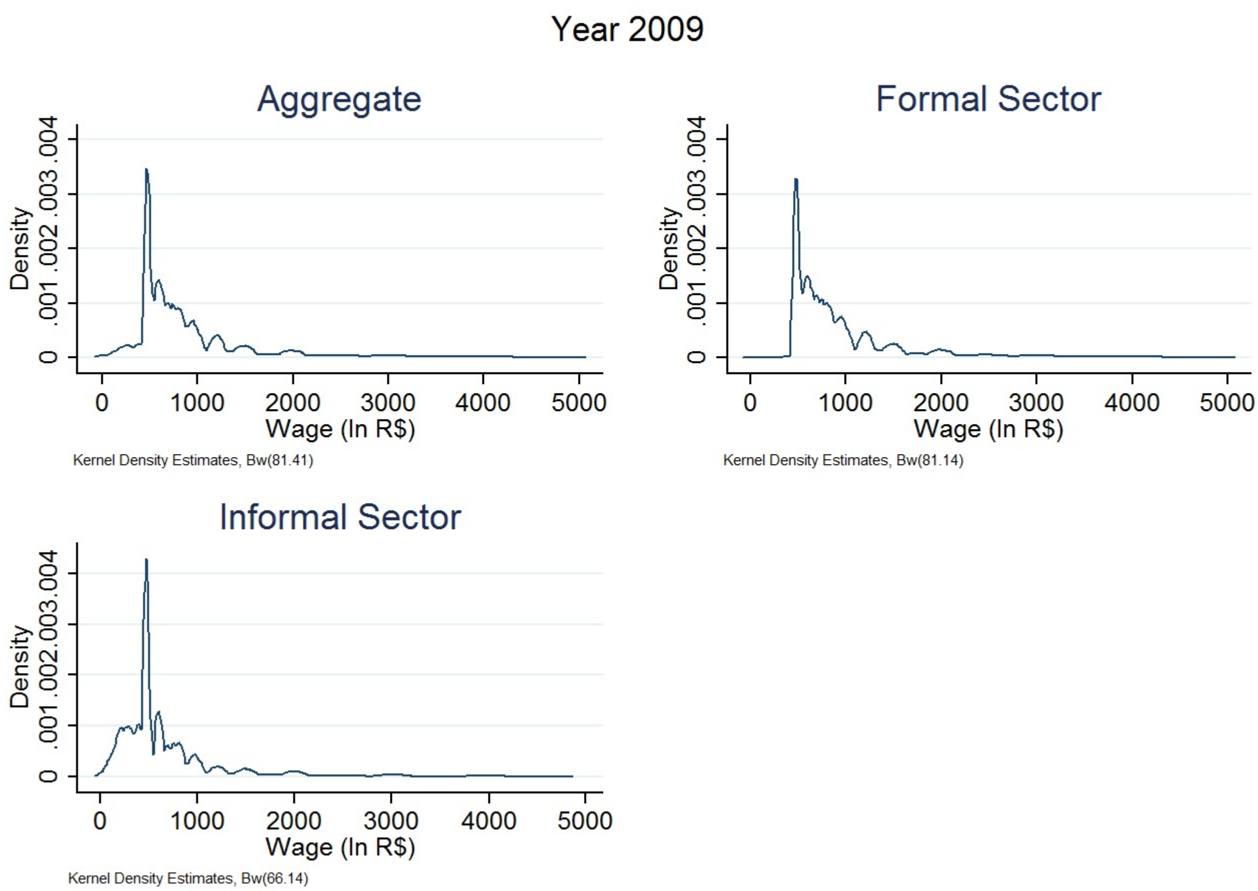


Figure 5: Kernel Density Estimates of the Observed and Latent Wage Distributions

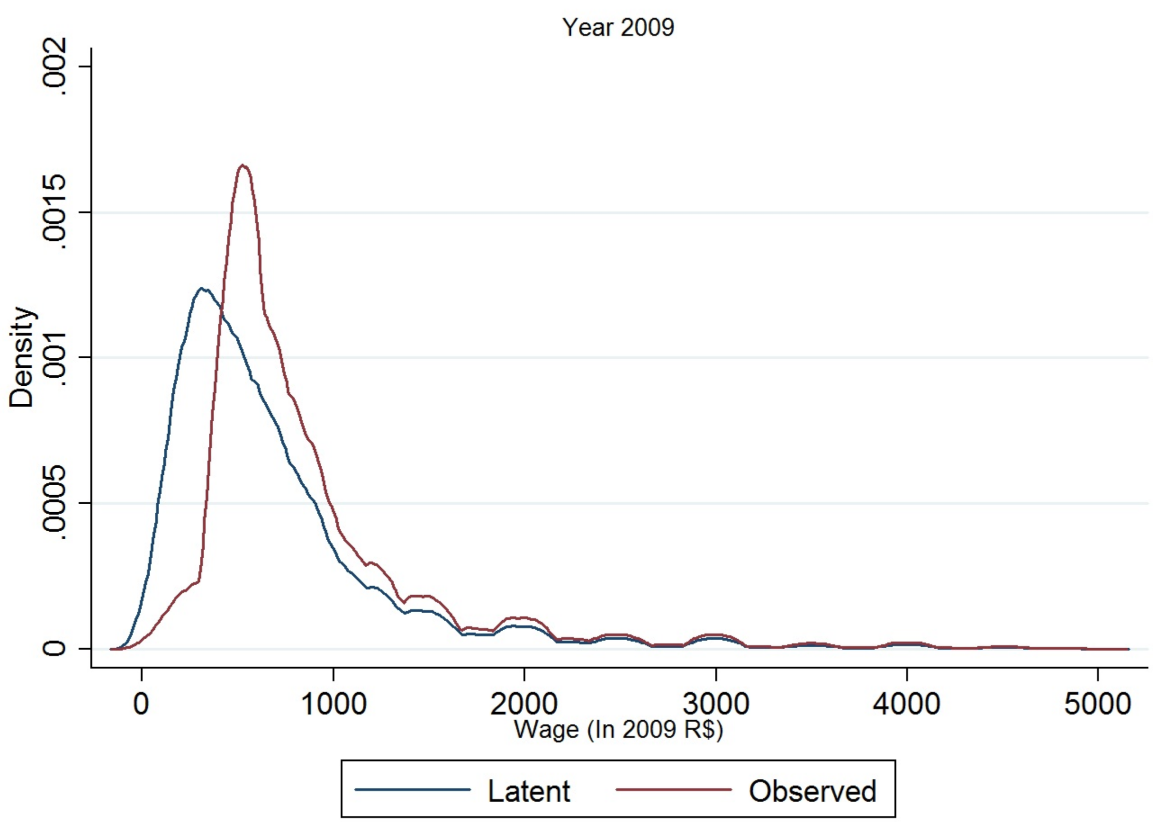


Figure 6: Formality vs. Wages

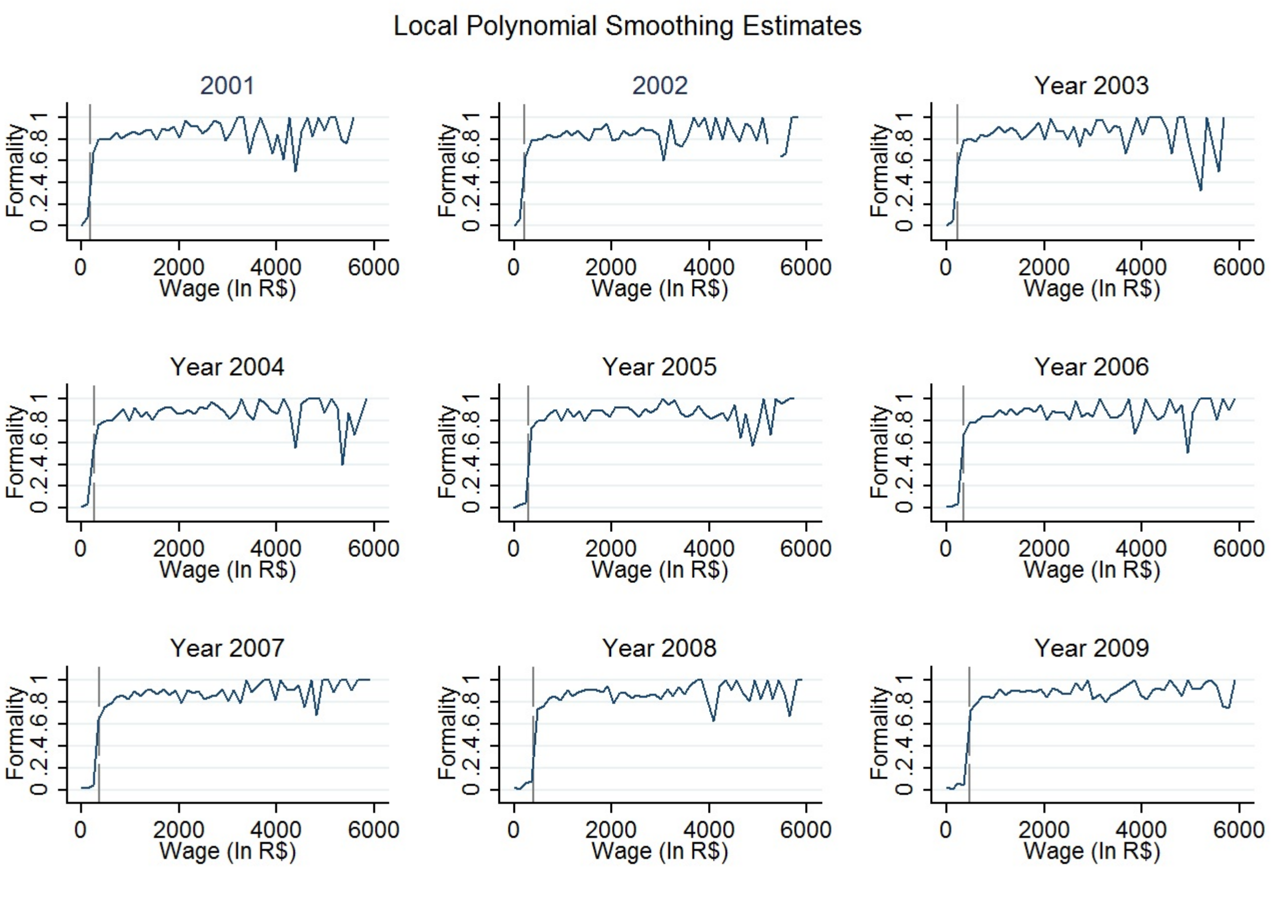
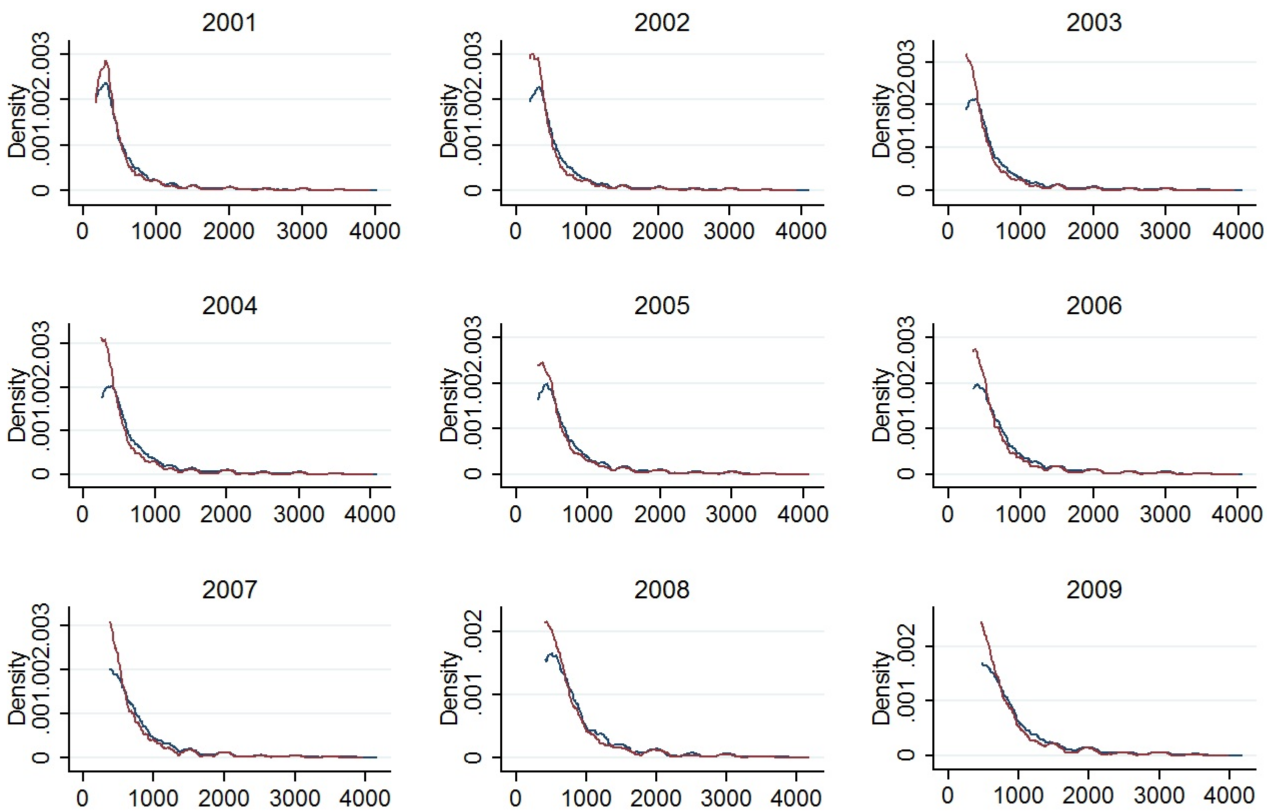


Figure 7: Wage Distribution above the Minimum Wage by Sector



Note: Kernel Density Estimates with boundary correction at the Minimum Wage. Bandwhidth 1.8 times the Silverman's rule of thumb.

Table 1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Wage	579252	746.1369	1083.035	1	350000
Gender	579252	0.351909	0.4775662	0	1
White	579252	0.5115166	0.4998678	0	1
Education	579252	9.802977	3.87046	1	17
Experience	579252	4.12858	5.553721	0	53
Age	579252	32.84479	9.997799	19	59
Formal Sector	579252	0.7252405	0.4463934	0	1
Mw	579252	317.86	92.8839	180	465

Source: PNAD (2001 to 2009 years).

Table 2: Descriptive Statistics by Sector

	Formal Sector	Informal Sector	Difference
Wage	827.63	530.98	296.652***
	(1.613)	(2.694)	(3.164)
Gender	0.36	0.34	0.022***
	(0.001)	(0.001)	(0.001)
White	0.54	0.44	0.097***
	(0.001)	(0.001)	(0.001)
Education	10.21	8.81	1.399***
	(0.006)	(0.010)	(0.011)
Experience	4.52	3.13	1.392***
	(0.009)	(0.014)	(0.016)
Age	33.26	31.85	1.405***
	(0.015)	(0.025)	(0.029)
Minimum Wage Worker	0.12	0.16	-0.040***
	(0.001)	(0.001)	(0.001)
N	420,097	159,155	579,252

Source: PNAD (2001 to 2009). Heteroskedasticity Robust errors in parenthesis.

Table 3: Minimum and Subminimum Wage Conditional Probabilities

	Pr(W=Mw)	Pr(W<Mw)
Unconditional		
	0.1339	0.0798
Conditional on Gender		
Male	0.1162	0.0719
Female	0.1667	0.0942
Conditional on Race		
White	0.0942	0.0510
Nonwhite	0.1519	0.0897
Conditional on Education		
Less than 5 years	0.1988	0.1849
Less than 12 years	0.1627	0.1155
More than 12 years	0.0527	0.0287
Conditional on Region		
South	0.0577	0.0374
Southeast	0.0913	0.0438
Center-West	0.1233	0.0479
North	0.1874	0.1045
Northeast	0.2384	0.1692

Source: From 2001 to 2009 PNADs. N=579252.

Table 4: Model Parameters Estimates by Year

Coef/Se	2001	2002	2003	2004	2005	2006	2007	2008	2009
P_1	0.202***	0.217***	0.206***	0.232***	0.180***	0.157***	0.113***	0.192***	0.121***
	(0.006)	(0.009)	(0.005)	(0.006)	(0.006)	(0.004)	(0.003)	(0.005)	(0.003)
P_2	0.256***	0.356***	0.289***	0.293***	0.349***	0.262***	0.176***	0.304***	0.208***
	(0.007)	(0.015)	(0.007)	(0.008)	(0.011)	(0.006)	(0.004)	(0.007)	(0.005)
F(Mw)	0.253***	0.260***	0.311***	0.291***	0.328***	0.400***	0.446***	0.345***	0.434***
	(0.005)	(0.008)	(0.004)	(0.005)	(0.007)	(0.005)	(0.005)	(0.004)	(0.005)
P_1^f	0.106***	0.222***	0.139***	0.179***	0.165***	0.077***	-0.015***	0.131***	0.023***
	(0.013)	(0.021)	(0.010)	(0.012)	(0.012)	(0.007)	(0.005)	(0.008)	(0.005)
P_2^f	0.191***	0.231***	0.223***	0.226***	0.258***	0.213***	0.152***	0.264***	0.186***
	(0.006)	(0.011)	(0.006)	(0.006)	(0.009)	(0.005)	(0.003)	(0.006)	(0.004)
P_1^i	0.525***	0.199***	0.461***	0.443***	0.244***	0.516***	0.711***	0.494***	0.669***
	(0.018)	(0.038)	(0.017)	(0.019)	(0.026)	(0.013)	(0.008)	(0.016)	(0.011)
P_2^i	0.475***	0.801***	0.539***	0.557***	0.756***	0.484***	0.289***	0.506***	0.331***
	(0.018)	(0.038)	(0.017)	(0.019)	(0.026)	(0.013)	(0.008)	(0.016)	(0.011)
U	0.543***	0.427***	0.506***	0.475***	0.471***	0.581***	0.711***	0.504***	0.670***
	(0.012)	(0.024)	(0.011)	(0.013)	(0.016)	(0.010)	(0.006)	(0.010)	(0.007)
ρ	0.772***	0.781***	0.792***	0.798***	0.816***	0.819***	0.823***	0.832***	0.847***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Sample Size:	71,397	71,051	68,319	68,196	65,755	62,587	58,269	58,241	55,502

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors computed by 100 bootstrap replications.

Table 5: Distributional Effects of the Minimum Wage

	2001	2002	2003	2004	2005	2006	2007	2008	2009
E(lw)									
Observed	5.994***	6.054***	6.152***	6.220***	6.322***	6.400***	6.489***	6.578***	6.658***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Latent	5.793***	5.874***	5.926***	6.020***	6.090***	6.093***	6.097***	6.363***	6.310***
	(0.009)	(0.014)	(0.008)	(0.009)	(0.012)	(0.010)	(0.010)	(0.008)	(0.008)
Mw Effect	0.201***	0.180***	0.225***	0.200***	0.231***	0.307***	0.392***	0.215***	0.348***
	(0.008)	(0.012)	(0.007)	(0.008)	(0.011)	(0.009)	(0.009)	(0.007)	(0.007)
Sd(Lw)									
Observed	0.769***	0.774***	0.753***	0.737***	0.719***	0.699***	0.693***	0.684***	0.661***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Latent	0.916***	0.929***	0.916***	0.885***	0.901***	0.881***	0.897***	0.846***	0.851***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.009)	(0.006)	(0.006)
Mw Effect	-0.147***	-0.155***	-0.163***	-0.149***	-0.182***	-0.182***	-0.204***	-0.162***	-0.190***
	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)	(0.004)	(0.006)	(0.004)	(0.004)
$q^{80}(lw)-q^{20}(lw)$									
Observed	1.157***	1.112***	1.124***	1.015***	1.099***	1.050***	0.916***	1.062***	0.948***
	(0.003)	(0.020)	(0.012)	(0.016)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Latent	1.419***	1.476***	1.267***	1.386***	1.386***	1.386***	1.447***	1.204***	1.204***
	(0.022)	(0.045)	(0.019)	(0.000)	(0.012)	(0.021)	(0.015)	(0.010)	(0.012)
Mw Effect	-0.262***	-0.364***	-0.143***	-0.372***	-0.288***	-0.336***	-0.531***	-0.142***	-0.256***
	(0.022)	(0.052)	(0.022)	(0.016)	(0.012)	(0.021)	(0.015)	(0.010)	(0.012)
Gini									
Observed	0.069***	0.069***	0.065***	0.063***	0.061***	0.058***	0.056***	0.055***	0.052***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Latent	0.087***	0.087***	0.085***	0.081***	0.081***	0.080***	0.080***	0.072***	0.074***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Mw Effect	-0.018***	-0.018***	-0.019***	-0.017***	-0.021***	-0.022***	-0.024***	-0.017***	-0.022***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Sample Size:	71,397	71,051	68,319	68,196	65,755	62,587	58,269	58,241	55,502

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors computed by 100 bootstrap replications.

Table 6: Role of Covariates: Estimates of the parameters by sub-groups

Parameters\Conditional On:	Female	Male	Black	White	Educ<12	Educ>=12
P_1	0.099***	0.137***	0.110***	0.086***	0.158***	0.081***
	(0.004)	(0.004)	(0.006)	(0.003)	(0.005)	(0.003)
P_2	0.174***	0.231***	0.201***	0.166***	0.227***	0.184***
	(0.007)	(0.007)	(0.012)	(0.006)	(0.006)	(0.006)
F(Mw)	0.549***	0.364***	0.532***	0.392***	0.500***	0.388***
	(0.008)	(0.006)	(0.012)	(0.006)	(0.006)	(0.006)
P_1^f	0.008	0.039***	-0.007	-0.017***	0.015*	0.004
	(0.007)	(0.007)	(0.013)	(0.006)	(0.009)	(0.005)
P_2^f	0.155***	0.205***	0.190***	0.144***	0.213***	0.160***
	(0.006)	(0.007)	(0.011)	(0.005)	(0.006)	(0.005)
P_1^i	0.699***	0.638***	0.737***	0.705***	0.721***	0.641***
	(0.017)	(0.015)	(0.026)	(0.015)	(0.011)	(0.018)
P_2^i	0.301***	0.362***	0.263***	0.295***	0.279***	0.359***
	(0.017)	(0.015)	(0.026)	(0.015)	(0.011)	(0.018)
U	0.726***	0.632***	0.689***	0.749***	0.615***	0.735***
	(0.010)	(0.011)	(0.016)	(0.008)	(0.011)	(0.008)
ρ	0.868***	0.837***	0.843***	0.858***	0.798***	0.879***
	(0.003)	(0.002)	(0.007)	(0.002)	(0.003)	(0.002)
Sample Size:	26,030	45,367	6,176	34,530	40,013	31,384

note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors computed by a 100 bootstrap replications. Data from 2009 PNAD.

Table 7: Minimum Wage Effects on Labor Tax Revenues

	2001	2002	2003	2004	2005	2006	2007	2008	2009
R	0.977***	0.966***	0.954***	0.959***	0.929***	0.904***	0.866***	0.903***	0.863***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)
s/ρ	0.912***	0.892***	0.895***	0.891***	0.884***	0.884***	0.899***	0.889***	0.902***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
$E_h(w \mid s = 1)/E_f(w \mid s = 1)$	1.241***	1.219***	1.265***	1.249***	1.243***	1.334***	1.410***	1.229***	1.348***
	(0.008)	(0.010)	(0.007)	(0.007)	(0.010)	(0.008)	(0.010)	(0.007)	(0.015)
U.F(Mw)	0.137***	0.111***	0.157***	0.138***	0.154***	0.233***	0.317***	0.174***	0.291***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)	(0.006)
Sample Size:	71,397	71,051	68,319	68,196	65,755	62,587	58,269	58,241	55,502

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors computed by 100 bootstrap replications.

Table 8: Formality vs. Wages - Linear Regression Estimates

Conditional on:	w>0		w>Mw		w>1.5Mw		w>2Mw	
Year	Coef/Std. Error	N	Coef/Std. Error	N	Coef/Std. Error	N	Coef/Std. Error	N
2001	0.022***	55498	0.009***	47157	0.004***	36477	0.002***	25342
	(0.001)		(0.001)		(0.001)		(0.001)	
2002	0.024***	58327	0.007***	47135	0.002***	34279	0.000	23521
	(0.001)		(0.001)		(0.001)		(0.001)	
2003	0.030***	58268	0.010***	46225	0.004***	33170	0.003***	22785
	(0.001)		(0.001)		(0.001)		(0.001)	
2004	0.024***	62586	0.007***	50127	0.004***	36897	0.001*	23105
	(0.001)		(0.001)		(0.001)		(0.001)	
2005	0.031***	65755	0.005***	50560	0.001*	34693	-0.001	21276
	(0.001)		(0.001)		(0.001)		(0.001)	
2006	0.029***	68196	0.006***	51639	0.001*	32635	0.000	19997
	(0.001)		(0.001)		(0.001)		(0.001)	
2007	0.028***	68319	0.006***	53640	0.001*	34257	0.000	21543
	(0.001)		(0.001)		(0.001)		(0.001)	
2008	0.029***	71051	0.005***	54107	-0.000	34091	-0.001	21184
	(0.001)		(0.001)		(0.001)		(0.001)	
2009	0.012***	71397	0.002***	54918	-0.000	35804	-0.001*	20285
	(0.001)		(0.000)		(0.000)		(0.000)	

Note: *** p<0.01, ** p<0.05, * p<0.1. Coefficient reffers to the elasticity of formality with respect to wages measured at the minimum wage based on a linear probability model.

Table 9: Descriptive Statistics by Sector: The role of Minimum Wage

Parameters\Conditional On:	W>0			W>Mw		
	Formal Sector	Informal Sector	Difference	Formal Sector	Informal Sector	Difference
Wage	827.631	530.979	296.652***	900.693	760.900	139.793***
	(1.613)	(2.694)	(3.164)	(1.820)	(4.980)	(4.429)
Gender	0.358	0.336	0.022***	0.343	0.281	0.062***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
White	0.539	0.442	0.097***	0.565	0.519	0.046***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Education	10.211	8.812	1.399***	10.386	9.406	0.980***
	(0.006)	(0.010)	(0.011)	(0.006)	(0.014)	(0.014)
Experience	4.525	3.133	1.392***	4.730	3.606	1.123***
	(0.009)	(0.014)	(0.016)	(0.010)	(0.018)	(0.021)
Age	33.256	31.851	1.405***	33.539	33.196	0.342***
	(0.015)	(0.025)	(0.029)	(0.016)	(0.035)	(0.037)
N	420,097	159,155	579,252	366,065	89,397	455,462

Source: PNAD (2001 to 2009). Heteroskedasticity Robust errors in parenthesis.

Table 10: Placebo Tests: Discontinuity estimates using minimum wage values of other years

Year	2001	2002	2003	2004	2005	2006	2007	2008
	Coef/se	Coef/se	Coef/se	Coef/se	Coef/se	Coef/se	Coef/se	Coef/se
Mw02	1.060***							
	(0.063)							
Mw03	0.173***	0.647***						
	(0.062)	(0.068)						
Mw04	0.814***	0.490***	1.562***					
	(0.036)	(0.171)	(0.081)					
Mw05	0.611***	0.546***	0.089***	1.100***				
	(0.016)	(0.016)	(0.015)	(0.026)				
Mw06	0.724***	0.673***	0.524***	0.238***	0.585***			
	(0.014)	(0.015)	(0.013)	(0.029)	(0.145)			
Mw07	1.066***	0.821***	0.601***	0.432***	-0.055	1.478***		
	(0.019)	(0.014)	(0.021)	(0.018)	(0.204)	(0.084)		
Mw08	1.671***	1.578***	1.305***	1.233***	1.249***	1.280***	1.963***	
	(0.033)	(0.026)	(0.022)	(0.027)	(0.029)	(0.194)	(0.103)	
Mw09	1.291***	1.283***	1.000***	0.868***	1.053***	0.658***	0.798***	1.009**
	(0.024)	(0.023)	(0.020)	(0.013)	(0.018)	(0.024)	(0.023)	(0.489)
Average	0.926	0.862	0.847	0.774	0.708	1.139	1.380	1.009
Sample Size:	71,397	71,051	68,319	68,196	65,755	62,587	58,269	58,241

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors computed by 100 bootstrap replications.

Table 11: Robustness - Mccary's Density Discontinuity Estimator

Coef/Se	2001	2002	2003	2004	2005	2006	2007	2008	2009
P_1	0.134***	0.163***	0.163***	0.152***	0.137***	0.132***	0.108***	0.156***	0.114***
	(0.008)	(0.004)	(0.004)	(0.005)	(0.004)	(0.003)	(0.002)	(0.007)	(0.002)
P_2	0.171***	0.268***	0.229***	0.191***	0.267***	0.221***	0.169***	0.247***	0.196***
	(0.011)	(0.007)	(0.005)	(0.007)	(0.007)	(0.004)	(0.003)	(0.010)	(0.003)
F(Mw)	0.336***	0.318***	0.363***	0.386***	0.390***	0.442***	0.457***	0.394***	0.448***
	(0.012)	(0.005)	(0.005)	(0.007)	(0.005)	(0.004)	(0.003)	(0.009)	(0.003)
P_1^f	-0.028*	0.098***	0.056***	0.030***	0.073***	0.031***	-0.023***	0.068***	0.011***
	(0.017)	(0.009)	(0.008)	(0.009)	(0.008)	(0.005)	(0.004)	(0.011)	(0.004)
P_2^f	0.127***	0.174***	0.177***	0.148***	0.197***	0.179***	0.146***	0.214***	0.175***
	(0.009)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.009)	(0.003)
P_1^i	0.683***	0.397***	0.573***	0.635***	0.424***	0.592***	0.723***	0.590***	0.688***
	(0.021)	(0.019)	(0.014)	(0.014)	(0.018)	(0.011)	(0.007)	(0.017)	(0.009)
P_2^i	0.317***	0.603***	0.427***	0.365***	0.576***	0.408***	0.277***	0.410***	0.312***
	(0.021)	(0.019)	(0.014)	(0.014)	(0.018)	(0.011)	(0.007)	(0.017)	(0.009)
U	0.695***	0.568***	0.608***	0.657***	0.596***	0.647***	0.723***	0.598***	0.689***
	(0.019)	(0.010)	(0.008)	(0.011)	(0.010)	(0.006)	(0.004)	(0.016)	(0.005)
ρ	0.772***	0.781***	0.792***	0.798***	0.816***	0.819***	0.823***	0.832***	0.847***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Sample Size:	71,397	71,051	68,319	68,196	65,755	62,587	58,269	58,241	55,502

note: *** p<0.01, ** p<0.05, * p<0.1

Table 12: Labor Tax effects under a "No Unemployment" assumption

	2001	2002	2003	2004	2005	2006	2007	2008	2009
r	1.093***	1.077***	1.081***	1.059***	1.057***	1.052***	1.052***	1.037***	1.039***
	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)
s/ρ	0.912***	0.892***	0.895***	0.891***	0.884***	0.884***	0.899***	0.889***	0.902***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
$E_h(w \mid s = 1)/E_g(w \mid s = 1)$	1.198***	1.208***	1.208***	1.189***	1.195***	1.191***	1.170***	1.167***	1.152***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.008)
Sample Size:	71,397	71,051	68,319	68,196	65,755	62,587	58,269	58,241	55,502

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors computed by 100 bootstrap replications.

Appendix 1

Proof of lemmas 1 and 2:

On the following exposition I will assume: (i) the density and its first derivative are different from zero at the minimum wage and non parametrically identified everywhere. [17] (ii) $\rho'(Mw) \neq 0$ and (iii) f(Mw) $\neq 0$.

Given Assumptions 2 and 3, the relationship between the observed density and the latent one can be writen as:

$$h(w) = \begin{cases} \frac{P_1(w)f(w)}{D} \text{ if w < Mw} \\ \int^{Mw} \frac{P_2(w)f(w)}{D} dw \text{ if w = Mw} \\ \frac{f(\lambda(w))\lambda'(w)}{D} \text{ if w > Mw} \end{cases} \tag{3}$$

Where D is a rescaling factor that ensures both densities integrate to one.

To simplify the algebra, it is advisable to first remove the spillover effects, so the wages observed above the minimum wage are equal to the latent ones. So define:

$$\widetilde{w} = \begin{cases} w \text{ if } w \leq Mw \\ \lambda(w) \text{ if } w > Mw \end{cases}$$

I will abuse notation and denote by from now on h(w) the density of observed wages after removing spillovers and I will use w (not $\widetilde{w}$) to denote "spillover free" wages. This distinction is irrelevant in the absence of spillovers. So, given Assumptions 2 and 3, and relying on the transformation that removed any spillover effects, the latent wage density and the observed wage density are related through the following equation:

$$h(w) = \begin{cases} \frac{P_1(w)f(w)}{D} \text{ if w < Mw} \\ \int^{Mw} \frac{P_2(w)f(w)}{D} dw \text{ if w = Mw} \\ \frac{f(w)}{D} \text{ if w > Mw} \end{cases} \tag{4}$$

Given Assumptions 1, 2 and 3, the latent share the formal sector $\rho(w^*)$ is identified using the information above the minimum wage. This can be done by estimating a simple linear probability model, for example. Given Assumptions 2, 3 and 4, we have that:

$$P_1(Mw) = lim_{\epsilon \to 0^+} \frac{h(Mw - \epsilon)}{h(Mw + \epsilon)}$$

Also, looking at the derivative of the wage density, we have that:

$$h'(w) = \begin{cases} \frac{P_1'(w)f(w)}{D} + \frac{P_1(w)f'(w)}{D} \text{ if w < Mw} \\ \frac{f'(w)}{D} \text{ if w > Mw} \end{cases} \tag{5}$$

[17] Identification of the derivative of the wage density can be achieve by strengthening the continuity of the latent wage density to continuity and boundedness up to the third order derivative.

Then, it can be shown that:

$$P_1'(Mw) = lim_{\epsilon \to 0^+} \left(\frac{h'(Mw-\epsilon)}{h'(Mw+\epsilon)} - \frac{h(Mw-\epsilon)}{h(Mw+\epsilon)} \right) \cdot \frac{h'(Mw+\epsilon)}{h(Mw+\epsilon)}$$

Since the RHS of this equation contains only objects of the observed wage distribution, this implies that $P_1'(Mw)$ is identified. Together with the identification of $P_1(Mw)$, the function $P_1(w)$ is identified. Given that the functions $P_1(w)$ and $\rho(w)$ are identified, P_1^f and P_1^i are identified.

$$P_1^i = P_1(Mw) - \frac{\rho(Mw)}{\rho'(Mw)} \cdot P_1'(Mw)$$

$$P_1^f = [P_1(Mw) - (1-\rho(Mw)) \cdot P_1^i] \cdot \rho(Mw)^{-1}$$

Since:

$$P_1(w) = P_1^f \rho(w) + P_1^i (1-\rho(w))$$

Then, the function $P_1(w)$ is identified.

Inverting the relationship between the observed and latent wage densities, we have that:

$$f(w) = \begin{cases} \frac{h(w) \cdot D}{P_1(w)} \text{ if w < Mw} \\ h(w) \cdot D \text{ if w} \geq \text{Mw} \end{cases} \tag{6}$$

Which implies that:

$$D = [\int^{Mw} \frac{h(w)}{P_1(w)} dw + 1 - H(Mw)]^{-1}$$

Since the function $P_1(w)$ is already identified and H(Mw) is just the fraction of workers on the observed wage distribution that earns less or equal then the minimum wage, D is identified. This implies the identification of the entire latent wage distribution, $f(w)$. Using the latent wage density and together with the function $\rho(w)$ allows identification of the latent densities of the formal and informal sectors and finally the remaining parameters P_2^f and U_f.

$$f(w^*|s^*=1) = \frac{Pr(s^*=1|w^*) \cdot f(w^*)}{Pr(s^*=1)} = \frac{\rho(w^*) \cdot f(w^*)}{\int \rho(u) f(u) du}$$

$$f(w^*|s^*=0) = \frac{Pr(s^*=0|w^*) \cdot f(w^*)}{Pr(s^*=0)} = \frac{(1-\rho(w^*)) \cdot f(w^*)}{\int (1-\rho(u)) f(u) du}$$

$$P_2^f = \frac{Pr(w=Mw|s=1)}{1-Pr(w=Mw|s=1)} \cdot \frac{1-F(Mw|s^*=1)}{F(Mw|s^*=1)}$$

$$U^f = 1 - P_1^f - P_f^2$$

q.e.d.

It is important to remark that the identification result remains even if one assume that P_2^f and U_f are non specified functions of the latent wage, as long as P_1^f remains constant. In this scenario, the parameters recovered above are expectations - $E(P^{2f})$ and $E(U^f)$ - over the distribution of workers whose latent wages are below the minimum wage. The assumption of constant probabilities is maintained just to simplify the exposition.

Also, it should be notice that this proof does not need the wage distribution to peak above the Minimum Wage. In fact, one can actually identify the effects of minimum wage regardless of where in the latent wage distribution the minimum wage happens to be set, as long as the density of wages is greater than zero at the minimum wage, P_f^1 and P_i^1 are constants and either one of them is greater than zero. Is important to remark this feature since Doyle (2007) argues that the wage distribution peaks well above the minimum wage as evidence that the shape of the observed distribution of wages is informative of the shape of the latent one, which is not actually needed for identification.

Proof of Corollary 4.3:

Let $\nu(.)$ be a functional of a distribution, such as the gini coefficient. Then the functional treatment effect of the minimum wage (TE), the Sector Specific Treatment Effect of minimum wage on wages (STE_ν) and the effect of minimum wages on the share of the formal sector, ATE_{s,s^*} are identified. The identification of TE_ν, STE_ν and ATE_{s,s^*} follows directly from the identification of the **joint** distribution of observed and latent sector and wages from i.i.d data on (w_i, s_i).

Appendix 2 - Identification of the Restricted Version of the Model

Assumption 12. *Latent wage and sector distributions:*

$$w^* \sim F$$
$$s^* \sim B(\rho)$$
$$s^* \perp w^*$$

Assumption 13. *No spillover:*

$$w^* > Mw \rightarrow w = w^*$$

Given that the economy is governed by the laws described by assumptions 10, 11, 2 and 4 the aggregate wage density will look like this:

$$h(w) = \begin{cases} \frac{P_1 f(w)}{D} \text{ if w} < \text{Mw} \\ \frac{P_2 F(mw)}{D} \text{ if w} = \text{Mw} \\ \frac{f(w)}{D} \text{ if w} > \text{Mw} \end{cases} \tag{7}$$

Where $D \equiv 1 - F(Mw)(1 - P1 - P2)$ so both densities integrate into one. This is exactly the one-sector version of this model, as proposed by Doyle (2007). This means that at least the aggregate parameters P_1, P_2 and U are identified.

$$P_1 = lim_{\epsilon \rightarrow 0} \frac{h(Mw - \epsilon)}{h(Mw + \epsilon)}$$

To identify P_2 one just needs to verify that:

$$P_2 = P_1 \cdot \frac{Pr(w = Mw)}{Pr(w < Mw)}$$

Given P_1, $F(Mw)$ can identified by the following (see Doyle, 2007):

$$F(Mw) = \frac{Pr(w < Mw)}{P_1(1 - H(Mw)) + Pr(w < Mw)}$$

Now, to recover the sector-specific parameters, first ρ needs to estimated. This can be easily done by using the sample proportion truncated above the minimum wage:

$$\rho = N^{-1}(1 - H(Mw))^{-1} \sum_{i=1}^{N} s_i \mathbb{I}\{w_i > Mw\}$$

The relationship between the aggregate data parameters P_1 and P_2 and the sector-specific model parameters can be easily derived as:

$$\begin{aligned}
&P_1 = \rho P_1^f + (1-\rho)P_1^i \\
&P_2 = \rho P_2^f + (1-\rho)P_2^i \\
&U = \rho U^f \\
&P_1^f + P_2^f + U_f = 1 \\
&P_1^i + P_2^i = 1
\end{aligned}$$

Notice that, given the aggregate data parameters and ρ, this is a system of five equations and five unknowns. Unfortunately, the system is rank deficient, so one extra equation needs to be added to back up the sector-specific parameters.

Relying on the identification of ρ , U^f is identified by:

$$U_f = \frac{U}{\rho} = \frac{1 - P_1 - P_2}{\rho}$$

To recover P_2^f , it is necessary to look at the formal sector density:

$$h^f(w) = \begin{cases} 0 \text{ if w} < \text{Mw} \\ \frac{P_2^f F(mw)}{D^f} \text{ if w} = \text{Mw} \\ \frac{f(w)}{D^f} \text{ if w} > \text{Mw} \end{cases} \tag{8}$$

Where $D^f = 1 - F(Mw)(1 - P_f^2)$ is a scaling factor so both densities integrate into one. The key feature of the formal sector that allows for identification of P_2^f is that since the density is zero below the minimum wage, the scaling factor on the denominator is a function of only one unknown parameter (notice that F(Mw) is already identified). Finally, using:

$$Pr(w = Mw|s = 1) = P_2^f F(Mw)/D^f$$

It is possible to show that:

$$P_2^f = \frac{Pr(w = Mw|s = 1)}{1 - Pr(w = Mw|s = 1)} \cdot \frac{1 - F(Mw)}{F(Mw)}$$

The right-hand side of this equation consists only of quantities that are already identified. With an estimate of P_2^f based on the expression above, we can now go back to the system and recover all the other parameters:

$$P_2^i = \frac{P_2 - \rho P_2^f}{1 - \rho}$$

Finally:

$$P_1^i = 1 - P_2^i$$

And:

$$P_1^f = 1 - P_2^f - U/\rho$$

The latent wage density can be recovered in the same way as in the baseline model, that is, by inverting the relationship and using the fact that D and P_1 were already identified:

$$f(w) = \begin{cases} \frac{h(w) \cdot D}{P_1} \text{ if w} < \text{Mw} \\ h(w) \cdot D \text{ if w} \geq \text{Mw} \end{cases} \tag{9}$$

Appendix 3 - Tax Efffects of minimum wage under alternative hypothesis

To give an idea of the importance of the unemployment effects on the matter, I will also compute the effects of minimum wages on tax based on a different model. In this version, I will force the unemployment effects to be equal to zero. By doing so, it is not possible to find a continuous latent wage distribution that generates the data. A discontinuous latent wage distribution can be nevertheless estimated. Formally, the model works as follows. First, I will keep the first and third assumptions. The second assumption will be replaced by the following:

Assumption 14. *No Unemployment Effects*

Under the minimum wage a fraction P^1 of workers will earn the same wage as in the latent wage distribution. The remaining fraction will earn the minimum wage. These fractions can be sector-specific as in the baseline model.

Notice that there is no Assumption 4 (continuity) in this case. Under these assumptions, the observed wage density will relate to the latent density by the following equation:

$$h(w) = \begin{cases} P_1 g(w) \text{ if w} < \text{Mw} \\ (1 - P_1) G(mw) \text{ if w} = \text{Mw} \\ g(w) \text{ if w} > \text{Mw} \end{cases} \tag{10}$$

Where g(w) is the latent wage distribution based on this different set of assumptions. In this case, we only need to estimate P_1. One way to do it is by recognizing that in this case:

$$P_1 = \frac{Pr(w < Mw)}{Pr(w < Mw) + Pr(w = Mw)}$$

Therefore, a consistent estimator can be constructed by plugging in the maximum likelihood estimator of the respective quantities. With an estimate of P_1 the latent wage density can be easily estimated by properly reweighting the observed wage density. Next, the tax effects of minimum wages can be computed under the "No Unemployment" assumption. And it will be given by:

$$r \equiv \frac{t^1}{t^0} = \frac{s}{\rho} \cdot \frac{E_h(w|s=1)}{E_g(w|s=1)}$$

This is exactly the same expression as before without the unemployment component. Importantly, the expected wages under the latent distribution also change, since the estimate of the latent distribution is different under this different set of assumptions. To make this distinction clear, I denote the latent wage density under this model by $g(w)$ and the distribution by G. Table 12 shows the estimates of r for the years of 2001 to 2009. Minimum wages have a sufficiently strong effect on average wages to compensate for the reduction in the formal sector share due to sector transition. This can be seen by positive MW effects on tax revenues based on the no unemployment assumption.