Title
Essays on Labor Supply and Family Risk Sharing, & Understanding Corruption in Developing Countries

Author
Beylis, Guillermo Raul

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Labor Supply and Family Risk Sharing, & Understanding Corruption in Developing Countries

A dissertation submitted in partial satisfaction of
the requirements for the degree
Doctor of Philosophy in Economics

by

Guillermo Raul Beylis

2013

Abstract of the Dissertation

Essays on Labor Supply and Family Risk Sharing, & Corruption in Developing Countries

by

Guillermo Raul Beylis

Doctor of Philosophy in Economics

University of California, Los Angeles, 2013

Professor Maurizio Mazzocco, Chair

The added worker effect refers to the notion that married women increase their labor supply in response to an unemployment event by their husband. Previous literature has focused on developed economies and has generally found small or insignificant effects.

In chapter 1, I provide reduced form evidence of a large and significant added worker effect in the context of a developing country: Mexico. Wives whose husband lost their job are 11 percentage points more likely to enter the labor force, an entry rate that is 60% higher than wives whose husbands did not lose their jobs. Furthermore, the evidence suggests that entry is transitory as over 70% of wives exit the labor force within a year. Additionally, descriptive statistics suggest that wives who enter transitorily tend to choose low-skill, low-entry cost occupations such as domestic employees, street vending, etc... In this chapter I also develop and estimate a structural model of household time allocation decisions that captures the key determinants of household labor supply. The estimated model is then used to perform two counterfactual policies; one policy offers unemployment insurance (UI) at varying replacement rates, and highlights the important crowd out effects of UI on the added worker effect. The second policy studies the effects of shutting down access to low entry cost jobs for families with different levels of savings, and highlights the heterogeneous value of this sector for different families.

Chapter 2 provides a framework for understanding the decision of politicians to engage in corrupt activities. We developed and estimate a structural model of a politician's decisions to provide public goods and engage in corruption over the span of his political career. While our model is general, we develop and estimate it in the context of municipal governments in Brazil. Overall, we make two main contributions to the existing literature. First, the proposed framework captures many of the various mechanisms by which politicians choose to engage in corruption, which enables us to assess empirically the relative importance of these different mechanisms. This lies in contrast to the previous literature whose empirical evidence has been mostly based on simple correlations. As a second contribution, we use the estimated model to evaluate the effectiveness of anti-corruption policies that increase politicians' wages, induce a higher probability of being audited by a central authority, or increase term limits.

The dissertation of Guillermo Raul Beylis is approved.

Maria Casanova

Hugo Andres Hopenhayn

Sarah Reber

Maurizio Mazzocco, Committee Chair

University of California, Los Angeles

2013

Contents

ACKNOWLEDGEMENTS

I am indebted to my committee chair, Maurizio Mazzocco, for his invaluable advice, continuous support, and constant encouragement. I am grateful to my committee members Maria Casanova, Hugo Hopenhayn, and Sarah Reber for their advice and guidance. I would also like to thank Moshe Buchinsky, Dayanand Manoli and Sandra Black for comments that have enriched my research and nurtured my professional development.

Chapter 2 is a version of the working paper "Understanding Corruption: Theory and Evidence from the Audits of Local Governments" and is a joint paper with Maurizio Mazzocco and Frederico Finan.

I would also like to acknowledge my graduate student colleagues who have shared the pains and pleasures of the program. A special gratitude to Wilko Ziggy Shulz Mahlendorf, Anthony Keats and Arturo Harker, who have enriched my research and provided much needed relief and support during the hardest times of my career. I am grateful as well to all of my instructors, professors, teaching assistants, as well as the many undergraduate students whom I had the pleasure of teaching.

VITA

B.S. Economics, ***Universidad Torcuato di Tella*** **2003**

M.A. Economics, ***Universidad Torcuato di Tella*** **2005**

M.A. Economics, ***UCLA*** **2009**

Chapter 1

Female Labor Supply and Family Risk Sharing

1.1 Introduction

The main goal of this paper is to uncover the existence and magnitude of the added worker effect, the notion that married women increase their labor supply in response to an unemployment event by their husband, in the context of a developing country. The added worker effect (AWE) has been studied as a potentially important component of the female labor supply as well as a self-insurance mechanism available to households. In developing countries this is especially important as female labor participation rates are low, access to credit is limited and social safety nets are weak and virtually non-existent. Understanding the AWE is therefore of significance to policy makers designing labor market interventions such as unemployment insurance, as well as aiding policy makers in developing countries aiming at increasing the labor force participation rates of women[1].

The main contribution of this paper is to provide evidence of a large and significant added worker effect in a developing country. The paper presents evidence and implications of this finding in three parts. The first part of the paper presents reduced form evidence on the nature of the added worker effect. The analysis is done on Mexican urban labor markets using the 1987 - 2004 waves of the nationally representative survey of employment in urban labor markets (ENEU for its spanish acronym). Mexico presents an ideal setting for the study of the AWE because of the economic environment that surrounds household labor supply decisions. As in most developing countries, access to credit is very limited, so households are unable to smooth consumption by borrowing against future income. Second, there is no unemployment insurance and social safety nets are weak and virtually non-existent, thus, this eliminates concerns of crowd out by social insurance

[1] For example, in 2007 the Mexican government introduced a child care program seeking to increase the female labor force participation: Estancias Infantiles para Apoyar a Madres Trabajadoras

programs. Third, the relatively high macroeconomic volatility endemic to developing countries translates into an environment of uncertainty in the outcomes of the labor market. Finally, Mexican households seem to conform to the theoretical structure of primary earners and secondary workers as over 60% of households have only one spouse (generally the men) in the labor force, and the other spouse (generally the women) actively engaged in household production services.

The paper presents linear probability estimates (OLS) that suggest that wives are 11 percentage points more likely to enter the labor force as a result of their husband losing their job; an entry rate that is 60% higher relative to wives whose husband did not lose their jobs. Moreover, the evidence suggests that the entry of wives is largely transitory, as over 75% of wives that entered the labor force are observed to exit within a year. Exit of the labor force appears to be voluntary, as less than 10% of wives that exited claimed to have lost their job involuntarily, and the overwhelming majority of women respond that they are not looking for employment in the following periods due to household chores. Additionally, the paper presents some descriptive evidence on the characteristics of employment for the "added worker wives"; the jobs they perform are heavily tilted towards low skill jobs with low entry costs.

To address the potential endogeneity issues associated with the study of unemployment, the paper provides estimates on the sample of families whose primary earner lost their job due to plant or business shutdown or relocation. This is a commonly used instrument in the literature as it provides an unbiased sample of workers. Plant closings affect all workers in the plant or business independently of their observed and unobserved characteristics. Results suggest that OLS estimates may be downward biased as the measured added worker effect is 20% higher for workers who lost their job due to plant or business closing. This is in line with the argument in previous studies that suggest that workers who are more likely to be unemployed due to some unobserved characteristic (e.g. tastes for leisure) are married to women who share that characteristic and thus are less likely to enter the labor force themselves.

The second part of the paper develops and estimates a structural model that aims to capture the trade-offs faced by families. This approach will allow me to quantify the value of insurance provided by spouses, as well as to evaluate the impact of different labor market interventions and policies that are being debated in developing countries, such as unemployment insurance. The general framework is one of full commitment with no divorce; a household is composed of two members who make decisions regarding consumption, savings, leisure, household production services, labor supply and sector of employment that maximizes the total expected welfare of the household. The model has the following main features. First, the model includes a household public good that provides utility to household members but demands time for its production, this feature captures the trade-offs families face between allocating time to leisure, labor supply and household production services. Second, the framework incorporates a process of human capital accumulation that captures the incentives for specialization within the household. Third, the model incorporates savings decisions and a borrowing constraint that captures the importance of self-insurance mechanisms in this environment. Fourth, to reflect the uncertainty that households face in the labor market, the model includes wage shocks that describe the volatile wage processes observed in the data and an exogenous job destruction shock. Finally, the model tries to depict the job choice patterns observed in the data by offering workers the choice between two sectors: structured and unstructured. The structured sector reflects the traditional sector of employment; workers who participate in this sector accumulate human capital and their earnings are a function of their level of human capital. However, this sector has an entry cost; workers must spend a period in full-time job search to receive a wage offer in the subsequent period. Jobs in the unstructured sector have no entry costs, do not accumulate human capital and have lower levels of earnings that are independent of the level of human capital. The model is then estimated by standard dynamic programming tools and indirect inference.

The third part of the paper uses the estimated model to perform counter factual policy simulations aimed at quantifying the value of the self-insurance mechanisms available to households. One policy simulation offers unemployment insurance to households with varying replacement rates, that range from 5% to 60% of the husband's expected earnings and measures the crowd out effect of UI on the added worker effect. Analysis on the earning patterns within the household suggests that wives are able to gain about 50% to 60% of the income loss due to the unemployment event of the husband. The results of the policy simulations suggest that unemployment insurance would have strong crowd out effects; an unemployment scheme that offers households 45% of the husband's expected earnings would virtually wipe out the measured added worker effect. This result is in line with the argument in Cullen and Gruber (2000), who also find very strong crowd out effects for the U.S. where the average replacement rate of UI is about 46%. The second counter factual simulation is related to policies that are often debated in developing countries regarding the large informal sector. The policy simulates the effects of shutting down access to jobs that are informal, low skilled and have low entry costs (e.g. domestic employees, street vending). Notice that this is only part of the informal sector; however, the types of jobs considered are those where monitoring costs are low and implementation of the policy is feasible and less expensive. The paper evaluates the impact of this type of policy on different households according to their level of wealth. The availability of jobs in the unstructured sector plays an important and heterogeneous role for households; lower wealth families benefit the most from the presence of this type of jobs as they participate often and for longer periods in the labor force. Middle-wealth households value this type of jobs as they provide the opportunity to increase their labor supply transitorily without high adjustment costs. Higher wealth households find little value in this sector as they rely mostly on savings to smooth consumption when faced with adverse income shocks.

The results presented in this paper are in sharp contrast to previous literature which has focused on developed countries. Empirical evidence regarding the existence and magnitude of the AWE for developed countries has been conflicting and inconclusive. Most studies have found no significant effects (Pencavel (1982), Maloney (1987, 1991), Spletzer (1997), Layard et al. (1980)), and the studies which have uncovered some evidence (Mincer (1962), Heckman and MaCurdy (1980), Lundberg(1985)) show very small effects. Gruber and Cullen (2000) make an important point that both the theoretical and empirical studies have largely ignored the potentially important role of unemployment insurance. Unemployment insurance lessens the loss of household income, and thus may be crowding out the added worker effect. Exploiting state level differences in the generosity of UI programs, they find evidence of a strong crowd out effect of UI, and their estimates suggests that in the absence of unemployment insurance wives total work hours would increase by 30%.

The paper is organized in the following manner. Section 2 provides a description of the data and the methodology used in estimating the added worker effect. Section 3 presents the reduced form results on the measured added worker effect, the transitory nature of the entry of wives, and descriptive statistics on the characteristics of the jobs performed by the 'AWE wives'. Section 4 presents the structural model of household labor supply decisions. Section 5 describes the computation of the model as well as the methodology used for estimation. Section 6 presents the functional form assumptions to operationalize the model, and discusses the moments selected to identify the parameters of the model. Results of the structural estimation jointly with the results of the policy simulations are presented in section 7. Finally, section 8 concludes.

1.2 Data and Empirical Strategy

1.2.1 Data

The empirical analysis is done using the information available in the National Urban Employment Survey (ENEU for it's Spanish acronym) collected by the Mexican Statistical Institute (INEGI) for the sample years 1987 -2004. The survey is designed as a five quarter rotating panels, such that we have five observations on each household for a fifteen month period at 3 month intervals. The survey provides information on demographics, labor market participation, unemployment, job characteristics, earnings, hours of work, and time devoted to household production for all household members who are 12 years and older. For the younger members only their age and gender is recorded. Although the panel is quite short time wise, it allows for the analysis of short-term responses of family labor supply to labor market shocks. Furthermore, since households are interviewed every three months, recall bias common in retrospective surveys is expected to be minor.

The sample consists of households with the family head and spouse presently living in the household, where both members are between 20 and 60 years of age. I will further restrict the sample to those households that have observations in all five periods. Additionally, as I am interested in measuring the AWE on the extensive margin, the major part of the analysis will be carried out on households where the wife is out of the labor force (OLF) and the husband is employed at the time of the first interview. As is done in the previous literature I exclude from the analysis households where the husband transitions voluntarily out of employment (quits) and temporary workers. Although they don't represent a significant proportion of the sample, the behavior of this type of families may be very different. The ENEU actually started in 1987, however, starting in the third quarter of 1994, the phrasing of some questions changed significantly. Additionally the geographical coverage of the survey expanded significantly. Thus, there is "seam" problem when trying to join the two data sets. Hence I separate the two waves of the survey for independent estimation. Following the criteria stated above I constructed panels starting from 1987:1 to 1994:2 and 1994:3 to 2003:4. Therefore we have 26 panels for the first years of the survey which are pooled together for estimation adding up to 69,315 households and 38 panels for the second part of the survey which are pooled together for estimation with 161,975 observations.

For the analysis we need to define the three states of labor force status: employed, unemployed and out of the labor force. Definitions are standard; employed individuals are those that were working during the last week or were temporarily not working but held a job (e.g. vacations, sick leave). Unemployed are those who searched for a job during the last four weeks, or if they were on temporary lay-off or those waiting to report to a job within thirty days or more. Individuals who are not employed or unemployed are considered to be out of the labor force.

1.2.2 Descriptive Statistics

Table 1 presents some descriptive statistics on the sample of married couples with ages between 20 and 60. Women tend to be slightly younger than men, and they are less educated on average. However, education levels are low for both men and women; almost 35% of men have attained primary education or less, while 40% of women have attained primary education or less. Furthermore, less than 25% of men and less than 13% of women have gone beyond high school education.

The most salient feature of the descriptive statistics is the large difference that arises between the two measures of labor force participation for women (rows 3 and 4 of Table 1). The traditional cross-section measure of labor force participation suggests that about 36% of women were either employed or unemployed.

However, the panel nature of the data set allows us to ask what percentage of women have participated in the labor force in the course of 15 months. By this measure, almost 53% of women have participated in the labor force, a 17 percentage point difference. These two statistics put together suggest that women are entering and exiting the labor force frequently. Moreover, when compared to men, it suggests that women have very different labor force patterns than men. This evidence is consistent with the notion of households having a 'primary earner', generally the men, who are actively and consistently participating in the labor force; and a 'secondary earner', generally the wives, who participate in the labor force occasionally and transitorily.

Delving deeper into the labor supply patterns of women, Figure 1 plots the number of quarters that women participate in the labor force. The figure uncovers the heterogeneity in the labor force participation patterns of women. Almost 50% of women are OLF for all five quarters, reflecting the cross-section measure of LFP. However, there is also a significant proportion of women (20%) who seem to be actively engaged in the workforce, participating all five quarters and acting similarly to men. But the graph helps to illustrate the differences found between the static measures of LFP and the dynamic measures; about 30% of women appear to be participating in the labor force only transitorily, entering and exiting within a 15 month period.

The descriptive statistics reveal interesting patterns in the labor force supply of women. A large proportion of women appear to act as 'secondary earners' of the household, with lower participation rates than men, and entering into the labor force occasionally and transitorily. This is an important aspect of the female labor supply and one that motivates the analysis below. As 'secondary earners', are women providing insurance within the household, supplying hours of labor as a response to adverse shocks to household income? Specifically, I will focus on measuring the response of wives that are out of the labor force to unemployment events of the husband.

1.2.3 Empirical Strategy

The methodology used to estimate the added worker effect (AWE) on the extensive margin follows the previous literature. First, the AWE will be defined as the response of wives to an unemployment event suffered by the primary earner of the household. However, as Maloney (1991) points out, there are several margins where we could expect there to be a response of the spouse. For example, the labor supply of the husband may be constrained; husbands may have less shifts than desired, or they may have lost a second job, or may not be able to get overtime.

The second important aspect to note is that the sample is conditioned on the labor force status of both members of the household. In particular, at the time of the first interview wives must be out of the labor force and the husband must be employed. Notice however, that this is the starting point for over 60% of the sample. Therefore, the analysis will compare the entry rates of wives whose husband's were unemployed at any point in the following four quarters versus the entry rates of wives whose husbands were employed continuously over the following year.

For the analysis, it is important to incorporate as many periods as possible as entry of wives need not be contemporaneous to the unemployment event. Stephens (2002) makes the important point that the timing of entry depends on the perceived probability of the husband suffering unemployment. Wives who believe that unemployment of the husband is very likely will perhaps adjust their labor supply even before the actual unemployment event. On the other hand, wives may need time to adjust their labor supply, hence entry of wives may be one or more periods after the job loss actually occurs. Therefore, I will incorporate all the relevant information in the five interviews and collapse it into one observation per household. Notice as well, that the intention is to estimate the response of wives who are OLF to the unemployment event of their husbands. As such, the relevant source of variation or 'treatment' in the data is the unemployment event of the husband. The time scope of the survey is that of 15 months and over 85% of the population only suffers one unemployment event in this time frame. Thus, estimating the AWE by collapsing the information in the five interviews into one observation is essentially the same as doing the analysis exploiting the panel nature of the data.

The econometric framework will be based on estimating the following equation:

$$Y_i = \alpha + \beta H_i + \gamma X_i + \eta_i$$

The dependent variable Y_i measures the entry of the wife into the labor force. Entry can be either directly into employment or into unemployment. Accordingly, the dummy variable Y_i is equal to one if the wife enters the labor force in any of the four quarters following the first one, and is equal to zero if she does not. Similarly, our main variable of interest H_i, is a dummy variable that takes the value of 1 if the husband suffers an unemployment event in any of the four quarters following the first one and zero if he never transitions into unemployment. Consequently, the parameter β is the main object of interest, the measure of the added worker effect. Specifically, the coefficient β is interpreted as the marginal effect of the husband's unemployment on the wives' probability of entering the labor force.

X_i refers to a set of control variables that are included sequentially as to examine the sensitivity of the estimated AWE to the inclusion of additional regressors. Control variables include individual characteristics of the wife, such as age, age^2, years of education, and past labor force experience; individual characteristics of the husband, such as age, age^2, years of education, and industry of employment; household characteristics, namely the presence of children under six years of age, children between 6 and 12, and children between 12 and 18. I also include the unemployment rate in the county of residence as to proxy for local labor market conditions. Finally, all specifications include year and quarter fixed effects as well as state fixed effects.

Previous literature have identified three potential issues when estimating the AWE; first, is the issue of the discouraged worker effect (DWE). The discouraged worker effect is the notion that if the unemployment of the husband was caused by a general economic downturn, wives may be discouraged to enter the labor force because they perceive the probability of employment to be very low. Thus, to the extent that there is a significant discouraged worker effect, estimates of the added worker effect will be downward bias. The second issue is one raised by Spletzer (1997), who finds that the significant AWE found in descriptive statistics essentially disappears once he controls for the past labor force experience of wives. The argument suggests that couples may be positively matched on labor force volatility, that is, husbands who are more likely to enter into unemployment are married to wives that are entering and exiting the labor force frequently and hence, the observed AWE is actually just spurious correlation. This kind of endogeneity would actually overestimate the actual AWE. Finally, the third issue identified in the literature is one that would bias the results against finding a significant AWE. The argument is that if husbands are more likely to be unemployed because of some unobservable characteristic (e.g. low tastes for work, unobserved human capital) and they are married to wives who also share that characteristic, those wives may receive lower wage offers or have higher reservation wages than the average woman in the population. Thus, the AWE would be underestimated.

To address the issue of the discouraged worker effect, the researcher would want to control for the correlation between the husbands unemployment and the job prospects of the wives. In this paper, an effort is done to address this issue in two ways. First, I include the average unemployment rate in the county of residence as to proxy for local labor market conditions. Additionally, I estimate the added worker effect on two different samples as done in Skoufias and Parker (2004): one sample comprises the recession suffered in Mexico during the last quarters of 1994 and the beginning of 1995 ('Tequila Crisis') when unemployment soared to over 8%. The second sample comprises the boom of the business cycle during the year 1999 when unemployment fell below 3%.

To deal with the potential endogeneity issues raised in the previous literature I propose to use the sample of unemployed workers that lost their job due to plant/business shutdown or relocation. Plant shutdown has been used often in the labor literature because it provides an exogenous sample of unemployed workers. The argument proposes that a plant shutdown causes all workers in the plant, independently of their observed and unobserved characteristics to be unemployed and hence provides an unbiased sample of workers. However, using this sample of unemployed workers is not without cost; one issue that arises is that of representativeness. Workers who have suffered unemployment due to plant or business shutdown are disproportionately from the manufacturing, commerce and service industries. Therefore, to have an adequate control group, I will restrict the analysis to those households where the primary earner is employed in any of these three industries

(55.23% of the sample). Another potential issue concerns the self-employed; business shutdown for the self-employed may be directly related to their entrepreneurial ability, hence I will restrict the sample to salaried workers in the selected industries (41.73% of the sample).

Unfortunately the design of the survey leads to a potential issue of sample selection. Only workers who are unemployed at the time of the interview provide a response for the reason they lost their job. Therefore, workers who found a job between the time of plant shutdown and the survey interview are not going to appear in the relevant sample. If the timing of re-employment is correlated to some characteristics of the workers then the sample would suffer from selection bias. Although the time frame for this selection process is rather short (three months), I will examine the sensitivity of the estimates to the timing of the plant shutdown. Specifically, I will further restrict the sample to those unemployed workers that lost their job due to plant shutdown within one month of the interview.

1.3 Results

1.3.1 Linear Probability Model (OLS)

Table 2 presents the results of the linear probability model. As noted above, controls for the characteristics of the wives, household structure, husbands characteristics, and local labor market conditions are introduced sequentially to examine the sensitivity of estimates to the inclusion of controls. The results show a remarkably stable coefficient. In all specifications both the magnitude and the significance of the AWE is relatively unchanged around 11.2 percentage points. The first column, without any controls, allows to interpret the beta coefficient as the additional entry rate of wives who suffered an unemployment event relative to wives who didn't. The constant reflects the average entry rate of wives that did not suffer an unemployment event and is 19.1%. The coefficient of 0.11 implies that the entry rate is 58% higher for wives whose husband's suffered unemployment. This is an important result as it contrasts sharply with the results found in the previous literature focused on developed countries.

The addition of control variables affects the estimated AWE only slightly. Noteworthy is the effect of the presence of children under 6 years of age, who seem to restrict the entry rates of all wives. This is consistent with findings in the literature that the presence of young children limits the ability of households to supply more labor to the market. The coefficients on the age variables reflect that younger households tend to have higher rates of entry than older couples. This is consistent with the life cycle model predictions that older couples tend to have more savings and thus can smooth consumption through spending down savings rather than supplying additional hours to the market. Finally, the average unemployment rate in the county of residence has a significant and large effect on the entry rate of wives. The result again suggests that the discouraged worker effect is not strong enough to counteract the AWE, even during economic downturns.

In what follows, I examine how the presence of children of different ages affect the ability of households to respond to income shocks. To the basic regression described above I included interaction terms for each age group of the children with the unemployment event of the husband. This allows for within group comparisons of families who have children in the same age group. Results are reported in Table 3. Although the presence of young children restricts the entry rate of wives, it does not appear to limit the ability of wives to respond to the unemployment of the husband. For households who have children between the ages of 6 and 12, the entry rate is similar to households without children, but they are 5.65 percentage points more likely to enter the labor force in response to an unemployment spell of the husband. This effect is large and statistically significant at the 1% level. Finally, the presence of children between 12 and 18 years has a positive but insignificant effect on the measured added worker effect. A possible interpretation of these differences is that children between 12 and 18 are viable candidates to enter the labor force themselves, thus substituting the entry of wives. Considering the low levels of education in the population, young adults may be dropping out of high school to enter the labor force in response to the unemployment event of the household breadwinner. These results are interesting and warrant further research on the role of children in family labor supply decisions.

1.3.2 Plant or Business Shutdown

In this section I present the results on the sample of unemployed workers that lost their job due to plant or business shutdown/relocation. As noted above, unemployment due to plant or business closing comes disproportionately from the manufacturing, commerce, and service industries. Additionally, self-employed workers who have lost their job due to business closing are potentially a selected sample as their entrepreneurial ability may be correlated to the business failing. So, in the following analysis I restrict the sample to workers who were salaried and working in the manufacturing, commerce or service industries. This represents over 41% of the original sample.

The analysis that follows will measure the added worker effect for workers who were unemployed due to business/plant shutdown. This sample is arguably an unbiased sample of workers as plant closings affects all workers in the plant independently of their observed or unobserved characteristics. It is important to note that the control group includes both workers who did not lose their jobs as well as workers who were unemployed due to reasons other than plant closing (e.g. directed firings, sickness). It is important to include all other unemployed workers, as some of the workers that lost their jobs due to plant closing would have lost their jobs regardless. Hence, including both workers that did not lose their job, and those that were directly fired will represent the true distribution of unobserved characteristics in the population.

The first column of Table 4 presents the linear probability estimates (OLS) on this selected sample. Results show that for this sample the added worker effect is slightly smaller than for the rest of the population (9.38 vs 11 percentage points). Column II of Table 4 presents the estimates for the sample of workers who were unemployed due to plant or business closing or relocating. The coefficient implies that wives whose husband lost their job due to plant closing are 11.6 percentage points more likely to enter the labor force relative to all other wives (those whose husband did not lose their job and wives whose husband was fired). Hence, results suggest that the OLS coefficients may be biased downward. This bias is in line with the argument in previous studies that there may be some unobserved characteristic (e.g. tastes for leisure) shared by wife and husband that makes the husband more likely to be unemployed and the wife less likely to enter the labor force. Although the bias is substantial, the coefficient on the sample of plant closing workers is 23% higher, it is not clear that the estimates for the added worker effect in developed countries would change substantially.

The magnitude of the coefficients on the other explanatory variables change slightly but reflect the same patterns than in the original sample. Women who are more educated are more likely to enter the labor force, the presence of children under six years of age restrict the labor supply of wives, as well as kids between 12 and 18 appear to increase the entry rates of wives. Additionally, the higher educational attainment of husbands decreases the likelihood of wives entering the labor force. Finally, the county unemployment rate has a positive effect on wives' entry rates, again suggesting that the DWE is not an important source of bias.

Finally, in the third column of Table 4, I present the results for the sample of unemployed workers who lost their job due to plant or business shutdown within one month of the interview. The one month window is likely to reduce the potential selection bias that affects the estimates in column II. However, to the extent that higher quality workers are able to find jobs within a month, the selection bias will still be present. The coefficient of 11.9 percentage points is slightly higher but not statistically different from the coefficient in column II. This result provides some suggestive evidence that the potential selection bias is not very strong for this sample.

1.3.3 Transitory Entry

The results presented above suggest that the AWE is large and significant. Wives are more likely to enter the labor force as a response to the unemployment event of the husband. The value of the insurance provided by wives to the household depends on the level of lost earnings they are able to replace. In turn, this value depends on the wages they are able to obtain as well as on the amount of hours they supply to market activities. In this subsection I analyze the transitory nature of the entry of wives, and in the following subsection I study the level of earnings that wives are able to obtain as well as the types of jobs they are taking. Unfortunately, the short span of the panel does not allow for a complete analysis as we will not be able to observe the exact number of periods in the labor force for all wives. However, an effort is done to analyze the available information in this limited environment.

As in the study of the added worker effect, theory does not provide unambiguous predictions regarding the permanent or transitory nature of the added worker effect. On one hand, there is a strong motive to smooth labor supply over the life cycle, suggesting that wives may increase their labor supply over all future periods. Additionally, in the presence of high entry costs to the labor force, wives may find it optimal to remain inside the labor force once they were 'shocked' into it. This is especially true if they perceive the likelihood of future unemployment events to be high. On the other hand, if the income loss is relatively small or if entry costs are low, wives may find it optimal to enter the labor force transitorily.

As mentioned above, the short span of the panel survey limits our ability to analyze the duration of wives' participation in the labor force. Observations will be 'right censored' in the sense that we can not observe the behavior of wives after the fifth interview. For example, for wives that are observed to enter the labor force in the last interview, we will not be able to say anything about the duration inside the labor force. However, descriptive statistics suggests that for a majority of women, entry appears to be transitory in nature. For example, 62% of the women observed to enter the labor force are also observed to exit within the 5 quarters. The figure is about 60% for wives whose husband was unemployed and re-employed within this period. Limiting our observations to those wives who are observed to enter the labor force by the second interview (thus maximizing the number of periods after wife entry) 77.5% of them are observed to exit, with about 32% of them participating for one quarter, 24% for two quarters and 22% for three quarters. For wives who entered by the second interview and their husband's lost their job, almost 75% are observed to exit, with 22.7% participating for only one period, 25.7% for two periods, and 26.4% for three periods.

Using regression analysis we can study the effect of husband's re-entry into employment on the probability that wives exit the labor force. Hence, we restrict the sample to those wives who entered the labor force in response to an unemployment event of the husband. For this exercise, the treatment is the re-employment of the husband, and we evaluate the effect on the probability that wives exit the labor force. The regression we estimate is then:

$$WifeExit_i = \alpha + \beta \text{HusbandRe-employed}_i + \gamma X_i + \eta_i,$$

where $WifeExit_i$ takes the value of one if the wife exits the labor force and 0 otherwise, and HusbandRe-

employed_i takes the value of 1 if the husband is employed following the period of unemployment and 0 otherwise. It is important to note at this point, that most unemployment spells are short, with 83% of the husbands who were unemployed are seen to be employed in the subsequent quarter. The results are presented in Table 5. The effect of the husband finding employment on the probability that wives exit the labor force is positive and highly significant. Wives are 12 percentage points more likely to exit the labor force in response to the husband finding employment; this translates into an exit rate that is 27% higher than those wives whose husband did not find a job in the observable time frame.

It is important to note that exit of wives refers to exiting the labor force, not employment. Therefore, wives are observed to exit the labor force altogether and they don't appear to be looking for work during the rest of the sample period. Furthermore, over 80% wives who exited cited personal or family reasons for exiting, and less than 10% claim that their jobs were lost involuntarily. Finally, when asked why they are not looking for a job during the last four weeks over 95% respond that they do not want to work or they do not have time. Thus, the evidence suggests that wives are entering the labor force transitorily and they are exiting voluntarily.

1.3.4 Point of Entry

Finally, in this section I explore the characteristics of the jobs that wives take when entering the labor force transitorily. Table 6 presents some measures of job characteristics and labor supply patterns of wives that entered the labor force ('entry wives'), 'AWE wives' (those who entered as a response to the unemployment event of the husband), wives who are engaged in the labor force for all five quarters (labeled working wives), and compares them with men.

The statistics reflect meaningful differences between wives who appear to be 'secondary earners' and working wives. Working wives, although supplying less hours than men on average, are supplying significantly more hours than 'entry wives' and 'awe wives'. On the other hand, wives who enter transitorily, are allocating more time to home production activities. Transitory entry is significantly weighted towards self-employed activities, and jobs than don't offer benefits. The earnings wives obtained are significantly lower than that of men and working wives, this is due in part because they supply less hours to the market, but also because their hourly wages are significantly lower.

A detailed examination of the occupations performed by women that enter the labor force transitorily reveals interesting patterns. The most common occupation is that of domestic employees, followed closely by street vending, sales agents in retail shops and in-home production (foods and mending clothes). When compared to working wives, the differences in the patterns of occupation are very revealing. Working wives tend to be professionals, technicians, teachers and education professionals, secretaries and administrative staff. For example, only 5% of working wives and over 15% of 'awe wives' are domestic employees. About 10% of working wives are involved in the retail sector, while the figure is over 22% for entry wives. The figures for street vending and in-home production are negligible for working wives, but comprise a significant proportion (17%) of the wives who are entering transitorily.

This observed patterns motivate the idea that wives who are seeking to enter the labor force only transitorily, are taking occupations which are very different from the occupations performed by men and by working wives. The characteristics of their occupations suggest that they are taking jobs which don't require high level of skills or education, that are readily accessible (don't have significant entry costs), and have flexible work schedules.

1.3.5 Robustness checks

In this section I present some extensions to the basic framework as robustness checks of the main results. To address the issue of couples being matched on labor force volatility, I present in Table 7 the results adding controls for the past labor force experience of wives. I divided the sample according to the past labor force experience of wives into five groups: wives that worked within the year previous to the first interview are classified as 'volatile', wives whose last job was between a year and three years before the first interview (omitted group), wives whose last job was 3 to 5 years before the interview, wives who held a job 5 years before the interview, and finally wives that claim to have never worked. The concern is that the large estimated AWE may be due to spurious correlation. Husbands who tend to lose their jobs frequently are married to wives who are also entering and exiting the labor force frequently. However, controlling for past labor force experience reveals that this source of bias does not seem to be an important issue in this context. The coefficients reveal that 'volatile' wives are in fact more likely to be entering the labor force, however they don't significantly bias the results as they are only a small proportion of the total sample (6.6%). The coefficients also reveal that the longer the time since the woman had a job, the less likely they are to enter the labor force. Additionally, as expected, wives that never worked are also less likely to enter the labor force. Noteworthy as well, is the stability of the measured added worker effect, which remains large and significant at 11 percentage points. Adding interactions as to study the AWE within groups (i.e. comparing wives that have similar labor market experience) reveals that the measured AWE is significant and of similar magnitude for all groups. Reported coefficients in Table 8 suggest that for wives that never worked and for wives that held a job more than five years previous to the first interview, the AWE is about 1 percentage point smaller, but the difference is not statistically significant. Additionally, for 'volatile' wives, the measured added worker effect is higher by 1.25 percentage points, but again the coefficient is not significant.

An additional test for spurious correlation has been suggested in the literature. If the measured AWE is in fact due to spurious correlation, because husbands and wives are entering and exiting the labor force frequently, then we should also expect to find a positive significant correlation between the likelihood of wives exiting the labor force when the husband loses his job. Results of this test are presented in Table 9. I find that once you add controls to the regressions there is actually a negative and significant effect, wives are less likely to exit the labor force if the husband has lost his job. This evidence further suggests that spurious correlation is not driving the results. Moreover, the negative correlation between wife exit and husband unemployment may be interpreted as another margin of the AWE, where wives stay in the labor force longer if the husband faces unemployment[2].

Another extension is done by measuring the added worker effect for the lowest decile of the income distribution. If husbands who are more likely to become unemployed due to some unobserved characteristics (e.g. high tastes for leisure, unobservable human capital) are married to wives who also share that characteristics, then the added worker effect may be biased downward. However, these kind of workers should also be less educated and should be earning less than the average man in the population. Thus, I re-estimate the regression presented above only for those workers at the bottom decile of the income distribution. The results are presented in Table 10. The findings suggest that for this group the added worker effect is in fact very similar to that of the whole population. It is important to note, that the observed unemployment is present across all deciles of the income distribution. Although workers in the bottom deciles of the earnings distribution are slightly more likely to suffer an unemployment event (2.88), we observe unemployment across all deciles of the income distribution. For example, of the workers in the top deciles of the wage distribution 2.10% suffered an unemployment spell. Hence, the similarity of the results is not driven by unemployment events occurring only to low wage workers.

As noted by Maloney (1991), there are several margins where we could expect the AWE to be present. Even if husbands are not unemployed, their labor supply may be constrained. Workers may want to work more hours but may be unable to so; reduced shifts, reduction in overtime, loss of a second job, etc... In this section, I present one possible additional margin of the AWE: under employment of the husband. Under employment is defined as those workers who are supplying less than 35 hours to the market for involuntary reasons[3]. Results, presented in Table 11, show that wives are 3.5 percentage points more likely to enter the labor force in response to an under-employment event of the husband. The estimated effect is significant and

[2] I thank Leah Platt-Boustan for this comment

[3] Involuntary reasons include: production crisis, lack of clients, lack of financing, lack of inputs or machinery breakdown. Excluded are those who cite the following reasons: characteristic of the job, vacations, family reasons, or doesn't want/need to work more hours.

remarkably stables across specifications. Hence, it appears that the AWE is present along several margins.

Finally, I present estimates conditioning the sample on the duration of the unemployment event of the husband. As noted above, over 83% of unemployment events last only one quarter, however 13% of husbands are seen to be unemployed two quarters and 3% are unemployed for 3 quarters and less than 1% are observed to be unemployed for 4 quarters. As there may be selection bias of the workers who are unemployed for longer periods, I present in the first column of Table 12 the estimated added worker effect for those husbands who had an unemployment event that lasted one quarter. In the second column, I add three dummy variables that indicate those households where the unemployment event lasts 2, 3, and 4 quarters. For those husbands who had a short unemployment spell, the measured added worker effect falls slightly to 10.5 percentage points. The duration of the unemployment event seems to have a positive effect on the entry rates of wives; for workers who were unemployed for two quarters the additional AWE is 1.31 percentage points higher, although not statistically significant. For workers who were unemployed for three quarters, the additional added worker effect is 2.48 percentage points higher, but again not statistically significant. Even though the estimates are not significantly different, it is interesting to see that as expected, the measured added worker effect increases with the duration of the unemployment spell. On a final note, for those workers who are observed to be unemployed for four quarters, there is a huge and significant added worker effect, wives are 42.6 percentage points more likely to enter the labor force. However, it should be noted that there are only 17 observations in this category, less than 0.01 percent of the sample.

1.4 Model

In this section I develop a model of labor supply and time allocation decisions within a household. The general framework of the model is a household decision model with full commitment and no divorce. Two agents, m and f conform a household and they make decisions regarding consumption, savings, leisure, time devoted to producing a public household good, hours supplied to the labor market, and they can choose the sector of employment. In this framework, the joint decisions maximize the total expected welfare of the household. Due to the low divorce rates observed in the data, the model does not allow for households to dissolve and there is no renegotiation of the terms of the marriage.

Overall, the framework captures five important factors that influence the labor supply decisions within the household. First, the model captures the trade-offs families face when deciding how to allocate time to leisure, labor market supply and household good production services. Second, the incentives for specialization within the household. Third, the limited ability of families to smooth income shocks through credit. Fourth, the uncertainty that households face in the labor market. Finally, the framework allows me to capture the different job opportunities available to workers. In the next three subsections I describe how I incorporate this features into the household decision framework.

1.4.1 Preferences and Technology

The starting point of the model is that members of the households receive utility from three sources: leisure (l_t^i), consumption (c_t^i), and the level of household public good (Q_t). The preferences are described by the following general utility function:

$$U(c_t^i, l_t^i, Q_t)$$

To produce the public good agents must allocate time to household services (d_t^m, d_t^f). In general, the household good production function takes the form:

$$Q_t = q(d_t^m, d_t^f)$$

The function $q(d_t^m, d_t^f)$ is assumed to be increasing and concave in both arguments. Furthermore, I will assume that d_t^m and d_t^f are perfect substitutes, hence only the total amount $d_t^m + d_t^f$ will determine the level of the household public good.

Every period t each agent has a limited amount of time τ that they can allocate to leisure, household production, or labor supply (h_t^i). Thus, each agent within the household faces the following binding time constraint:

$$\tau = h_t^i + d_t^i + l_t^i \text{ for } i = m, f$$

These features of the model reveal the essential trade-offs families face when allocating the available time within the household. Household will weigh the benefit of enjoying an additional unit of leisure, against the value of having higher levels of public good, and against the value of increasing their income by means of supplying more hours to the labor market.

1.4.2 Borrowing Constraints and Savings

Limited access to credit is a salient characteristic of developing economies. As families can not borrow against future income, they must instead rely on self-insurance mechanisms such as holding savings or increasing labor supply of household members. The model captures the importance of these mechanisms, by explicitly allowing households to choose the amount of family savings (b_{t+1}) every period and the amount of hours devoted to market activities.

In the model, households face borrowing constraints that reflect the lack of access to credit. Every period families must have a non-negative amount of savings:

$$b_{t+1} \geq 0$$

Incorporating decisions on savings in the model plays an important role. As mentioned above, households can not borrow against future income, but they can self-insure against future income shocks by holding assets. Labor market decisions and outcomes will determine the level of resources available to households; decisions regarding the level of assets to be held by the family will determine the need for increasing labor supply when faced with an adverse income shock. Families with higher levels of savings may not need to increase the total labor supply and instead choose to smooth consumption by spending their savings. On the other hand, families with low levels of saving may be forced to increase their labor supply.

1.4.3 Labor Market Environment

The labor market characterization in the model is meant to capture the patterns of job choice observed in the data. Spouses who are considered to be 'primary earners' are generally involved in the traditional or structured sector of the labor market. They are either salaried workers in established firms, or they are self-employed with registered businesses and formal places of business (workshops, stores, etc...). Secondary earners on the other hand, seem to be in the labor force transitorily and they choose different types of jobs; these are jobs that are considered to be part of the unstructured sector (street vendors, domestic employees, etc...). Jobs in the unstructured sector seem to be always accessible and to provide workers with flexible schedules. However, earnings in this sector are substantially lower, and they don't appear to increase with experience.

The model tries to depict this pattern by offering workers the choice between two sectors: the structured and the unstructured. Three main attributes distinguish the two sectors: Human capital accumulation, entry costs, and work hours. In the following subsections I describe each sector in detail.

Structured Sector

The main attribute of the structured sector is that workers accumulate human capital. Every period involved in the structured sector is rewarded with an additional unit of human capital. Human capital accumulation or experience evolves according to the following rule:

$$HK^i_{t+1} = \left\{ \begin{array}{c} HK^i_t + 1 \text{ if i works in the structured sector} \\ HK^i_t \text{ if not} \end{array} \right\}$$

Human capital is valued in this sector and the wage offered to workers is determined in part by their level of human capital. Thus, participating in this sector not only provides higher earnings in the current period but it also increases expected wages in all future periods. Wages are determined according to the following wage process:

$$\varpi_{it}^{structured} = \alpha_1 + \alpha_2 * \; HK_t^i + \alpha_3 * \; (HK_t^i)^2 + \varepsilon_{it}$$

Human capital accumulation, jointly with the need to devote time to household production, characterize the incentives for specialization within a household. The member of the household with higher level of human capital will have stronger incentives to enter the structured sector. As he or she accumulates more human capital over time, the incentives for specialization only get stronger.

Another important feature is the presence of an entry cost. Workers must spend one period in full-time job search to access the structured sector. There is no uncertainty regarding job search, once workers have paid the entry cost they receive a wage offer with certainty. This particular attribute is meant to capture the different incentives for primary earners and secondary earners. Workers who intend to participate continuously in the labor force are willing to pay the entry cost as there is a very high value of holding a job in this sector. On the other hand, for workers who intend to supply labor market hours only transitorily, the value of participating in this sector is much lower, and they may be unwilling to pay the entry cost.

Finally, jobs in this sector are full-time only. This reflects the patterns observed in the data where primary earners are overwhelmingly full-time workers. However it provides an additional disincentive for secondary workers who may want to provide hours to household production as well as to the labor market.

Unstructured Sector

The unstructured sector differs from the structured sector by not accumulating human capital. Moreover, earnings are not conditional on the level of human capital of it's workers. The type of jobs included in this sector are those that don't require a significant level of skills and that don't increase significantly with experience. The wage process is given by:

$$\varpi_{it}^{unstructured} = \beta_1 + \epsilon_{it}$$

Another characteristic of this sector is that there is no entry cost to participate. This sector is accessible at any time for any agent who is willing to participate. This is meant to capture the patterns of entry of wives in the data; wives do not spend a lot of time in unemployment, but rather they immediately respond to adverse income shocks by working in the unstructured sector.

Finally, the unstructured sector provides only part-time jobs. This simplification is due to the observed choice of hours by wives in the data. Most workers in this sector seem to choose to allocate only part of their time to labor market activities. Thus, this sector provides the flexibility for workers to earn some income but it also allows for them to devote time to the production of the household good.

Labor Market Risk

In this section I describe the uncertainty that households face in the labor market. There are two types of shocks that characterize the labor market uncertainty; an exogenous job destruction shock for workers in the structured sector, and wage shocks in both sectors.

Workers in the structured sector face an exogenous job destruction shock λ. That is, with probability λ the job is destroyed, workers are unemployed and they must pay the entry cost to re-enter the structured sector. Additionally, workers face wage shocks ε_{it} every period. ε_{it} is an i.i.d. shock from a $N(0, \eta)$ distribution. Wage shocks are meant to capture the volatility of the income stream of workers. We can also interpret the wage shocks as the loss of income that occurs when workers are sub-employed, they lose overtime hours, reduced shifts, etc...

Workers in the unstructured sector also faces wage shocks every period, but their jobs can not be destroyed. The wage shocks are meant to capture the different levels of earnings observed in the data. Furthermore, it also allows the model some flexibility as the magnitude and sign of the wage shock affect the incentives for wives to participate in the labor force. Therefore, entry of wives is not only motivated by an adverse income shock of the husband (the insurance motif) but also by good labor market opportunities that may present themselves.

1.4.4 The Household Decision Process

In this section I will formally describe the decision process of the household. It is important to note from the start that the choice of sector of employment is conditional on the realization of the job destruction shock as well as on the past labor force status of the agents. In terms of the model there are four possible states of labor force status: employed in the structured sector, employed in the unstructured sector, unemployed, and out of the labor force (OLF) . The unemployed status refers to those workers who were not employed but spent their last period in full-time job search (i.e. they paid the entry cost into the structured sector). Out of the labor force is the status assigned to those agents who were not employed and did not pay the entry cost.

At the beginning of any period t, labor market shocks are realized. For those who are part of the structured sector, they will observe if their job was destroyed or not. For the workers whose job was not destroyed, they observe the wage shocks and are presented with wage offers from both sectors. They can now decide to continue working in the structured sector, switch to the unstructured sector or drop out of the labor force. For those workers whose job was destroyed, they can now choose between being unemployed and paying the entry cost to the structured sector (full-time job search), working in the unstructured sector, or staying out of the labor force. Agents who were out of the labor force or working in the unstructured sector have the same choice set; they can choose to pay the entry cost in the structured sector, work in unstructured sector or remain out of the labor force.

Formally, the choice set for employment E_t^i is conditional on the realization of the job destruction shock λ_t^i, and the past labor force status, defined as ss_{t-1}. Define $j_t = 1$ if the job was destroyed and $j_t = 0$ if the job was not destroyed, and let the abbreviations str,uns,olf,une signify structured,unstructured,out of the labor force, and unemployed respectively. Then, the choice for employment is defined as:

$$E_t^i = \left\{ \begin{array}{c} \text{str,uns,olf if } ss_{t-1}^i = str \text{ and } j_t = 0 \\ \text{une,uns,olf if } (ss_{t-1}^i = str \ \& \ j_t = 1) \text{ or } ss_{t-1}^i \neq str \end{array} \right\} \text{for } i = m, f$$

After labor market shocks are realized and employment choice sets are defined, households will make their decisions regarding labor force status, consumption, savings, leisure and hours in home production that maximizes the total expected household welfare. Formally, the relevant choice set defined by D_t^{hh} is,

$$D_t^{hh} = \{b_{t+1}, c_t^m, c_t^f, d_t^m, d_t^f, l_t^m, l_t^f, ss_t^m, ss_t^f | (E_t^m, E_t^f)\}$$

Let M be the Pareto weight assigned to the head of the household. As noted above, the full commitment model does not allow for renegotiation of the Pareto weight in the household welfare function, thus M is fixed over time. Additionally, let us define $\beta \in (0, 1)$ as the common discount factor and R as the real rate of return for assets held by the households.

Formally, household hh is composed of two agents, m and f. A household is defined by the level of assets they hold, the human capital of each member, the past labor force status of each member and the realization of the labor market shocks. The state space that defines a household in period t is then:

$$S_t^{hh}(\text{deterministic}) = \{b_t, HK_t^m, HK_t^f, ss_{t-1}^m, ss_{t-1}^f\}$$

$$S_t^{hh}(\text{stochastic}) = \left\{ \begin{array}{c} (\lambda_t^i, \varepsilon_t^i) \text{ if } ss_{t-1}^i = structured \\ (\varepsilon_t^i) \text{ if } ss_{t-1}^i \neq structured \end{array} \right\} \text{for } i = m, f$$

Agents m and f jointly make decisions that maximize the total expected welfare of the household,

$$E\left[\sum_{t=1}^{T} \beta^t \left(M * U(c_t^m, l_t^m, Q_t) + (1 - M) * U(c_t^f, l_t^f, Q_t)\right)\right]$$

subject to the constraint that in each period and state of nature expenditures on consumption plus savings must equal the available resources,

$$c_t^m + c_t^f + b_{t+1} = b_t * R + h_t^m * \varpi_{tm}^{ss} + h_t^f * \varpi_{tf}^{ss}$$

and subject to the time constraints for each member

$$\tau^i = l_t^i + h_t^i + d_t^i \text{ for } i = m, f$$

as well as satisfying the borrowing constraint every period

$$b_{t+1} \geq 0$$

and the technology constraints implied by the production function for the household public good

$$Q = q(d_t^m, d_t^f)$$

and finally the law of motion of human capital

$$HK_{t+1}^i = \left\{ \begin{array}{c} HK_t^i + 1 \text{ if } ss_t = structured \\ HK_t^i \text{ if } ss_t \neq structured \end{array} \right\} \text{for } i = m, f$$

1.5 Computation and Estimation

In this section I describe the approach used to estimate the model. In the estimation I use standard dynamic programming tools and indirect inference. Specifically, the estimation is performed in two steps. For a given set of parameters that characterize the model, we first simulate the individual decisions. We then match some of the statistical moments that characterize the data with the corresponding moments obtained using the simulated data. The estimated parameters are obtained by minimizing a function of the distance between the simulated and data moments required by indirect inference.

The simulation of the model requires the derivation of its recursive formulation and of the corresponding values functions. For exposition purposes it will be convenient to separate the problem in two steps. First, households choose consumption, leisure, household production and savings conditional on the choice of labor force sector (ss_t^m, ss_t^f). The value function for this step will be denoted $V_{t,ss_t^m,ss_t^f}^{hh}(S_t^{hh}, t)$. In the second step, household choose the labor force sector combination (ss_t^m, ss_t^f) that maximizes total expected household welfare. The final value function is thus denoted as $V_t^{hh}(S_t^{hh}, t)$.

As before, let S_t^{hh} be the set of state variables at time t for household hh. The problem that households face in the first step of the problem can be stated as follows:

$$V_{t,ss_t^m,ss_t^f}^{hh}(S_t^{hh}, t) = \max M * U(c_t^m, l_t^m, Q_t) + (1 - M) * U(c_t^f, l_t^f, Q_t) + \beta * E\left[V_{t+1}^{hh}(S_{t+1}^{hh}, t+1) | S_{t+1}^{hh}\right]$$

$s.t.$

1. $c_t^m + c_t^f + b_{t+1} = b_t * R + h_t^m * \varpi_{t,m}^{ss} + h_t^f * \varpi_{t,f}^{ss}$
2. $\tau^i = l_t^i + h_t^i + d_t^i$ for $i = m, f$
3. $b_{t+1} \geq 0$
4. $Q = q(d_t^m, d_t^f)$
5. $HK_{t+1}^i = \left\{ \begin{array}{c} HK_t^i + 1 \text{ if } ss_t = structured \\ HK_t^i \text{ if } ss_t \neq structured \end{array} \right\}$ for $i = m, f$

The second step of the problem is then choosing the maximum between all possible combinations of labor force sector. It is important to remember that the possible choices of sector are conditional on past labor force sector and the realization of the job destruction shock. Thus, we have to condition on the relevant choice set of each agent E_t^m and E_t^f:

$$V_t^{hh}(S_t^{hh}, t) = \max_{ss_t^m, ss_t^m}(V_{t,ss_t^m,ss_t^f}^{hh}(S_t^{hh}, t)|E_t^m, E_t^f)$$

The value function for each household hh is computed starting from the last period and moving backwards in two steps following Keane and Wolpin(1994). In the first step, the state space is "discretized". Then using the corresponding grid, the value functions are computed for each period and each point of the state space in the grid. Finally, using the probability distribution for the discretized state space we can compute for each period the expected value functions conditional on the set of state variables $E\left[V^{hh}(S_t^{hh}, t)|S^{hh}\right]$ .In the second step, the expected value functions are approximated using non-parametric methods. In practice, I regress the values of $E\left[V^{hh}(S_t^{hh}, t)|S^{hh}\right]$ obtained for each point in the grid on a polynomial of the discretized state variables. The corresponding coefficients are then used to construct the expected value functions for each period and value of the state space. Once the expected value functions are known, we can simulate the decisions of the households observed in the data for different values of the parameters that characterize the model. The parameters of the model are then estimated using indirect inference.

1.6 Empirical Specification and Moment Selection

In this section I will discuss the empirical implementation of the model presented above. First, I will present the functional form assumptions that are needed to operationalize the model: 1) utility function; 2) household public good production function; and 3) wage process for each sector

1.6.1 Functional Form Assumptions

Utility function: Each member of the household hh has preferences over consumption (c_t^i), leisure (l_t^i) and the household public good (Q_t) that are described by the following CES utility function:

$$U(c_t^i, l_t^i, Q_t) = \frac{[(c_t^i)^{\sigma} * (l_t^i)^{(1-\sigma)}]^{(1-\gamma)}}{(1-\gamma)} + \theta * Q_t$$

where $\gamma > 0$, and $0 < \sigma < 1$.

The parameter γ captures the inter-temporal aspects of individual preferences. In particular $-1/\gamma$ is interpreted as the inter-temporal elasticity of substitution. The parameter σ captures the intra-period features of individual preferences and it measures in each period the fraction of expenditure assigned to agent i which is allocated to consumption. The parameter θ captures the agent's taste for the household public good relative to the composite leisure-consumption private good.

Household good production function: The production function for the household good takes as inputs the time of each member of the household.

$$Q_t = \mu * \ln(d_t^m + d_t^f)$$

The most salient feature of this production function is the assumption that d_t^m and d_t^f are perfect substitutes. The parameter μ is interpreted as a measure of the productivity of the total hours devoted by the household.

Wage process: The model assumes a different wage process for each sector. In the structured sector, wages are conditional on the level of human capital while earnings in the unstructured sector are independent of the level of human capital. Thus we have two different wage processes:

$$\varpi_{it}^{structured} = \alpha_1 + \alpha_2 * \ HK_t^i + \alpha_3 * \ (HK_t^i)^2 + \varepsilon_{it} \text{ for the structured sector}$$

where ε_{it} is normally distributed with mean zero and variance $\eta_{structured}$.

$$\varpi_{it}^{unstructured} = \beta_1 + \epsilon_{it} \text{ for the unstructured sector}$$

where ϵ_{it} is normally distributed with mean zero and variance $\eta_{unstructured}$.

It should be noted that even though the model doesn't explicitly incorporate education as part of the human capital of agents, in practice the education level of the agents is incorporated by increasing the initial level of human capital. The returns to an additional year of education are transformed into equivalent years of additional experience. In this way, agents with higher levels of education start with higher levels of human capital in the model, and will therefore receive higher wages.

1.6.2 Moment Selection and Parameters to Estimate

The model is completely characterized by the parameters of the utility function and household welfare function $\{M, \beta, \sigma, \gamma, \theta\}$, the productivity parameter in the household good production function $\{\mu\}$, the parameters in the wage process $\{\alpha_{1,}\alpha_2, \alpha_{3,}\beta_1, \eta_{structured}, \eta_{unstructured}\}$, the job destruction rate $\{\lambda\}$, and some miscellaneous parameters that define the time horizon, the available time each period and the real rate of return on assets $\{T, \tau, R\}$.

Ideally, all the parameters of the model would be estimated. However, due to lack of data and identification issues not all of the parameters are going to be estimated within the model. Some of the parameters will be fixed, and the values of the estimated parameters will be conditional on the value of the fixed parameters. Parameters that are not estimated, are set to values that were estimated outside the model, come from alternative data sources, or are estimates previously used in the literature.

The quarterly real rate of return on assets is set to 1.012 with an implied annualized rate of 5%. This value was set to the average real rate of return calculated by the World Bank Development Indicators. The discount factor β is set to 0.9878 per period, which implies an annualized discount factor of about 0.95, a common assumption on the labor literature. The time horizon T is set to 72 years, the life expectancy in 1995 as calculated by the World Bank, and the available time per period τ is set to 16 hours per day for a five day week. The Pareto weight in the household welfare function is assumed to take the value of 0.5, implying that husbands' and wives' utility are equally weighted. Due to lack of data on the household allocation of consumption, I can not adequately estimate the Pareto weights of the household welfare function. Thus, I try to take a neutral position and assume that members are equally important to the household welfare. In the same line, the parameters of the utility function will not be estimated within the model but are taken from the estimates of Mazzocco (2013), and take the value of 2 for γ and the value of 0.5 for σ.

It has been noted in the previous literature on household decision models that the parameters θ and

μ can not be separately identified. That is, the preference parameter for the household good can not be separately identified from the productivity parameter. Hence, only $\pi = \theta * \mu$ will be estimated within the model.

Thus, there are 8 parameters that will be estimated by indirect inference: the job destruction rate λ, the household good composite parameter π, the parameters in the wage process for the structured sector $(\alpha_1, \alpha_2, \alpha_3, \eta_{structured})$ and the parameters in the wage process of the unstructured sector $(\beta_1, \eta_{unstructured})$

Job destruction rate λ: To estimate the probability that a job in the structured sector is destroyed I use as a moment the separation rates from the structured sector. The intuition is straightforward, as the destruction rate increases we should observe higher separation rates from the structured sector.

Preference for household good (composite) π: To estimate the relative taste for the household public good I chose the average hours devoted to home production as a moment. As households increase the relative taste for the household good, intuitively it follows that they will increase the hours devoted to household good production.

Labor market parameters $(\alpha_1, \alpha_2, \alpha_3, \eta_{structured})$ and $(\beta_1, \eta_{unstructured})$: The parameters of the wage offer distribution for workers in the structured (unstructured) sector are estimated by matching the parameters obtained by regressing the log wages of workers in the structured (unstructured) sector on the variables that characterize the wage function. Namely, for the structured sector, log wages are regressed on a constant and the level of human capital of the worker. For the unstructured sector log wages are regressed on a constant. Additionally, the variance of the residuals of the regression are matched to the variance parameters: $\eta_{structured}, \eta_{unstructured}$.

Additional Moment: As a consistency check of the model, I allow for an additional moment to be matched by the model. The moment I selected is the percentage of workers participating in the unstructured sector.

1.7 Results of the Model

The estimated values for a subset of the model's parameters are reported in Table 13. The model performs quite well in matching the selected moments of the data. The model does not match the moments exactly because it has an additional moment that is not tied to any parameter. The variance of the residuals of the log wage regressions for both sectors are slightly underestimated; this is likely due to the homogeneity of the agents. In particular, the wages in the unstructured sector only differ by their observed shock $\eta_{unstructured}$. For the structured sector, although education is implicitly incorporated in the model, agents do not differ in their ability, thus limiting the heterogeneity that is likely to be present in the data. The log wage regression for the structured sector reveals that the returns to experience are slightly overestimated while the constant is underestimated. This is likely due to the fact that workers choose to stay in the structured sector even for wage shocks that are very negative, because the value of remaining employed in the structured sector is very high. Since wage shocks are i.i.d. and there is a significant entry cost to this sector, many workers find the value of remaining inside the structured sector with low wage shocks preferable to being unemployed and having to pay the entry cost in the future. This result also explains why the job destruction rate is overestimated in the model. Finally, the model matches the average hours devoted to household production quite well. The slightly overestimated number of hours is likely due to the fixed hours of employment imposed in the model for simplicity. Workers do not choose the exact number of hours they work but rather they are assigned 40 hours of work per week for full-time jobs and part-time workers are assigned 25 hours of work. Obviously, the data presents a lot more variation in the choice of hours of work, and this could explain the differences in the estimated hours of work.

Finally, the model overestimates the proportion of workers in the unstructured sector. This moment is not tied to any parameter and is chosen as a consistency check of the model. The higher proportion of workers in the unstructured sector is explained by the sector choices of women. Women in the poorest households (measured by their level of savings) devote more time to labor market activities relative to wealthier households. These type of women are employed more frequently and for a longer time; however they choose the unstructured sector as their sector of employment for two reasons: first, they are devoting time to household production as well as market activities, second, they have very low levels of human capital (both education and experience). Thus, the full-time only jobs in the structured sector is a strong disincentive for women and they choose to participate in the unstructured sector as it provides the flexibility to allocate time to household production. Additionally, since they have very low levels of human capital, the earnings they receive in the unstructured sector are comparable to the earnings they would receive in the structured sector, lessening the incentives to pay the entry cost to the structured sector. Finally, even though they participate in the labor force more often and for more periods, when they reach a higher level of savings they tend to drop out of the labor force, hence, the value of accumulating human capital is very low for these women.

1.7.1 Validation of the Model

In this subsection I will now describe the validation test of the model. The model is meant to capture the incentives of households to specialize in labor market activities and in household good production. Additionally, the model captures the uncertainty households face through the inherent risk in the labor market and their limited ability to smooth income shocks through credit. Finally the framework allows me to capture the different value to workers of jobs with different characteristics. The estimated parameters of the model describe the trade-offs that families face when deciding their optimal allocation of resources. However, none of the moments used in estimation are directly related to the added worker effect. Thus, it seems a sensible idea to test the model in this dimension. I present below the estimation procedure of the added worker effect on the simulated data.

The estimation procedure in the simulated data is designed to be as comparable as possible to the regressions done on the actual survey data. First of all, I restrict the sample to young workers, that is, household heads between the ages of 25 to 40. To replicate the structure of the survey data, I gather the relevant information of the simulated agents for five consecutive quarters and ignore the information of the following periods. Then, I condition the sample of simulated agents in the same way that is done for the estimation of the AWE in the data. Namely, I condition on the sample on wives that are OLF and husbands that are employed in the first period. As in the reduced form evidence I collapse all the pertinent information into one observation per household, and finally I perform the same regression as in the reduced form evidence. Results are presented in Table 14.

The model performs relatively well capturing the actual AWE, that is, the effect of husband unemployment on the marginal probability of wives entering the labor force. However, it under predicts the entry rates of wives whose husbands did not become unemployed.

One additional result presented in this section, is the estimation of the AWE on the simulated data including additional explanatory variables. In this regression I can explicitly include savings as a regressor and infer the role it plays in substituting (or crowding out) the AWE. Results of this regression are presented in Table 15.

The results of this regression highlight the important role that savings play as a self-insurance mechanism for the household. For every additional unit of savings (1000's of 2000 Mexican Peso) the entry rate of wives decreases by 2.3%. Thus, households with higher savings are less likely to have a secondary worker entering the labor force. This reflects not only that wives are less likely to enter when husbands suffer some negative income shock, but also that they are less likely to enter when offered an attractive wage in the unstructured sector.

1.7.2 Policy Simulations

In this section, I use the initial parameter estimates to discuss two policy simulations. One policy simulation will offer unemployment insurance to workers who had their job destroyed in the structured sector. I will vary the replacement rates from 5% to 60% of the expected earnings of the husband. This policy simulation provides further evidence to the argument in Cullen and Gruber (2000) highlighting the important crowd-out effects of unemployment insurance on the measured added worker effect. The second policy will shut down the unstructured sector, and evaluate the welfare effects on different segments of the population according to their level of assets. This type of policy is often debated in developing countries which seek to reduce the level of informality in the economy. The informal sector comprises jobs that fall into the structured sector as well as the unstructured sector; many factories and small businesses avoid paying taxes for a large part of their workforce. However, monitoring informality in the structured sector is much more costly and difficult to implement than monitoring of the unstructured sector. Due to this difference in monitoring costs, governments have often resorted to policies that police and punish workers in the unstructured sector.

Unemployment Insurance

The first policy simulation refers to the provision of unemployment insurance at varying levels of replacement rates. In essence, I use the preliminary estimated parameters of the model and simulate the decisions of households when workers of the structured sector are offered unemployment insurance when their jobs are destroyed. This policy simulation allows me to measure the effect of unemployment insurance on the labor supply decisions of the households. In particular, it allows me to measure the crowd-out effect of unemployment insurance on the measured AWE.

Specifically, the policy consists of offering unemployment insurance to workers who lost their jobs in the structured sector. I vary the replacement rates of UI, from 5% to 60% of the predicted wage ($UI^i = x * (\alpha_1 + \alpha_2 * HK_t^i + \alpha_3 * (HK_t^i)^2)$ for $x = 0.05, 0.10, ..., 0.60$). For each level of generosity of the unemployment insurance, I re-estimate the AWE regression in the same way as done in the validation exercise. Figure 2 presents the results.

Results clearly show there is a significant crowd-out effect of UI. Already at 45% replacement rate, the AWE becomes statistically insignificant. Further increasing the generosity only drives the AWE closer to zero. Also noteworthy is that for replacement rates below 10%, unemployment insurance has almost no effect. However between 10 and 60 percent, there is clear and monotonic drop in the measured AWE as I increase the generosity of the UI. For comparison, consider that the average replacement rates for the U.S. in 2010 is 46.2%, although there exists significant variation by states; Alaska has the lowest replacement rates at 33% and Hawaii has the highest at 56.5%[4].

The results lend support to the argument provided in Cullen and Gruber (2000). Unemployment Insurance plays a very important role in the labor supply decisions of households. Another interesting point we can take from this policy simulation is the level of insurance that is provided by the secondary workers of the household. Back of the envelope calculations using the reduced form results, suggested that wives were providing about 30 to 40 percent of the earnings lost by the husband in the period of unemployment. The results of the policy simulation indicate that when offered 20 to 45 percent of the past earnings, the AWE is dramatically reduced. This result seems to be consistent with the implied level of insurance of the reduced

[4] Statistics from the Department of Labor for 2010. Available at http://www.doleta.gov/unemploy/chartbook.cfm

form results. Moreover, wives are willing to accept lower levels of income through UI, because wives do not have to enter the labor force, allocating time to labor market activities, and instead they can allocate time to increase the level of household public good and consume more leisure.

Shutting Down the Unstructured Sector

The second policy simulation tries to capture the value of the unstructured sector to households. This sector provides households with accessible and flexible job opportunities. Both of this characteristics are an important part in determining the ability of households to self-insure through increased labor supply. The fact that there is no entry cost allows secondary workers to respond promptly when faced with an adverse shock to household income. The part-time schedule allows secondary earners the ability to continue providing household public good and supply labor market hours.

However, the value of the unstructured sector is not limited to the insurance value they provide to households. For families with low level of savings and low levels of human capital (low levels of earnings) the unstructured sector is a stable source of employment for secondary earners. Poorer households have both members participating actively and continuously in the labor force. Thus, the value of the unstructured sector for these types of families is more important than just providing insurance.

Finally, there is an indirect insurance value provided by the unstructured sector. Occasionally, workers are presented with good job opportunities in the unstructured sector. This refers to high wage offers resulting from a large and positive wage shock $\eta_{unstructured}$. This good opportunities allow the households to sporadically increase the labor supply and increase family savings. As mentioned above, holding assets is a self-insurance mechanism in itself.

The simulation results I present considers the effects of a policy that would shutdown the unstructured sector. By shutting down the unstructured sector, families will be affected in their ability to self-insure through increased labor supply. However, as noted above, the impact of this policy is likely to be very different across households. To capture the heterogeneity of the effects I divide the sample of simulated agents into three groups according to their levels of household savings. To measure the value provided by the unstructured sector, I solve the model and simulate the household decisions without the unstructured sector and measure the average welfare for each group for a five year period. Then, I solve for the level of income transfers for each group (transfers every period for three years) that equalize the average welfare in each group. This provides a rough measure of the value that the unstructured sector provides to each group.

Results are presented in Table 16. I present the resulting income transfer as a percentage of the expected wage in the unstructured sector (i.e. Income transfer $= x * \varpi_{it}^{unstructured}$)

The results highlight the heterogenous role that the unstructured sector provides to households. For the poorest households, the income transfer needed to equate the welfare with and without the unstructured are very large. Essentially, more than 3/4 of the expected earnings in the unstructured sector have to be transferred to the poorest families every period to make them indifferent. This is due to the fact that poor households have both members of the household supplying labor market hours. Generally the husband is employed in the structured sector and the wife is employed in the unstructured sector. Hence, for a large part of this group, the value of the unstructured is more than just insurance, it is an important part of their livelihood.

For the "middle wealth" group, the role of the unstructured sector is vastly different. It is in this group that the AWE and sporadic entry is most prevalent. Wives in this group are generally out of the labor force specializing in household services and enter only occasionally into the structured sector. Entry is motivated by adverse income shocks to the husband's earnings as well as by good job opportunities presented in the unstructured sector. The income transfer is only 27% of the wage they would expect to earn in the unstructured sector.

Finally, for the higher wealth group, the income transfer is about 8%. Families in this group generally resort to savings when faced with negative income shocks. Wives are generally out of the labor force continuously. However, the income transfer is not 0% for two reasons. First, families around the 66th percentile (bottom of the wealthiest households) actually still benefit from the insurance value of the unstructured sector, similar to the middle wealth group. Second, families at the very top of the savings distribution have very different labor supply decisions. Specifically, in the richest households, husbands drop out of the structured sector and work in the unstructured sector. The reversal in the labor supply decisions has to do with the inflexibility of work hours in the structured sector. Families with very high level of savings have low marginal utilities of consumption, and thus are in less need of increasing their income or their savings. However, the marginal utility of leisure and household good are unaffected by the level of savings, therefore, husbands are willing to sacrifice higher levels of earnings in the structured sector for the increase in available hours when working in the unstructured sector.

1.8 Conclusions

This paper examines the added worker effect in the context of a developing country. Mexico presents the ideal conditions for studying the response of wives to unemployment of the husbands for four main reasons. First, households conform to the traditional view of primary and secondary workers within a household, 60% of households have the household head participating in the labor force actively and permanently, and wives are generally out of the labor devoting time to household good services. Second, as in most developing countries, there is limited access to credit thereby limiting the ability of households to smooth consumption by borrowing against future income. Third, there is no unemployment insurance in Mexico, and thus there is no concern of social insurance programs crowding out the added worker effect. Finally, the high macroeconomic volatility of the economy, translates into uncertain labor market outcomes for households.

The reduced form evidence presented in the paper, uncovers a large and significant added worker effect. Wives whose husband suffered an unemployment event, are 11 percentage points more likely to enter the labor force than wives whose husband did not lose his job. The entry rate is thus 60% higher as a result of the unemployment spell. Additionally, the results on the sample of workers that lost their job due to plant closing suggest that OLS results may be downward biased, and the actual AWE may be 20% larger, around 12 percentage points. Furthermore, the analysis revealed the transitory nature of the added worker effect, where 75% of the wives that entered appear to voluntarily exit the labor force. Another interesting feature of the analysis, is that wives who enter transitorily tend to have jobs in what is deemed as the unstructured sector. That is, occupations such as domestic employees, street vendors, in-home production of food and clothes; jobs that require low levels of skills, have low entry costs, and are readily available to wives facing tough economic conditions.

The model developed and estimated in this paper captures five important factors that influence the labor supply decisions within the household. First, the model captures the trade-offs families face when deciding how to allocate time to leisure, labor market supply and household good production services. Second, the model includes human capital accumulation which jointly with the demand for time from the household good production technology characterizes the incentives for specialization within the household. Third, by explicitly modelling savings decisions and borrowing constraints, the model is able to capture the limited ability of families to smooth income shocks through credit. Fourth, the uncertainty that households face in the labor market are captured in the model by having job destruction shocks as well as shocks to the wages offered to workers. Finally, the framework includes a choice of sector of employment that allows me to capture the different value of the job opportunities available to the different types of workers.

The model is estimated by Indirect Inference and the model is then validated through an out of sample test; the model accurately captures the marginal effect of unemployment on the wives' probability of entering the labor force (the AWE), but it under predicts the entry rate of wives whose husband didn't suffer an unemployment spell. The estimated parameters of the model are used to perform two counter factual policies: offering unemployment insurance to those workers in the structured sector, and second, shutting down the unstructured sector. Results of the counter factual policies suggest that in accordance to the point made by Cullen and Gruber (2000), there is a strong crowd out effect of unemployment insurance. The AWE measured in the simulated data, virtually disappears when the unemployment insurance reaches a 45% replacement rate. On the other hand, the value of the unstructured sector is revealed to be heterogenous across families. For households with low level of savings, the unstructured sector is frequently used to increase the level of earnings in the household and thus is highly valuable. For families in the middle third of the savings distribution, the value of the unstructured sector is much lower, as wives are generally out of the labor force and only participate when faced with an unemployment event or a sharp reduction in earnings of the primary earner of the household. Finally, for households at the top of the savings distribution, the unstructured sector has very little value, as these families mostly rely on savings to smooth consumption when faced with adverse income shocks.

Table 1: Summary Statistics on the Sample of Married Women and Men

	Men	Women
Age	39.88	37.15
Education years	9.03	8.30
LFP rates (cross-section)	95.17	35.99
LFP rates (longitudinal)	98.38	52.69
Hours of Work	44.80	12.19
Hours of Work (conditional)	47.16	34.99
Hours in Home Production	6.38	41.60

Table 2: Marginal Effect of Husband's Unemployment on the Probability of Wives Entering the Labor Force

	(I) Wife Entry	(II) Wife Entry	(III) Wife Entry	(IV) Wife Entry	(V) Wife Entry
Husband Unemployed	0.110***	0.114***	0.114***	0.113***	0.112***
	(0.00766)	(0.00760)	(0.00758)	(0.00757)	(0.00756)
Wife Age		0.0178***	0.00832***	0.0150***	0.0151***
		(0.000843)	(0.000998)	(0.00132)	(0.00131)
Wife $Age^2/100$		-0.0267***	-0.0160***	-0.0221***	-0.0223***
		(0.00108)	(0.00127)	(0.00166)	(0.00166)
Wife Education		-0.000403	0.00220	0.0163***	0.0164***
		(0.00133)	(0.00134)	(0.00163)	(0.00163)
I(children under 6 yrs)			-0.0309***	-0.0344***	-0.0342***
			(0.00261)	(0.00261)	(0.00261)
I(children 6-12 yrs)			0.00584**	0.00471**	0.00480**
			(0.00237)	(0.00238)	(0.00238)
I(children12-18 yrs)			0.0369***	0.0370***	0.0369***
			(0.00258)	(0.00260)	(0.00260)
Husband Age				-0.00696***	-0.00706***
				(0.00135)	(0.00135)
Husband $Age^2/100$				0.00559***	0.00570***
				(0.00162)	(0.00162)
Husband Education				-0.0215***	-0.0216***
				(0.00142)	(0.00142)
Av. Unemployment					0.793***
					(0.0595)
Constant	0.191***	-0.0787***	0.106***	0.146***	0.0750***
	(0.00952)	(0.0186)	(0.0207)	(0.0241)	(0.0247)
Observations	161,975	161,975	161,975	161,975	161,975

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Marginal Effect of Husband's Unemployment on the Probability of Wives Entering the Labor Force
Effect of Children's Age on the Added Worker Effect

	Wife Entry
Husband Unemployed	0.0754***
	(0.0143)
Wife Age	0.0151***
	(0.00131)
Wife Age^2/100	-0.0223***
	(0.00166)
Wife Education	0.0164***
	(0.00163)
I(children under 6 yrs)	-0.0341***
	(0.00263)
I(children 6-12 yrs)	0.00342
	(0.00240)
I(children12-18 yrs)	0.0363***
	(0.00262)
I(children under 6 yrs)$*$ Husband Unemployed	-0.00362
	(0.0159)
I(children 6-12 yrs)$*$ Husband Unemployed	0.0565***
	(0.0153)
I(children 12-18 yrs)$*$ Husband Unemployed	0.0220
	(0.0161)
Husband Age	-0.00705***
	(0.00135)
Husband Age^2/100	0.00570***
	(0.00162)
Husband Education	-0.0216***
	(0.00142)
Av. Unemployment	0.793***
	(0.0595)
Constant	0.0752***
	(0.0247)
Observations	161,975

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Marginal Effect of Husband Unemployment due to Plant Closing on the Probability of Wives Entering the Labor Force

	(I) Wife Entry	(II) Wife Entry	(III) Wife Entry
Husband Unemployed	0.0938***		
	(0.0169)		
Plant Closing		0.116***	
		(0.0364)	
Plant Closing (1 month)			0.119**
			(0.0545)
Wife Age	0.00450	0.00448	0.00452
	(0.00286)	(0.00286)	(0.00286)
Wife $Age^2/100$	-0.0101***	-0.00999***	-0.0101***
	(0.00368)	(0.00367)	(0.00367)
Wife Education	0.0145***	0.0145***	0.0144***
	(0.00290)	(0.00291)	(0.00291)
I(children under 6 yrs)	-0.0226***	-0.0225***	-0.0224***
	(0.00580)	(0.00580)	(0.00580)
I(children 6-12 yrs)	0.00235	0.00254	0.00246
	(0.00510)	(0.00511)	(0.00511)
I(children12-18 yrs)	0.0366***	0.0364***	0.0364***
	(0.00565)	(0.00565)	(0.00565)
Husband Age	-0.000204	-0.000428	-0.000452
	(0.00286)	(0.00286)	(0.00286)
Husband $Age^2/100$	-0.000211	9.73e-05	0.000145
	(0.00349)	(0.00349)	(0.00349)
Husband Education	-0.0158***	-0.0159***	-0.0159***
	(0.00283)	(0.00283)	(0.00283)
Av. Unemployment	0.525***	0.538***	0.537***
	(0.163)	(0.163)	(0.163)
Constant	0.0927*	0.0964*	0.0960*
	(0.0499)	(0.0499)	(0.0499)
Observations	29,123	29,123	29,123

Robust standard errors in parentheses

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

Table 5: Marginal Effect of Husband's Re-employment on the Probability of Wives Exiting the Labor Force

	(I) Wife Exit	(II) Wife Exit	(III) Wife Exit	(IV) Wife Exit	(V) Wife Exit
Husband Re-employed	0.121***	0.119***	0.120***	0.124***	0.127***
	(0.0293)	(0.0291)	(0.0291)	(0.0293)	(0.0292)
Wife Age		-0.0119	0.00602	0.00350	0.00436
		(0.0114)	(0.0133)	(0.0171)	(0.0172)
Wife Age^2/100		0.0217	-0.00112	-0.000837	-0.00197
		(0.0147)	(0.0169)	(0.0213)	(0.0214)
Wife Education		-0.0182	-0.0267	-0.0232	-0.0222
		(0.0167)	(0.0168)	(0.0208)	(0.0208)
I(children under 6 yrs)			0.000473	-0.00211	-0.00333
			(0.0314)	(0.0317)	(0.0316)
I(children 6-12 yrs)			-0.0428	-0.0443	-0.0447
			(0.0288)	(0.0294)	(0.0294)
I(children12-18 yrs)			-0.0715**	-0.0686**	-0.0714**
			(0.0308)	(0.0311)	(0.0312)
Husband Age				0.00224	0.00295
				(0.0164)	(0.0165)
Husband Age^2/100				0.000227	-0.000686
				(0.0196)	(0.0196)
Husband Education				-0.00376	-0.00385
				(0.0179)	(0.0179)
Av. Unemployment					1.012
					(0.628)
Constant	0.446***	0.612**	0.372	0.520*	0.384
	(0.107)	(0.244)	(0.272)	(0.306)	(0.321)
Observations	1,458	1,458	1,458	1,458	1,458

Robust standard errors in parentheses

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

Table 6: Summary Statistics on Characteristics of Employment

	Men	Working wives	Entry wives	AWE wives
Hours	44.82	37.84	24.61	23.92
Home Prod.	6.08	28.50	36.33	36.86
Earnings	3825	3686	1540	1516
Hourly Wage	22.71	23.40	16.14	15.97
Type of Jobs				
Business Owner	8.91	2.56	2.39	0.72
Self-Employed	16.31	13.46	42.57	43.28
Pieceworkers	8.72	2.99	12.20	11.75
Salaried	66.06	81.00	42.90	44.25
% Formal	66.89	77.38	24.21	19.26

Table 7: Marginal Effect of Husband's Unemployment on the Probability of Wives Entering the Labor Force Controlling for Past Labor Force Experience of Wives

	(I) Wife Entry	(II) Wife Entry	(III) Wife Entry	(IV) Wife Entry	(V) Wife Entry
Husband Unemployed	0.109***	0.111***	0.111***	0.111***	0.109***
	(0.00755)	(0.00750)	(0.00748)	(0.00747)	(0.00746)
I(Never Worked)	-0.115***	-0.117***	-0.121***	-0.120***	-0.119***
	(0.00396)	(0.00405)	(0.00405)	(0.00406)	(0.00405)
I(Last job within 1 year)	0.147***	0.144***	0.140***	0.139***	0.139***
	(0.00597)	(0.00595)	(0.00595)	(0.00594)	(0.00594)
I(Last job btwn 3-5 years)	-0.0690***	-0.0689***	-0.0643***	-0.0632***	-0.0631***
	(0.00539)	(0.00537)	(0.00537)	(0.00536)	(0.00536)
I(Last job over 5 years)	-0.114***	-0.120***	-0.124***	-0.121***	-0.121***
	(0.00391)	(0.00399)	(0.00401)	(0.00401)	(0.00401)
Wife Age		0.0243***	0.0134***	0.0172***	0.0174***
		(0.000846)	(0.000993)	(0.00130)	(0.00130)
Wife Age^2/100		-0.0332***	-0.0202***	-0.0234***	-0.0236***
		(0.00108)	(0.00126)	(0.00164)	(0.00164)
Wife Education		-0.00395***	-0.000875	0.0121***	0.0121***
		(0.00132)	(0.00134)	(0.00161)	(0.00161)
I(children under 6 yrs)			-0.0257***	-0.0286***	-0.0284***
			(0.00258)	(0.00259)	(0.00258)
I(children 6-12 yrs)			0.0201***	0.0184***	0.0185***
			(0.00237)	(0.00238)	(0.00238)
I(children12-18 yrs)			0.0378***	0.0371***	0.0371***
			(0.00255)	(0.00257)	(0.00257)
Husband Age				-0.00323**	-0.00334**
				(0.00134)	(0.00134)
Husband Age^2/100				0.00191	0.00202
				(0.00160)	(0.00160)
Husband Education				-0.0197***	-0.0198***
				(0.00140)	(0.00140)
Av. Unemployment					0.785***
					(0.0589)
Constant	0.282***	-0.122***	0.0726***	0.0902***	0.0198
	(0.0100)	(0.0186)	(0.0207)	(0.0241)	(0.0246)
Observations	161,975	161,975	161,975	161,975	161,975

Robust standard errors in parentheses

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

Table 8: Marginal Effect of Husband's Unemployment on the Probability of Wives Entering the Labor Force Within Group Comparison by Past Labor Force Experience

	Wife Entry
Husband Unemployed	0.111***
	(0.0236)
I(Never Worked)	-0.119***
	(0.00410)
I(Last job within 1 year)	0.138***
	(0.00602)
I(Last job btwn 3-5 years)	-0.0646***
	(0.00541)
I(Last job over 5 years)	-0.121***
	(0.00405)
I(Never Worked)$*$ Husb. Unemployed	-0.0144
	(0.0266)
I(Last job within 1 year)$*$ Husb. Unemployed	0.0126
	(0.0375)
I(Last job btwn 3-5 years)$*$ Husb. Unemployed	0.0715*
	(0.0390)
I(Last job over 5 years)$*$ Husb. Unemployed	-0.00584
	(0.0263)
Wife Age	0.0174***
	(0.00130)
Wife Age^2/100	-0.0236***
	(0.00164)
Wife Education	0.0121***
	(0.00161)
I(children under 6 yrs)	-0.0284***
	(0.00258)
I(children 6-12 yrs)	0.0185***
	(0.00238)
I(children12-18 yrs)	0.0371***
	(0.00257)
Husband Age	-0.00332**
	(0.00134)
Husband Age^2/100	0.00200
	(0.00160)
Husband Education	-0.0198***
	(0.00140)
Av. Unemployment	0.785***
	(0.0589)
Constant	0.0201
	(0.0246)
Observations	161,975

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Marginal Effect of Husband's Unemployment on the Probability of Wives Exiting the Labor Force
Test for Spurious Correlation

	(I) Wife Exit	(II) Wife Exit	(III) Wife Exit	(IV) Wife Exit	(V) Wife Exit
Husband Unemployed	-0.000200	-0.0211***	-0.0212***	-0.0215***	-0.0208***
	(0.00286)	(0.00275)	(0.00274)	(0.00275)	(0.00275)
Wife Age		-0.0192***	-0.0159***	-0.0228***	-0.0228***
		(0.000249)	(0.000278)	(0.000478)	(0.000478)
Wife Age^2/100		0.0216***	0.0186***	0.0233***	0.0233***
		(0.000262)	(0.000290)	(0.000493)	(0.000493)
Wife Education		-0.113***	-0.114***	-0.110***	-0.110***
		(0.000792)	(0.000797)	(0.000982)	(0.000982)
I(children under 6 yrs)			0.0299***	0.0323***	0.0323***
			(0.00161)	(0.00161)	(0.00161)
I(children 6-12 yrs)			0.00110	0.00144	0.00144
			(0.00144)	(0.00143)	(0.00143)
I(children12-18 yrs)			-0.0178***	-0.0201***	-0.0201***
			(0.00151)	(0.00151)	(0.00151)
Husband Age				0.00807***	0.00807***
				(0.000476)	(0.000476)
Husband Age^2/100				-0.00487***	-0.00487***
				(0.000462)	(0.000462)
Husband Education				0.00252***	0.00254***
				(0.000871)	(0.000871)
Av. Unemployment					-0.196***
					(0.0348)
Constant	0.869***	1.508***	1.424***	1.325***	1.342***
	(0.00616)	(0.00827)	(0.00893)	(0.00951)	(0.0100)
Observations	313,610	313,610	313,610	313,610	313,610

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Marginal Effect of Husband's Unemployment on the Probability of Wives Entering the Labor Force
Estimates for Workers in the Lowest Decile of the Income Distribution

	(I)	(II)	(III)	(IV)	(V)
	Wife Entry	Wife Entry	Wife Entry	Wife Entry	Wife Entry
Husband Unemployed	0.113***	0.112***	0.110***	0.111***	0.109***
	(0.0227)	(0.0225)	(0.0225)	(0.0225)	(0.0225)
I(Never Worked)	-0.120***	-0.121***	-0.125***	-0.126***	-0.126***
	(0.0127)	(0.0129)	(0.0129)	(0.0129)	(0.0129)
I(Last job within 1 year)	0.136***	0.133***	0.128***	0.127***	0.128***
	(0.0184)	(0.0183)	(0.0183)	(0.0183)	(0.0183)
I(Last job btwn 3-5 years)	-0.0646***	-0.0610***	-0.0546***	-0.0543***	-0.0541***
	(0.0181)	(0.0180)	(0.0180)	(0.0180)	(0.0180)
I(Last job over 5 years)	-0.104***	-0.112***	-0.115***	-0.116***	-0.116***
	(0.0129)	(0.0130)	(0.0131)	(0.0131)	(0.0131)
Wife Age		0.0273***	0.0169***	0.0126***	0.0129***
		(0.00252)	(0.00295)	(0.00388)	(0.00388)
Wife Age^2/100		-0.0361***	-0.0236***	-0.0184***	-0.0187***
		(0.00319)	(0.00372)	(0.00482)	(0.00482)
Wife Education		0.00305	0.00575	0.00624	0.00618
		(0.00478)	(0.00482)	(0.00533)	(0.00532)
I(children under 6 yrs)			-0.0315***	-0.0315***	-0.0311***
			(0.00824)	(0.00826)	(0.00826)
I(children 6-12 yrs)			0.0309***	0.0298***	0.0295***
			(0.00767)	(0.00771)	(0.00771)
I(children12-18 yrs)			0.0306***	0.0291***	0.0286***
			(0.00841)	(0.00847)	(0.00846)
Husband Age				0.00695*	0.00695*
				(0.00392)	(0.00392)
Husband Age^2/100				-0.00811*	-0.00811*
				(0.00466)	(0.00466)
Husband Education				0.000529	0.000446
				(0.00506)	(0.00505)
Av. Unemployment					0.954***
					(0.183)
Constant	0.327***	-0.156***	0.0283	-0.0283	-0.120
	(0.0329)	(0.0582)	(0.0641)	(0.0729)	(0.0748)
Observations	16,227	16,227	16,227	16,227	16,227

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Marginal Effect of Husband's Under-Employment on the Probability of Wives Entering the Labor Force

	(I) Wife Entry	(II) Wife Entry	(III) Wife Entry	(IV) Wife Entry	(V) Wife Entry
Husband Under-employed	0.0348***	0.0411***	0.0412***	0.0326***	0.0319***
	(0.00450)	(0.00451)	(0.00451)	(0.00455)	(0.00455)
Wife Age		0.0177***	0.00833***	0.0150***	0.0152***
		(0.000844)	(0.000998)	(0.00132)	(0.00132)
Wife $Age^2/100$		-0.0266***	-0.0160***	-0.0221***	-0.0223***
		(0.00108)	(0.00127)	(0.00166)	(0.00166)
Wife Education		0.000720	0.00328**	0.0165***	0.0166***
		(0.00133)	(0.00135)	(0.00163)	(0.00163)
I(children under 6 yrs)			-0.0313***	-0.0347***	-0.0345***
			(0.00261)	(0.00261)	(0.00261)
I(children 6-12 yrs)			0.00567**	0.00472**	0.00481**
			(0.00237)	(0.00238)	(0.00238)
I(children12-18 yrs)			0.0366***	0.0368***	0.0368***
			(0.00258)	(0.00260)	(0.00260)
Husband Age				-0.00711***	-0.00721***
				(0.00135)	(0.00135)
Husband $Age^2/100$				0.00578***	0.00589***
				(0.00162)	(0.00162)
Husband Education				-0.0206***	-0.0208***
				(0.00142)	(0.00142)
Av. Unemployment					0.804***
					(0.0595)
Constant	0.195***	-0.0767***	0.108***	0.149***	0.0765***
	(0.00953)	(0.0186)	(0.0207)	(0.0241)	(0.0247)
Observations	161,975	161,975	161,975	161,975	161,975

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 12: Marginal Effect of Husband's Unemployment on the Probability of Wives Entering the Labor Force Controlling for the Duration of the Unemployment Spell

	(I) Wife Entry	(II) Wife Entry
Husband Unemployed (1 quarter)	0.105***	0.105***
	(0.00814)	(0.00814)
I(Never Worked)	-0.119***	-0.119***
	(0.00406)	(0.00405)
I(Last job within 1 year)	0.139***	0.139***
	(0.00595)	(0.00594)
I(Last job btwn 3-5 years)	-0.0636***	-0.0631***
	(0.00537)	(0.00536)
I(Last job over 5 years)	-0.121***	-0.121***
	(0.00401)	(0.00401)
Wife Age	0.0173***	0.0174***
	(0.00130)	(0.00130)
Wife Age^2/100	-0.0235***	-0.0236***
	(0.00164)	(0.00164)
Wife Education	0.0121***	0.0121***
	(0.00161)	(0.00161)
I(children under 6 yrs)	-0.0283***	-0.0284***
	(0.00259)	(0.00258)
I(children 6-12 yrs)	0.0182***	0.0185***
	(0.00238)	(0.00238)
I(children12-18 yrs)	0.0373***	0.0371***
	(0.00257)	(0.00257)
Husband Age	-0.00333**	-0.00333**
	(0.00134)	(0.00134)
Husband Age^2/100	0.00203	0.00202
	(0.00160)	(0.00160)
Husband Education	-0.0198***	-0.0198***
	(0.00140)	(0.00140)
Husband Unemployed (2 quarters)		0.0131
		(0.0218)
Husband Unemployed (3 quarter)		0.0248
		(0.0427)
Husband Unemployed (4 quarter)		0.426***
		(0.104)
Av. Unemployment	0.784***	0.784***
	(0.0590)	(0.0589)
Constant	0.0231	0.0201
	(0.0247)	(0.0246)
Observations	161,295	161,975

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Parameter Estimates and Matched Moments

Parameter	Moment	Model	Data
λ	% job destroyed	8.61	7.58
π	av. hours in HP	53.03	47.95
$\varpi_{it}^{\text{unstr}}$	constant in wage regression	13.12	11.57
$\eta_{unstructured}$	var of residuals of wage regression	7.34	13.42
α_1	constant in wage regression	45.12	52.65
α_2	return to experience in wage regression	0.66	0.52
α_3	return to experience^2 in wage regression	-0.019	-0.03
$\eta_{structured}$	var of residuals of wage regression	16.15	23.42
	% of workers in the unstructured sector	18.12	11.57

Table 14: Validation of the Model: Linear Probability Estimates of the AWE: Simulation vs Data

	Simulation Wife Entry	Data Wife Entry
Husband Unemployed	0.142***	0.115***
	(0.021)	(0.007)
Constant	0.041***	0.188***
	(0.003)	(0.013)
Observations	11,374	137,859

Table 15: AWE Estimates using Simulated Data
The Role of Household Savings

	Wife Entry
Husband Unemployed	0.073**
	(0.037)
Age	-0.013
	(0.012)
Age^2	-0.001
	(0.002)
HH savings / 1000	-0.023***
	(0.001)
Constant	0.253***
	(0.14)
Observations	11,374

Table 16: Policy Simulation Results: Income transfers that equate average welfare with and without the unstructured sector

	Average Transfer
Poorest 1/3	77%
Middle Wealth	27%
Wealthiest 1/3	8%

Figure 1: Patterns of Labor Force Participation of Married Women
Number of Quarters in the Labor Force

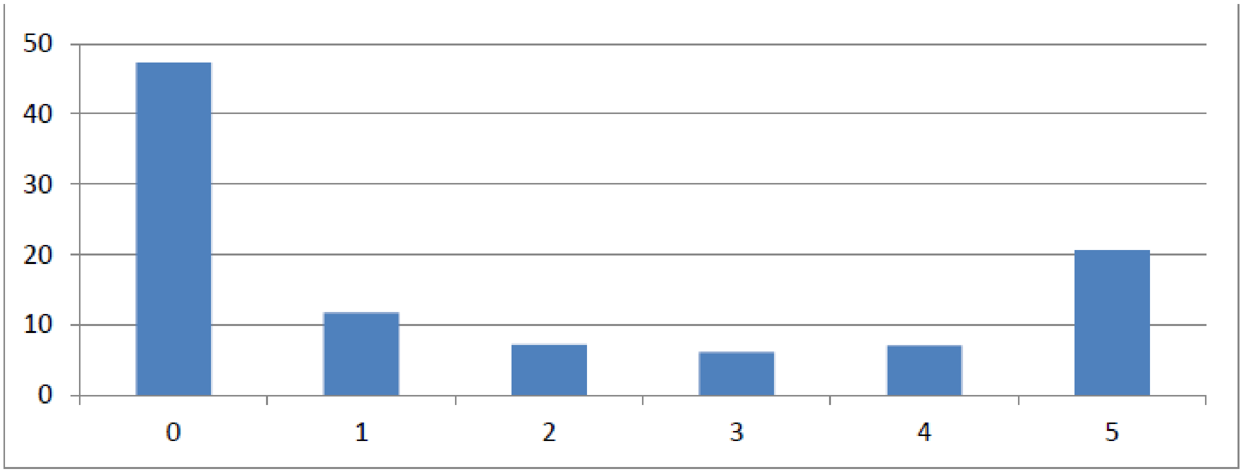
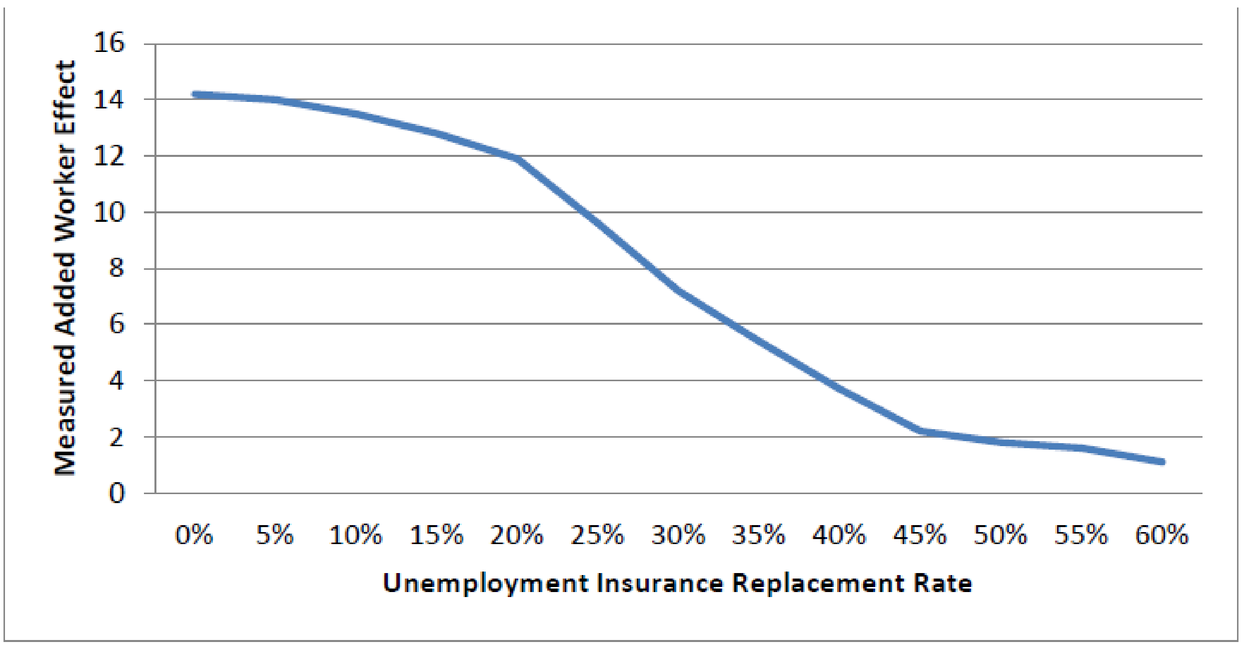


Figure 2: The Crowd Out Effect of Unemployment Insurance on the Added Worker Effect



1.9 References

1. Ashenfelter, O., (1980) "Unemployment as Disequilibrium in a Model of Aggregate labor Supply," *Econometrica*, 48:547-64

2. Cullen, J. and J. Gruber (2000) "Does Unemployment Insurance Crowd out Spousal Labor Supply?," *Journal of Labor Economics*, 18(3):546-572

3. Cunningham, W., "Breadwinner or Caregiver", Policy Research Working Paper #2743, The World Bank

4. Dynarski, S., and J. Gruber (1997) "Can Families Smooth Variable Earnings?," *Brooking Papers on Economic Activity*, 1:1997, 229-84

5. Heckman, J. and T. MaCurdy (1980) "A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 47, 1, 47-74

6. Heckman, J. and T. MaCurdy (1982) "Corrigendum on A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 49,659-660

7. Horvath, F. (1982) "Forgotten Unemployment: recall bias in retrospective data," *Monthly Labour Review*, March 40-44

8. Humphrey, D (1940) "Alleged 'Additional Workers' in the Measurement of Unemployment," *Journal of Political Economy*, 30:412-419

9. Keane, M. and K. Wolpin (1994). "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics,* 76 (4): 648-672

10. Killingsworth, M. and J. Heckman (1986) "Female Labor Supply: A Survey," in Handbook of Labor Economics, Volume I, eds. Orley Ashenfelter and Richard Layard. Amsterdam: Elsevier Science Publishers BV

11. Juhn, C., and K. Murphy (1997) "Wage Inequality and Family Labor Supply," *Journal of Labor Economics*, 15, 1, 72-97

12. Juhn, C., and S. Potter (2007) "Is there Still an Added Worker Effect?," Federal Reserve Bank of New York. Staff Report No. 310

13. Layard, R., M. Barton, and A. Zabalza (1980) "Married Women's Participation and Hours," *Economica*, 47, 51-72

14. Lundberg, S., (1985) "The Added Worker Effect," *Journal of Labor Economics*, 3, 1, 11-37

15. MaCurdy, T. (1981) "An Empirical Model of labor Supply in a Life Cycle Setting," *Journal of Political Economy*, 89, 6, 1059-1085

16. Maloney, T., (1987) "Employment Constraints and the Labor Supply of Married Women: A Reexamination of the Added Worker Effect," *Journal of Human Resources*, 22, 1, 51-61

17. Maloney, T., (1991) "Unobserved Variables and the Elusive Added Worker Effect," *Economica*, 58, 173-187

18. Maloney, W., (1998) "The Structure of Labor Markets in Developing Countries: Time series Evidence on Competing Views," Policy Research Working Paper #1940, The World Bank, Washington, D.C.

19. Mazzocco, M., "Individual Rather than Household Euler Equations: Identification and Estimation of Individual Preferences Using Household Data," Revise and Resubmit, *Journal of Political Economy*

20. Mincer, J., (1962) "Labor Force Participation of Married Women," in Aspects of Labor Economics, ed. H.G. Lewis. Princeton, N.J.: Princeton University Press.

21. Parker, S., and E. Skoufias (2004) "The Added Worker Effect over the Business Cycle: Evidence from Urban Mexico," *Applied Economic Letters*, 11:10, 625-630

22. Shore, S., (2007) "For better, For Worse: Intra-Household Risk Sharing over the Business Cycle," University of Pennsylvania Working Paper

23. Spletzer, J., (1992) "Reexamining the Added Worker Effect," *Economic Inquiry*, 35: 417-427

24. Stephens, M., (2002) "Worker Displacement and the Added Worker Effect," *Journal of Labor Economics*, 20:3, 504-537

25. Woytinsky, W.S. (1942) "Three Aspects of Labor Dynamics," Washington, D.C.: Committee on Social Security, Social Science Research Council

Chapter 2

Understanding Corruption

Theory and Evidence from the Audits of Local Governments

2.1 Introduction

The abuse of entrusted power by politicians through rent-seeking and corruption is a serious concern in much of the developing world. Developing countries have provided numerous examples of political elites diverting funds intended for basic public services such as health, schools, and roads.[1] As a result, corruption is often considered the single greatest obstacle for economic and social development, and several studies have documented a strong negative relationship between corruption and measures such as investment and growth (Mauro 1995). Yet despite its costs and the potentially large welfare losses associated with it, our understanding of what determines corruption and how we can reduce it remains rather limited.

The goal of this paper is to provide a framework for understanding the decision of politicians to engage in corrupt activities. We developed and estimate a structural model of a politician's decisions to provide public goods and engage in corruption over the span of his political career. By capturing the incentives and constraints that politicians face, our model allows us to distinguish between at least two key explanations

[1] For instance, the former President of Zaire, Mobutu Sese Seko, is believed to have embezzled $5 billion before being ousted in 1997. Also, it has been suggested that the former president of the Philippines Ferdinand Marcos has stolen over $35 billion dollars. See also Di Tella and Schargrodsky (2003) and Olken (2007).

for the large variation in corruption that we observe today: the utility that politicians derive for engaging in corruption and the punishment politicians receive if caught for corruption. Using this model, we evaluate the effectiveness of anti-corruption policies that increase politicians' wages, induce a higher probability of being audited by a central authority, or increases term limits.

While our model is general, we develop and estimate it in the context of municipal governments in Brazil. Local governments in Brazil provide an ideal institutional setting to understand how corruption is determined for at least two reasons. First, in many countries the most corrupt governments seem to be at the local level, where governments are under the control of narrow elites that use the apparatus for personal gain (Rose-Ackerman 1999). For instance, as reported in Table 1, municipalities in Brazil receive on average 1,497,998 Reals per year in order to provide such public services as education, health, and sanitation. With the large influx of federal funds and the potential for local capture, it is not surprising that corruption at the municipal level is now an overarching concern (VEJA 2004). According to our data on Brazil, corruption was discovered in 74 percent of all municipalities, where on average almost 8 percent of these federal funds were diverted (see rows 2 and 3 of Table 1). This translates into losses of approximately $600 millions in local governments per year. Second, in 2003 the Brazilian government introduced an anti-corruption program that randomly audited municipal governments for their use of federal funds. Based on these detailed audit reports, we estimate our model using objective measures of corruption for local governments, thus overcoming an important obstacle that has plagued much the existing literature on corruption (Svensson 2005).[2]

Using the estimated model we have simulated the effect of the three anti-corruption policies described above: (i) a raise in the salaries of politicians; (ii) an increase in term limits; (iii) an increase in the probability that a municipality is audited. Our main findings are as follows. The most effective anti-corruption policies are the term limits policy and the policy that increases the probability of an audit. These policies have larger effects than increasing the politicians' salaries for all municipalities. On average, the term limits policy generates a larger reduction in corruption than the audit policy. But their effect varies depending on the level of corruption. The increase in the probability of an audit is more effective in municipalities with high levels of corruption, whereas the term limit policy has better results in places where there is a limited

[2] Due to the inherent difficulties in measuring corruption, many previous studies have had to rely on subjective measures of corruption based on either perceptions or self-reported information. Unfortunately, these data rarely provide unbiased estimates and are often influenced by the characteristics of the survey respondents.

level of corruption.

Overall, our paper makes two main contributions to the existing literature. First, the proposed framework captures many of the various mechanisms by which politicians choose to engage in corruption, which enables us to assess empirically the relative importance of these different mechanisms. This lies in contrast to the previous literature whose empirical evidence has been mostly based on simple correlations. These estimates are often confounded by other unobserved determinants of corruption and provide only limited insights into the mechanisms that produce these associations. But it is precisely the understanding of these mechanisms that is critical for the design of policy as a redress for corruption. As a second contribution, we use the estimated model to evaluate the impact of policies that have been recently proposed by policy makers and economists as potential instruments for curbing corruption. While Ferraz and Finan (2008), in a related study, have estimated the effects Brazil's audit policy on the re-election rates of incumbent mayors, the study is unable to determine whether the policy actually reduced corruption - the program's principal objective. Using this framework, this paper overcomes this limitation.

There is a growing literature that has analyzed corruption and decisions by politicians.[3] Our paper is one of the first to use a structural approach to investigate these two issues. An important exception is the paper by Diermeir, Keane, and Merlo (2005), who estimate a structural model to quantify the returns to a career in the United States Congress.

2.2 Descriptive Evidence

In this section we describe the evidence on mayors' decisions that can be gathered from the data. This evidence is useful for two reasons. First, it enables us to clarify which factors have an impact on the behavior of politicians. Second, it can be used to test the model developed in the next section.

We start by describing the evidence on corruption. Corruption is evident in the data. In the period under investigation, on average mayors received $1,497,998$ Reals in funds and stole $85,908$ of them, which corresponds to 5.8% of the total. The data also indicate that there is heterogeneity in corruptions and that the distribution of stealing is skewed to the left. About 26% of audited mayors did not steal. The $50th$, $75th$,

[3] See for instance the surveys by Svensson (2005), Rose-Ackerman (2009), and Besley (2006).

and 90th percentiles of the amount stolen correspond to 24, 244, 96, 704, and 230, 408 Reals, respectively.

We will now discuss how stealing and the production of public consumption varies with the number of terms. Our findings indicate that there is a positive correlation between the number of terms and the amount of corruption. In the first term, average stealing is below the unconditional mean at 76, 973 Reals, whereas average stealing is significantly higher in the second term at 97, 790 Reals. The fraction of mayors who stole a positive amount of funds is also significantly different between the two terms. In the first term, this fraction is 72% and it increases at 77% in the second term. We observe similar patterns if we condition on the allocated funds, population size, private inputs, and education. Specifically, when we regress the logarithm of stealing on these variables and on the number of terms, we find that the coefficient on the number of terms is 0.395 with standard errors equal to 0.198. The production of public consumption also changes with the number of terms. When we regress the logarithm of public consumption on funds, population size, private inputs, education, and the number of terms, the coefficient on this last variable is -0.078 with standard errors equal to 0.019. This indicates that public consumption decreases on average by about 8% in the second term.

The data also indicate that corruption and public consumption vary with the size of the municipality. To report our findings we construct three groups of municipalities: small municipalities which have fewer than 10,000 inhabitants, medium municipalities which have between 10,000 and 60,000 inhabitants, and large municipalities which have more than 60,000 inhabitants. Using these three groups, we find that, after controlling for the amount of funds allocated to a municipality, on average mayors in small municipalities steal 31% less than mayors in large municipalities and mayors in medium municipalities steal 13% less than their counterpart in large municipalities. The production of public consumption also changes with the size of the municipality. Conditional on funds, small municipalities produce 71% more per-capita public consumption than large municipalities and medium municipalities produce 14% more than large municipalities.

We now provide evidence on the effect that stealing and the production of public consumption have on the electoral outcome. We do this by estimating a probit model in which we have as dependent variable the outcome of the election and as independent variables population size, age of the mayor, and the campaign contributions of the mayor relative to the challengers in addition to a dummy equal to one if the mayor has stolen and the amount of public consumption. We find that stealing has a negative and significant effect on the probability of being reelected. It reduces the probability of reelection by 17%. Per-capita consumption has the opposite effect. The coefficient on this variable is positive and statistically significant. For example, for a fifty years old mayor in a larger municipality who was not audited the probability of reelection increases by 8 percentage points if per-capita public consumption increases from the value that corresponds to the $25th$ percentile to the value that corresponds to the $75th$ percentile.

We conclude this section by describing the decision of incumbent mayors on whether to run for reelection. We find that a large number of mayors choose not to run. In our data, 28% of mayors decided not to run for a second term. Of the mayors that choose to run for reelection, 57.3% were reelected. This result suggests that there is an incumbent bias among voters which is not too large. We also looked at which variables have an impact on the mayor's decision of not running for a second term and we found that one variables has the largest effect: the amount stolen in the current term. As one might expect, the larger the amount stolen the less likely it is that the incumbent runs for reelection.

2.3 Model

In this section, we develop a model of a politician's decision to engage in corruption. Although our model is quite general, given the empirical analysis that follows, we consider a particular type of politician: a mayor. Overall, our framework captures six important factors that we believe determines the level of corruption in a particular economy. First, politicians that care more about public consumption are generally less corrupt. Second, the level of corruption depends on the return of one additional dollar of public funds invested in the production of public consumption. Corruption is generally lower if the marginal return is higher. Third, experience as a mayor generally has a positive effect on future wage offers unless the mayor was found to be corrupt. Fourth, wealthier mayors generally steal less. Fifth, fines and jail terms deter corruption. Lastly, mayors that plan to run for reelection steal less. In the next three subsections, we describe how we incorporate these insights into a model of mayors' decisions.

2.3.1 Preferences and Technology

The starting point for our model is that individuals care about the amount of public good available in a particular municipality. Some examples of public goods provided by municipalities include schools, police force, parks, and roads. The amount of public consumption produced in a municipality depends on the fraction of public funds invested in its production. This implies that, everything else equal, in municipalities with more corruption less public consumption is produced and therefore its residents experience lower levels of welfare.

Consider a municipality m populated by n individuals, all of whom live for T periods and are potential politicians. In each period t there is uncertainty in the municipality which is denoted by the state of nature ω. Each individual i is characterized by a common discount factor $\beta \in (0,1)$ and by preferences over a private good c, a public good Q, and a dummy equal to one if she is currently in power dp^i. The last variable enables us to consider the possibility that an individual derives utility from being in power. These preferences vary across individuals and are described by the following utility function:

$$u^i \left(c_t^i, Q_t, dp^i \right).$$

The heterogeneity in the utility functions is introduced to allow individuals to differ in their preferences for public consumption relative to private consumption. This feature of the model addresses the first determinant of corruption: municipalities governed by individuals that care more about public consumption relative to private consumption should enjoy higher levels of public good.

The public good is produced according to a municipal-specific production technology. It depends on inputs from the private sector z^{pr}, the amount of public funds invested in public consumption z^{pu}, the ability of the mayor governing the municipality a, as well as a vector X^m of municipal characteristics. One example of municipal characteristics that affect the production of Q_t is the size of the municipality. We denote the production function for public consumption by

$$f_m \left(z^{pu}, z^{pr}, a, X_m \right)$$

The production function captures the second determinant of corruption described above. The level of corruption depends on the marginal product of the public inputs.

Each individual owns $\bar{h}$ units of labor which they supply inelastically in return for a wage w. If an individual becomes a mayor, he receives a deterministic wage $\bar{w}$ set by the central government. Otherwise, wages are drawn from the distributions $f_{pm}(w|Z)$ if the individual had been a mayor in the past or $f_{nm}(w|Z)$ if the individual had never been a mayor, where Z denotes a vector of individual and municipal characteristics that determine local wages. The different wage process for past mayors is meant to incorporate the possibility that individuals that have served as mayors generally receive better wage offers. To capture the additional insight that past mayors who were found to be corrupt receive potentially lower wage offers, we will allow the mean of the wage distribution for past mayors to depend on the amount stolen if the mayor was caught stealing.

Individuals possess non-labor income, Y, and can save or borrow an amount b at an interest rate R. Since wages, non-labor income, and savings determine the wealth of mayors, these aspects of the model capture the idea that wealthier mayors are less likely to steal.

2.3.2 Mayor's Decisions

Mayors make two types of decisions. They first decide the amount of public funds z^{pu} to invest in the production of public consumption Q and the amount they intend to steal s. They then decide how to allocate their private resources, which include the amount stolen, between consumption c and savings b.

Each municipality is audited with probability p. To indicate that a municipality was audited we set the variable δ equal to 1. If the municipality is audited and the mayor has engaged in corruption, the amount stolen becomes public knowledge. In addition, the mayor must pay a penalty which is given by the fine schedule $g(s)$, where $g(s)$ is increasing in the amount stolen. We do not explicitly model jail terms. They are transformed into monetary payments and added to the fine schedule. The fines are meant to capture an additional potential determinant of corruption: lower fines increase corruption.

The mayor's decisions influence his probability of being re-elected in two ways. First, if a municipality was audited the voters know whether a mayor is corrupt and the amount stolen. As a result, voters are less likely to vote for the incumbent, where the probability of voting for the incumbent is decreasing in the amount stolen. Second, if the municipality was not audited, the voters only observe the level of public goods provision during the term that precedes the elections. The level of the public good is used to infer the ability of the incumbent and his preferences for public relative to private consumption. As is common in the political economy literature (e.g. Barro (1970), Ferejohn (1986)), we assume that voters adopt a retrospective voting strategy, whereby incumbents who provide public consumption above a particular threshold, Q^*, are rewarded with re-election. The threshold is determined endogenously and is a function of the following variables: the amount of public good produced by the municipality during the term; whether a mayor was audited and, conditional on being audited, the amount stolen; the campaign contribution received by the challengers relative to the one received by the incumbent r_c; an error term ϵ which captures the residual randomness. To model this election rule, we use a reduced-form formulation that incorporates these main features, i.e.

$$\text{the incumbent is reelected if} \qquad Q_t \geq Q_t^*,$$

where $Q^* = h(\delta, s, r_c, \epsilon)$. If individual i is elected we set the variable ρ_i to 1. This electoral rule captures the last determinant of corruption that we intend to model: mayors that plan to run for reelection steal less.

2.3.3 The Individual Decision Process

We are now ready to formally describe the decision process of individual i in municipality m. Individual i chooses the amount of private consumption and savings that maximizes his lifetime utility

$$E\left[\sum_{t=1}^{T} \beta_i^t u^i\left(c_t^i, Q_t\right)\right],$$

subject to the constraint that in each period and state of nature expenditure on consumption plus savings must equal the available resources,

$$c_t^i + b_t^i = w_t^i \bar{h}_t + 1_{\{\rho_t^i=1\}} s_t^i + R_t b_{t-1}^i + 1_{\{\rho_{t-1}^i=1, \delta_{t-1}^i=1\}} g\left(s_{t-1}^i\right) \qquad \text{for each } t \text{ and } \omega.$$

If individual i is the mayor, he also decides how much to steal, how much to invest in public consumption, and whether to run for mayor at the end of the current term. Moreover, his decision problem must satisfy two additional constraints. First, in each period and state of nature the resources stolen plus the resources invested in the production of public consumption must equal public funds[4], f_t^{pu},

$$z_t^{pu} + s_t^i = f_t^{pu} \qquad \text{for each } t \text{ and } \omega.$$

Second, the production function determines the amount of public consumption provided to the municipality,

$$Q_t = f\left(z_t^{pu}, z_t^{pr}, a, X_t^m\right) \qquad \text{for each } t \text{ and } \omega$$

Some remarks are in order. First, individual i can be fined in period t only if he was the mayor during the previous period and he was audited. Second, the sources of uncertainty faced by individual i depends on whether he is the current mayor. If he is, they correspond to the amount of funds the municipality will receive from the central government, whether he will be audited, and whether he will be reelected. If individual i is not the current mayor, the sources of uncertainty are his wage, the ability and the preferences of the current mayor.

2.4 Computation and Estimation

In this section we describe the approach used to estimate the model. In the estimation we use standard dynamic programming tools and indirect inference (Smith (1993), Gourieroux, Monfort, and Renault (1993), Gourieroux and Monfort (1996)). Specifically, the estimation is performed in two steps. For a given set of parameters that characterize the model, we first simulate the individual decisions. We then match some of the statistical moments that characterize the data with the corresponding moments obtained using the simulated data. The estimated parameters are obtained by minimizing a function of the distance between the simulated and data moments required by indirect inference.

[4] We do not model local taxes because in Brazil 85 percent of a municipality's receipts are transfers from the central government.

The simulation of the model requires the derivation of its recursive formulation and of the corresponding value functions. To recover the recursive formulation it is important to describe the timing of events and decisions. At the beginning of term t, the current mayor decides whether to run for reelection. If he decides to run, he faces a challenger who is selected randomly from the population of the municipality. If not, two challengers face each other in the election. Elections take place in each municipality and their outcomes determine the mayor that will govern each municipality for term t. Wages, public funds, and private inputs are then realized. The central government collects fines from mayors that were caught stealing in the previous period. The mayors then choose the fraction of public funds to invest in public inputs, the fraction to steal, consumption, and savings. At the end of the period, a fraction of municipalities is audited where the fraction is exogenously determined.

We can now describe how the value function for each individual i can be computed. There are two different value functions that we have to calculate to determine the optimal decisions for each individual: the value function of current mayors and the value function of past mayors. To understand why both value functions are needed, observe that we are interested in the decisions of individual i only insofar as they provide information about the amount of corruption and the amount of public consumption that characterizes the corresponding municipality. We are therefore only interested in the decisions of individuals who are current mayors. When making decisions, mayors take into account the effects they have on their future welfare. As a consequence, to determine their optimal decisions for term t one needs to know their expected value function for term $t+1$. The expected value function at $t+1$ is a combination of two parts: the expected value conditional on still being the mayor, $E\left[V_M\right]$, and the expected value conditional on not being in power, $E\left[V_{PM}\right]$. The value function of past mayors is therefore part of the computation.

We are now in the position to describe the recursive formulation of the problem for mayors and past mayors. Let S_t^M and S_t^{PM} be the set of state variables at t for, respectively, a current and past mayor. Since only 3 percent of past mayors have run again for election after leaving office for at least one term, we assume that individuals can be mayor only once in their life. The decision problem of a past mayor for term t can then be written in the following form:

$$V_{PM}^{i}\left(S_{t}^{PM},t\right)=\max\ u^{i}\left(c_{t}^{i},Q_{t}\right)+\beta E\left[V_{PM}^{i}\left(S_{t+1}^{PM},t+1\right)\right]$$

$$s.t.\ \ c_{t}^{i}+b_{t}^{i}=w_{t}^{i}\bar{h}+R_{t}b_{t-1}^{i}+1_{\{\rho_{t-1}^{i}=1,\delta_{t-1}^{i}=1\}}g\left(s_{t-1}^{i}\right).$$

where as mentioned above ρ_{t-1}^{i} equals 1 if individual i was a mayor in the previous period and δ_{t-1}^{i} equals 1 if the municipality was audited in the previous period.

We can now describe the recursive formulation of the decision process for a current mayor. Let $p\left(S_{t}^{M}\right)$ be the probability that the incumbent is elected conditional on the state variables. One of the decisions of the current mayor is whether to run for reelection in the current term. If he decides to run, he wins the election with probability $p\left(S_{t}^{M}\right)$ and then solves the following problem:

$$V_{WM}^{i}\left(S_{t}^{M},t\right)=\max\ u^{i}\left(c_{t}^{i},Q_{t}\right)+\beta E\left[V_{M}^{i}\left(S_{t+1}^{M},t+1\right)\right]$$

$$s.t.\ \ c_{t}^{i}+b_{t}^{i}=w_{t}^{i}\bar{h}_{t}+1_{\{\rho_{t}^{i}=1\}}s_{t}^{i}+R_{t}b_{t-1}^{i}+1_{\{\rho_{t-1}^{i}=1,\delta_{t-1}^{i}=1\}}g\left(s_{t-1}^{i}\right)$$

$$z_{t}^{pu}+s_{t}^{i}=f_{t}^{pu}$$

$$Q_{t}=f\left(z_{t}^{pr},z_{t}^{pu},a,X_{t}^{m}\right).$$

With probability $1-p\left(S_{t}^{M}\right)$ the election is won by the challenger, in which case the mayor's value function corresponds to the value function of a past mayor. The value function of an incumbent that chooses to run for reelection can therefore be computed as follows:

$$V_{RM}^{i}\left(S_{t}^{M},t\right)=p\left(S_{t}^{M}\right)V_{WM}^{i}\left(S_{t}^{M},t\right)+\left(1-p\left(S_{t}^{M}\right)\right)V_{PM}^{i}\left(S_{t}^{PM},t\right).$$

If the mayor decides not to run for reelection, her value function corresponds to the one of a past mayor, i.e. $V_{NRM}^{i}\left(S_{t}^{M},t\right)=V_{PM}^{i}\left(S_{t}^{PM},t\right)$. Note that in our model it may be optimal for a mayor to not run for reelection if the incumbent receives an attractive wage offer as a past mayor. For instance he may receive an attractive offer from a law firm or he may receive the offer to run for governor. Finally, the value function of the current mayor can be computed as the maximum of V^i_{RM} and V^i_{NRM}:

$$V^i_M\left(S^M_t, t\right) = \max\ \left\{V^i_{RM}\left(S^M_t, t\right), V^i_{NRM}\left(S^M_t, t\right)\right\}.$$

The set of state variables depends on whether individual i is the current mayor. If he is, S^M_t includes the following variables: the number of terms individual i has been in power; the municipality in which the mayor is in power; individual i's education; his preferences for public consumption relative to private consumption; individual i's ability; the amount of public good produced by the mayor in the previous period; the amount stolen by the mayor in the previous period; individual i's wealth; the probability the municipality will be audited. A past mayor has a larger set of state variables. It includes the state variables of a current mayor except ability and the current term. It also includes the following variables which characterize the mayor currently in power in her municipality for term t: the number of terms the current mayor has been in power; the amount of public good produced by the mayor in the previous period; the amount stolen by the mayor in the previous period; the mayor's wealth, preferences for public consumption and ability. The variables of the current mayor affect the past mayor decisions because they determine the amount of public good produced in the municipality.

The value function for each individual i is computed starting from the last period and moving backward in two steps following Keane and Wolpin (1994). In the first step, the state space is "discretized". Then using the corresponding grid, the value functions are computed for each period and each point of the state space in the grid. Finally, using the probability distribution for the discretized state space we can compute for each period the expected value functions conditional on the set of state variables $E\left[V|S\right]$. In the second step, the expected value functions are approximated using non-parametric methods. In practice, we regress the values of $E\left[V|S\right]$ obtained for each point in the grid on a polynomial of the discretized state variables. The corresponding coefficients are then used to construct the expected value functions for each period and value of the state space. Once the expected value functions are known, we can simulate the decisions of individuals in the municipalities observed in the data for different values of the parameters that characterize the model. The parameters of the model are then estimated using indirect inference.

2.5 Empirical Specification and Moment Selection

The model presented in Section analyzes a mayor's decision to engage in corruption. In this section, we discuss the specifications used to estimate the structural parameters of the model using municipal-level data. To operationalize our model, functional form assumptions are needed for four key aspects of the model: 1) utility function; 2) public good production function; and 3) electoral rule; 4) wage process.

2.5.1 Functional Form Assumptions

Utility: In our model, a citizen i living in municipality m in period t has preferences over private consumption $c_t^{i,m}$ and per-capita public consumption Q_t^m/pop that are described by the following utility function:

$$u^i\left(c_t^{i,m}, Q_t^m, dp_t^{i,m}\right) = \frac{(c^{i,m})^{1-\gamma}}{1-\gamma} + \theta_i \frac{(Q^m/pop)^{1-\delta}}{1-\delta} + \psi dp_t^{i,m},$$

where *pop* denotes the population size of the municipality. The parameter θ_i captures a citizen's taste for public consumption relative to private consumption. Through θ_i we introduce an important source of heterogeneity into the model, as municipalities that are governed by mayors with a higher θ_i should experience less corruption, all else equal. In the estimation of the model, we assume that θ_i takes on two values $\{\theta_L, \theta_H\}$, where $\theta_L < \theta_H$. In each municipality there is a fraction π_θ of high types.

Public consumption: Local public goods are produced according to a municipal-specific Cobb-Douglas production function:

$$Q_t^m = (z_t^{pu})^{\alpha_1}(z_t^{pr})^{\alpha_2} exp^{\alpha_3+\alpha_4 a_i+\alpha_5 E^m+\alpha_6 P^m}$$

where z_t^{pr} and z_t^{pu} represent the amount of private and public sector funds, respectively, that are invested into public consumption. Mayors influence the local public goods provision both through the amount of public funds invested but also by their ability a_i, where $a_i \in \{0, 1\}$. Individuals with high ability are more able and can produce more public goods with the same level of inputs. The proportion of high ability-types in the population is denoted by π_A. We also allow for the mayor's education level, E^m, and the size of the municipality, P^m, to affect public goods production.

Electoral rule: As mentioned in the model section, individuals decide whether to vote for the incumbent by adopting a retrospective voting strategy, whereby incumbents who provide public consumption above a particular threshold, Q^*, are rewarded with re-election. We assume that the threshold is a function of the following variables: the amount of public good produced by the municipality during the term; whether a mayor was audited, *audit*, and, conditional on being audited, the amount stolen, s; the mayor's age, a_m; the campaign contribution received by the challengers relative to the one received by the incumbent r_c; one dummy for small municipalities, d_{sm}, and one for medium municipalities, d_{mm}; an error term $\varepsilon \sim N(0,1)$ which captures the residual randomness. This assumption implies the following reduced-form formulation for the electoral rule:

$$Q_t - Q_t^* = \lambda_1 + \lambda_2 \text{audit}_t + \lambda_3(\text{audit}_t \times s_t) + \lambda_4 Q_t + \lambda_5 a_{m,t} + \lambda_6 r_{c,t} + \lambda_7 d_{sm,t} + \lambda_7 d_{mm,t} + \varepsilon_t$$

Wage Process for Past Mayors: We assume that the wage offers for past mayors are drawn from the following process:

$$\ln w_t^i = \delta_1 + \delta_2 t + \delta_3 E + \delta_4 P^m + \eta_t$$

where t, E, and P^m are a time trend, the education of the past mayor, and the size of the municipality, and η_t is normally distributed with mean zero and variance σ_η^2.

2.5.2 Moment Selection

Given these functional form assumptions, we estimate the model using the method of indirect inference. Using this method, we estimate five sets of parameters. In this section, we discuss the moments used to estimate the model.

Production Function Parameters ($\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$): Given the functional form assumption of the production function, the parameters on the observable inputs of production ($\alpha_1, \alpha_2, \alpha_3, \alpha_5, \alpha_6$) are estimated by matching the parameters obtained from a regression of the log of public consumption on the log of private and public inputs, the mayor's education, and the size of the municipality. In the data, public consumption is measured as municipal-level GDP. The parameter on ability, α_4, the unobserved component of the production function, is estimated by matching the variance of the residuals obtained from the regression mentioned above.

Preference for public consumption parameters (θ_H, θ_L): To estimate the relative taste for public consumption for a good mayor, θ_H, and a bad mayor, θ_L, we use average stealing and average stealing conditional on mayors that are in their second term.

Preference for being in power parameter (ψ): To estimate the taste for being in power, ψ, we use the fraction of individuals that choose to run for a reelection.

Probability parameters (π_θ, π_A): To estimate the probability of observing a mayor with a high taste for public consumption, π_θ, we use as a moment the probability of observing stealing greater than zero conditional on being audited. To estimate the probability of observing a high ability mayor, π_A, we use the second moment of the residuals obtained by estimating the production function without ability.

Labor market parameters: The parameters of the wage offer distribution for past mayors are estimated by matching the parameters obtained by regressing log wages of past mayors on the variables that characterize the wage function. We are currently in the field collecting the data on wages of past mayors.

Electoral Rule Parameters ($\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7$): The parameters of the electoral rule are estimated by matching the parameters obtained from a probit model in which the dependent variable is the outcome of the election and independent variables are the audit dummy, the interaction between this dummy and the amount stolen by the mayor, per-capita public consumption, the mayor's age, the mayor's campaign contributions relative to the challenger with the highest level of contributions, one dummy for small municipalities, and one for medium municipalities.

2.6 Data

This section describes the data used in the analysis. We first describe the approach used to measure corruption. We finish by providing summary statistics of the main variables used to obtain the preliminary estimates for Brazil.

As with any illegal activity, obtaining data on corruption is a difficult task. Several empirical studies that focus on illegal behavior have used indirect evidence to analyze its determinants and consequences. However, a small, but growing body of literature has tried to assess corruption more directly by focusing on two forms: bribery of public officials and the theft of public resources (see for instance Svensson (2005) and Olken (2007)).

For our analysis, we exploit the data used by Ferraz and Finan (2008). Their approach, although related to the studies cited above, uses a new methodology made possible by the availability of audit reports from Brazil's and Puerto Rico's anti-corruption programs. Contained in each report is the total amount of federal funds audited for the current administration, as well as, an itemized list describing each irregularity and in most cases the amount of funds involved. Each report classifies the irregularities found by auditors into corruption indicators and estimates the amount of public resources misappropriated for each irregularity. Based on the coding of these reports, our principal measure of corruption is the total amount of resources related to corrupt activities, expressed as a share of the total amount of resources audited.[5] We complement these corruption measures with a rich data set that combines information on various socioeconomic characteristics of the municipality with both local public finance data and election results.

We now describe in more detail the Brazilian data. The public finance data, which is collected from Brazil's National Treasury, are used to construct a measure of annual intergovernmental transfers received by municipalities from 1997-2005. From the electronic files of Brazil's electoral commission, we obtain results from the 1996, 2000, and 2004 mayor elections. These data contain vote totals for each candidate by municipality, along with various individual characteristics such as the candidate's gender, education,

[5] Political corruption is defined to be any irregularity associated with fraud in procurements, diversion of public funds, and over-invoicing. Specifically, we define a *procurement to be irregular* if: i) a required procurement was not executed; ii) the minimum number of bids was not attained; iii) there was evidence of fraud in the procurement process (e.g. use of bids from non-existing firms). *Diversion of public funds* is defined as any expenditure without proof of purchase or provision and/or direct evidence of diversion provided by Brazil's auditing agency. Finally, *over-invoicing* is defined as any evidence that public goods and services were bought for a value above the market price.

occupation, and campaign contributions. We use this information to account for various individual mayor characteristics that might affect corruption. The final piece of data come Brazil's statistical office, which provide municipal-level GDP, as well as, private GDP for 2001-2005. Finally, according to the Brazilian law, a mayor who is caught stealing must return the amount stolen and pay a fine that is equal to 1.5 the amount stolen. We choose the fine schedule used in the estimation accordingly.

Table 2 presents summary statistics for Brazil for some of the main variables used in the analysis. According the audit reports, municipalities diverted six percent of all funds that were transferred from the federal government. Corruption is also higher among second-term mayors compared to first-term mayors, which is consistent with re-election incentives. Based on the election data, we find that re-election rates are about 40 percent and that incumbent mayors receive over twice as much campaign contributions as the challenger.

2.7 Results

In this section we discuss the estimation results, the ability of the model to match the main features of the data, and the performance of three different policies designed to reduce corruption. In Tables 3, 4 and 5 we report the estimated coefficients. The parameters of the wage distribution have been fixed, since the wages of past mayors are still in the process of being collected. The only result that is worth mentioning is that the estimated relative presences for public consumption are on the high side at 16.5 and 16.6. The reason for this high coefficient estimates is that the model has to explain why in the data on average mayors steal just 6% of the available funds.

We now describe how well the model can match the data moments selected for the estimation of the coefficients. In Table 6, we compare the simulated and data moments for public consumption and stealing. We match these moments reasonably well with the exception of the fraction of mayors who choose not to run for re-election which is underestimated by the model. In Table 7, we report the simulated and data moments used for the estimation of the production function parameters. All the simulated moments are very close to the data moments. Finally, Table 8 describes the simulated and data moments selected for the estimation of the electoral rule. As for the previous set of moments, we match reasonably well the electoral rule moments.

We now move to test the performance of the estimated model by comparing the features of the data described in section 2 with the same features obtained from the simulation of the estimated model. It is worth remarking that these comparisons are strong tests of the model since we do not match many features of the data when we estimate the model.

We start by discussing the performance of the model in terms of matching the unconditional distribution of stealing. The corresponding simulated and data moments are reported in Table 9. We match closely the mean of the distribution, but this should not be surprising since this is one of the moments used in the simulation. We also match the fact that in the data stealing is skewed to the right. The degree of skewness, however, is slightly larger in the data. In our simulations, we find that 35% of mayors do not steal against 26% in the data. The $50th$, $75th$, and $90th$ percentiles are equal to $49,327$, $101,447$, and $161,629$ in the simulations and to $24,244$, $96,704$, and $230,408$ in the data.

When we condition stealing on the number of terms, the model does a good job at matching the data. The simulated and data moments are documented in Table 10 and 11. We match the fact that in the data stealing is smaller in the first term. In our simulations stealing in the first term is $51,057$, whereas in the second term it is $98,036$. In the data, the two numbers are $76,973$ and $97,790$. We also match the observed pattern that the percentage of mayors stealing is smaller in the first term. In our simulations, 62% of mayors steal in the first term and 69% in the second term. In the data, we observe 72% of mayors that steal in the first term and 77% in the second term. We also match the sign of the coefficient obtained by regressing the logarithm of stealing on the number of terms after controlling for the allocated funds, private inputs, population size, and education of the mayor. But we are off in terms of magnitude since the coefficient is equal to 2.785 in the simulations and to 0.360 in the data. We do better in matching the coefficient obtained by regressing the logarithm of public consumption on terms after controlling for the same variables. Our model generate a coefficient that is equal to -0.010, whereas in the data the coefficient is equal to -0.078.

Table 12 reports the performance of the model when we condition on municipality size. After controlling for the amount of funds, private inputs, population size, and education of the mayor, our model predicts that mayors in small and medium municipalities steal less than mayors in large ones. The coefficient obtained by regressing log stealing on the small municipality dummy is negative and equal to -0.237, whereas in the data it is equal to -0.619. For medium municipalities the coefficient is estimated to be -0.140 in the simulated data and -0.468 in the actual data. We also do a good job in matching the differences in the production of public consumption by municipality size. We can replicate the result observed in the data that small municipalities produce more per-capital public consumption than medium and large municipalities and the result according to which medium municipalities produce more per-capita consumption than large municipalities. Specifically, when we regress the logarithm of per-capita public consumption on a small municipality dummy after controlling for the same set of variables we obtain a coefficient which is equal to 0.508 using the simulated data and to 0.972 using the actual data. The same coefficient for medium municipalities is equal to 0.249 in the simulated data and to 0.510 in the actual data.

We now test whether the model is able to match the effect of stealing and per-capita public consumption on the electoral outcome observed in the data. When we estimate a probit model where the dependent variable is the outcome of the election and the set of independent variables is composed of stealing, per-capita public consumption, population size, age of the mayor, and the campaign contributions of the mayor relative to the challengers using the simulated data, we observe that stealing has a negative effect and per-capita consumption a positive effect on reelection, which is consistent with the pattern observed in the data. Using the simulated data, we can also compute the incumbent bias in elections. The model generates a positive incumbent bias similarly to the data, but the size is too large with 95.7% of incumbents that are reelected in the model against 57.3% in the data.

In the last part of this section, we use the parameter estimates to discuss three policy simulations. The discussion highlights the potential usefulness of this model for informing policy. Figure 1 presents the relationship between corruption and the amount of federal funds received by the municipality based on simulated data from the model. In addition to the base case simulation (depicted by the solid line), the figure plots results of three policy simulations. The base case is computed using the estimated parameters and the probability of being audited before the Lula's anti-corruption program was introduced.

The first simulation considers the effects of a policy that would double a mayor's wage. By increasing the opportunity cost of stealing, corruption will presumably decrease. The second simulation considers the effects of a policy that increases term limits from 2, which is the current limit in Brazil, to 3. One would expect that mayors who face reelection incentives are significantly less corrupt than mayors who are unable to run for reelection. By increasing an elected official's political horizon, the incentive to engage in corruption decreases. The final simulation considers the anti-corruption policy implemented by Lula's government. In this simulation, we increase the probability of being audited from 5%, which was the probability before the Lula's program, to 17%, which is the current probability. The audit policy has two separate effects. First, the audit reveals the politician's type, thus reducing the probability of a bad type getting re-elected. Second, mayors that have been audited are forced to pay a significant fine.

As expected, the figure depicts a positive relationship between the amount of funds received and the amount stolen, which explains why, as Brazil has become more decentralized, local corruption has become such a serious problem. Given our current set of parameter values, our model also suggests that while increasing the wages of politicians may encourage mayors to engage in less corruption, its effects are small relative to the other two policies. Although these results are still preliminary, they question the effectiveness of a policy which is frequently endorsed by both academic and policy circles alike. The preliminary results suggests that the most effective manner of combating corruption is to either increase the likelihood of getting audited or to increase the term limits. On average the term limit policy generate the largest decrease in corruption. But the policy that increases the probability of being audited is more effective at reducing corruption in municipalities with high level of stealing.

2.8 Conclusion

In this paper, we provide a framework for understanding the decisions of politicians to engage in corrupt activities. We develop a structural model of a politician's decisions to provide public goods and engage in corruption over the span of his political career. We then estimate the model using objective measures of local corruption based on the audit reports of an anti-corruption program conducted in Brazil in 2003. Using this framework, we explore three policies for reducing corruption: 1) increasing the probability of a municipality being audited, 2) increasing politicians' wages; and 3) increasing term limits from 2 terms to 4 terms. Our results suggest that increases the probability of being audited and/or the term limits greatly reduces corruption in local governments. For the audit policy, the reduction in corruption stems from the anticipated effects of not only losing re-election but also having to face prosecution and paying a fine. For the term limit policy, the decrease is due to the increase in an elected official's political horizon.

The framework proposed in the paper, while applied to local governments in Brazil, is quite general. Future research will explore the corruption decisions of mayors in Puerto Rico, and thus provide an interesting point of comparison.

Table 1: Description of the Audit Reports Data, Brazil

Amount of Resources Transferred from the Federal Government	7,987,323
Proportion of municipalities with at least one irregularity	0.740
Share of audited resources related to corruption conditional on at least one irregularity	0.078

Table 2: Summary Statistics

	Mean	Standard Deviation
Average GDP (2002-2005)	84,651.0	276,003.1
Average Private GDP (2001-2004)	35,695.6	122,513.5
Federal transfers	1,498.0	2,602.4
Population	21,786.8	39,262.1
Mayor's education level	12.76	4.17
Mayor's age	48.0	7.9
Relative campaign contributions	2.14	2.55
Re-election rates 2004	0.57	0.50
Share of resources found to be corrupt	0.062	0.099
Among first-term mayors	0.057	0.100
Among second-term mayors	0.069	0.095

Table 3: Estimated Parameters: Probabilities of Types and Preferences

Parameters	Estimates	Standard Errors
Probability of Low Ability	0.755	[0.101]
Probability of Bad Type	0.050	[0.041]
Relative Taste for Q Bad Type	16.5	[2.392]
Relative Taste for Q Good Type	33.1	[1.949]
Utility from Being in Power	53.7	[2.69]

Table 4: Estimated Parameters: Production Function

Parameters	Estimates	Standard Errors
Public Inputs	0.379	[0.142]
Private Inputs	1.513	[0.185]
Constant	0.586	[0.158]
Ability	1.280	[0.372]
Education	0.147	[0.023]
Small Municipality	-0.114	[0.045]
Medium Municipality	0.029	[0.038]

Table 5: Estimated Parameters: Electoral Rule

Parameters	Estimates	Standard Errors
Constant	1.457	[0.328]
Audit	0.131	[0.059]
Audit*Dummy Stealing	-1.043	[0.194]
Public Consumption	0.197	[0.043]
Mayor's Age	-0.089	[0.014]
Relative Contributions	0.164	[0.014]
Small Municipality Dummy	-0.062	[0.103]
Medium Municipality Dummy	-0.186	[0.092]

Table 6: Moments: Public Consumption and Stealing

Moment	Model	Data
Average Per-capita Public Consumption	3.611	3.167
Average Stealing	72.457	85.908
Average Stealing, Second Term	98.037	97.790
Fraction of Audited Mayors Caught Stealing	0.652	0.742
Fraction of Audited Mayors Caught Stealing, Second Term	0.691	0.768

Table 7: Moments: Production Function

Moment	Model	Data
OLS Coefficient on Public Inputs	0.369	0.424
OLS Coefficient on Private Inputs	1.521	1.441
OLS Coefficient on Education	0.150	0.112
OLS Coefficient on Small Municipality Dummy	-0.096	-0.097
OLS Coefficient on Medium Municipality Dummy	3.746	0.297
OLS Constant	0.850	0.444
Second Moment of the Production Function Residuals	0.305	0.302

Table 8: Moments: Electoral Rule and Choice of Incumbent

Moment	Model	Data
Audit Dummy	0.125	0.243
Audit*(Stealing greater than zero)	-1.036	-0.430
Public Consumption	0.201	0.138
Mayor's Age	-0.092	-0.109
Relative Contributions	0.167	0.194
Small Municipality Dummy	-0.065	-0.071
Medium Municipality Dummy	-0.193	-0.359
Constant	1.478	1.318
Incumbent Not Running for Reelect.	10.5%	28.2%

Table 9: Simulated and Actual Data: Stealing

	Model	Data
Average Stealing	72,457	85,908
Fraction of Audited Mayors Caught Stealing	0.652	0.742
Median Stealing	49,327	24,244
75th Percentile For Stealing	101,447	96,704
90th Percentile For Stealing	161,629	230,408

Table 10: Simulated and Actual Data: By Term

	Model	Data
Av. Stealing, First Term	51,057	76,973
Av. Stealing, Second Term	98,036	97,790
% Stealing, First Term	62%	72%
% Stealing, Second Term	69%	77%

Table 11: Simulated and Actual Data: By Term

	Model	Data
Coeff. Log Stealing on Terms	2.785	0.360
Coeff. Log Per-capita Public Cons. on Terms	-0.010	-0.078

Note: Controlling for Funds, Private Inputs, Population Size, and Education

Table 12: Simulated and Actual Data: By Municipality Size

	Model	Data
Coeff. of Log Stealing on Small Mun.	-0.237	-0.619
Coeff. of Log Stealing on Medium Mun.	-0.140	-0.468
Coeff. of Log Per-Capita Public Cons. on Small Mun.	0.508	0.972
Coeff. of Log Per-Capita Public Cons. on Medium Mun.	0.249	0.510

Note: Controlling for Funds, Private Inputs, Population Size, and Education

Table 13: Simulated and Actual Data: Electoral Rule

	Model	Data
Effect of Positive Amount of Stealing	-1.036	-0.430
Effect of Per-capita Public Consumption	0.201	0.138
Incumbent Reelected If Running	95.7%	57.3%

Figure 1: Policy Simulations: All Policies

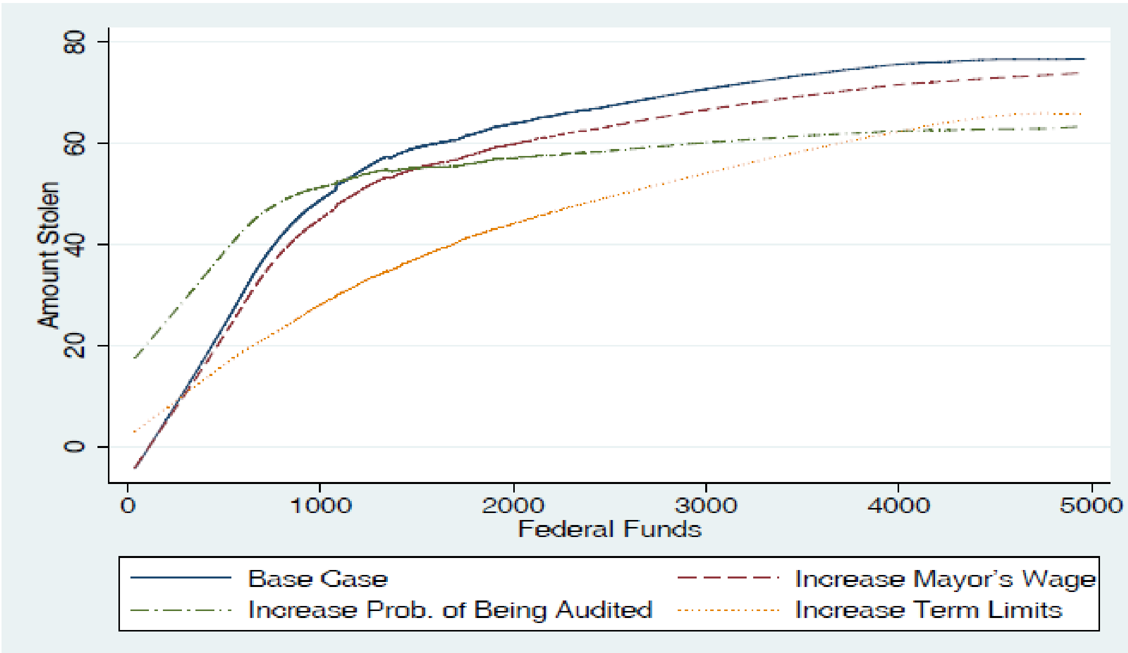


Figure 2: Policy Simulations: By Term

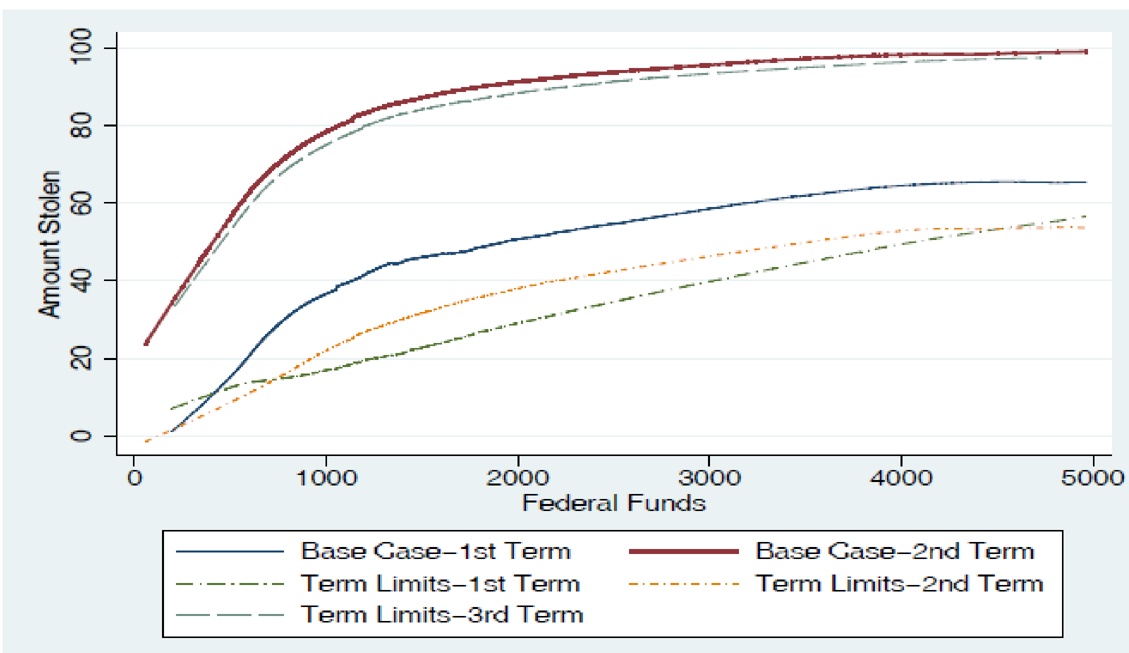
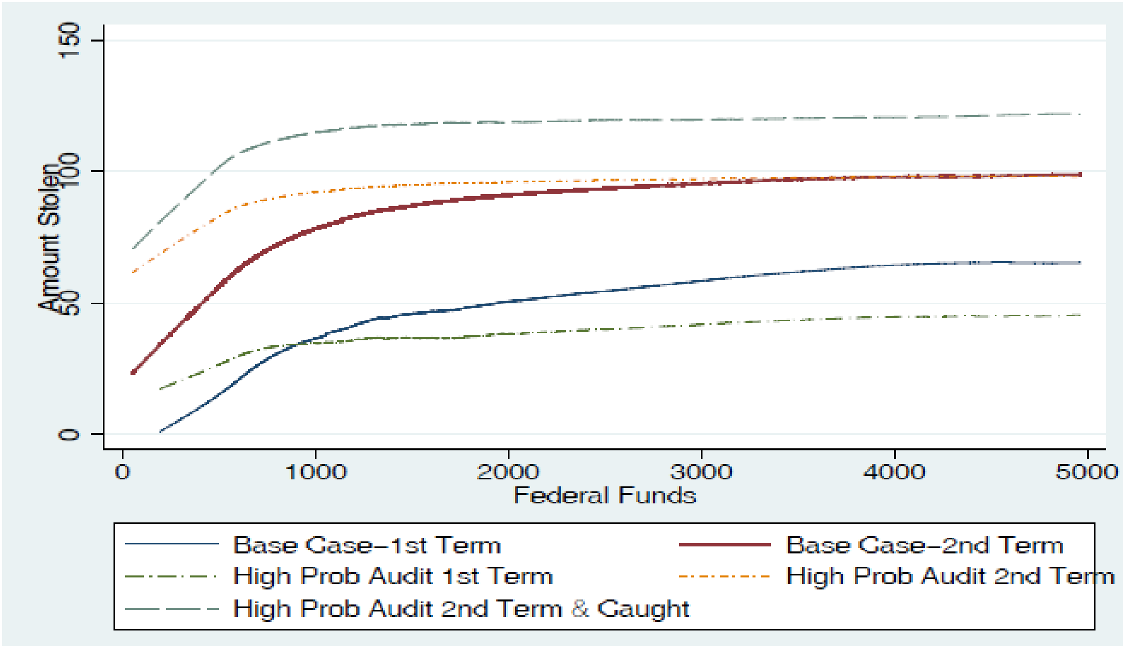


Figure 3: Policy Simulations: By Audit Probability



2.9 References

1. Barro, Robert (1970) "The Control of Politicians: An Economic Model," *Public Choice,*14:19-42.

2. Besley, Timothy (2006) "Principled Agents? The Political Economy of Good Government," Oxford, UK: The Lindahl Lectures, Oxford University Press.

3. Diermeier, D., M. Keane, and A. Merlo (2005) "A Political Economy Model of Congressional Careers," *American Economic Review,* 95 (1): 347-373.

4. Di Tella, Rafael, and E. Schargrodsky (2003) "The Role of Wages and Auditing During a Crackdown on Corruption in the City of Buenos Aires," *Journal of Law and Economics,* 46 (1): 269.

5. Ferejohn, J. (1986) "Incumbent Performance and Electoral Control." *Public Choice* 50: 5-25.

6. Ferraz, C., and F. Finan (2008) "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," *Quarterly Journal of Economics,* 123 (2): 703745.

7. .Ferraz, C., and F. Finan (Forthcoming) "Electoral Accountability and Corruption in Local Governments: Evidence from the Audit Reports," *American Economic Review.*

8. Gourieroux, C., and A. Monfort (1996 "Simulation-Based Econometric Methods," New York: Oxford University Press.

9. Gourieroux, C., A. Monfort, and E. Renault (1993) "Indirect Inference," *Journal of Applied Econometrics,* 8 (December): S85-118.

10. Keane, M., and K. Wolpin (1994) "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics,* 76 (4): 648-672.

11. Mauro, P., (1995) "Corruption and Growth," *Quarterly Journal of Economics,*110: 681-712.

12. Olken, B., (2007) "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy,* 115, no. 2.

13. Rose-Ackerman, S., (1999) "Corruption and Government: Causes, Consequences, and Reform," New York: Cambridge University Press.

14. Smith, A., (1993) "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics,* 8: S63-84.

15. Svensson, J., (2005) "Eight questions about Corruption," *Journal of Economic Perspectives,* 19 (3): 19-42.

16. VEJA (2004) "Pragas Urbanas: Desperdicio, Desvio e Corrupcao".