Illness Shocks, Implicit Contracts and Labor Outcomes in Mexico*

Federico H. Gutierrez†

Vanderbilt University

Abstract

Acute morbidity is one of the most frequent productivity shocks in developing countries. Despite the numerous studies on consumption smoothing, little is known about the role of labor contracts in helping workers cope with it. This paper implements a new test for an implicit contract equilibrium against a spot labor market equilibrium to analyze to what extent labor contracts are used as a form of insurance against health shocks. Results consistently indicate that contractual arrangements in Mexico insure wage workers against idiosyncratic productivity fluctuations associated with illness episodes. To validate the methodology, the test is also implemented on a sample of self-employed workers. In this case, it correctly rejects the contract market. Results of the paper are important to better understand how health shocks affect wages, and workers choose occupations in developing countries.

JEL: O120, O170, J410, J240

Keywords: implicit contracts, acute morbidity, illness, health shocks, Mexico, occupational choice.

*I thank Mark Rosenzweig, T. Paul Schultz and Christopher Udry for their advice and support. I also thank Andrea Moro for valuable suggestions. All remaining errors are my own.

†Vanderbilt University, Economics Department, 317 Calhoun Hall, Box 1819 Sta B, 2115 West End Ave 37240, Nashville, TN. Email: federico.gutierrez@vanderbilt.edu

1 Introduction

Acute morbidity is one of the most frequent productivity shocks in developing countries. Despite the numerous studies on consumption smoothing, little is known about the role of labor contracts in helping workers cope with this risk. This paper attempts to fill this gap in the literature. It tests to what extent labor market contracts are used as a form of insurance against health shocks.

Contract theory indicates that one of the reasons for an employer and an employee to have a contract is to efficiently allocate risk. This hypothesis, also know as 'implicit contract theory'(Azariadis [1975], Baily [1974], Holmstrom [1981]), states that, competing for workers, risk-neutral firms are willing to provide insurance to risk-averse workers in the form of labor contracts with salaries that are invariant to short-run productivity fluctuations.[1] In contrast to a spot labor market, in the contract market, wages do not necessarily equate the marginal product, and the individual labor supply tends to fluctuate more in the face of productivity shocks.

In order to explore if workers in Mexico are insured against productivity fluctuations associated with acute morbidity, I implement a new test for an implicit contract equilibrium against a Walrasian or spot market equilibrium that relies on the response of labor outcomes to illness shocks. The test is derived from a simple contract model and the intuition is straightforward. As previously mentioned, in an implicit-contract market, the remuneration of a worker should be invariant *in the short-run* to productivity fluctuations, in this case illness shocks. Nonetheless, *in the long-run*, the frequency the worker gets sick affects his average productivity. Consequently, when an employer hires an employee, it has the incentive to correctly estimate how often the worker will miss days at work to offer him the appropriate compensation. So, a testable identification condition is that in an implicit contract equilibrium, the probability of getting ill, but not illness, is a determinant of the worker's remuneration. On the other hand, the opposite is true in the spot market where wages equal the marginal product at all times. In this case, an illness episode is expected to have a short-run impact on the worker's remuneration, but how often the worker gets sick is irrelevant to explain his current productivity level and hence, his current wage.

With the same argument, the model I present predicts that the interaction of labor contracts and the sexual division of labor within the household should explain part of the gender wage gap, adding

[1] The name 'implicit contract theory' may create confusion. The contract may be implicit or explicit. What is generally implicit is the insurance component that emerges by separating earnings from productivity fluctuations.

an extra testable condition. Women tend to be in charge of house duties, which likely make them miss more days at work than men. If this is the case, contracts are expected to internalize these additional responsibilities of women. In particular, if the division of tasks within the household is such that when a child - or any other member of the family - gets sick is the mother the person in charge of staying at home to take care of him, then the employer, in anticipation to the missing days at work, is expected to offer her a relatively low-salary contract. Once again, this characteristic is specific to contracts that smooth income and is not present in a spot labor market.

In line with the literature, in this paper, a contract means any labor agreement, even if it is not written. The reason is that being written or not is a characteristic that is neither necessary nor sufficent to validate the implicit contract hypothesis. As Malcomson [1999] indicates, there are other reasons to write contracts than the allocation of risk. So, there may be no insurance in formal labor arrangements. On the other hand, the lack of a written contracts is also not evidence that the insurance component is absent in labor arrangements. Thomas and Worrall [1988] theoretically show that contracts that insure workers against productivity fluctuations can be self-enforcing in the sense that both the employer and the employee have no incentives to renege. So, under certain conditions there is no need to write contracts to ensure their fulfillment. As a result, the only way to test for the existence of insurance in contracts is by analyzing labor outcomes, as this paper does.

The implicit contract theory has received a lot of attention from macroeconomists since it explains important regularities in the labor market, particularly wage rigidities. Surprisingly, it has been barely studied by development economists at the micro level. If empirically verified in low- and middle-income countries, the implicit contract theory can significantly improve the understanding of the different mechanisms to smooth consumption in countries characterized by high levels of risk. Precisely, the main contribution of this paper is to test the implicit contract theory for a particular risk that is especially relevant in developing countries, the risk of illness shocks.

Results of this paper indicate that wage workers in Mexico are in labor arrangements that insure them against idiosyncratic productivity shocks associated with acute morbidity. Consistent with the predictions of the contract model, the wage of the average worker does not decrease with illness shocks. But, the probability of getting sick seems to be internalized in contracts and significantly reduces hourly earnings. The response of the individual labor supply to illness shocks is also consistent with the predictions of the contract model. To validate the methodology, I also implement the test on self-employed workers. In this case, the test correctly rejects the implicit contract hypothesis.

This paper lies in the intersection of three literatures: the empirical evidence of implicit contracts, consumption smoothing mechanisms in developing countries, and the impact of health on labor outcomes. Beaudry and DiNardo [1991] test the implicit contract hypothesis in the U.S. using different stages of the business cycle. They argue that in a spot market, only current labor market conditions should explain wages (i.e., unemployment rate) but in an implicit contract equilibrium, past labor conditions should be more relevant (i.e., unemployment rate at the moment of signing the contract and the minimum unemployment rate during the life of the contract). More recently, Ham and Reilly [2002] also use the unemployment rate in the U.S. to test the implicit contract hypothesis. In this paper, I adopt a different and more direct approach. I use illness as a measure of productivity fluctuation and study the behavior of labor outcomes in response to it.

With respect to consumption smoothing mechanisms, there is a growing number of studies that focus on *ex-post* institutions. That is, on how households smooth consumption after an income shock. Some of the papers analyze particular mechanisms such as the sale of investment assets in moments of financial distress (Rosenzweig and Wolpin [1993]), migration and marriage (Rosenzweig and Stark [1989]) or informal credit markets (Udry [1994]), while others focus on the combined effectiveness of all available mechanisms (Townsend [1994], Ethan Ligon and Worrall [2002], among others). On the other hand, *ex-ante* mechanisms - those meant to reduce the volatility of income - have also received some attention. Rosenzweig and Biswanger [1993] analyze how farmers in India choose the composition of productive assets in a way that the variance of profits is reduced at the expense of lowering its expected value. Kochar [1999] studies how workers in rural India shift labor from farm to off-farm employment when they face crop shocks. Notably, the use of labor contracts as a mechanism to smooth income has not been thoroughly studied in developing countries, at least in an economy-wide setting. Closely related, the literature has focused on explaining the determinants of long-term contract workers versus short-term casual workers (Bardhan [1984], Mukherjee and Ray [1995]). But, no attempt has been made to test if insurance against productivity shocks is a predominant characteristic of labor arrangements in low and middle-income countries. This paper contributes to eliminate this deficiency in the literature.

This paper also relates to the literature on the relation between health and labor outcomes

(Schultz and Tansel [1997], Pitt and Rosenzweig [1986], among others). Strauss and Thomas [1998] emphasize that the "absence of a consistent impact [of morbidity] on wages ... may be because the health indicators used ... tend to reflect shorter-term health problems ... but wages tend to adjust relatively slowly"p. 802. Here, not only do I show evidence in favor of this statement, but I also present a clear explanation. I argue that the impact of illness on wages cannot be correctly understood without taking into consideration labor markets institutions such as the presence of implicit contracts. To my knowledge, this is the first study that explains how acute morbidity affects wages in the long-run but not in the short-run.

It is important to mention that illness shocks generate two types of costs. A direct cost derived from medications and doctors' services and an indirect one derived from a decrease in the worker's productivity. This paper only deals with the indirect cost. It studies if employers use contracts to implicitly insure the worker against productivity fluctuations by offering him a stable remuneration. Health insurance, meant to cover part of the direct cost, is not studied here.

The rest of the paper is organized as follows. In section 2, I develop a model comparing the outcomes of an implicit contract market versus a spot market, and derive testable implications with respect to earnings, wages and hours worked. In section 3, I present the data, variable definitions and descriptive statistics.

In section 4, I estimate the impact of illness on the time allocation, including the labor supply. I show that, consistent with the implicit contract hypothesis, the response of hours worked to illness is larger in absolute value for wage workers than for self-employed workers.

In section 5, I present baseline results. In section 6, I deal with occupational choices of workers. I show that, in agreement with the model, the availability of ex-post mechanisms to smooth consumption is negatively associated with workers choosing to be wage earners. In section 7, I analyze how robust results are with respect to migration, firm location, firm size and the potential endogeneity of illness. Finally, in section 8, I conclude.

2 Theoretical framework

In this section, I present a simple model from which I derive the testable implications of the paper.

Assume that workers in the economy may work either in the spot labor market, where wages adjust contemporaneously to the productivity of workers, or in the contract market. In this latter case, firms offer contracts (E_s, h_s) that specify the earnings and the number of hours required to work in each possible event s; where $s = 0$ if the worker is healthy and $s = 1$ if he is ill.

Contracts should be such that firms do not lose money. So, the expected value V of a contract cannot be negative and is given by:

$$V = (1 - q) \ \left[f'_{h0} h_0 - E_0\right] + q \ \left[f'_{h1} h_1 - E_1\right] \geq 0 \tag{1}$$

If the worker is healthy, that occurs with probability $(1 - q)$, the firm gets the productivity of labor f'_{h0} times the number of hours worked by the employee h_0, and it has to pay him E_0. With probability q the worker is sick, and an analogous situation occurs in this event although the productivity of labor, hours worked and earnings may be different in this state. If the market is competitive and firms are free to enter, the equilibrium expected value of the contract should be zero ($V = 0$). Equation (1) assumes no firm heterogeneity. But, in subsequent sections, I will incorporate the size of the establishment in the analysis.

Workers have three sources of income in each state s: earnings E_s, non-labor income y, and possibly some form of insurance b against income fluctuations. So, the budget constraint in each state is given by:

$$C_0 = E_0 + y - qb \quad ; \qquad b \leq \bar{b} \tag{2}$$

$$C_1 = E_1 + y + (1 - q)b \tag{3}$$

The insurance operates through, but it is not restricted to, a network of friends and relatives that the worker can ask for money when needed. Then, the amount qb in (2), which plays the role of a premium, is paid by the worker when he is healthy to other members of the network that are in need. When the worker gets sick and is unable to work normally, he receives $(1 - q)b$ from the network. I assume that the insurance is actuarially fair. Nothing fundamental is based on this assumption, although it is reasonable for the network to be sustainable. More importantly, I assume that b is bounded from above ($b \leq \bar{b}$) at least for some workers, indicating that ex-post mechanisms to smooth consumption are incomplete leaving room for ex-ante ones such as those embedded in labor contracts.

In each state $s \in \{0, 1\}$ the utility of the worker depends on consumption C_s, leisure l_s and illness s, which takes the value 0 if he is healthy and 1 if he is sick.

$$\mathcal{U} = (1-q)\ U(C_0, l_0, 0) + q\ U(C_1, l_1, 1) \tag{4}$$

In a competitive market, in order to attract workers risk neutral firms have the incentive to offer employees the best possible contract. Then, the optimal contract for the average representative worker maximizes utility (4) subject to (1), (2), (3), non-negativity constraints and the time constraint $T = l_s + h_s$, where T is total hours available.

The necessary first order conditions are:

$$U'_{c_s} - \lambda \quad = 0 \quad , \ s = 0,1 \tag{5}$$

$$U'_{l_s} - \lambda f'_{h_s} \geq 0 \quad , (= 0 \text{ if } l_s < T) \ \ , \quad s = 0,1 \tag{6}$$

$$-U'_{c_0} + U'_{c_1} \geq 0 \quad , (= 0 \text{ if } b < \bar{b}) \tag{7}$$

Equation (5) indicates that the optimal contract eliminates any fluctuation in the marginal utility of consumption across states of nature. This is precisely the insurance component in this type of contract. Illness shocks reduce the productivity of workers, but labor contracts efficiently allocate the risk of productivity fluctuations away from workers to firms.

The insurance component in labor contracts and the ex-post mechanisms to smooth consumption (i.e., insurance network) are perfect substitutes. This can be seen in the polar situation where the worker has no binding limit to borrow money from the network ($b < \bar{b}$). In this case, equations (7) and (5) yield the same condition, and the insurance component in labor contracts provides no additional benefit for the worker. However, if ex-post mechanisms are insufficient, as I assume here, there is value in the insurance embedded in labor contracts.

Assuming that the utility function is additively separable in consumption and hours worked (i.e., $U(C_s, h_s, s) = U^c(C_s) + U^l(l_s, s)$), I obtain from equation (5) implication 1.

Implication 1 $\qquad E_0 = E_1$

The optimal contract provides full insurance. Earnings are independent of illness shocks.

Equation (5) indicates that the optimal contract guarantees the worker the same level of consumption in each state (i.e., $(U'_{c_0} = U'_{c_1}) \Rightarrow (C_0 = C_1)$). A sufficient condition for this to occur is that the employer pays the worker the same amount of money in each state ($E_0 = E_1$). However, this is not a necessary condition unless the worker has no other means to smooth consumption (i.e., $\bar{b} = 0$). But, since the employer is a risk-neutral agent, there is no cost in providing full insurance. Thus, it will do so in order to offer the best contract irrespectively of the worker's access to ex-post insurance mechanisms.[2]

With respect to labor supply, equation (6) gives the relation between hours worked in sickness and in health.

$$\frac{U'_{l_1}}{U'_{l_0}} \geq \frac{f'_{h_1}}{f'_{h_0}} \qquad \text{, with equality if } h_1 > 0 \tag{8}$$

I assume that the marginal disutility of work increases with illness (i.e., marginal utility of leisure increases, $U'_l|_{s=1} > U'_l|_{s=0},\ \forall l$). This is a simple way to model the fact that the body uses part of the energy to fight the disease leaving less energy to the rest of the activities. Consistent with Becker [1985], it implies that even if the marginal product remains constant ($f'_{h_1} = f'_{h_0}$), energy intensive activities - work in this case - decrease with illness. If, in addition, the marginal product decreases with illness ($f'_{h_1} < f'_{h_0}$), the response of the labor supply is even larger (if the substitution effect dominates).

In relation to the spot market where wages equal the marginal product, the response of labor supply to illness is larger in the contract market.

Implication 2 $\underbrace{(h_0 - h_1)}_{\text{contract}} \geq \underbrace{(h_0 - h_1)}_{\text{spot}}$

The response of the labor supply to illness is (weakly) larger in the contract market than in the spot market.

The intuition of implication 2 is that, in the contract market, the worker is able to decrease the hours worked in the event of illness without affecting his earnings as shown in implication 1. On the contrary, in the spot market, where the worker receives his productivity in each state of nature, any reduction in the number of hours worked due to illness reduces his earnings. So, the worker has less incentive to reduce the labor supply in the spot market than in the contract market.[3]

[2] We can think that employers do not observe the worker's ex-post mechanisms to smooth consumption. Then, offering a contract that provides full insurance guarantees that it is the optimal one even for those that are well insured by other means.

[3] This can be shown as follows.

The first order condition in the spot market is the standard one derived from the leisure-consumption model

Implication 3 $\frac{E}{h_0} = w = f'_{h_0}(1-q)$; (in contract market)	
In the contract market, but not in the spot market, the wage rate depends negatively on the *probability* of getting sick. On the other hand, only in the spot market illness shocks negatively affect wages.	

For the sake of simplicity, I assume that equation (8) holds with strict inequality. This means that contracts allow workers to stay at home when they get sick. Since earnings in both states are the same $E_0 = E_1$, equation (1) reduces to implication 3.[4]

It is important to emphasize the different predictions for the probability of getting sick and for illness itself. The probability of getting sick (variable q) is a determinant of wages only in the contract market, not in the spot market where wages equal the marginal product at all times ($w_s = f'_{h_s}$). On the other hand, illness shocks decrease wages only in the spot market. In the contract market workers are fully insured.

2.1 Women and the care of others

In relation to men, women have traditionally devoted a larger proportion of their time to housework and childrearing. As Becker [1981] explains, these gender differences in the allocation of time can be explained not only by cultural elements in the society, but also by a comparative advantage of women in these activities. So, in this context it is reasonable to expect that when a child, or any other member of the family gets sick, the mother stays at home to take care of him/her. Evidence

---

($U'_{c_s} w_s \leq U'_{l_s}, \quad s \in \{0,1\}$, with equality if $l_s < T$). When the worker is employed ($l_0 < T$), it implies:

$$\frac{U'_{l_1}}{U'_{l_0}} \geq \frac{U'_{c_1}}{U'_{c_0}} \frac{f'_{h_1}}{f'_{h_0}} \qquad \text{,with equality if } l_1 < T \tag{9}$$

In (9) I use the fact that in the spot market wages equal the marginal product. Condition (9) is more likely to hold with equality than condition (8) because $\frac{U'_{c_1}}{U'_{c_0}} \geq 1$ for workers in the spot market. Then,

$$\underbrace{\frac{U'_{l_1}}{U'_{l_0}}}_{spot} \geq \underbrace{\frac{U'_{l_1}}{U'_{l_0}}}_{contract} \tag{10}$$

Assuming that the utility function in each state is $U(C_s, h_s, s) = U^c(C_s) - a^s exp(l_s) \quad a > 1$, equation (10) entails implication 2.

[4]Nothing fundamental is missing with this assumption. If the optimal contract specifies that a positive number of hours should be worked when an employee is sick, the condition $\frac{\partial w}{\partial q} < 0$ still holds, which is the relevant aspect of implication 3 for the empirical part of the paper.

of this statement will be shown elsewhere in this study. In this section, I will take this claim as an assumption.

Assumption 1: *If a child gets sick, the mother but not the father will stay at home to take care of him/her.*

If assumption 1 holds, the interaction of labor contracts and the sexual division of labor in the household is expected to increase the gender gap. Implication 4 summarizes this claim.

Implication 4 $\frac{E}{h_0} = w_{women} = f'_{h_0}(1-q)^{n+1}$;	(in contract market)
In the contract market, but not in the spot market, if women miss more days at work than men to take care of others, then the gender wage gap should increase.	

Implication 4 is an extension of implication 3 accounting for the fact (for now, only an assumption) that women stay at home when other members of the family get sick. The variable q is the probability of getting sick. Until now, I have only considered the illness of the worker. Nonetheless, the concept can be extended to include illness of other members in the family. For example, if the probabilities of getting sick are independent (it can be easily generalized when they are not), a woman will attend work only if she and the 'n' members of the family she takes care of do not get sick. This event occurs with probability $(1-q)^{n+1}$. Women usually act as caretakers of family members when these fall sick. This may induce female workers to miss additional days at work. Employers in the contract market, taking this into account, are expected to offer women a contract with a lower wage rate in comparison to men.

Implication 4 links the interaction of labor contracts and the sexual division of tasks in the family with the gender wage gap. This relation is unique to this type of contract. In contrast, in a spot market the variable q is not a determinant of wages. So, the gender wage gap should not be affected by the different frequencies that women and men miss days at work.

2.2 Occupational Choice

I conclude the model with workers selecting either the contract market or the spot market. In order to do this, I introduce heterogeneity in health. I assume that each worker has a probability q_i of getting sick. The employer does not observe q_i directly. Instead, the employer estimates it based on observable characteristics of the worker ($\bar{q_i} = E[q_i|x_i]$), and offer a contract based on this information (i.e., a contract that solves problem (1)-(4) for $q = \bar{q}_i$). I assume that firms can offer only one contract per worker.

On the other hand, the worker compares the maximum utility he would obtain in the contract market and the maximum utility he would obtain in the spot market. Proposition 1 states that relatively healthy workers tend to choose the spot market and workers who often become sick tend to prefer the contract market.

Proposition 1	$\exists\, q^*$ s.t. if $q_i \geq q^*$	$\Rightarrow U(contract) \geq U(spot)$
	if $q_i < q^*$	$\Rightarrow U(contract) < U(spot)$

Proof: If $q_i = 0$ then $U(contract) \leq U(spot)$. If $q_i = \bar{q}$ then $U(contract) \geq U(spot)$. Because $G(q_i, .) = U(contract) - U(spot)$ is continuous and monotonic in q_i, by the intermediate value theorem $\exists\, q^* \in [0, \bar{q}]$ $s.t$ $G(q_i) = 0$. Q.E.D.

In words, in the spot market the worker maximizes the utility in each state. If q_i is zero, then only one state exists and the utility of the spot market is maximum. If the worker has a probability of getting sick equal to the average for a group with his observable characteristics, then the contract maximizes problem (1)-(4) and so maximizes the utility of the worker. Because the difference in utilities is continuous in q_i, then there exists a value that makes the worker indifferent between choosing the spot market or the contract market.

Since there is adverse selection in the contract market, employers compute the expected value of q_i conditional on the pool of workers that remain in the market. This implies a lower expected value of q_i, which exacerbates the adverse selection problem. However, there is no possibility that the contract market disappears. To see this, let's assume that only the spot labor market exists. Firms, knowing the distribution of q_i in the population, have the incentive to offer a contract that solves problem (1)-(4) for $q = q_{min}$ and attract at least one worker from the spot market.

On the other hand, the assumption about employers not observing the probability of getting sick is necessary for some workers to choose the spot market. Otherwise, the employer would offer a contract customized to each worker such that his utility is maximum. However, if the probability of getting sick is private information, as I assume here, other variables influence the occupational choice. For example, the availability of other means to smooth consumption affects the threshold q^* in proposition 1. Thus, two workers with the same probability of getting sick may take different decisions, one choosing the spot market and the other one the contract market.

Understanding the occupational selection process is important for two reasons. Firstly, when testing implications 3 and 4, I will empirically analyze wage determinants for groups of workers with different occupations. This may introduce econometric problems if selection is not considered. Secondly and most importantly, workers are expected to choose their occupation based on how much they value the insurance component in labor contracts, which gives additional testable implications.

Formally the occupational decision is based on the difference between the indirect utility functions of the worker choosing either the contract market or the spot market.

$$G = \mathcal{V}^{\text{cont}}(q_i, \bar{q}_i, f'_{h0}, f'_{h1}, y, \bar{b}) - \mathcal{V}^{\text{spot}}(q_i, f'_{h0}, f'_{h1}, y, \bar{b}) \tag{11}$$

The worker will choose the contract market if $G \geq 0$, and the spot market otherwise. In addition to the individual probability of getting sick, the decision will depend on the marginal productivities when healthy and sick, non-labor income y, the probability of getting sick of other workers that are observationally similar to him ($\bar{q}_i$), and the availability of ex-post mechanism to smooth consumption ($\bar{b}$). Of all these elements, only the productivities and $\bar{q}_i$ (in the contract market) affect wages. The individual probability of getting sick, although affecting the average productivity of the worker, is not included in the contract since it is not directly observed by the employer.

In the following sections, I test implications 1 through 4 with special emphasis on the last two. They provide a test for the existence of insurance in labor contracts (i.e., implicit contract equilibrium against a spot market one) and the connection between labor contracts and the gender wage gap. Although reasonable, *Assumption 1* should not be taken for granted. I will provide evidence of it before testing implications 1 to 4.

3 Data and descriptive statistics

The source of data is the Mexican Family Life Survey (MxFLS)(Rubalcava and Teruel [2007a,b]). It is a longitudinal multi-thematic survey representative of the Mexican population. So far, there are two rounds available. The first one took place in 2002 and the second one in 2005 ending in 2006.

The sample size for the first round consists of 8,440 households and 35,679 individuals distributed in 150 urban and rural localities throughout Mexico. The second round tracks first round individuals regardless of change of residence, household division, or household formation. The re-contact rate is over 90%.

Table 1 shows the most important variables used in the analysis together with descriptive statistics. The measure of illness for adults is an indicator that takes the value one if the person stopped doing any of his daily activities or work due to illness during the four weeks previous to the interview. For children, it is similar. It takes the value one if the child was at least one day inactive because of illness for the same interval of time.

There are some concerns about the interpretation of this measure (see Strauss and Thomas [1998, 2008] for a discussion). First, it is not a perfect measure of health because it may be associated with the type of activity and the opportunity cost of missing days at work. So, self-employed and wage workers are expected to have different incentives to miss days at work. In particular, it is possible that self-employed workers do not report illness, according to the definition I use, in situations where wage workers generally do. If this is the case, the consequence for the empirical analysis is that the response of labor supply to illness is upwardly biased (in absolute terms) for self-employed workers in relation to wage workers. This goes in opposite direction to implication 2. Hence, if implication 2 is verified despite this bias, the evidence in favor of an implicit contract is even stronger. Nonetheless, an instrumental variable approach will be used to eliminate the labor supply component of the illness definition.

A second concern about this measure of illness is that it is subjective. So, what one person considers illness from a clinical point of view, another person may not. In table 2, we see that the prevalence of illness is significantly higher for women than for men. It could be that they tend to get sick more often but, it could also be that women are more sensitive to symptoms. This is not a concern as long as what women consider illness affects their labor supply. Employers only care if the worker misses days at work or not. Even more importantly, the subjective component in this definition of illness does not affect the differential impact of illness on other family members. For example, it does not affect if a child illness affects differently the labor supply of the mother and the father.

The first panel of table 2 shows the prevalence of illness among men, women and children. On average, 5% of men are sick in a given point in time. Women report being sick more often. Averaging the two years, close to 10% of them report this condition. The lower panel of table 2 shows the distribution of adults among types of occupation. Conditional on being in the labor force, most of the workers are wage workers, 69% of male workers and 63% of female worker in 2002.

Table 3 shows the fraction of workers in agricultural activities, the average hours worked in the month of the interview and the geographical distribution (in terms of location size) of the workers.

4 Earnings and labor supply

In this section, I empirically analyze the response of earnings and hours worked to illness shocks. According to the model presented in section 2, for workers either in the spot market or in the contract market, the hours worked h_{it} in each period t are given by (12). They depend on the marginal productivity $f_i'(s_{it})$, which depends on whether the worker is ill or healthy, illness itself s_{it}, ex-post insurance $\bar{b}$, and non-labor income y_{it}. For those in the contract market, the labor supply also depends on $\bar{q}_i$, the probability of getting ill estimated by the employer based on observable characteristics of workers.

$$h_{it} = h(f_i'(s_{it}), s_{it}, q_i, \bar{q}_i, \bar{b}_i, y_{it}) \tag{12}$$

To be more realistic I extend the model in two dimensions. First, workers live in households with several members where decisions are likely to be coordinated and outcomes of one member may effect others. Hence, hours worked should also depend on the marginal productivity, illness, non-labor income and the rest of the determinants in (12) of all members in the household. Second, I include characteristics of the community X_{jt} that may affect the time allocation of workers. For example, improvements in infrastructure or weather shocks may affect the demand of labor at a locality level.

Considering these extensions, I take a linear approximation and the first difference across time (years 2005 and 2002) of (12) and obtain the following estimating equation:

$$\Delta h_{ij} = \beta_0^j + \beta_1 \Delta s_{ij} + \beta_2 \Delta s_{-ij} + \beta_3 \Delta y_{ij} + \Delta \epsilon_{ij} \tag{13}$$

The change in the labor supply of worker i in community j is a function of changes in his or her illness, changes in the illness status of other members in the family, and changes in non-labor income. Equation (13) is similar to the specification used in Gertler and Gruber [2002]. It differs in the inclusion of the health condition of other members in the household as additional regressors. In (13), taking variables in first difference eliminates the potential endogeneity bias due to preferences and other time invariant characteristics of the worker. It also eliminates the individual probability of getting sick q_i, the probability of getting ill of observationally similar workers $\bar{q}_i$, and the availability of ex-post mechanisms to smooth consumption $\bar{b}$ for all members in the family, also assumed to be time invariant. The coefficient β_1 in (13) captures the direct impact of illness as well as the indirect impact that it exerts through changes in the productivity. The vector β_2 captures the response of labor supply to the illness of other family members. Since the survey provides information about non-labor income and assets at the household level, I assume that resources are pooled in the household and its effect is captured in β_3. Finally, changes in variables at the community level X_{jt}, which likely constitute one of the most important sources of endogeneity (Gertler and Gruber [2002] p. 57), are eliminated by the community fixed effects (β_0^j).

Identification in (13) requires strict exogeneity of regressors (see Wooldridge [2010]). This condition is violated if temporary shocks to the labor supply in the first period (e.g. involuntary unemployment) affect regressors in the second period (e.g., illness status).[5] Nonetheless, this is unlikely in this case because three years separate the first and the second period. A temporary shock to the labor supply in the interview week of 2002 is unlikely to affect the illness condition of the worker when he or she was reinterviewed in 2005.

Implication 2 of the model presented in section 2 states that the response of labor supply to illness is larger for workers in the implicit contract market than for workers in the spot market. In tables 4 and 5, I show the results of estimating equation (13) for wage workers and self-employed workers separately. As previously explained, those in this last group should behave as if they were employed in the spot market in the sense that they cannot separate earnings from productivity shocks (i.e., the insurance component of labor contracts is not possible for them). Consequently, I use these workers as a comparison group.

There may be some concerns with respect to selection bias in estimating (13) for different groups of workers. Nonetheless, taking the first difference eliminates any unobserved time invariant characteristic of the worker that would induce selection bias if the regression were estimated in levels (e.g., health endowment of workers). Furthermore, in subsequent sections I show that self-selection is not an important source of bias.

In column 1 of tables 4 and 5, the specification includes only the worker's own illness. Column 2

[5] Permanent shocks are not a concern since the first difference eliminates them.

adds the health status of his or her children, measured as the fraction of sons and daughters that are ill, and column 3 includes the health condition of the spouse. On average, wage workers decrease their labor supply seven hours per week when they get sick. On the other hand, self-employed workers show no decrease in the labor supply for similar specifications. These results suggest that workers are insured against productivity shocks.

The two groups, self-employed and wage workers, may be involved in different types of tasks. This may create different incentives to stay at home when sick that are confounded with the effect of insurance in contracts. For this reason, in column 4 tables 4 and 5, I include dummy variables for the 18 occupational categories coded in the data.[6] As results show, conclusions are unaffected. The point estimation indicates that wage workers reduce the labor supply in almost six hours more in comparison to self-employers.

Since the distribution of hours worked is left censored at zero, in column 5 I estimate equation (13) using Honoré [1992] method for censored regression models with fixed effects. Results are almost identical to those previously shown.

As previously explained, the way illness is defined may create problems comparing self-employed and wage workers. Illness is reported if it was severe enough for the worker to stop doing any daily activity, including work. In order to eliminate the labor supply component from the variable illness, I report results of estimating (13) instrumenting Δs with ΔFlu in column 6 of tables 4 and 5. Although the flu is not the only illness that may affect the worker's productivity, it is highly correlated with the variable *illness* but measured independently of the worker's occupation, which make it a good instrument to eliminate the labor supply component.[7][8] The estimated coefficients of the IV approach are significantly larger, probably due to the elimination of the measurement error bias. But, consistent with implication 2, the response of labor supply to illness is significantly

[6] The occupations are: professionals; technicians; educators; workers in art, shows and sports; officials and directors in public, private and social sectors; workers in agriculture, forestry and fishing; bosses and supervisors; artisans and workers in production, repair and maintenance; operators of fixed machinery; assistant and laborers in industrial production; drivers of mobile machinery and transport vehicles; department heads, coordinators and supervisors in administrative activities; administrative support staff, merchants and sales representatives; traveling salespeople; workers in the service industry; domestic workers, safety and security personnel.

[7] Schultz and Tansel [1997] also use an instrumental variable approach to eliminate the labor supply component from their acute morbidity measure.

[8] The MxFLS also provides information about symptoms, such as joint pain, difficulty breathing, etc., But, they are not used in the IV approach because F-test in the first stage indicates that they are weak instrument.

larger for wage worker.

The empirical evidence in tables 4 and 5 is consistent with the idea that wage workers are in an implicit contract equilibrium as implication 2 in the model suggests. The larger elasticity of labor supply to illness for workers in the implicit contract market is a consequence that for them, but not for workers in the spot market, earnings are invariant if they miss days at work due to illness. The decrease in productivity should be borne by the employer who acts as insurer. Evidence of this will be presented in subsequent sections.

Now, I estimate the impact of illness on earnings. If employers insure employees, then earnings should be unaffected by illness shocks despite the significant reduction in the labor supply. I use the same specification as before but instead of hours I use earnings as the dependent variable.

$$\Delta E_{ij} = \beta_0^j + \beta_1 \Delta s_{ij} + \beta_2 \Delta s_{-ij} + \beta_3 \Delta y_{ij} + \Delta \epsilon_{ij} \tag{14}$$

Table 6 shows the results of estimating equation (14) for wage workers. As we can see, neither own illness nor illness of other members in the family affects earnings despite the fact that labor supply is significantly reduced as a consequence of these events. This lack of response of earnings to productivity shocks is consistent with the implicit contract hypothesis as implication 1 in section 2 states.

With respect to self-employed workers, since there is no significant impact of illness on the hours worked, earnings should not decrease unless the efficiency of labor declines as a consequence of the illness. In the last two columns of table 6, I pool the sample of wage workers and self-employed workers to compare the two groups. The impact of illness on earnings is negative and larger for self-employed workers. Although, the magnitude of the illness coefficient is economically important, it very imprecisely estimated. The reason for the lack of precision is that the number of observation for self-employed workers who reported all the information needed for this regression is small.

4.1 Time allocation and the division of labor in the family

Assumption 1 in section 2 indicates that only women miss days at work when other members of the family get sick. However, this assumption should not be taken for granted. It is necessary to show some empirical evidence.

Before moving to regression analysis, in figure 1, I show the identity of the person who takes care of each child under 12 years old. Not surprisingly, in the vast majority of the cases it is the mother. This is true even considering only children whose mothers are employed, suggesting the possibility that women stay at home and miss days at work if children require special care, for instance, in the event of illness. In order to investigate this, I estimate the impact of own and other member's illness on the time allocation of men and women separately.

Table 7 shows results for women. When the labor supply equation is estimated in first differences, results indicate that an average mother reduces 2.5 hours worked per week when their children get sick, which represents a 21% reduction. Since the fraction of women working zero hours is not negligible, it is appropriate to address the censoring of hours worked. In columns 4, 5 and 6, the labor supply equation is estimated using Honoré [1992] method. The coefficient of child illness increases in absolute terms to 9 hours per week. Also, when the husband gets sick, female workers also significantly reduce hours worked in market activities.

Table 8 is the analogous of table 7 although for men. In this case, the first difference method indicates that own illness decreases hours worked in more than seven hours per week which represents an average reduction of 17%. Contrary to women, men seems not to reduce the hours worked when another member of the family gets sick. The average drop in male labor supply when children get sick is two hours (and it is not statistically significant), which represents only a 4.8% decline in hours worked in contrasts to the 21% previously mentioned for women. Results from Honoré-Tobit model show slightly higher absolute values. But, when compared to women, the 3.5 hours reduction for men is significantly lower than the 9 hours for women, and the gap is even larger in percentage terms.

Using the fact that primarily women, but not men, miss days at work when other members of the family get sick, in the following section I estimate a wage equation that tests for the presence of insurance in labor agreements, and show how this characteristic of contracts affects the gender wage gap.

5 Test for insurance in labor contracts

Testable implications with respect to the behavior of wages in the contract market and in the spot market can be combined in a single equation. In the spot market, wages are expected to fluctuate with contemporaneous productivity shocks, such as illness. On the other hand, wages in the contract market are expected to be determined by the ex-ante probability of getting sick, as explained in section 2. So, the two hypotheses about the labor market are nested in the following estimating equation:

$$log(w_{ijt}) = log(f'_{h0\ ijt}) + \beta_1\ male_{ij} + \beta_2\ sick_{ijt} + \beta_3\ psick_{ij} + \beta_4\ psick_{ij} * male_{ij} + \delta_t + \mu_{ijt} \quad (15)$$

Expression (15) for the (log)wage of worker i in locality j in period t is obtained by taking the log of the equations contained in implications 3 and 4, and including the contemporaneous illness status of the worker (variable *sick*). The log marginal product (f'_{h0}) is approximated with years of formal education and potential experience. I also include the binary variable *male* to capture standard discrimination, and possibly unobserved productivity differences between men and women. The variable *psick* is a measure of the probability of getting sick (q in implications 3 and 4). It is included in levels and interacted with *male* because employers are expected to offer different contracts to men and women, even when they have the same probability of getting sick. The reason for this is that contracts are expected to internalize the fact that women tend to stay at home when other family members get sick. Finally, δ_t is a time effect, and μ_{ijt} is the error term.

The probability of getting sick is not directly observable. Nonetheless, we expect that employers estimate it based on the characteristics of the worker, conditional on the pool of applicants. For this reason, I compute the probability of getting sick as the prevalence of illness among wage workers in the locality with observationally similar characteristics - age, sex and education. The reason for using only wage workers and not all workers or all members of the community is because workers choose occupation based on their probability of getting sick, so we can expect differences between wage workers and the rest of the members in the locality.[9]

The test for implicit contracts is based on the coefficients of illness, the probability of getting sick, and the probability of getting sick interacted with the gender dummy. If workers are in the spot market, the expected value of the parameters are $\beta_3 = 0$, $\beta_4 = 0$ and $\beta_2 < 0$. That is, wages equal the current marginal product, which decreases when the worker is sick. On the other hand, if workers are in the implicit contract market the expected value of the coefficients are $\beta_3 < 0$, and $\beta_4 > 0$. That is, the higher the probability of getting sick the lower the wage rate. Additionally,

[9] *psick* is computed as the predicted value of regressing *sick* on age, sex, education and a set of locality dummy variables for the sample of wage workers. None of the results of this paper significantly change if *psick* is computed with non-linear models such as logit or probit.

the higher the probability of getting sick the higher the gender gap since women spend on average more time at home taking care of other members.

As a baseline, I estimate equation (15) ignoring workers self-selecting into occupations and potential econometric problems related to migration and firm location, firm heterogeneity and the possible endogeneity of illness. I will address these topics subsequently.

Table 9 shows the results of estimating equation (15) pooling the 2002 and 2005 surveys. Columns 1 and 2 are the standard Mincer equations for wage workers and self-employed workers respectively. As expected, log hourly earnings increases with education, shows an inverted U-shape in potential experience and a positive male wage gap.

In columns 3 and 4, I include the probability of getting sick, the probability of getting sick interacted with a gender dummy, and the current health condition of the worker. Columns 5 and 6 add community characteristics among regressors (population size and the fraction of the locality labor force in agricultural activities). Clustered standard errors are computed allowing for arbitrary correlation of errors at the community level in each year.

Results for wage workers are consistent with the implicit contract hypothesis: higher probability of getting sick implies lower wages. Additionally, the higher the probability of getting sick, the wider the gender wage gap because women miss days at work more frequently in relation to men to take care of other members in the family when sick.

In order to validate the test, I estimate equation (15) for self-employed workers. The test should fail to reject the spot market since these workers cannot separate productivity shocks from earnings. The test does not require the comparison of wage workers and self-employed worker (the test is based on the exclusion of the *psick* from the wage function as explained above.). However, it is an important way of assessing the validity of the methodology.

Results for self-employed workers correctly fail to reject the spot market equilibrium. Neither locality illness nor locality illness interacted with the gender dummy are significantly different from zero. Moreover, when community controls are included (column 6), the current health condition of the worker (i.e., own illness) negatively affects the wage rate as the model predicts for those in the spot market.[10]

[10] In the regression for self-employed workers the prevalence of illness is computed including only self-employed. However, results are robust to including all workers in the locality.

6 Occupational Choice

In this section, I incorporate the worker's decision to become either a self-employed or a wage worker. This is important for two reasons. Firstly, the estimation of the hourly earnings equation ignoring this decision may yield biased results if there are unobservable characteristics of the worker that affect simultaneously the occupational choice and the potential wage rate in each of the occupations. Secondly, the model presented in section 2 gives unambiguous predictions about occupational choice determinants that should be empirically verified. Specifically, those workers with relatively less access to ex-post mechanisms to smooth consumption are expected to be more likely to become wage workers since they value the insurance embedded in labor contracts more.

Section 2.2 gives the determinants of the occupational choice. Equation (11) simply indicates that the worker chooses to be a wage worker if the indirect utility function in this occupation is higher than the indirect utility function of being self-employed. Taking a log linear approximation of (11) I obtain the following latent variable equation.

$$G^* = log(f'_{h0}) + \beta_1\ male + \beta_2\ sick + \beta_3\ psick + \beta_4\ psick * male + \beta_5\ log(y) + \beta_6\ log(\bar{b}) + \eta \quad (16)$$

$$G = \begin{cases} 1 & \text{if} \quad G^* \geq 0 \\ 0 & \text{if} \quad G^* < 0 \end{cases}$$

The marginal product is again approximated with education and potential experience as in (15). The variable *sick* is introduced to capture the difference between the marginal product when the worker is sick and when he is healthy. It does not have a direct effect on the occupational choice. As before, *psick* is estimated probability of getting sick that is used by firms to determine the salary in contracts. The interaction with the gender dummy captures the fact that $\bar{q}_i$ in (11) differs by gender. The individual probability of getting sick q_i is not observed. So, it is contained in the error term.

There are two observable or at least partially observable variables suggested by the model that affect the occupational choice but not the hourly earnings equation. These are non-labor income (y) and the availability of ex-post mechanisms to smooth consumption ($\bar{b}$).

With respect to non-labor income, I only use inheritance for being the one excludable from the hourly earnings equation. Other measures of income are more likely to be influenced by current and past decisions of the worker. The model predicts that workers with higher non-labor income are more likely to become self-employed because they can use it as a safety net in times of meager labor income. So, wealthy workers are more willing to choose a risky activity.[11] Another reason for the association between wealth and occupation, not included in the model but recognized in the literature, is investment indivisibilities. Becoming self-employed usually involves buying machinery or making another type of lump-sum investment that requires the wealth of the worker to be above a certain threshold. Either for self-insurance or for investment indivisibilities motive, a positive income shock such as inheritance is likely to nudge the worker to switch from being an employee to a self-employed worker.

As the model predicts, another factor that increases the probability of becoming self-employed is the availability of ex-post mechanisms to smooth consumption. In developing countries, a very important one is insurance networks. Consumption fluctuations are mitigated with the use of loans and gifts from friends and relatives. Although it is impossible to know how many people the worker can ask for money when needed, the MxFLS survey provides information on whether the worker has a relative in the U.S. We can expect that households connected with residents in the U.S. use these contacts to smooth consumption. I show evidence of this in table 10.

Table 10, column 1, shows the regression of remittances, computed as a binary variable, on illness, on whether any member of the household has a relative in the U.S. and on the interaction of the two. The regression includes time-locality fixed effects to compare workers in the same community in the same year. Results show that workers with relatives in the U.S. are more likely to receive transfers in the event of illness confirming that this is a mechanism to smooth consumption. Columns 2 of the same table shows the same specification with log total expenditure excluding health expenditure as the dependent variable. Consumption does not decrease in the event of illness showing that, on average, workers are insured against this event. Finally, column 3 shows a regression with health expenditure as a dependent variable.

Back to hourly earnings, equations (15) and (16) constitutes a model with endogenous selection. The potential bias from estimating (15) without considering the occupational choice, as I did in the previous section, comes from the possible correlation between the error terms in the two equations ($corr(\eta, \mu) \neq 0$). An element that may induce this correlation is the unobservable part of the worker's productivity not captured by education and potential experience. It is important to emphasize that the unobserved individual probability of getting sick q_i cannot be responsible for

[11] This is true if the utility of the worker exhibits decreasing absolute risk aversion

the correlation between the error terms μ and η. It is a determinant of the occupational choice but it is not in the error term of the hourly earnings equation because this variable is unknown to the employer and cannot be not internalized in labor contracts.

In table 11, columns 1 and 2, I estimate simultaneously equations (15) and (16) for wage workers using a maximum likelihood approach assuming that the error terms are jointly normally distributed. As the model predicts, inheritance and having a relative in the U.S. negatively affect the probability of being a wage worker. With respect to the wage equation, results are almost unaffected by making occupational choice endogenous. This is verified with the likelihood ratio test. This test cannot reject the null hypothesis that the selection equation and the wage equation are independent. Columns 3 and 4 include locality characteristics among regressors with no significant impact on results. In all the cases, inheritance and having a relative in the U.S. are statistically significant and with the expected sign.

In the last two columns of table 11, I relax the parametric assumption of the errors and estimate the model semiparametrically. The method consists of two steps. In the first step, I estimate the selection equation (16) with a flexible functional form.

$$G = \Phi(X\beta) + \gamma_1\Phi^2(X\beta) + \gamma_2\Phi^3(X\beta) + ... + \gamma_{j-1}\Phi^1(X\beta) + u \tag{17}$$

I select the number of terms in equation (17) using generalized cross validation. $\Phi(.)$ is the normal distribution, so if only one term is selected the model reduces to a standard probit model. In this way, the functional form is selected such that a normality test is nested in it.

In the second step I estimate the wage equation (15) using a control function approach that consists of including the single index $X\beta$ of the first stage in a flexible form. In particular I use a polynomial approximation (see Vella [1998])

The last columns of table 11 shows the results of estimating the model semiparametrically. Cross validation indicates that the best fit of (17) is with two terms. The parameter γ_1 is significantly different from zero rejecting normality of unobservables in the selection equation. Nonetheless, the wage equation shows almost identical results compared to the maximum likelihood approach. The control function is statistically significant but it does not impact the estimations.

7 Other results

This section deals with potential identification problems. In particular, it analyzes the potential endogeneity of illness, migration and firm location, and firm heterogeneity.

7.1 Endogeneity of illness

The causality between illness and wages may go in the opposite direction introducing bias to regression (15). That is, very low wages may affect the intake of food debilitating the immune system and increasing illness. Although this is a possibility in low-income countries, it is unlikely in Mexico. Figures 2 and 3 show the density of the body mass index (BMI) for wage workers and self-employed workers respectively. Only a very small fraction of them are undernourished suggesting that reverse causality is probably not important for the population of interest.

Nonetheless, in table 12, I run the wage equation dropping 10% of the sample with lower wages.[12] In case of reverse causality these workers are the most likely to be responsible for it. When people are very poor, an increase in wages is likely to improve nutrition and consequently reduce illness. But, after certain threshold is reached, an increase in wages creates no further improvement on nutrition.

Results in table 12 shows that conclusions are unaltered. Locality illness, and locality illness interacted with gender continue to explain hourly wages for wage workers but not for self-employed workers.

Another potential source of endogeneity is the possibility that illness is correlated with health endowment and other unobservable determinants of the worker's productivity. For example, workers who are currently ill may be less energetic all the time, which introduces correlation between the error term μ and *illness* in (15). Nonetheless, for workers in the implicit contract market these elements are determinants of wages only if they are observable by the employer. In section 2.2, the model indicates that the individual probability of getting sick does not affect wages since it is unknown by the employer.

In order to control for the part of the productivity that is unobserved by the econometrician but not by the employer, I estimate equation (15) using individual fixed effects. With this specification I

[12] Since I truncate the sample based on the dependent variables, instead of OLS I use truncated regression assuming that the error term is normally distributed.

can still identify β_2 (and only β_2). If the unobservable elements in the error term are not significantly correlated with illness, then the estimated value of this coefficient should be similar in the fixed effect and in the OLS specification.

Table 13 shows the result of comparing the estimated impact of illness on wages using OLS and individual fixed effects. The Hausman test fails to reject the hypothesis that they are equal suggesting that endogeneity is not a concern. This is true for both wage workers and self-employed workers, although in this latter case the precision of the estimation with fixed effects is low as a consequence of the reduction in the sample size.

7.2 Migration and firm location

In regression (15) the probability of getting sick, which is computed as the prevalence of illness in the locality among the relevant group of workers, may be correlated with other characteristics in the community affecting wages.

Roback [1982] shows that with perfect mobility of capital and labor, wages in local labor markets are determined by amenities in the community. Some amenities enter the utility function (e.g., good climate conditions) generating immigration, increasing the supply of workers and consequently decreasing wages. On the other hand, other amenities enter the production function (e.g., a river to drain residuals) attracting firms, increasing the capital per worker ratio and hence increasing wages.

A concern in regression (15) is that these amenities may be correlated with the prevalence of illness in the community. For example, in localities where floods occur frequently, the prevalence of illness may be relatively high but outmigration may also be high.

However, if these unobserved locality characteristics have the same impact on all workers in the community, then they cannot explain the difference between wage workers and self-employed workers found in table 9. Nonetheless, it is possible that some amenities affect these two groups of workers differently since they are likely to perform different tasks. For this reason I perform the following robustness check.

Formally, the error term in (15) can be decomposed into S_j that represents characteristics or 'amenities' in locality j and ϵ_{ijt}, a worker i.i.d. shock, ($\mu_{ijt} = S_j + \epsilon_{ijt}$). Following Roback [1982], I assume that amenities are a time-invariant characteristics of the locality. As explained above, S_j may affect wages through migration or firm location and also be correlated with *psick*.

$$log(w_{ijt}) = log(f'_{h0\ ijt}) + \beta_1\ male_{ij} + \beta_2\ sick_{ijt} + \beta_3\ psick_{ij} + \beta_4\ psick_{ij} * male_{ij} + \delta_t + S_j + \epsilon_{ijt} \quad (18)$$

The solution is the estimation of (18) with community fixed effects. It eliminates S_j and the potential endogeneity for omitting this variable. However, the cost of doing this is that β_3 is not identified since *psick* is computed as the average illness of workers in the locality with observationally similar characteristics. But, this methodology still allows the estimation of β_4 which is also a key parameter to identify an implicit contract equilibrium from a spot market one.

In table 14, I show the results of estimating equation (18) including locality fixed effects that eliminate S_j. The interaction of locality illness and the male dummy is positive and statistically significant for wage workers. Moreover, the magnitude of the estimation is similar to the one found in column 5 table 9, the regression that controls for community characteristics. When (18) is estimated using the self-employed sample, the interaction of locality illness and the male dummy is statistically zero. Hence, results confirm that hourly earnings of wage workers behave consistently with the implicit contract hypothesis but not the hourly earnings of self-employed workers.

7.3 Firm size

Large and small firms may be affected differently when an employee gets sick. In large firms, the absence of an employee is expected to have a minor impact on production. Some of the tasks assigned to a worker can be temporarily performed by others within the establishment. But, in small firms, each worker represents an important part of the workforce. His absence may significantly hurt the production process. Thus, large firms are expected to 'penalize' workers less when the probability of getting sick is higher offering contracts with a relatively low salary.

In table 15, I estimate equation (15) including firm size in levels, and interacted with illness and with the probability of getting sick. Wages tend to increase with firm size, which is consistent with the empirical evidence in the literature (Oi and Idson [1999]), and, as expected, with the interaction of firm size and the probability of getting sick.

Another possible effect of firm size is on the risk aversion of employers. The assumption about risk neutral firms is incorrect if entrepreneurs care not only about the level but also about the volatility of profits. This is likely to be true for owners of small firms. They may be less wealthy and may have less access to credit markets than owners of large firms. If this is the case, both employers and employees in small firms share part of the risk. Thus, wages should fluctuate with productivity shocks. In table 15, the interaction of firm size with *psick* is statistically insignificant suggesting that even small firms fully insure workers.

8 Summary and conclusions

Illness episodes are extremely common in low- and middle-income countries. So, it is of crucial importance to understand how workers smooth consumption in the face of productivity fluctuations associated with acute morbidity. This paper focuses on a particular mechanism to smooth consumption that has not been thoroughly studied in developing countries. Using data from Mexico, it explores how labor contracts are used a form of insurance against health shocks.

Clear testable implications are derived from a simple model. In a contract market, the probability of getting ill, but *not* illness episodes, should decrease the worker's remuneration. This is because the employer is expected to fully insure the worker paying him a salary invariant to illness shocks. But, at the moment of offering a contract, the employer has the incentive to estimate how often the worker will miss days at work due to illness to adjust his salary accordingly. This argument is particularly important for female workers since they tend to miss additional days at work to take care of other family members who get sick. On the other hand, in a spot labor market where the worker's wage equates his contemporaneous productivity, illness episodes are expected to decrease the worker's remuneration, but not his probability of getting ill. Consequently, the test for insurance in labor contracts presented in this paper relies on the different responses, in the contract and the spot market, of hourly earnings to illness shocks, and to the probability of getting ill.

Results consistently indicate that wage workers in Mexico are well insured against productivity fluctuations associated with acute morbidity. Illness episodes do not decrease the worker's remuneration. But, the estimated probability of getting sick significantly decreases hourly earnings. At the moment of negotiating the salary, women are penalized more heavily since they usually act as caregivers of other family members. To validate the methodology, the test is also implemented on the sample of self-employed workers. In this case, the test correctly rejects the contract market. Other results of the paper are also consistent with wage workers being in a contract market that insure them against idiosyncratic productivity shocks, such as the response of the individual labor supply to illness shocks, and the importance of ex-post mechanisms to smooth consumption as determinants of the worker's occupational choice.

This paper significantly contributes to the understanding of how workers smooth consumption in developing countries. The literature has not previously tested if contracts that smooth income are predominant in low- and middle income countries. This paper fills this gap.

Another important contribution of this paper is that it sheds light on how illness episodes affect hours worked, and wages. It emphasizes that these relations cannot be correctly understood without considering labor market institutions, such as the presence of 'implicit contracts' in the economy.

References

C. Azariadis. Implicit contracts and underemployment equilibria. *The Journal of Political Economy*, 83(6):pp. 1183–1202, Dec. 1975.

M. N. Baily. Wages and employment under uncertain demand. *The Review of Economic Studies*, 41(1):pp. 37–50, Jan. 1974.

P. Bardhan. *Land, Labor and Rural Poverty*, chapter 4. Columbia University Press, 1984.

P. Beaudry and J. DiNardo. The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data. *The Journal of Political Economy*, 99(4):pp. 665–688, Aug. 1991.

G. S. Becker. *A Teatise on the Family*, chapter 2. Harvard University Press, 1981.

G. S. Becker. Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, 3(1):pp. S33–S58, 1985.

J. P. T. Ethan Ligon and T. Worrall. Informal insurance arrangements with limited commitment: Theory and evidence from village economies. *Review of Economic Studies*, 69(1):pp. 209–244, 2002.

P. Gertler and J. Gruber. Insuring consumption against illness. *The American Economic Review*, 92(1):pp. 51–70, Mar. 2002.

J. C. Ham and K. T. Reilly. Testing intertemporal substitution, implicit contracts, and hours restriction models of the labor market using micro data. *The American Economic Review*, pages pp. 905–927, Sep. 2002.

B. Holmstrom. Contractual models of the labor market. *The American Economic Review. Papers and Proceedings of the Ninety-Third Annual Meeting of the American Economic Association*, 71 (2):pp. 308–313, May 1981.

B. Honoré. Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica*, 60(3):pp. 533–565, May 1992.

A. Kochar. Smoothing consumption by smoothing income: Hours-of-work responses to idiosyncratic agricultural shocks in rural india. *Review of Economics and Statistics*, 81:pp. 50–61, Feb. 1999.

J. M. Malcomson. Individual employment contracts. In *Handbook of Labor Economics*, volume 3, chapter 35, pages 2292–2365. Russell Sage Foundation. Blau and Ehrendberg editors, 1999.

A. Mukherjee and D. Ray. labor tying. *Journal of Development Economics*, 47, 1995.

W. Oi and T. Idson. Firm size and wages. In *Handbook of Labor Economics*, volume 3B, chapter 33. Orley Ashenfelter and David Card, 1999.

M. M. Pitt and M. R. Rosenzweig. Agricultural prices, food consumption, and the health and productivity of indonesian farmers. chapter 5, pages pp. 153–182. in Inderjit Singh, Lyn Squire, and John Strauss, 1986.

J. Roback. Wages, rents, and the quality of life. *The Journal of Political Economy*, 90(6):pp. 1257–1278, Dec. 1982.

M. R. Rosenzweig and H. Biswanger. Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal*, Jan. 1993.

M. R. Rosenzweig and O. Stark. Consumption smoothing, migration and marriage: Evidence from rural india. *The Journal of Political Economy*, 97(4):pp. 905–926, Aug. 1989.

M. R. Rosenzweig and K. Wolpin. Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investment in bullocks in india. *The Journal of Political Economy*, 101(2):pp. 223–244, Apr. 1993.

L. Rubalcava and G. Teruel. *The Mexican Family Life Survey: First Wave*, Feb. 2007a. www.mxfls.uia.mx.

L. Rubalcava and G. Teruel. *Users Guide for The Mexican Family Life Survey: Second Round*, Oct. 2007b. www.mxfls.uia.mx.

T. P. Schultz and A. Tansel. Wage and labor supply effects of illness in cote d'ivoire and ghana: Instrumental variable estimates for days disabled. *Journal of Development Economics*, 53:pp. 251–286, 1997.

J. Strauss and D. Thomas. Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):pp. 766–817, Jun. 1998.

J. Strauss and D. Thomas. Health over the life course. In *Handbook of Development Economics*, volume 4, chapter 54. T. Paul Schultz and John Strauss editors, 2008.

J. Thomas and T. Worrall. Self-enforcing wage contracts. *The Review of Economic Studies*, 55(4): pp. 541–553, Oct. 1988.

R. Townsend. Risk and insurance in village india. *Econometrica*, 62:pp. 539–592, 1994.

C. Udry. Risk and insurance in a rural credit market: An empirical investigation in northern nigeria. *Review of Economic Studies*, 63(3)(no. 208):pp. 495–526, 1994.

F. Vella. Estimating models with sample selection bias: A survey. *The Journal of Human Resources*, 33(1):pp. 127–169, 1998.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*, chapter Ch 10. MIT press, 2010.

A Tables

Table 1:

		2002		2005	
Illness		mean	s.d.	mean	s.d.
Adult illness	In the last 4 weeks, did you stop doing any of your daily activities or work, due to any illness?	0.0867	0.2814	0.0713	0.2574
Children illness	In the last 4 weeks, at least one day the boy/girl was inactive because of any illness	0.0782	0.2684	0.0435	0.2041
Hours worked	Hours worked in market activities during last week	21.954	25.488	19.624	25.162
log earnings					
Wage workers	(log) Earnings last month, all jobs	7.8686	0.8835	8.092	0.8164
Self-employed	(log) Net profits last month, self reported	7.4787	1.2432	7.8346	1.2416
log wages					
Wage workers	(log) [earnings / (4*hours worked)]	2.7769	0.9335	2.9897	0.8727
Self-employed	(log) [net profits/ (4*hours worked)]	2.6006	1.2149	2.9576	1.1989
remittances	Money, aid, donations, or gifts given by any relative or friend who lives in Mexico or abroad	0.0595	0.2366	0.0211	0.1438
relative in the US	Do you have any relative living in the US?	0.352	0.4776	0.3329	0.4713

All statistics computed for 21 to 55 year-olds

Table 2: Descriptive statistics

Illness

2002			2005		
Men	Women	Children	Men	Women	Children
0.055	0.111	0.078	0.052	0.086	0.044
(0.227)	(0.314)	(0.268)	(0.221)	(0.28)	(0.204)

s.d. in parenthesis

distribution by type of occupation (21 to 55 years old)

	2002		2005	
	Men	Women	Men	Women
wage worker	63.92	25.48	63.78	24.73
self-employed	25.68	11.33	21.39	9.43
without compensatior	3.06	3.57	4.68	3.47
non-labor force	7.34	59.63	10.15	62.38

Table 3: Descriptive statistics

fraction in agriculture (21 to 55 years old)

	2002		2005	
	Men	Women	Men	Women
wage worker	0.21	0.05	0.22	0.04
self-employed	0.35	0.03	0.36	0.04

hours worked (21 to 55 years old)

	2002		2005	
	Men	Women	Men	Women
wage worker	45.83	38.96	46.64	39.47
self-employed	44.73	30.30	44.18	31.81

locality size (21 to 55 years old)

	2002	2005
more than 100,000	42.77	41.03
15,000 to 100,000	9.24	10.27
2,500 to 15,000	10.52	12.23
less than 2,500	37.48	36.47

Table 4: Impact of illness on hours worked (zeros included): wage workers

Sampe: wage workers

Dep Var: hours worked

VARIABLES	First difference i	ii	iii	First diff iv	Honore-Tobit v	IV approach vi
own illness	-6.933***	-6.917***	-6.915***	-6.955***	-7.007***	-32.47*
	(1.683)	(1.683)	(1.685)	(1.702)	(1.923)	(17.33)
children illness		-1.586	-1.574			
		(2.170)	(2.172)			
spouse illness			-1.107			
			(1.673)			
no child		-1.414	-1.477			
		(1.388)	(1.398)			
no spouse			0.399			
			(1.396)			
non-labor income	4.760	5.679	6.060	6.649	-1.433	14.64
	(21.10)	(21.12)	(21.18)	(21.20)	(9.053)	(23.74)
assets	1.287*	1.320*	1.316*	1.306*	1.294*	1.652**
	(0.724)	(0.724)	(0.725)	(0.727)	(0.749)	(0.823)
Constant	0.378	0.609	0.535	4.466		2.869
	(0.515)	(0.580)	(0.631)	`		(4.354)
Occupation dummies	No	No	No	Yes	No	Yes
Observations	1,620	1,620	1,620	1615	3,240	1,602
Number of localities	143	143	143	143	143	143
mean hours 2002	44.35	44.35	44.35	44.3	44.35	44.31

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 5: Impact of illness on hours worked (zeros included): self-employed workers

Sampe: self-employed workers

Dep Var: hours worked

	First difference			First diff	Honore-Tobit	IV approach
VARIABLES	i	ii	iii	iv	v	vi
own illness	0.0229	0.172	0.138	-0.933	-0.0196	-9.601
	(3.576)	(3.582)	(3.616)	(3.672)	(3.643)	(20.61)
children illness		-2.741	-3.145			
		(5.427)	(5.460)			
spouse illness			-1.189			
			(4.084)			
no child		4.160	3.913			
		(3.647)	(3.667)			
no spouse			2.809			
			(3.437)			
non-labor income	16.95	17.35	16.49	18.61	17.75**	17.13
	(20.41)	(20.43)	(20.50)	(20.52)	(8.002)	(21.07)
assets	0.713	0.687	0.771	1.089	0.756	0.682
	(2.755)	(2.758)	(2.798)	(2.959)	(0.974)	(3.118)
Constant	1.583	0.828	0.303	9.987		10.82
	(1.142)	(1.304)	(1.456)	(9.300)		(9.575)
Occupation dummies	No	No	No	Yes		Yes
Observations	463	463	463	460	926	457
Number of localities	126	126	126	126	126	126
mean hours 2002	41.63	41.63	41.63	41.53	41.63	41.55

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 6: Impact of illness on earnings

method: first difference, locality fixed effects

Dep Var: log earnings

	Wage workers			Full sample	
VARIABLES	i	ii	iii	iv	v
own illness	-0.0187	-0.0189	-0.0164	-0.303	-0.251
	(0.0699)	(0.0699)	(0.0701)	(0.269)	(0.271)
children illness		-0.109	-0.112		
		(0.0928)	(0.0930)		
spouse illness			-0.0272		
			(0.0709)		
no child		-0.0169	-0.0231		
		(0.0588)	(0.0593)		
no spouse			0.0515		
			(0.0593)		
wage worker				-0.135	-0.0619
				(0.0921)	(0.101)
wage worker * own illness				0.283	0.246
				(0.279)	(0.282)
non-labor income	0.268	0.289	0.340	0.976	0.821
	(0.772)	(0.773)	(0.776)	(0.731)	(0.730)
assets	0.0309	0.0322	0.0323	0.0367	0.0393
	(0.0298)	(0.0298)	(0.0299)	(0.0304)	(0.0306)
Constant	0.289***	0.290***	0.281***	0.422***	0.616***
	(0.0216)	(0.0243)	(0.0265)	(0.0884)	(0.189)
Occupation dummies	No	No	No	No	Yes
Observations	1,194	1,194	1,194	1,283	1,277
R-squared	0.001	0.003	0.003	0.006	0.054
Number of localities	138	138	138	139	139

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Women's labor supply

Female sample

Dep. Var.: hours worked

	First difference			Honore Tobit		
VARIABLES	i	ii	iii	iv	v	vi
own illness	-2.001**	-1.929**	-1.822**	-4.748**	-4.725**	-4.407*
	(0.908)	(0.909)	(0.907)	(2.363)	(2.382)	(2.470)
children illness		-2.504*	-2.582*		-9.058***	-9.615***
		(1.441)	(1.437)		(3.243)	(3.361)
spouse illness			-2.373			-7.928*
			(1.518)			(4.691)
no child		0.174	0.103		-0.298	-0.928
		(0.967)	(0.965)		(2.444)	(2.495)
no spouse			3.043***			6.984***
			(0.756)			(2.136)
non-labor income	17.46	17.25	18.51*	43.47	43.57	40.56
	(10.83)	(10.83)	(10.81)	(28.36)	(27.86)	(28.58)
assets	-0.223	-0.196	-0.203	-0.470	-0.450	-0.469
	(0.280)	(0.280)	(0.280)	(0.915)	(0.792)	(0.799)
Constant	0.833**	0.732*	-0.523			
	(0.358)	(0.398)	(0.505)			
Observations	3,722	3,722	3,722	7,444	7,444	7,444
Number of localities	150	150	150	150	150	150
mean dep var in 2002	11.43	11.43	11.43	11.43	11.43	11.43

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Men's labor supply

Male sample

method: first difference, locality fixed effects

	First difference			Honore Tobit		
VARIABLES	i	ii	iii	iv	v	vi
own illness	-7.471***	-7.463***	-7.429***	-8.045***	-8.090***	-8.030***
	(1.844)	(1.844)	(1.844)	(2.105)	(2.089)	(2.109)
children illness		-1.872	-2.032		-3.525	-3.716
		(2.313)	(2.311)		(2.479)	(2.437)
spouse illness			2.210			2.563*
			(1.530)			(1.522)
no child		-0.706	-0.0286		-0.817	-0.0326
		(1.498)	(1.518)		(1.554)	(1.580)
no spouse			-6.272**			-7.173**
			(2.483)			(2.962)
non-labor income	0.671	0.971	-0.139	4.468	4.942	3.718
	(16.42)	(16.43)	(16.41)	(14.56)	(14.32)	(14.63)
assets	1.468	1.509	1.512	1.513*	1.582*	1.582*
	(1.131)	(1.132)	(1.131)	(0.834)	(0.849)	(0.862)
Constant	-0.312	-0.242	0.0153			
	(0.549)	(0.616)	(0.623)			
Observations	2,233	2,233	2,233	4,466	4,466	4,466
Number of localities	150	150	150	150	150	150
mean dep var in 2002	43.96	43.96	43.96	43.96	43.96	43.96

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 9: Baseline results

Dep Var: log hourly earnings

VARIABLES	wage worker i	self-employed ii	wage worker iii	self-employed iv	wage worker v	self-employed vi
secondary	0.466***	0.554***	0.464***	0.550***	0.376***	0.405***
	(0.0238)	(0.0666)	(0.0333)	(0.0732)	(0.0292)	(0.0649)
college	1.172***	1.280***	1.165***	1.277***	1.040***	1.058***
	(0.0342)	(0.0987)	(0.0452)	(0.116)	(0.0409)	(0.105)
potential experience	0.0412***	0.0292**	0.0411***	0.0302**	0.0405***	0.0287**
	(0.00353)	(0.0120)	(0.00321)	(0.0124)	(0.00322)	(0.0121)
potential experience sq/100	-0.0695***	-0.0414	-0.0691***	-0.0437*	-0.0702***	-0.0403
	(0.00838)	(0.0253)	(0.00770)	(0.0249)	(0.00752)	(0.0246)
male	0.0336*	0.203***	-0.0976*	0.268***	-0.0550	0.326***
	(0.0196)	(0.0575)	(0.0565)	(0.0798)	(0.0537)	(0.0794)
year 2005	0.191***	0.306***	0.192***	0.309***	0.208***	0.282***
	(0.0169)	(0.0547)	(0.0375)	(0.0842)	(0.0320)	(0.0707)
psick			-1.496*	0.644	-1.678**	0.492
			(0.763)	(0.666)	(0.689)	(0.685)
male * psick			2.052***	-0.731	1.698**	-0.501
			(0.783)	(0.642)	(0.726)	(0.657)
own illness			0.000122	-0.184	0.01000	-0.228**
			(0.0426)	(0.117)	(0.0420)	(0.108)
Constant	1.914***	1.722***	2.024***	1.660***	2.209***	1.988***
	(0.0419)	(0.147)	(0.0777)	(0.166)	(0.0736)	(0.166)
locality characteristics	No	No	No	No	Yes	Yes
Observations	8,602	1,869	8,523	1,846	8,245	1,812
R-squared	0.170	0.132	0.170	0.134	0.195	0.175

Clustered standard error in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

locality characteristics include population size and fraction of labor force in agriculture.

Table 10: Relative in the US as mechanism to smooth consumption

method: locality-year fixed effect

VARIABLES	remittances i	log cons ii	health expend. ii
log permanent labor income++	-0.0362***	0.747***	0.182***
	(0.00941)	(0.0371)	(0.0235)
relative in US	0.0278***	0.0997***	0.0487***
	(0.00523)	(0.0208)	(0.0132)
relative in US * hhold head sick	0.0304*	-0.0314	0.00293
	(0.0179)	(0.0711)	(0.0450)
hhold head sick	0.00606	0.0746	0.111***
	(0.0135)	(0.0537)	(0.0340)
log household size	-0.00510	0.184***	0.0399***
	(0.00523)	(0.0208)	(0.0132)
Constant	0.322***	1.965***	-1.199***
	(0.0741)	(0.292)	(0.185)
Observations	7084	7173	7176
Number of year-localities	300	300	300

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

++ computed as the prediction of log labor income on education, experience and gender

log cons excludes health expenditures

Table 11: Endogenous occupational choice

Dep. Var.: log hourly earnings

	ML estimation +		ML estimation ++		Semiparametric +++	
	wage equation+	Selection equation	wage equation++	Selection equation	wage equation	Selection equation
VARIABLES	i	ii	iii	iv	v	vi
secondary	0.472***	0.147***	0.379***	0.110***	0.533***	0.232***
	(0.0234)	(0.0377)	(0.0237)	(0.0391)	(0.039)	(0.067)
college	1.172***	0.156***	1.044***	0.102*	1.255***	0.336***
	(0.0310)	(0.0534)	(0.0316)	(0.0554)	(0.050)	(0.116)
potential experience	0.0402***	-0.0432***	0.0406***	-0.0432***	0.025***	-0.061***
	(0.00372)	(0.00657)	(0.00368)	(0.00658)	(0.007)	(0.013)
potential experience sq/100	-0.0675***	0.0484***	-0.0707***	0.0476***	-0.049***	0.075***
	(0.00865)	(0.0146)	(0.00853)	(0.0147)	(0.011)	(0.026)
male	-0.0780*	-0.0290	-0.0508	-0.0180	-0.069	0.029
	(0.0410)	(0.0673)	(0.0405)	(0.0671)	(0.054)	(0.113)
year 2005	0.191***	0.409***	0.209***	0.415***	0.334***	0.545***
	(0.0217)	(0.0314)	(0.0216)	(0.0314)	(0.069)	(0.071)
psick	-1.372***	-0.842	-1.643***	-0.990	-1.503***	-0.623
	(0.464)	(0.729)	(0.460)	(0.727)	(0.636)	(1.15)
male * psick	1.792***	1.848**	1.641***	1.847**	2.377***	2.431
	(0.558)	(0.880)	(0.551)	(0.874)	(0.796)	(1.785)
own illness	0.00276	-0.150**	0.00687	-0.152**	-0.034	-0.166*
	(0.0428)	(0.0653)	(0.0422)	(0.0653)	(0.054)	(0.096)
inheritance/10,000		-0.0144*		-0.0144*		-0.016
		(0.00851)		(0.00856)		(0.029)
relative in US		-0.0949***		-0.0879***		-0.099***
		(0.0328)		(0.0332)		(0.047)
xb					7.555***	
					(2.689)	
xb2					-5.66***	
					(1.963)	
gamma						-0.047***
						(0.015)
Constant	2.009***	1.365***	2.200***	1.459***	-0.225	1.599***
	(0.0578)	(0.0975)	(0.0580)	(0.100)	(0.896)	(0.189)
locality characteristics	N	N	Y	Y	N	N
Observations	9720	9720	9720	9720	9720	9720

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

+ LR test of indep. eqns. (rho = 0): chi2(1) = 0.16 Prob > chi2 = 0.6868

++ LR test of indep. eqns. (rho = 0): chi2(1) = 0.10 Prob > chi2 = 0.7525

+++ Bootstrap standard errors, 800 replications

Table 12: Truncated regression

method: truncated regression; lower 10% of sample dropped

Dep Var: log hourly earnings

	wage worker	self-employed	wage worker	self-employed
VARIABLES	i	ii	iii	iv
secondary	0.614***	0.576***	0.543***	0.432***
	(0.0362)	(0.0860)	(0.0368)	(0.0813)
college	1.485***	1.320***	1.380***	1.114***
	(0.0449)	(0.115)	(0.0454)	(0.110)
potential experience	0.0525***	0.0352**	0.0515***	0.0363**
	(0.00474)	(0.0147)	(0.00473)	(0.0147)
potential experience sc	-0.0804***	-0.0570*	-0.0795***	-0.0588*
	(0.0117)	(0.0320)	(0.0116)	(0.0315)
male	-0.165***	0.175	-0.143**	0.240**
	(0.0584)	(0.118)	(0.0612)	(0.115)
year 2005	0.174***	0.272***	0.187***	0.244***
	(0.0250)	(0.0709)	(0.0248)	(0.0677)
psick	-2.127***	0.904	-2.630***	0.877
	(0.689)	(0.887)	(0.735)	(0.864)
male * psick	3.540***	-0.960	3.348***	-0.821
	(0.837)	(1.028)	(0.881)	(1.014)
own illness	0.0516	-0.161	0.0709	-0.209
	(0.0602)	(0.163)	(0.0608)	(0.135)
Constant	1.648***	1.661***	1.838***	1.944***
	(0.0823)	(0.204)	(0.0840)	(0.206)
locality characteristics	N	N	Y	Y
Observations	7,673	1,666	7,415	1,634

robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

locality characteristics include size and fraction of labor force in agriculture.

Table 13: Hausman test

wage workers

	FE	OLS	Difference	Hausman Test	
own illness	-0.085037	0.0001223	-0.08516	chi2(1)	1.6
s.e.	0.0788764	0.0409703	0.0674012	Prob>chi2 =	0.2064

self employed

	FE	OLS	Difference	Hausman Test	
own illness	0.1705317	-0.184095	0.3546267	chi2(1)	0.98
s.e.	0.374254	0.1068359	0.3586812	Prob>chi2 =	0.3228

Table 14: Locality fixed effect

method: locality fixed effect

Dep Var: log hourly earnings

	wage worker	self-employed
VARIABLES	i	ii
secondary	0.357***	0.261***
	(0.0314)	(0.0733)
college	1.006***	0.896***
	(0.0468)	(0.115)
potential experience	0.0401***	0.0219*
	(0.00317)	(0.0126)
potential experience sq/100	-0.0715***	-0.0322
	(0.00725)	(0.0261)
male	0.0237	0.225***
	(0.0334)	(0.0545)
year 2005	0.219***	0.349***
	(0.0230)	(0.0671)
male * psick	1.507**	0.223
	(0.659)	(0.625)
own illness	0.000309	-0.170*
	(0.0428)	(0.0954)
Constant	1.977***	1.970***
	(0.0465)	(0.147)
Observations	8523	1846
R-squared	0.131	0.091
Number of localities	150	150

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 15: Impact of firm size

Dep Var: log hourly earnings

	wage worker	
VARIABLES	i	ii
secondary	0.361***	0.448***
	(0.0304)	(0.0337)
college	1.012***	1.131***
	(0.0423)	(0.0452)
potential experience	0.0412***	0.0407***
	(0.00333)	(0.00338)
potential experience sq/100	-0.0720***	-0.0682***
	(0.00771)	(0.00792)
male	-0.0667	-0.0988*
	(0.0553)	(0.0590)
year 2005	0.218***	0.201***
	(0.0328)	(0.0380)
psick	-2.421***	-1.957**
	(0.794)	(0.921)
male * psick	1.668*	1.395
	(0.875)	(0.969)
own illness	-0.00624	0.00602
	(0.0785)	(0.0782)
log firm size	0.0210*	0.0285**
	(0.0116)	(0.0131)
log firm size * psick	0.293*	0.205
	(0.159)	(0.180)
log firm size * male * psick	0.0577	0.264
	(0.210)	(0.226)
log firm size * own illness	0.0108	0.00639
	(0.0327)	(0.0324)
Constant	2.168***	1.968***
	(0.0785)	(0.0866)
locality characteristics	Y	N
Observations	7782	7782
R-squared	0.202	0.180

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

locality characteristics include size and fraction of labor force in agriculture.

B Figures

Figure 1: Person taking care of the child

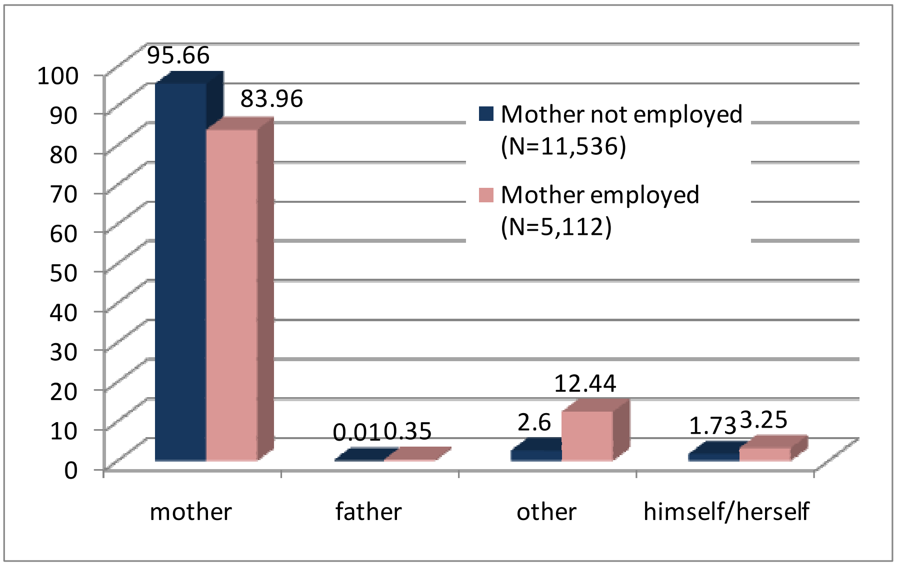

Figure 2: BMI wage workers

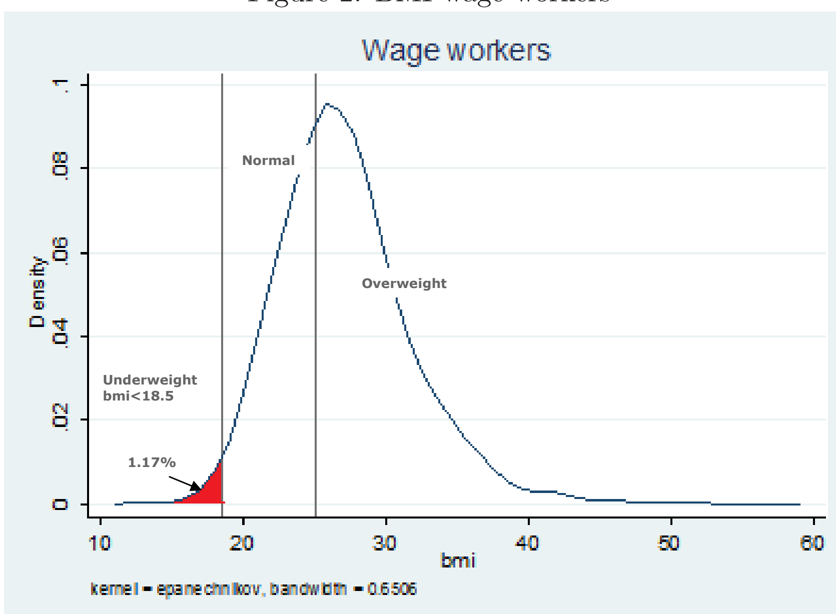

Figure 3: BMI self-employed workers