Structural Modelling of Female Labour Participation and Occupation Decisions

Rafael E. De Hoyos

February 2006

CWPE 0611

Not to be quoted without permission

Structural Modelling of Female Labour Participation and Occupation Decisions

Rafael E. De Hoyos*

February 2, 2006

Abstract

The objective of this paper is to estimate the parameters defining female labour participation and occupation decisions. Departing from a theoretical framework, we use micro data to estimate the wage-participation elasticity in Mexico. Consistency between the selectivity-adjusted wages and the multinomial participation equations is achieved via a two-step estimation procedure following Lee (1984). We use the results of our model to test and quantify three hypotheses explaining recent increases in female labour participation in Mexico. Our results show that the observed 12 per cent increase in female labour participation in Mexico between 1994 and 2000 is explained by the combination of a negative income shock caused by the 1994-95 Peso crisis, the increase in expected wages taking place in the manufacturing sector during the post-NAFTA period and a reduction in the female reservation wage.

JEL Classification: C34, J23, J24, J31,
Keywords: Participation, wage differentials, microsimulation, Mexico

*Faculty of Economics, University of Cambridge. Email: red29@cam.ac.uk. I am grateful for the comments made by Ajit Singh, Jaime Ruiz-Tagle, Hamish Low and seminar participants at EDGE 2005, Universita' Bocconi, Milan.

1 Introduction

Increasing female labour participation is an important aspect that is contributing to the development process in many emerging economies. Identifying the determinants of female labour participation and occupation decisions improves our understanding of the dynamics of labour supply and its interaction with economic development. The objective of the present study is to estimate the determinants of female labour participation and occupation decisions within a structural, utility-maximizing framework, and to use them to explain the recent increase in female participation in the Mexican labour market.

Based on the relationships described by the structural model, the paper develops a microeconometric model to obtain the determinants of labour participation and occupation decisions. We depart from a utility maximizing framework where individuals' choice depends on a set of comparisons between *expected* market wages and a subjective *reservation* wage. Although the agent's choice depends, ultimately, on personal and household characteristics and a subjective valuation of leisure, we estimate the way in which participation reacts to changes in expected wages (partcipation-wage elasticity). Given that participation/occupation decisions are the outcome of a non-random utility-maximizing process, we model *expected* wages taking selectivity into account; therefore, selectivity-adjusted wages and labour particiaption/occupation functions are estimated using a two-step procedure.

The structural models of labour occupation developed in Heckman and Sedlacek (1985) and Heckman and Honore (1990) are based on Roy's (1951) concept of comparative advantage. The papers show that agents' occupation decisions are not entirely determined by differences in market wages; personal preferences also play a significant role in the *selection* process. These findings, together with the fact that the econometrician only observes market wages and not personal preferences, have led almost all labour supply studies to use reduced-form estimations where wages are substituted by their determinants (i.e. observed personal characteristics within a human capital framework). An exception to this is Gong and van Soest (2002) who develop a model that explicitly links expected wages with labour participation and occupation decisions. Using data for Mexico, the authors find that wages have a positive effect upon women's participation and their occupational choices.

The present study adds to the literature on methodological and empirical grounds. Our approach is novel in three respects: (1) the microeconometric framework derives from a structural model with expected wages explicitly determining labour participation/occupation decisions, (2) using micro data for several years we uncover the dynamics of intra-household female labour participation dependence and (3) the participation and occupation effects of exogenous shocks (e.g. currency crisis) are evaluated using microsimulation techniques. One advantage of estimating a structural model is that its results can be used to evaluate out-of-sample response to, for example, income tax changes. Furthermore, our focus solely on participation and occupation decisions rather than on endogenising hours of work is based on the institutional rigidities present in many developing countries, where working hours are not freely chosen.[1]

The model is estimated using biannual Mexican household data for the years 1994 to 2000. This period allows us to test changes in female labour participation/occupation wage elasticities during a period of liberalizing reforms. Female labour participation in urban Mexico increased by 12 per cent between 1994 and 2000, representing a total of more than 700,000 new entrants during the first six years of NAFTA. Given the simultaneity of these two events, it is tempting to posit a causal relationship, explaining the increase in female participation as a result of NAFTA. Our model allows us to create a *hypothetical* economy where the participation/occupation structure is free of 'market-induced' changes. Hence, using microsimulation techniques, we can quantify how much of the increase in female participation was brought about by changes in expected wages (market conditions) versus changes explained by shifts in structural parameters determining female *reservation wage function*.

The paper is organized in the following way. In the next section, we develop the model, stressing the necessary assumptions to identify the parameters in our empirical strategy. Section 3 shows the results of the model using Mexican household data for the period 1994 to 2000. In section 4, we carry out a microsimulation analysis to test the impact of exogenous changes in parameters upon the employment and the occupation structure. Finally, a summary and conclusions are shown in the last section.

[1]As stated by Heckman: 'Participation (or employment) decisions generally manifest greater responsiveness to wage and income variation than do hours-of-work equations for workers' (Heckman, 1993, pg. 117).

2 The Model

2.1 Theoretical Framework

Following Heckman (1974), we assume that an agent's participation decision is determined by the difference between market wages and a reservation wage function (what Heckman calls shadow prices of female time). Heckman develops a model for the binomial choice problem, however it can easily be extended to a multinomial one. Define a reservation wage, w^*_{ij}, as the minimum wage required to observe individual i working in occupation j. Such a reservation wage will be determined by the agent's personal and household characteristics:

$$w^*_{ij} = w^*_j(\boldsymbol{X}_i, \boldsymbol{Z}_i) \qquad \forall\, j = 1 \ldots J \tag{1}$$

Where $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ are vectors of personal and household characteristics of individual i, respectively. Instead of having a single reservation wage function for each individual (as in Heckman's model), we have $(J-1)$ of them plus an implicit one, i.e. as many as the number of choices.[2] The $(J-1)$ reservation wages is derived from the assumption that occupations have different characteristics, apart from monetary ones, that have a value for the individuals. Allowing for different reservation wages across choice is justified on the basis of differences in observable characteristics (e.g. working conditions) and unobservable ones (e.g. an individual preference for a particular occupation) across occupations. Therefore, individuals attach a different personal valuation to each occupation. A well-documented example of this is the institutional rigidities present in the labour market, where the lack of working hours flexibility can be substituted by occupational choices.[3] The utility valuation given to the different occupational characteristics is captured by vectors $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$. On the other hand, expected market wages, following conventional human capital theory, are defined by the well-known function $\hat{w}_{ij} = \boldsymbol{X}_i\hat{\boldsymbol{\beta}}_j$.

Once a reservation and an expected market wage are defined for each occupation, agents' choices will be based on a series of pair comparisons between $\hat{w}_j$ and w^*_j. The utility maximizing choice will depend not merely on the level of these two components but on the difference between them, i.e. $(\hat{w}_j - w^*_j)$.

[2]The implicit one is the shadow price of leisure.

[3]See Deaton and Muellbauer, 1980, p. 86.

In this framework, the conventional reservation wage (i.e. whether an individual decides to work or not) is defined implicitly by the same set of pair comparisons. An individual participates in the labour market as long as one of the differences $\hat{w}_j - w_j^*$ is positive, but her final occupational choice will be the one which maximizes the gap between them. There is an implicit *utility* function embedded in this maximizing process which can be defined as:[4]

$$V_{ij} = V(\hat{w}_{ij} - w_{ij}^*), \qquad \forall\ j \tag{2}$$

A reduced form of (2) takes into account the fact the we do not observe w_j^* and hence we can only include the observable components that determine reservation wages, i.e. $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$. Assuming that *utility*, $V(\cdot)$, is a linear function of its arguments and adding a random component, it can be defined as:

$$V_{ij} = \lambda \hat{w}_{ij} - (\boldsymbol{X}_i \boldsymbol{\gamma}_j' + \boldsymbol{Z}_i \boldsymbol{\gamma}_j) + \eta_{ij} \tag{3}$$

where η_{ij} is a stochastic component. We are implicitly assuming that, controlling for differences in $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$, the marginal utility of monetary income, λ, is constant across individuals and occupations, therefore λ is a scalar parameter.[5] Since $\hat{w}_i$ is fully determined by $\boldsymbol{X}_i$, a major problem with equation (3) is that we cannot identify both sets of parameters, λ and γ_j' at the same time. Changes in $\boldsymbol{X}_i$ will have a double and simultaneous effect, on the one hand upon expected market wages and, on the other, upon reservation wages. To tackle this problem, as an alternative to (3), we define a less flexible but more parsimonious version of the *utility* function. Substitute the reduced form version of the expected wage function ($\boldsymbol{X}_i \hat{\boldsymbol{\beta}}_j$) into (3):

$$V_{ij} = \lambda(\boldsymbol{X}_i \hat{\boldsymbol{\beta}}_j) - (\boldsymbol{X}_i \gamma_j' + \boldsymbol{Z}_i \boldsymbol{\gamma}_j) + \eta_{ij} \tag{4}$$

[4]Notice that while V is defined by elements $\hat{w}$ and w^*, reservation wages will depend, in turn, on individual preferences; *utility* is thus ultimately defined by monetary income, a personal valuation of it and individual preferences. The term utility used here should be taken with caution since probably it embeds demand-side restrictions in the labour market, hence the observed *choice* might not be entirely the outcome of a personal utility-maximizing process.

[5]To clarify the notation, λ is a scalar, $\boldsymbol{\gamma}_j'$ and $\boldsymbol{\gamma}_j$ are vectors of size K_1 and K_2 reflecting the effects of personal and household characteristics upon reservation wages measured in utility units.

Simplify:

$$V_{ij} = (\lambda - \gamma_j'/\hat{\boldsymbol{\beta}}_j)\boldsymbol{X}_i\hat{\boldsymbol{\beta}}_j - \boldsymbol{Z}_i\boldsymbol{\gamma}_j + \eta_{ij} \tag{5}$$

Define $\delta_j = (\lambda - \gamma_j'/\hat{\boldsymbol{\beta}}_j)$:

$$V_{ij} = \delta_j \hat{w}_{ij} - \boldsymbol{Z}_i\boldsymbol{\gamma}_j + \eta_{ij} \tag{6}$$

Notice that the wage-participation parameter, δ_j, will be positive if and only if $\lambda^*\hat{\boldsymbol{\beta}}_j > \gamma_j'$. Therefore, $(\lambda^*\hat{\boldsymbol{\beta}}_j)$ and (γ_j') can be interpreted as the substitution and income effects of changes in personal endowments $\boldsymbol{X}_i$, respectively. Participation will increase as a result of higher expected wages as long as the substitution effect is larger than the income effect. Equation (6) allows the marginal utility of fitted wages, $\hat{w}_i$, to differ across choices, capturing the unobservable effects deriving from the reservation wage function (γ_j') and the remunerations of personal characteristics across occupations $(\hat{\boldsymbol{\beta}}_j)$. The first element of $\boldsymbol{Z}_i$ is a column of ones (i.e. there is a different intercept for each occupation), accounting for the utility effects of occupation-specific attributes. Based on specification (6), individual i will choose occupation j if and only if:

$$V_{ij} > \max_{m \neq j}\{V_{im}\} \; \forall \; j \tag{7}$$

Framework (1)-(7) departs from Heckman's (1974) reservation wage concept and arrives to McFadden's (1974) utility maximization criteria. Unifying both approaches help us understand the dynamic processes that might lie behind the data we observe.

2.2 Empirical Strategy

This section elaborates on the aspects that we have to take into account in order to obtain a set of equations which are suitable for estimation. The advantage of having a structural model behind the estimations is that we can interpret the parameters in a way that is consistent with the theoretical framework. For example, from (3) we know that a change in one of the elements of $\boldsymbol{X}_i$ will have a double—and possibly opposing—effect upon the probability of participating in the labour market. On the one hand, an increase in $\boldsymbol{X}_i$ will tend to increase the agent's expected wage and this might have a positive effect upon participation. On the other hand, the same increase in $\boldsymbol{X}_i$ can increase the agent's reservation wage and hence reduce her participation probability. Although we do not observe this second effect, we could estimate a specification like (3) and try to identify both effects. However, as we have already pointed out, we cannot identify the parameters on $\hat{w}_i$ and $\boldsymbol{X}_i$ simultaneously. Furthermore, even using a parsimonious specification like (6), the interpretation of δ_j is not straightforward. Given the normalization assumption that is necessary to estimate the probability of participation and occupation based on the criteria described in (7), allowing the parameters on expected wages to vary across all outcomes will be misleading in terms of our theoretical model.[6] Instead of allowing expected wages to enter (6) as if they were characteristics of the individuals, we restrict δ_j to be the same for all occupations (δ). Thus we interpret expected wages as an *attribute* of the occupation rather than a characteristic of the individuals.[7]

Before estimating the model, we need one further assumption. The random components of (6), η_{ij}, can follow many distributions, e.g. normal, poisson, extreme value or a combination of various distributions (logit kernel or mixed logit). For simplicity, we assume that η_{ij} are i.i.d. with extreme value distribution. With all our assumptions at hand, the probability that agent i will choose occupation s is defined as:

$$Pr(i = s) = \frac{exp[\delta \hat{w}_{is} + \boldsymbol{Z}_i \boldsymbol{\gamma}_s]}{\sum_{j=1}^{J} exp[\delta \hat{w}_{ij} + \boldsymbol{Z}_i \boldsymbol{\gamma}_j]} \tag{8}$$

Equation 8 combines *attributes* of the occupations, $\hat{w}_i$, with *characteristics* of the individual $\boldsymbol{Z}_i$. The specification is a combination of a conditional and

[6]Say that we normalize by making the parameters of outcome 'not active' equal to zero. For every possible occupation, we will have $(J-1)$ expected wages, but the interpretation of the parameters for all of them would be in terms of the base category (not active). In terms of our structural model, an increase in the expected wage in occupation j does not have an effect upon the probability of participating in occupation j' **relative to being not active**, therefore there is no basis for including all $(J-1)$ expected wages as if they were characteristics of the individuals.

[7]A sufficient assumption to have a single parameter for expected wages across all outcomes is: that $\gamma_j'/\hat{\boldsymbol{\beta}}_j = c \;\; \forall \, j$ where c is a constant (see equation 5). This is equivalent to imposing a constant ratio of marginal market *price* of characteristics $\boldsymbol{X}_i$ relative to its subjective valuation (in terms of reservation wage) across occupations. In other words, every time personal characteristics increase their market remuneration in a particular occupation, individuals will increase, in the same proportion, their subjective valuation of them. This implies, obviously, a constant wage-participation elasticity across all individuals and occupations; this is certainly a restrictive assumption.

a multinomial logit with $\hat{w}$ varying across individuals and occupations and $\boldsymbol{Z}$ varying only across individuals (see Maddala, 1983, p. 44).

2.2.1 Selectivity and Expected Wages

Let us define the log of hourly wages, w_i, as a linear function of formal education, education interacting with a dummy variable for higher education, experience, experience squared and a regional dummy variable. These variables plus a constant are the elements of $\boldsymbol{X}_i$. We allow for different parameters across occupations, estimating a separate wage equation for each of them assuming that their residuals are only related via the selection criteria (7). Our working age population is defined as women between 12 and 65 years old without a physical impediment to work and not being full-time students. Women within this classification face the following set of choices: to participate in the labour market as self-employed or informal worker[8], work in the manufacturing sector, work in other formal sectors or not to participate in the labour market at all (not active).[9]

As we have already specified, the workers observed in each sector are not the outcome of a random process, indeed they follow criteria (7). Therefore, the wage equations in each of the three remunerated sectors have to account for selectivity. Following (8), we can obtain the conditional probabilities of participation for each sector and, given a parametrization rule, include them in the wage equation to control for selectivity.[10] The problem is that, as we can see from (8), the conditional probabilities obtained from the multinomial logit, $Pr(\cdot)$, are themselves a function of expected wages, therefore we have the following simultaneous equation model:

[8] We classify workers as being in the informal sector when they are non-professional self-employed workers. We exclude family workers who get no monetary remuneration (See Maloney 1999).

[9] Not active agents include women who were actively looking for a job (unemployed), not active housewives and 'other not active' (e.g. women such as pensioners and landladies). Housewives and 'other not active' women account for 95 per cent and 3 per cent of the total inactive female population, respectively. Less than 2 per cent of the female inactive population in 2000 was actively looking for a job, hence the use of a theoretical framework where one of the utility-maximizing choices is to be inactive is, at least, a plausible first approximation of the participation decision process.

[10] For a discussion of the advantages and disadvantages of the different ways to control for selectivity using a multinomial logit, see Bourguignon, Fournier and Gurgand (2004).

$$\hat{w}_{ij} = w[\boldsymbol{X}_i, Pr(\hat{w}_{ij}, \boldsymbol{Z}_i)] \tag{9}$$

To solve the simultaneity, we estimate (9) following a two-step procedure.[11] In the first step, we estimate $Pr(\cdot)$ using a reduced form of it with expected wages being substituted by its determinants $\boldsymbol{X}_i$: $Pr(\boldsymbol{X}_i, \boldsymbol{Z}_i)$, where $\boldsymbol{X}_i$ captures, in an indirect way, the wage effects upon $Pr(\cdot)$. In the second step, the conditional probabilities $Pr(\boldsymbol{X}_i, \boldsymbol{Z}_i)$ are included in the wage equations based on the results by Lee (1984). Define z_i as a vector containing $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$; selectivity-adjusted wage equations are estimated in the following way (we exclude the individual subindex for clarity):

$$w_s = \boldsymbol{X}\beta_s + \sigma_s\rho_s\left(\frac{\phi(H_s(z\gamma_s))}{F_s(z\gamma_s)}\right) + \varepsilon_s \tag{10}$$

where $\sigma_s\rho_s$ are parameters capturing the covariance between the wage and selection equations; $H_s(z\gamma_s)$ is a transformation of the multinomial logit index, $z\gamma_s$, into a standard normal distribution; ϕ is the standard normal density function and $F_s(z\gamma_s)$ is the marginal distribution of the multinomial logit residuals. Wage equations and the conditional probabilities from the selection equation share vector $\boldsymbol{X}$, therefore the identifying variables are contained in $\boldsymbol{Z}$. We define $\boldsymbol{Z}$ bearing in mind that its components must affect reservation wages but not market wages. In the case of female labourers, $\boldsymbol{Z}$ includes the number of children in the household (less than 7 years old); a dummy variable equal to one if the woman is not head of the household and the head is active; the income of all other household members and its quadratic form. As has been argued in Attanasio, Low and Sanchez-Marcos (2004), female labour participation is closely linked to household income variability and economy-wide shocks. To take this into account, the last element of $\boldsymbol{Z}$ is a measure of the variation of all other household members' income.[12] All these variables are expected to have a significant effect upon female reservation wages without affecting their expected market remuneration.

[11] We could have estimated (9) using maximum likelihood techniques; however, since we are interested in the robustness of our results to different selectivity-parameterization rules we opt for the two-step procedure. Given the two-step nature of the procedure, all the standard errors presented in Section 4 are corrected via bootstrapping methods.

[12] To construct this variable, we segmented the population into different labour cohorts (education, experience and working position); we used this information to compute the variance of all other household members' income (see Appendix A for details.)

Estimation of informal sector 'wages' using specification (10) implicitly assumes that this sector is *complete* and therefore it remunerates the marginal productivity of labour as the outcome of personal characteristics. Mercouiller, Ruiz and Woodruff (1997) find that returns to personal characteristics in the Mexican informal sector behave quite like those in the formal sector. The same study and those by Maloney (1999) and Gong, van Soest and Villagomez (2000) suggest that, controlling for personal characteristics, the informal sector in Mexico is a desirable destination rather than an inferior forced option. A special feature that is present in the informal sector that might have an advantage over its formal counterpart is the flexibility in working hours. To account for this occupational *attribute*, we include the standard deviation of working hours ($\tilde{h}$) in each sector as a determinant of participation and occupation. A note of caution is necessary at this point. Notice that $\tilde{h}$ will only vary across occupations but not between individuals. The same can be said about the intercept in (6)—the first element of $\boldsymbol{Z}$. Therefore including $\tilde{h}$ and allowing for a different intercept for each occupation will result in perfect multicolinearity. To avoid this problem, our estimations assume that the intercept for the informal sector equation is equal to zero. In other words, all particular attributes attached to the informal sector (apart from $\hat{w}$) will be captured by $\tilde{h}$.

Finally, the structural participation and occupation function is estimated using a *generalized* multinomial logit[13] using $\hat{w}_i$ as the fitted values of (10) for each sector and including $\tilde{h}$ and $\boldsymbol{Z}_i$ as regressors:

$$Prob(i = s) = \frac{exp[\delta \hat{w}_{is} + \varphi \tilde{h}_s + \boldsymbol{Z}_i \boldsymbol{\gamma}_s]}{\sum_{j=1}^{J} exp[\delta \hat{w}_{ij} + \varphi \tilde{h}_j + \boldsymbol{Z}_i \boldsymbol{\gamma}_j]} \tag{11}$$

3 Data

The model described in section 2 is estimated using Mexican household survey data (ENIGH) for households located in urban areas (communities with 15,000 inhabitants or more) for the years 1994, 1996, 1998 and 2000. Between 1994 and 1996, Mexico experienced great macroeconomic turbulence as a result of the Peso crisis that erupted in December 1994. In 1994, the country embarked on a free trade agreement with Canada and the US. The

[13] In our case, a combination of a conditional and a multinomial logit. Maddala (1983) shows that this two models are mathematically the same, hence I will simply refer to it as a multinomial logit (MNL).

years between 1996 and 2000 were a time of economic recovery, with high rates of growth mainly boosted by manufacturing exports. All these changes could have had a significant impact upon female labour participation and occupation decisions.

To summarize the most important changes taking place in the Mexican female labour market, in Figure 1 we show the annual percentage changes in labour participation, real wages, average formal education and average age.[14] We can see that female participation increased during the period of analysis with the proportion of active women rising from 41.6 per cent in 1994 to 46.6 per cent in 2000, representing a 12 per cent increase throughout the period. This might seem to be a small change, but when we consider the number of total women entering the labour market during those years, the increase is far from being trivial. An increase in female participation of 5 percentage points of the 1994 female working age population represents a total of 707,993 female labourers entering the market over and above the effects due to demographic trends.[15] Of the total amount, around 338,794 of the new entrants took place in the manufacturing sector, 244,684 new labourers went into other formal sectors, and 124,514 ended up in the informal sector. As we can see from the upper right part of Figure 1, these changes translate into an increase in the proportion of total female labourers in the manufacturing sector.

After the Peso crisis (1994-1995), real average wages for women working in urban areas decreased 30 per cent. Although urban areas average wages remained practically unchanged between 1996 and 1998, wages in the manufacturing sector rose 13 per cent during the same period. Manufacturing wages kept rising between 1998 and 2000, time during which wages in other formal sectors began to recover. Real wages in the informal sector showed no constant trend, with a positive change between 1996 and 1998 and an unexpected negative shift between 1998 and 2000.

In the lower right part of Figure 1, we show the average years of schooling for urban women within working age. We find that, on average, the level of formal education rose steadily between 1994 and 2000. Despite this overall

[14] Since ENIGH is not a probabilistic survey, we account for sampling design taking expansion factors, stratification and clustering into account. All the statistical analysis carried out throughout the paper accounts for survey design. See De Hoyos (2005) for details.

[15] The actual increase in female labour participation observed between 1994 and 2000 is 1,779,105 new entrants. The difference between the net entrants (707,993) and the actual one (1,779,105) is explained by an increase in the base population, i.e. by demographic and population changes as well as rural-urban migration during those years.

increase, the average education of female workers in the manufacturing sector, as opposed to the increase experienced in all the other sectors, decreased between 1994 and 1996 and remained below the 1994 level throughout the period. The large increase in female participation, together with a decrease in the educational level observed in the manufacturing sector, makes us suspect that during the years after NAFTA, this sector was absorbing the relatively unskilled female labourers who were entering the labour market. This could be a sign that the boom in the manufacturing sector after the Peso devaluation and the enactment of NAFTA made many unskilled Mexican women more likely to participate in the labour market. Finally, in the lower right part of Figure 1, we can see that entrants into the informal sector were younger, on average, than the incumbents.

Figure 1: Female Participation, Real Wages, Education and Age

Despite the opposite trends in real wages between the periods 1994-96 and 1996-2000, participation in the manufacturing sector always showed positive shifts. The explanation for the increasing participation in the manufacturing sector during the periods 1994-96 and 1996-2000 might be different. Perhaps women's participation between 1994 and 1996 was a response to the large negative income effect brought about by the Peso crisis; on the other hand, during the recovery period 1996-2000, increases in participation could be explained by the rise in real wages in the manufacturing sector.

A third hypothesis explaining the observed increase in participation is related to changes in women's *willingness* to work. Changes in female *shadow price of leisure*, e.g. reductions in 'reservation wage', play a significant role in participation decisions. For example, in a relatively traditionalist society like the Mexican one, female labour participation is still quite dependent upon the husband's labour status, i.e. if the husband (head of the household) is working, the probability of observing an active spouse is lower than if the head was not active. If this type of *dependency*, or any other factor affecting the female reservation wage function, is changing throughout the period, then we could conclude that the observed increase in female labour participation in Mexico between 1994 and 2000 can be explained by three hypotheses: (a) the negative income effect brought about by the crisis; (b) a positive substitution effect (increase in relative wages) after 1996, explained by the increase in manufacturing exports and (c) changes in female reservation wage functions. To be able to test these hypotheses and quantify their effects, a structural model linking market wages, exogenous income effects and female *dependency* with labour participation decisions must be estimated. If hypotheses (a) to (c) are true, the data should support the following statements: income effects are positive, i.e. participation probability decreases as income rises; wage-participation elasticity is positive and finally, the female reservation wage is decreasing over time. To test the validity of these statements, in the following sections we will show the model estimates and, based on this, we will quantify the effects of our three main hypotheses using microsimulation analysis.

4 Empirical Estimates of the Structural Model

In this section, we present the estimation results for the selectivity-adjusted wages (10) and the participation and occupation equation (11) for the years 1994, 1996, 1998 and 2000.[16] Given the large amount of results, the tables

[16]Equation (10) is estimated using our own Stata command, *svyselmlog*. *svyselmlog* is the survey version of the original *selmlog*. The command estimates the parameters of the main equation (in this case wages) correcting for selectivity using a multinomial logit and accounting for survey design effects; several forms of selectivity correction are available. *svyselmlog* is available from the SSC (Boston College) archives.

with the wage equation's estimates for each sector are placed in Appendix B.[17] An important result to notice from the wage equations (Tables 3 to 5 in Appendix B) is that the average return to schooling in the manufacturing sector is lower than the estimated for other formal sectors. This result suggests that the manufacturing sector in Mexico demands relatively less skilled female labourers (measured in years of formal education) compared to the skills demanded in other formal sectors. Another important result comes from the dummy variable capturing wage differentials between the northern states and the rest of the country. Controlling for selectivity, education and experience, during the recovery years 1996 to 2000, female manufacturing labourers in manufacturing industries located in the north of Mexico earned, on average, 20 per cent more than their northern counterparts working in other formal sectors and 25 per cent more than manufacturing labourers in other regions. This result supports the hypothesis of an increase in female labour demand in the manufacturing industry explained, in turn, by the rapid growth in exports in this sector during those years.[18]

Table 1 shows the estimation results of the participation and occupation equation (11) taking 'not active' as the base category, therefore all the parameters of the multinomial part of the equation are interpreted as changes in the probability of choosing a particular occupation relative to not being active.[19] The first two rows contain the estimates of two occupation attributes, i.e. expected wages ($\hat{w}$) and the standard deviation of working hours ($\tilde{h}$). Regarding the latter, the results show that female workers (or possible ones) perceive working hours stability as a positive attribute, therefore a *ceteris paribus* increase in the variance of working hours in a particular occupation reduces its probability of being chosen. As we would have expected *a priori*, an increase in expected wages in a particular sector increases the likelihood of observing workers in that sector. Following our theoretical model, a value of δ of, say, 1 implies a value of λ, the marginal utility valuation of money income, greater than 1. Given an exogenous increase in $\hat{w}$, the probability of observing a worker participating increases (i.e. $\delta > 0$).

The wage-participation elasticity is obtained by computing the marginal effect of δ. The estimated wage-participation elasticity is quite stable over

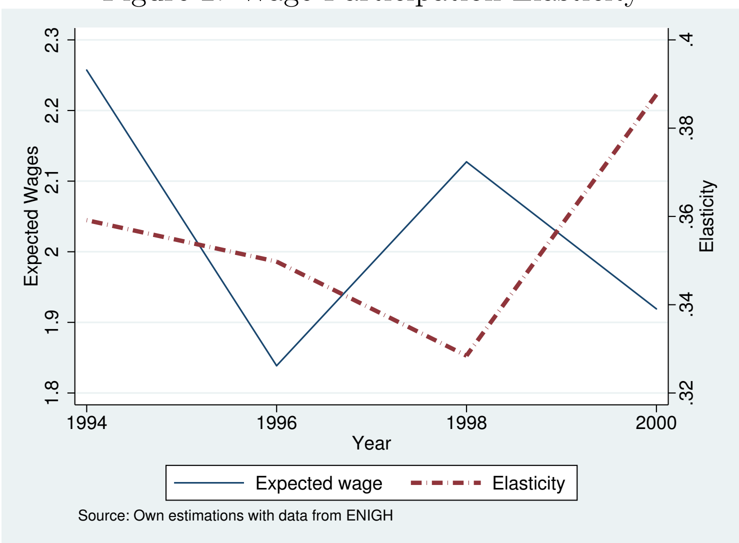
[17] Due to space limitations, we do not present the results from the first stage estimations, $Pr(\boldsymbol{X}, \boldsymbol{Z})$, however they are available from the author upon request.

[18] Most of the 'maquiladoras' (export processing zones) created after 1994 were located in the north of Mexico.

[19] In fact, the parameters in Table 1 are the marginal effects upon the latent function determining the participation probabilities—the utility function (6); the marginal effect of $\hat{w}$ upon the probability of being active is shown in Figure 2.

time, ranging from 0.33 in 1998 to 0.39 in 2000 (see Figure 2). This result supports hypothesis (b) postulated in Section 3, i.e. participation can be partly explained by positive changes in real wages occurring in the export-oriented manufacturing sector. Although a percentage increase in expected wages has a less than proportional increase in the probability of participation, women's responses to market incentives increased between 1998 and 2000. In Figure 2, we plotted the average fitted values of the log of hourly wages ($\hat{w}$) and the estimated wage-participation elasticity. An interesting point from Figure 2 is the large positive reaction of fitted values of wages between 1996 and 1998—although still below the pre-crisis level. In the next section, we will quantify the participation effects of these changes in expected wages.

Figure 2: Wage-Participation Elasticity



From the third row of Table 1 onwards, we show the participation and occupation effects of personal characteristics at the household level. Remember that all variables included in $\boldsymbol{Z}$ affect reservation wages without changing market wages. Hence $\boldsymbol{Z}$'s estimated parameters should be seen as the determinants of females' reservation wage function and orthogonal to $\hat{w}_i$.

We can see that, despite an expected *a priori* negative sign on the parameter relating participation to the number of children, we find that, for most years, this relationship is not significantly different from zero. A plausible explanation for this can lie in the strong family ties observed in Mexico, where the presence of grandparents in the household reduces (or eliminates) child care costs. A very interesting pattern arose in the parameters estimating labour participation dependence of female household members with respect to the labour status of the head of household. We allow for spouses and daughters to have a different response to the head of the household's participation decision. H_s^a and H_d^a are dummy variables taking a value of one when the women is a non-head of household (spouse or daughter respectively) and the head of the household is actively participating in the labour market. In the case of female spouses, the probability of participation decreases when their husband is working ($H_s^a < 0$). However, notice that this effect is decreasing over time, suggesting that women's participation decisions are becoming less dependent upon their husband's labour status. In the case of daughters, the story is completely different, with the probability of being employed in formal sectors increasing when the male head of the household is active ($H_d^a > 0$). These results suggest that the Mexican labour market is changing so that women's labour participation is less subject to their husband's labour status. Another way of interpreting this result is as a reduction in females' reservation wage or, equivalently, an increase in their *willingness* to work. Therefore hypothesis (c) suggested in section 3 also finds support in the data.

Table 1: Structural Model Results

	1994	1996	1998	2000
$\hat{w}$	1.296***	1.447***	1.112***	1.505***
$\tilde{h}$	-0.152***	-0.111***	-0.110***	-0.096***
Tradable Earner				
$Intercept$	-1.880***	-2.580***	-2.537***	-3.403***
$Children$	-0.373***	-0.179**	-0.14	-0.02
H_s^a	-1.133***	-1.137***	-1.011***	-0.767***
H_d^a	0.607***	0.776***	1.047***	1.456***
Y_m^0	-5.616***	-9.676***	-10.954***	-10.484***
$(Y_m^0)^2$	1.015	1.313	6.814*	5.884
$Var(Y_m^0)$	0.002	0.002	0.001	-0.033
Non-tradable Earner				
$Intercept$	-0.799***	-1.146***	-0.439***	-2.233***
$Children$	-0.161**	-0.230***	-0.126*	-0.206**
H_s^a	-1.140***	-1.141***	-0.949***	-0.967***
H_d^a	0.621***	0.173	0.641***	0.709***
Y_m^0	-3.911***	-5.017***	-7.615***	-5.330***
$(Y_m^0)^2$	0.699	0.563	6.005*	2.206
$Var(Y_m^0)$	0.002*	0.002	0.001**	0.001
Informal Sector				
$Children$	-0.12	-0.105	-0.118	-0.036
H_s^a	-0.676***	-0.668***	-0.604***	-0.524***
H_d^a	-0.970***	-0.707***	-0.577**	-0.203
Y_m^0	-9.929***	-13.188***	-11.670***	-11.398***
$(Y_m^0)^2$	1.799	1.929	6.894**	3.955
$Var(Y_m^0)$	0.002	0.002	0.002**	0.002
R^2	0.325	0.293	0.261	0.275
N	32284	36292	27492	24240

*,**,***, significant at the 10 per cent, 5 per cent and 1 per cent level respectively (with bootstrapped SE)

The last three controls included in our estimation, Y_m^0, $(Y_m^0)^2$ and $Var(Y_m^0)$, are capturing, respectively, the income effect, a quadratic form of it and the variance of all other household members' incomes.[20] As we can see from Table 1, the income effect is always positive, i.e. an increase in exogenous income reduces female participation, and with a not-significant second order contribution. Only in year 1998 the second order income effect ($(Y_m^0)^2$) is positive and significant, implying that a positive change in all other household members' income will decrease the probability of female participation and this will be stronger at lower levels of income. Finally, our variable capturing the variance of all other household members' income, $Var(Y_m^0)$, shows the expected positive sign although it is only significantly different from zero in 1998. During that year, female household members increased their participation as a reaction to an exogenous rise in the variance of household income. This could be seen as an optimum reaction in order to smooth consumption in a country with strong borrowing constraints like Mexico.

The empirical evidence presented so far shows strong support for the three hypotheses postulated in section 3. Given the estimated positive income effect, the negative shock caused by the Peso crisis resulted in more female labour participation. Our wage-participation estimates show that the observed increases in real wages in the manufacturing sector after 1996 also accounts for the increase in participation. Finally, reductions in the dependency between women's labour decisions and the head of household's labour status also play a significant role. We test the robustness of our results by using two other methods of selection-adjustment proposed in Dubin and McFadden (1984) and Bourguignon, Fournier and Gurgand (2004). Although the value of the parameters change, the qualitative results hold. The rest of the paper tries to quantify the *relative* importance of our three hypotheses in explaining the increase in labour participation. The results of our microsimulation analysis should be seen as a first approximation of the quantitative impact of exogenous shocks (e.g. crisis, trade liberalization, etc.) upon labour participation.

5 Microsimulation Analysis

How much of the total increase in female labour participation between 1994 and 2000 can we attribute to the negative income effect caused by the Peso

[20]For presentational purposes, all income variables were re-scaled to 1:100,000.

crisis of 1994-1995? What proportion of the increase in participation is explained by changes in female expected wages? How many of the net entrants in the different occupations reported in section 3 are the outcome of changes in females' reservation wage function? Using microsimulation techniques we are able to answer these important questions.[21]

Define $\boldsymbol{\Omega}_t$ as a vector containing all the estimated parameters from the participation function at time t:

$$\boldsymbol{\Omega}_t = (\hat{\delta}_t, \hat{\varphi}_t, \hat{\boldsymbol{\gamma}}_t)$$

Similarly, define $\boldsymbol{\chi}_t$ as a vector containing all variables explaining female labour participation and occupation at time t:

$$\boldsymbol{\chi}_t = (\hat{w}_t, \tilde{h}_t, \boldsymbol{Z}_t)$$

Finally introduce a time subindex in the random component of the *utility* function, η. Female labour participation and occupation decisions at time t are hence a function of the three components just defined:

$$Pr(\cdot)_t = Pr(\boldsymbol{\Omega}_t, \boldsymbol{\chi}_t, \eta_t) \tag{12}$$

Therefore, any change in the probability of female participation between t and t' can be decomposed into changes in parameters ($\boldsymbol{\Omega}$), exogenous variables ($\boldsymbol{\chi}$) and residuals (η).

As we discussed in sections 3 and 4, there are three strong and tested hypotheses explaining the increase in female labour participation: (a) a negative income shock caused by the Peso crisis, (b) a trade-induced positive shift in female labour demand, particularly in the manufacturing sector and (c) a long-run negative trend in women's reservation wage function. Simulating changes in the components of (12), we can quantify the *ceteris paribus* labour participation and occupation effects of our three hypotheses.

[21]For an explanation on the microsimulation technique used in this section see: Bourguignon and Ferreira (2004).

5.1 Measuring the Effects of Changes in $\hat{\beta}$

An exogenous macro shock (e.g. currency crisis, trade liberalization, etc.) will manifest itself as a change in the relative prices of the economy. In the labour market, the most important 'price' is the wage which, in turn, is defined by a set of 'prices' or returns to personal characteristics, $\hat{\boldsymbol{\beta}}$. Between 1994 and 1996, we would expect to observe a negative impact upon 'prices' of personal characteristics in all occupations as a consequence of the Peso crisis. If this is the case, then *expected* wages, $\hat{w}$, would decrease and, given a positive participation-wage elasticity, participation should be lower as a consequence of the crisis. On the other hand, the same negative shock upon the 'price' of personal characteristics, would reduce household incomes and this, in turn, increases the probability of female participation—if and only if income effects are positive, as we actually found out. To quantify these effects, in our labour participation function, exogenous changes in 'prices' will affect elements $\hat{w}$ and $\boldsymbol{Z}$ of component $\boldsymbol{\chi}$. Therefore the *value* of the independent variables defining female participation, $\boldsymbol{\chi}$, will be a function of the returns to personal characteristics, $\boldsymbol{\chi} = \chi(\hat{\boldsymbol{\beta}}, ...)$.[22] To account for overall $\hat{\boldsymbol{\beta}}$-induced changes in household incomes it is necessary to parameterize wages and participation decisions for men. This allows us to find out the *ceteris paribus* effect of changes in 'prices' upon variables: H_s^a, H_d^a, Y_m^0 and $(Y_m^0)^2$. Remember that H_s^a and H_d^a are dummy variables indicating the labour status of the head of the household and Y_m^0 measures all other household members' income. Thus variables H_s^a, H_d^a, Y_m^0 and $(Y_m^0)^2$ will be affected by changes in returns to personal characteristics both in the men's and women's labour market.[23] We therefore estimate men's wage equations and participation functions by following the same empirical strategy as we did for women.[24]

[22]The income variables used in the participation functions (Y_m^0) include the sum of income of *all other* household members; hence, although they are being parameterized here, they can be seen as being strictly exogenous for each particular individual.

[23]If our database was longitudinal, we wouldn't have to compute these hypothetical $\boldsymbol{Z}$ values, we could have use the *observed* change in income for each household. However, we cannot identify the same family in two different points in time, therefore we have to simulate the *exogenous* change in household income based on the observed changes in 'prices' of personal characteristics.

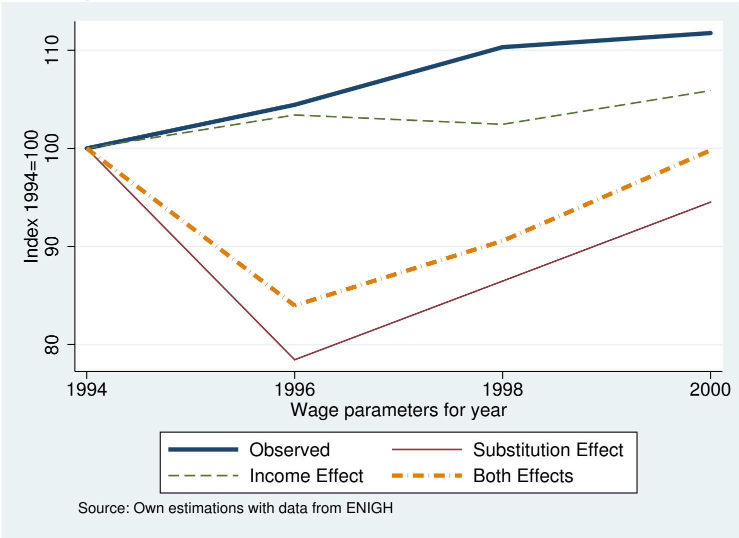
[24]The variables determining men's wages are the same as the ones used for women ($\boldsymbol{X}$). The identification variables $\boldsymbol{Z}$ for men include: the size of the household, and all other household members' income and its squared form. All the estimation results for men are available in De Hoyos (2005a).

A hypothetical or simulated set of expected wages and household characteristics, $\hat{w}^{\imath}$ and $\boldsymbol{Z}^{\imath}$, are created by substituting the estimated returns to personal characteristics for year t' ($\hat{\boldsymbol{\beta}}_{t'}$), into the database for year t. The new vector of returns to personal characteristics will have a direct household income effect via the change in labour income. Furthermore, whilst simulating household incomes, we allow the men in our sample to re-optimize their labour participation and occupation decisions given the new set of expected wages. The procedures we follow are as follows: (1) Estimate the model for men and women using the cross-sectional data for year t and t'. (2) 'Import' the wage parameters of year t' into the parameterized model for t. (3) Compute the new set of hypothetical expected wages and household incomes. (4) Allow male household members to change their occupational status given the new value of expected wages and household income. (5) Simulate household incomes using the hypothetical set of wages and men's occupational status. (6) Finally, we simulate the *ceteris paribus* female participation effects of changes in $\hat{\boldsymbol{\beta}}$ via the expected wages channel and the exogenous household income channel. This allow us to measure the impact—through its many channels—that an exogenous change in the market returns to personal characteristics ($\hat{\boldsymbol{\beta}}$) between t and t' has upon female labour participation:

$$Pr(\cdot)_t^{\imath} = Pr[\boldsymbol{\Omega}_t, \boldsymbol{\chi}_t^{\imath}, \eta_t] \tag{13}$$

Where $\boldsymbol{\chi}_t^{\imath} = \chi_t^{\imath}[\hat{w}^{\imath}(\hat{\boldsymbol{\beta}}_{t'}), \boldsymbol{Z}^{\imath}(\hat{\boldsymbol{\beta}}_{t'})...]$; $Pr(\cdot)_t^{\imath}$ is a *simulated* probability (since it is not observed) at the *micro* level (since we estimate one for each individual in the sample). Using the estimation results of our model plus the estimated parameters for men, the three components of (13), $\boldsymbol{\Omega}_t$, $\boldsymbol{\chi}_t^{\imath}$ and η_t, are observed. To summarize our procedure: we are 'importing' the estimated wage equation parameters ($\hat{\boldsymbol{\beta}}$) for year t' into the data set for year t. Once the parameters are in the database for year t, we simulate a hypothetical expected wage and household income ($\hat{w}^{\imath}$ and $\boldsymbol{Z}^{\imath}$) which we then use to construct $\boldsymbol{\chi}_t^{\imath}$. Finally we multiply $\boldsymbol{\chi}_t^{\imath}$ by $\boldsymbol{\Omega}_t$ and add the residual terms η_t. This will create a new *utility* maximizing decision and therefore a new set of participation/occupation probabilities for each woman of working age (actual and potential worker). We undertake three separate simulations, taking 1994 as the base year and 'importing' the estimated wage parameters for the years 1996, 1998 and 2000, respectively. The simulation results are summarized in Figure 3.

Figure 3: Simulated Participation Effect of $\Delta\hat{\beta}$



In Figure 3, we graph the time trend of the observed and simulated change in participation (with respect to 1994) due to a change in the wage equation parameters. To decompose the effect of $\Delta\hat{\boldsymbol{\beta}}$ even further, we perform three different simulations for each 'imported' $\hat{\boldsymbol{\beta}}$ we compute: (1) a simulation where only expected wages, $\hat{w}$ are allowed to change (continuous thin line); (2) a second one where only the household income elements of $\boldsymbol{Z}$ change (dashed thin line) and (3) a third one where both, $\hat{w}$ and $\boldsymbol{Z}$ are changing as a consequence of $\Delta\hat{\boldsymbol{\beta}}$ (dashed thick line).[25] We perform these separate simulations because we can interpret (1) and (2) as the *substitution* and *income effects*, respectively, of changes in returns or prices of personal characteristics.

For the moment, let us concentrate on the changes taking place between 1994 and 1996. The observed change in participation (the continuous thick line in Figure 3) is positive. Accoring to our model, if expected wages were the only element changing during those years, female participation would have been reduced by 22 per cent (continuous line). Given the positive wage-participation elasticity we found in section 4, this result is explained by the reduction in average $\hat{w}$ between 1994 and 1996 (see Figure 2). Concerning the

[25] Although the sum of the substitution and income effects are very close to the simulation where both effects are allowed, the decomposition methodology that we use does not show additive properties. Therefore the sum of the effects brought about by the different elements in equation 12 is not necessarily the equal to the total effect.

income effects of the crisis, the estimated reduction in payments to personal characteristics had a negative effect upon household's income (Y^0) and also upon men's participation decisions.[26] The *ceteris paribus* simulated participation effect of $\boldsymbol{Z}^\imath$ is shown in Figure 3 (dashed thin line). We can see that had the change in household incomes and head of the household participation decisions—as a consequence of a negative shock in $\hat{\boldsymbol{\beta}}$—been the only change taking place between 1994 and 1996, then female labour participation would have increased as much as the observed increase during those years. This is explained by the crisis' negative income effect which 'pushed' more women into the labour market. The final simulation presented in Figure 3 includes both changes, $\hat{w}^\imath$ and $\boldsymbol{Z}^\imath$ together in the same simulation. The simulated net participation effect is negative, in other words, had remunerations to personal characteristics decreased in the way they did between 1994 and 1996, female participation would have decreased 13 per cent, *ceteris paribus*.

Let us now analyze the changes that occurred during the recovery period 1996-2000. As $\hat{\boldsymbol{\beta}}$ experienced a positive change, household incomes increased and, given a positive income effect, the participation rate attributable to the income effect decreased, although it remained above the 1994 value (thin dashed line). Regarding wage effects, the increase in returns to personal characteristics taking place between 1996 and 1998 (see Figure 2) explains the simulated increase in female labour participation between 1996 and 1998 (continuous thin line) although still below the 1994 participation rate. These results show that, controlling for all other changes taking place in the economy between 1994 and 2000, shifts in the returns to personal characteristics would have *decreased* the participation rate by 0.2 per cent out of a total *increase* of 12 per cent observed during that period.[27]

So far, we have simulated the participation rate that we would have observed if returns to personal characteristics were the only elements changing in the economy, but what about the simulated occupation effects? In Appendix D we present the simulated percentage change in the female participation rate in the manufacturing, other formal and informal sectors. Surprisingly, the simulated change in participation attributable to shifts in $\hat{w}$ in the manufacturing sector between 1994 and 1996 is positive. Quite the opposite can be said for other formal sectors. These results suggest that, in the absence of labour rationing, the isolated participation effects of changes in $\hat{\boldsymbol{\beta}}$ would

[26] See Appendix C with the simulated mean household income brought about by $\Delta\hat{\boldsymbol{\beta}}$.

[27] This result depends very much on the non-rationed labour markets assumption, i.e. labour participation and occupation decisions are purely the outcome of a utility maximizing process and do not face labour demand restrictions.

have triggered female labour participation in the manufacturing sector and reduced participation in other formal sectors during the crisis years 1994-96. The effect is explained by the increase in $\hat{\boldsymbol{\beta}}$ taking place in the manufacturing sector even during a time of general contraction of the economy (1994-96).[28] The positive simulated participation trend in the manufacturing sector remains between the years 1996 and 1998 and then slows down in the period 1998-2000. Regarding the occupational changes brought about by the 1994-1996 negative income shock, notice how it has a much larger effect upon participation in the informal sector (a 20 per cent increase). This result suggests that in the presence of borrowing constraints, the informal sector acted as a 'cushion' absorbing female labourers, particularly young ones, that were 'pushed' into the labour market by the negative income shock.

To summarize, we have shown that the positive post-1994 trend in overall urban labour participation described in Figure 1 is explained partly by increases in manufacturing sector participation and, to a lesser extent by participation in the informal sector. The explanations behind these shifts are completely different. On the one hand, the positive trends in manufacturing participation are explained by positive shifts in $\hat{\boldsymbol{\beta}}$ for manufacturing female labourers. On the other hand, the increase in the informal sector participation is explained by a negative income shock. Therefore the data supports the hypothesis of a increase in female labour demand taking place in the manufacturing sector between 1994 and 2000 and an increase in particiaption as a concequence of the 1994-95 Peso crisis, primarily in the informal sector.

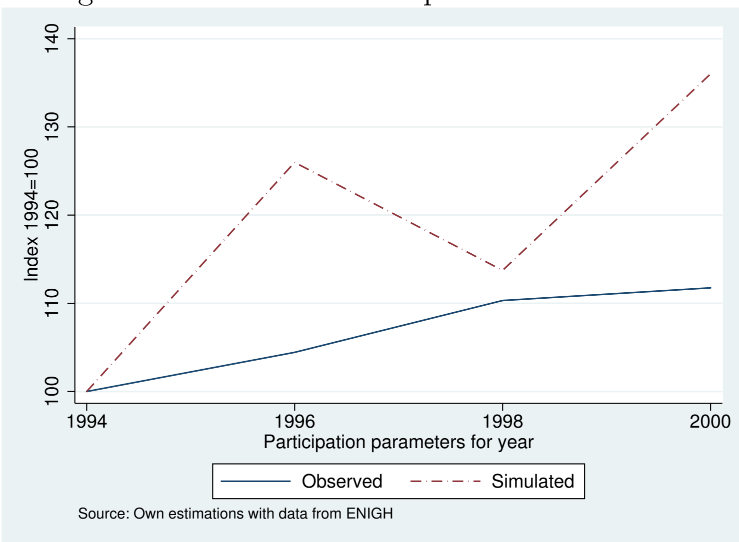
5.2 Measuring the Effects of Changes in $\boldsymbol{\Omega}$

The previous simulations uncovered the participation and occupation effects of exogenous changes in wage functions parameters which in turn affected the vector of explanatory variables, $\boldsymbol{\chi}$. All the counterfactuals constructed so far had evaluated the labour participation effects of changes in 'prices' of personal characteristics. In this section, we will simulate the participation and occupation effects that are attributable to changes in the participation equation parameters, $\boldsymbol{\Omega}$.

[28] At this point, is it worthwhile to remind the reader that, since we are controlling for selectivity, $\hat{\boldsymbol{\beta}}_j$ can be interpreted as a sector j-specific *treatment* effect. I.e. a relative increase in $\hat{\boldsymbol{\beta}}$ in the manufacturing sector, compared to other sectors, is indicating a manufacturing-specific wage premium.

Following the last section's counterfactual analysis, to capture the dynamics of changes in $\boldsymbol{\Omega}$, we take the year 1994 as the base and 'import', in separate simulations, the participation function parameters for the years 1996, 1998 and 2000. The way we interpret the results is similar to the previous simulations, i.e. the simulated participation and occupation decisions yield the *ceteris paribus* effect of changes in $\boldsymbol{\Omega}$ observed between t and t'. Based on our structural model, we can think of changes in $\boldsymbol{\Omega}$ as changes in women's reservation wage function parameters, or their subjective *willingness* to work. As we mentioned before, the parameters of the participation function (equation 11), should not be affected by changes in market conditions, therefore the simulations can be interpreted as capturing *exogenous* changes in women's willingness to work. The results of the simulations are shown in Figure 4.

Figure 4: Simulated Participation Effect of $\Delta\Omega$



From Figure 4 we can see that if the only changes observed between 1994 and 1996 had been the changes in participation function parameters, then we would have seen an increase in participation of 25 per cent. Given that the net effect of changes in $\boldsymbol{\chi}$ decreased participation during the crisis years 1994-95 (see Figure 3), we know that the increase in female participation observed after the crisis was, at least partly, the result of changes in the female reservation wage function. The positive participation effects remain between 1996 and 1998 though reducing in magnitude; between 1998 and 2000 it increases again to end up with a simulated increase of 35 per cent with respect to the observed rate for 1994.

In Appendix E, we present the simulated changes in occupation given the new set of parameters $\mathbf{\Omega}$. The results are revealing. The change in participation function parameters between 1994 and 1998 made participation in the non-manufacturing formal sectors more likely. This could be seen as a sign of a general increase in the preference for other formal sectors as opposed to the manufacturing sector.

The simulated participation capturing the effects of changes in $\mathbf{\Omega}$ revealed that changes in women's *willingness* to work would be enough to explain the total observed increase in the female participation rate between 1994 and 2000. If we interpret $\mathbf{\Omega}$ as being part of the function defining women's *price of leisure*, we can argue that female reservation wages decreased during the crisis and then returned to a higher value in 1998.

Therefore, we showed that the increase in participation which was observed between 1994 and 1996 is explained by the combination of a reduction in the female reservation wage and the negative household income effect brought about by the Peso crisis. During the recovery period, the change in participation is explained by an increase in expected wages, on the one hand, and an increase in women's reaction to market incentives (wage-participation elasticity), on the other. As we discussed in Section 5.1, the occupational effect of changes occurring during the crisis show that the trade-driven buoyant manufacturing sector was absorbing most of the new entrants throughout this period.

6 Summary and Conclusions

We developed a structural model describing female labour participation and occupation decisions. In our model, women's participation and occupation decisions are taken in a simultaneous way, with participation decisions being embedded in the occupational one. The structural model is used to define and interpret a microeconometric model suitable for estimation. We corrected for selectivity in the wage equations parameterizing the conditional probabilities deriving from a multinomial logit, as suggested by Lee (1984). Our model creates an explicit and causal relationship between expected selectivity-adjusted wages and partcipation/occupation decisions. Other factors such as the labour status of the head of the household, the number of children and all other household members' incomes are used as controls within a multinomial logit framework.

We apply the model to Mexican urban household data for the years 1994, 1996, 1998 and 2000. The estimated wage-participation elasticity fluctuates from a lower value of 0.33 to an upper one of 0.39. Between 1994 and 2000, female participation in Mexico rose by 12 per cent. Three hypotheses for the documented increase in participation have been postulated and tested in this paper: (a) a negative income shock caused by the Peso crisis of 1994-95, (b) a trade-induced increase in female labour demand in the manufacturing sector and (c) a change in the female reservation wage function. The results from our micro model support all three hypotheses, suggesting that the observed increase was the outcome of simultaneous and, sometimes, opposing effects being at work between 1994 and 2000.

Using microsimulation techniques, we were able to quantify the female participation/occupation effects attributable to changes in the parameters defining the participation and wage functions between 1994 and 2000. The results show that the increase in participation observed between 1994 and 1996 is only partly explained by the negative income shocks of the Peso crisis. Negative changes in female reservation wage function (the parameters in the participation equation) also help us explain most of the observed post-crisis increases in female participation. During the recovery period 1996-2000, increasing female participation is the outcome of relatively higher expected wages in the manufacturing sector and changes in the participation function parameters.

There is still plenty of scope for future increases in female labour participation in Mexico. We found a significant expansion in female labour market opportunities, primarily in the export oriented manufacturing sector. Non-market related changes in women's *willingness* to work also help to explain the recent positive trend. Given women's increasing reaction to market incentives (a higher wage-participation elasticity), a change towards a more *progressive* labour market should be supported and encouraged by the Mexican government via female training programs, dissemination of information about labour opportunities for women, promotion of export-oriented firms and, above all, the education of the female population.

References

[1] Attanasio, O., Low, H. and Sánchez-Marcos, V. (2004) 'Female labor supply as insurance against idiosyncratic risk', forthcoming, *Journal of the European Economic Association, Papers and Proceedings.*

[2] Bourguignon, F. and Ferreira, F. (2005) 'Decomposing changes in the distribution of household incomes: methodological aspects', in F. Bourguignon, F. Ferreira and N. Lustig, editors, *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, Oxford University Press.

[3] Bourguignon, F., Fournier, M. and Gurgand, M. (1998) 'Wage and labour force participation in Taiwan', *Mimeo* DELTA, Paris.

[4] Bourguignon, F., Fournier, M. and Gurgand, M. (2004) 'Selection bias correction based on the multinomial logit model: Monte-Carlo comparisons', *Mimeo* DELTA, Paris.

[5] De Hoyos, R.E. (2005a) 'The microeconomics of inequality, poverty and market liberalizing reforms', *UNU-WIDER* Resaerch Paper 2005/63, Helsinki.

[6] De Hoyos, R.E. (2005b) 'SVYSELMLOG: Stata module to compute selectivity adjustment based on the multinomial logit for survey design', *Statistical Software Components*, no. S454901, Boston College.

[7] Deaton, A. and Muellbauer, J. (1980) *'Economics and consumer behavior'*, New York: Cambridge University Press.

[8] Dubin, J.A. and McFadden, D. (1984) 'An econometric analysis of residential electric appliance holdings and consumption', *Econometrica*, vol. 52, 2.

[9] Gong, X. and van Soest, A. (2002) 'Wage differentials and mobility in the urban labor market: a panel data analysis for Mexico', *Labour Economics*, vol. 9.

[10] Gong, X., van Soest, A. and Villagomez, E. (2000) 'Mobility in the urban labor market: a panel data analysis for Mexico', *CentER discussion paper*, Tilburg University.

[11] Heckman, J. (1974) 'Shadow prices, market wages and labor supply', *Econometrica*, vol. 42, 4.

[12] Heckman, J. (1979) 'Sample selection bias as a specification error', *Econometrica*, vol. 47, 1.

[13] Heckman, J. (1990) 'Varieties of selection bias', *The American Economic Review*, vol. 80, 2.

[14] Heckman, J. (1993) 'What has been learned about labor supply in the past twenty years?', *The American Economic Review*, vol. 83, 2.

[15] Heckman, J. and Honore, B.E. (1990) ‘The empirical content of the Roy model’, *Econometrica*, vol. 58, 5.

[16] Heckman, J. and Sedlacek, G. (1985) ‘Heterogeneity, aggregation, and market wage functions: an empirical model of self-selelction in the labor market’, *Econometrica*, vol. 93, 6.

[17] Lee, L.F. (1983) ‘Generalized econometric models with selectivity’, *Econometrica*, vol. 51, pp. 507-512.

[18] Maddala, G. (1983) *‘Limited dependent and quantitative variables in econometrics’*, Cambridge: Cambridge University Press.

[19] Maloney, W. (1999) ‘Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico’, *The World Bank Economic Review*, vol. 13, 2.

[20] Marcouiller, D., Ruiz, V. and Woodruff, C. (1997) ‘Formal measures of the informal sector wage gap in Mexico, El Salvador, and Peru’, *Economic Development and Cultural Change*, vol. 45, 2.

[21] McFadden, D (1974), ‘The measurement of urban travel demand’, *Journal of Public Economics*, vol. 3.

[22] Roy, A.D. (1951) ‘Some thoughts on the distribution of earnings’, *Oxford Economic Papers*, vol. 3.

Appendix

A Measuring Household Income Variance

To estimate the way in which household income variations affect female labour participation we have to construct a variable which is able to capture variations in household incomes as *perceived* by **each** household member. Let us start by defining a measure of household income variance. In a cross-sectional framework, we cannot estimate the time-variation in household incomes, therefore we have to use the observed variation within socioeconomic and demographic groups. These groups were defined according to the following observable characteristics: gender, formal education, experience and position in their working place. The combination of these characteristics formed a total of 263 groups, each of them containing different average income (μ) and variance [$Var(Y)$]. Household income variance is computed using the within population cohorts variance. The way in which the population-cohort and the household variances are linked must be consistent with a measure of household variance showing some desired properties. Following Attanasio, Low and Sanchez-Marcos (2004), we define Axiom 1:

[**Axiom 1**] In the presence of borrowing constraints, where rational agents' optimal choice is to smooth consumption, women's participation probability increase as a response to a rise in the *expected* variance of household incomes. Therefore, the *observed* income variance of a household whose female member is active should be smaller than the one that we would have observed had participation not occurred.

Our preferred measure of household income should comply with Axiom 1. It turned out that the measure of household income variance that we propose here satisfies Axiom 1.

While forming income variance expectations, agents are aware of all other household members' characteristics, e.g. a wife knows the characteristics of her husband and children. Assume that the mean income and variance attached to each of the 262 population cohorts is constant over time and that this information is known by each of the household members. In the process of deciding whether to participate in the labour market or not, agents form their personal *perception* of what the variance of household income will be. This personal *perception* is the expected value of the variance of all other household members' income. For agents who are not participating in the labour market:

$$Var(Y_h^i) = \sum_j^m \frac{\mu_j}{\sum_j^m \mu_j} Var(Y_j) \quad \forall\, j = 1...m : \textit{Active Members} \tag{14}$$

Each non-active household member's expectation will be a weighted average of income variance of all active members. The weight assigned to each active member j is formed by the average income within j's population cohort divided by the sum of average incomes of all the population cohorts that the active household members belong to. In the case of active members, the expected variance will be formed by the variance that would be observed if they decided to abandon the labour market. This is equivalent to creating a household income variance with the following counterfactual: what would the variance have looked like had the agent decided not to participate. This definition of variance for the active members is comparable to the one computed for non-active members, since it is computing the household variance as if the active member were not active:

$$Var(Y_h^{i'}) = \sum_j^{m-1} \frac{\mu_j}{\sum_j^{m-1} \mu_j} Var(Y_j) \quad \forall\, j = 1...m : \textit{Active Members} \tag{15}$$

Given our definition of household income variance, a sufficient condition for it to satisfy Axiom 1 is:

[**A1**] The population cohort variance [$Var(Y)$] of the marginal women entering the labour market is smaller than the observed weighted average values for all other household members.

If **A1** is true, then the counterfactual variance—i.e. the variance in the absence of their participation—for participating women will be larger than the observed one. Equivalently, the total household income variance should decrease as an outcome of their participation. If **A1** is true and this information is known by female household members, then participating in the labour market is a consumption-smoothing decision. This is exactly the property stated by Axiom 1. For the Mexican labour market, **A1** turns out to be an empirical regularity, i.e. the observed income variance within female population subgroups is less than the observed statistic for male subgroups.

Table 2: Average Income and Variance

	Men	Women
μ	4,704	3,358
Var(Y)	5.24e+07	1.14e+07

B Selectivity-Adjusted Wages

Table 3: Wage Functions for the Manufacturing Sector

	1994	1996	1998	2000
$Schooling$	0.138***	0.105***	0.149***	0.111***
$Schooling * I(Ys > 11)$	-0.004	0.022*	0.003	0
$Experience$	0.068***	0.042***	0.071***	0.031**
$Experience^2$	-0.001***	-0.001**	-0.001***	0
$North$	0.074	0.141**	0.269***	0.333***
$Pr(manufacture)^{\dagger}$	0.275	0.142	-0.076	0.105
$Intercept$	0.355	0.579*	0.295	0.801**
R^2	0.267	0.246	0.28	0.236
N	491	609	511	428

Notes:

(1) *,**,***, significant at the 10 per cent, 5 per cent and 1 per cent levels respectively

(2) Bootstrap standard errors with 200 replications

(3) Data source: ENIGH 1994, 1996, 1998 and 2000

(4) $Pr(\cdot)^{\dagger}$ are computed according to equation 10

Table 4: Wage Functions for Other Earning Sectors

	1994	1996	1998	2000
$Schooling$	0.148***	0.136***	0.143***	0.131***
$Schooling * I(Ys > 11)$	0.020***	0.014***	0.021***	0.011*
$Experience$	0.077***	0.069***	0.060***	0.057***
$Experience^2$	-0.001***	-0.001***	-0.001***	-0.001***
$North$	-0.046	0.004	0.083*	0.105**
$Pr(other\ earner)^{\dagger}$	0.290***	0.245***	0.243*	0.065
$Intercept$	0.337**	0.21	0.059	0.637***
R^2	0.468	0.374	0.409	0.402
N	2213	2393	1950	1850

Notes:

(1) *,**,***, significant at the 10 per cent, 5 per cent and 1 per cent levels respectively

(2) Bootstrap standard errors with 200 replications

(3) Data source: ENIGH 1994, 1996, 1998 and 2000

(4) $Pr(\cdot)^{\dagger}$ are computed according to equation 10

Table 5: Wage Functions for the Informal Sector

	1994	1996	1998	2000
$Schooling$	0.081***	0.064***	0.052**	0.037
$Schooling * I(Ys > 11)$	0.013	0.004	0.034*	0.026
$Experience$	0.023	0.046***	0.033*	0.063**
$Experience^2$	0	-0.001**	0	-0.001*
$North$	-0.124	0.034	-0.076	-0.096
$Pr(informal)^{\dagger}$	0.062	0.364	0.143	0.701*
$Intercept$	1.368**	0.272	0.902	-0.23
R^2	0.053	0.062	0.052	0.084
N	620	857	663	581

Notes:

(1) *,**,***, significant at the 10 per cent, 5 per cent and 1 per cent levels respectively

(2) Bootstrap standard errors with 200 replications

(3) Data source: ENIGH 1994, 1996, 1998 and 2000

(4) $Pr(\cdot)^{\dagger}$ are computed according to equation 10

C Simulated Household Income Effects of $\Delta\hat{\boldsymbol{\beta}}$

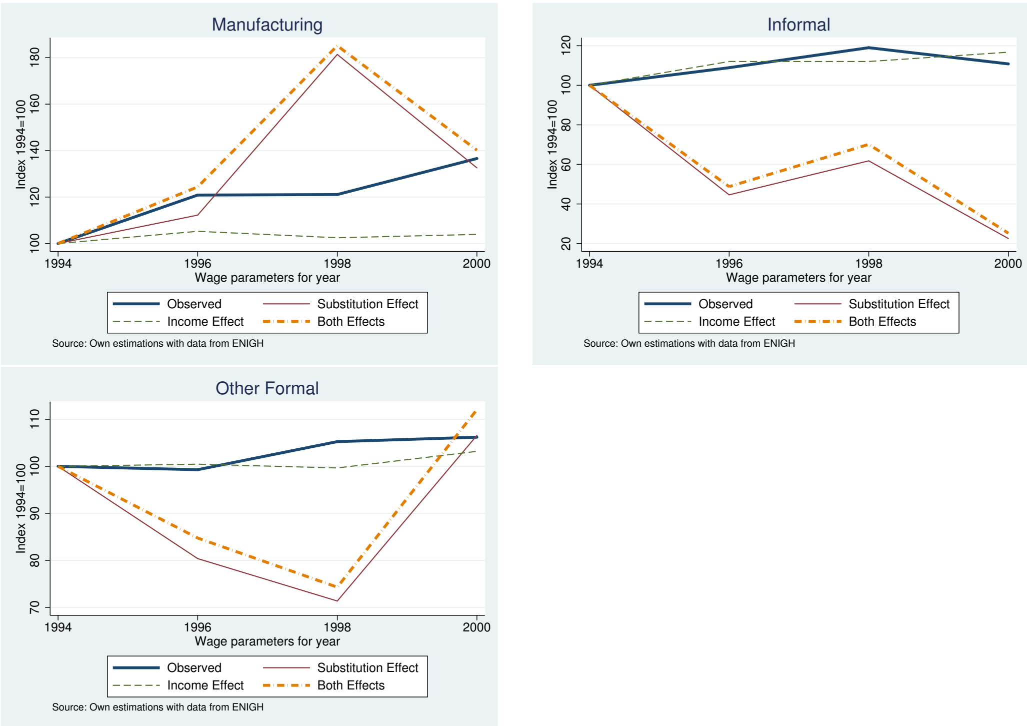
In order to capture within-household participation effects of $\Delta\hat{\boldsymbol{\beta}}$, we parameterise household incomes, including female and male members. We assume that men's participation decisions are independent from all other household members' labour status. Female household members, on the other hand, decide whether to enter the labour market or not taking into account all other household members' income (Y_m^0) and the head of the household labour status (H^a). H^a and Y_m^0 are endogenous once full household incomes have been parameterised. Hence an economy-wide shock upon $\hat{\boldsymbol{\beta}}$ will have an effect upon these two variables. In the next table, we show the simulated values of H^a and Y_m^0 for the different estimated values of β. The change in remunerations to personal characteristics resulted in an increase in male participation—the result is totally explained by increases in manufacturing expected wages. Therefore, had the change in β been the only change occurring between 1994 and 1996, the proportion of active men would have increased from 88.8 per cent to 92.1 per cent. The simulated values correspond to the observed increase in the male participation rate between 1994 and 1998. Regarding simulated Y_m^0, the trend follows very closely the observed path, with a huge negative shock between 1994 and 1996 and a gradual recovery thereafter. The advantage of the microsimulation over a simple distributional-neutral change in average incomes is that we can capture the changes in each and every household in our dataset. Hence we do not need to assume that over-all economy shocks have an homogenous effect upon every household.

	1994	1996	1998	2000
H^a				
Observed	0.888	0.889	0.902	0.920
Simulated	-	0.921	0.929	0.843
Y_m^0				
Observed	10,955	7,539	7,825	8,912
Simulated	-	7,091	7,950	8,807

Income figures are in real Pesos of August 2002

D Simulated Occupation Effect of $\Delta\hat{\boldsymbol{\beta}}$

Figure 5: Simulated Occupation Effect of $\Delta\hat{\boldsymbol{\beta}}$



E Simulated Occupation Effect of $\Delta\Omega$

Figure 6: Simulated Occupation Effect of $\Delta\Omega$