Completed fertility and the transition from low to high order parities: a double-hurdle approach

Alfonso Miranda*
Economics Department
University of Warwick

Email: ecrgw@warwick.ac.uk
Fax: 00 44 2476 523032

This version: June 2004

Abstract

This paper reports a study on the socio-economic determinants of completed fertility in Mexico. Special attention is given to study how socio-economic characteristics such as religion and ethnic group affect the likelihood of transition from low to high order parities. An innovative Double-Hurdle count model is developed for the analysis. Findings indicate that education and Catholicism are associated with reductions in the likelihood of transition from parities lower than four to high order parities. Being an indigenous language speaker, in contrast, increases the odds of a large family.

Keywords: Completed Fertility, Count Data, Double-Hurdle.

JEL Classification No.: J13, J15, C25.

---

*I am grateful to Wiji Arulampalam and Mark Stewart for their useful comments. I am indebted as well to the Mexican National Council for Science and Technology (CONACYT) for its financial support. A previous version of this paper was presented at the 2003 Annual Conference of the British Society for Population Studies, Bristol.
Correspondence Address: Department of Economics, University of Warwick, Coventry, CV4 7AL, UK. E-mail: ecrgw@warwick.ac.uk.

1. Introduction

The present paper presents a study on completed fertility in Mexico. As is well known, fertility data have special features that need explicit econometric modelling. In the case of the developed world, for instance, data often exhibit under-dispersion and a relative excess of zero and two counts. Data from developing countries like Mexico, in contrast, are commonly over-dispersed and do not contain a particularly large excess of two outcomes. This sort of data, however, poses other important challenges to the analyst. Namely, that a non-negligible proportion of cases are contributed by women who have a large number of children and who tend to move to high order parities without taking any action to limit their fertility. In fact, in the case of Mexico nearly 21% of women end their fertile life with more than six children (INEGI 1999) and use contraceptives much less intensively than women with fewer children (Gomez 1996).

Among other potential explanations, this sort of behaviour may be displayed because women with large families find themselves ‘locked’ in a regime in which the opportunity cost of extra children becomes particularly low. A large family, for example, may imply a permanent exit from the labour market and lead to further increases in family size. Clearly, some explicit account of this sort of behaviour is required when, as reported in Mexico, a good proportion of women give birth to a large number of children. Otherwise results will be difficult to interpret and most likely subject to serious bias.

Two main econometric avenues may be taken. One alternative would be to specify a Generalized Poisson Process, or *pure birth process*, as the main analysis technique and allow transition intensities to depend on women's accumulated stock of children – i.e., to introduce occurrence dependence in the stochastic process that generates completed fertility data (for further details see Winkelmann 2000). This possibility is exploited in Faddy (1997), Faddy and Bosch (2001), and Podlich, Faddy and Smyth (2004) in an extended count data framework. Applications, however, require the solution of a set of differential equations for which an exact analytical solution is not available. Numerical methods are needed and thus considerable computing power demanded.

An alternative approach would consider the assumption that low and high order fertility counts are drawn from different data generating mechanisms which do not exhibit occurrence dependence on their own. In such a context women move from one to another regime when their fertility crosses certain pre-established thresholds – say, zero and three children. Such an avenue, which is in line with the literature on hurdle count models (Mullahy 1986), is taken in the present work to develop a Double-Hurdle count model. The Double-Hurdle model is estimated by standard maximum likelihood techniques and can be easily extended to account for unobserved individual heterogeneity and endogenous switching across regimes. No special demands on computing power are involved. The Double-Hurdle model is used to study in detail how socio-economic characteristics such as religion and ethnic group affect the probability of transition from low to high order parities in Mexico.

The rest of the paper is organised as follows. Section two presents a brief discussion of the general institutional background on population issues in Mexico. Section three introduces the data. Section four discusses econometric issues and section five presents the empirical results. Finally, section six concludes.

2. Institutional background

In the last forty years consistent and significant reductions in total fertility rate (TFR) in Mexico have been registered - it went from 6.5 children per woman in 1970 to less than 3 children per woman in 2000 (INEGI 2000, INEGI 2001a). Among other factors the reduction in fertility is associated with an important decline in infant mortality, which in the period 1970-2000 passed from 68.5 to 17.5 deaths for each 1,000 births (INEGI 2000, INEGI 2001a). Other development indicators witness as well important improvements in the living standard of Mexican citizens. In fact, between 1970 and 1999 average education increased from 3.4 to 7.6 years and life expectancy went from 61 to 75 years (INEGI 2001b). During the same period of time real GDP per capita increased by 57% and urban population went from representing 60% to 75% of total population. Finally, female participation in the labour force (female workers /total women of working age) increased from 11% to 27% (World Bank 2001). All these aspects of modernization are likely to have influenced fertility reduction in Mexico.

Improvement in development indicators, however, is not homogeneous across broad ethnic groups. For instance, in the year 2000 the infant mortality rate among Mexican Indians was 1.2 times higher than the corresponding figure for Mexico as a whole (CONAPO 2002) Similarly, in 1997 average education in the indigenous population was reported to be three years lower than average education in the non-indigenous population (INEGI 1999). Obviously, differences in standards of living are reflected in differences in fertility rates. In fact, CONAPO (2001a) estimates that in the year 1996 Total Fertility Rate (TFR) for indigenous individuals was 67 percent higher than the corresponding figure for non-indigenous individuals – 4.7 children per Indian woman compared to 2.8 children per non-Indian woman.

Public policies are another important factor explaining fertility decline. In 1973 the Mexican government initiated for the first time a public programme to offer free contraceptives and to promote family reduction as a rational and responsible behaviour among Mexican citizens. Simultaneously, all previous legal restrictions on the sale of contraceptives were lifted. Between 1973 and 1979 these `family planning' campaigns targeted potential users of contraception in urban and sub-urban zones. But at the onset of the 1980s rural zones were also integrated into the campaigns (Cabrera 1994). During the last 20 years the geographical coverage of such campaigns increased significantly. However, universal access to modern contraceptives is still far from reality. Despite the failure to provide universal access to contraception, population policy in Mexico is widely considered a success, as the diffusion and adoption of modern contraceptives has increased dramatically in the past few decades. In fact, while in 1976 thirty percent of all married women - or those living in consensual union - were active users, in 1998 the figure was estimated to be seventy per cent (INEGI 2001b). Today, and since the late 1970s, the public sector constitutes the main source of contraceptives in the country though private supply remains important (INEGI 2001b).

A dramatic change in the composition of the demand for contraceptives is one of the most significant stylised facts of the last twenty years. Indeed, at the end of the 1970s nearly 35% of all users adopted the contraceptive pill, 19% IUD, and 9% permanent female sterilization (PFS). In contrast, in 1998 51% of users adopted PFS, 24% IUD, and 10% traditional methods. At this last date, the pill was selected by less than six per cent of all active users of contraception (INEGI 2001b). Gomez (1996) indicates that most young Mexican individuals do not adopt contraception before the arrival of a first or second child, and that many of them adopt PFS or IUD as their preferred method. In addition, the author finds that women with two or three children are responsible for most of the demand for contraceptives in Mexico. That is, the prevalence in the use of contraceptives among women with more than four children is much lower. Because of this, he concludes, women with more than three children are self-selected into a high-parity group, in contrast to women with less than four children who are self-selected into a low-parity group.

Demographers explain the observed trends in the demand for contraceptives as the outcome of various factors. They mention that the public health system in Mexico has undertaken a deliberate effort to promote the adoption of definitive natal control (definite contraception) among women who have three or more children. In fact, most of the 'delivery effort' of contraceptives has been concentrated on reaching women looking to initiate natal control after they reach their desired family size. According to Zavala de Cosio (1990), this policy has contributed to generate and to disseminate a new fertility norm among Mexicans, but at the same time it has bias the demand for contraceptives towards PFS and IUD. Lindstrom (1998) finds that Mexican women fear - many times on unfounded grounds - undesired side effects of hormonal contraceptives (such as cancer) and unwanted pregnancy due to their possible ineffectiveness. On the basis of these findings, the author suggests that fear to undesirable side effects of hormonal-based contraceptives is the main reason for the observed shift to PSF among Mexican women.

3. Data and Variable definition

Data from the National Survey of Demographic Dynamics 1997 (ENADID from its acronym in Spanish) is used. The ENADID is a micro-data set containing detailed economic and demographic information for 88,022 Mexican women aged between 15 and 54 years. Since completed fertility is the main concern of this study, a total of 19,477 cases of women aged 40 or over at the time of the ENADID interview (December 1997) are selected.

From a theoretical point of view it is not clear whether fertility decisions are taken in terms of lifetime number of pregnancies, lifetime number of live births, or lifetime number of surviving children. Obviously, lifetime number of pregnancies is the broadest concept as it is the cumulative sum of every conception a woman has during her fertile life. Number of live births excludes voluntary and involuntary miscarriages as well as stillbirths. Finally, number of surviving children removes infant deaths up to a certain age, say, age five. Most economic models of fertility choice consider that individuals decide in relation to the number of surviving children rather than over number of pregnancies or live births. That is, individuals choose the number of children they would like to have at the end of their fertile life, without regard to the number of pregnancies required to reach such a number of decedents (see for instance

Bergstrom 1989, Willis 1973). Hence, the death of a child is thought to induce a new pregnancy (or a series of failed pregnancies) such that final family size remains constant. In the same line of thought, unwanted children would be abandoned to die in the absence of better means of birth control.

In applied work, in contrast to the ideas mentioned above, the common practice is to define lifetime fertility as the number of children ever born live to a woman by the end of her childbearing period (see for instance Santos Silva and Covas 2000, Melkersson and Rooth 2000). The convention in applied work seems to be as arbitrary as the convention in theoretical work. Given that child mortality is not explicitly considered, the present work adopts the convention in theoretical literature. Therefore, completed fertility will be defined as total number of at least 5-years-old surviving children ever born to a woman during her lifetime, **children**. **Children** is the dependent variable. According to the descriptive statistics (see Table 1) **children** has mean 4.43 and variance 7.56. The data is therefore over-dispersed.

Figure 1 and Table 3 present details on the empirical distribution of **children**. For comparison proposes a theoretical Poisson distribution with mean 4.4 is also depicted. Notice first that, like data generated in developed countries, Mexican data exhibits an excess of zeroes relative to a Theoretical Poisson. This feature is found in most fertility data and various strategies for dealing with it have been introduced in the literature, including hurdle and zero-inflated count models (see the very informative surveys of Cameron and Trivedi 1986, Winkelmann 1995, Winkelmann 2000). Second, unlike data collected in developed countries, Mexican data do not contain a relative excess of one and/or two counts in reference to a Poisson distribution. Thus, there is no need here to inflate the probability of one and/or two counts. Finally, and more importantly, the Poisson distribution under-predicts the probability of observing counts 4, 5 and 6.

Looking closely at Figure 1 one may conclude that women who have more than three children seem to behave differently with respect to women who have a completed fertility of up to three. While women with less than four children, excluding zero outcomes, are well described by a standard Poisson, women with more than three children tend to transit to high parities more frequently than predicted. In fact, according to the data in Table 4, 53% of women who have more than three children transit to parities higher than five. And among those with more than five, 69% end fertile life with seven children or more. Intuitively, women who have four or more children may find themselves in a regime where the cost of an extra child is lower than the cost they would pay if their current fertility were lower than four. A fourth child could imply, for instance, a permanent exit from the labor market and a corresponding reduction in the opportunity cost of extra children. Although observed and unobserved heterogeneity are yet to be accounted for, these are relevant features of the data that the analyst should not neglect.

Controls for women's religion, ethnic group, education at age 12, cohort of age, and place of birth are included as explanatory variables (see table 1). The definition of these variables is as follows:

Catholic. Binary indicator that takes value one if the woman is catholic and zero otherwise. Defining two broad religious groups seems to be the finest sensible classification for Mexico given that nearly 90% of Mexicans are Catholics and a further 7% are Protestants.

Indspker. Dummy variable indicating whether an individual is able (indspker = 1) or unable (indspker = 0) to speak an indigenous language. **Indspker** proxies broad ethnic group (indigenous/mixed) rather than specific socio-cultural community. Clearly, neither indigenous nor mixed populations are homogeneous socio-cultural entities in Mexico. However, a broad ethnic-group classification seems to be sensible because attitudes towards contraception, family size, and female work are mostly traditional across indigenous groups (i.e., against remunerated female work and modern contraception), and contrast with modern attitudes commonly found among mixed individuals. **Indspker** presumes that indigenous individuals keep the ability to speak their own language and declared so to the ENADID interviewer. Obviously, in some cases an individual may have lost her indigenous-language skills but remains culturally indigenous. And some bilingual women may have hidden their language skills at the time of the ENADID survey. Therefore, **Indspker** is potentially recorded with measurement error. However, if present, such an error is likely to be small and non-correlated with observed and unobserved variables that may affect fertility – including **Indspker** itself.[1]

Edu12. Proxy variable for women's completed years of education at age 12. **Edu12** is an indicator of skills and human capital accumulated before the onset of reproductive life. Given that primary education in Mexico is composed of six compulsory grades and children initiate their instruction at age six, **Edu12** is bounded between zero and six and is not subject to individual choice. However, in rural and marginal urban zones there is a limited supply of education services and in some cases schools do not offer the six compulsory primary education grades. Long-term financial difficulties of the parental household may also result in a permanent dropout of their dependent children from primary education, especially in marginal zones where education law is not rigorously enforced. Temporary dropouts are unusual and course repetition is rarely extended beyond age 12. All these childhood 'contextual' factors induce variation in education at age 12 in Mexico. Clearly, though children have little influence on their early education there is still the possibility that **Edu12** may be endogenous. However, as is usual in most data sets, no valid instruments for education are available in the ENADID. Thus, **Edu12** is treated as an exogenous variable and the reader should interpret the results with due care.

Due to the lack of detailed information **Edu12** is built under a set of assumptions. First, as enforced by the federal law, it is supposed that all children initiate their primary education at age 6. Second, it is supposed that all children attend school continuously until the date of their definite dropout. Finally, it is assumed that none fails an attended course. These assumptions guarantee that completed years of education at age 12 may be calculated on the basis of information on women's date of birth and their current completed years of education – data indeed available in the ENADID. In practice, obviously, children may start education after age 6, drop out temporarily, and/or repeat some courses. **Edu12** thus contains some potential measurement error. This error, however, is likely to be small and, if present, it is supposed to be random and uncorrelated with all observed and unobserved explanatory variables (including **Edu12** itself). This is, once again, a strong assumption and results should properly be qualified.

Cohort of age. Using information on women's date of birth five cohorts can be defined, from 1940-1944 to 1955-1957. Four binary dummy variables indicating cohort of age are then generated (=1 if born in the corresponding 5-year period): **c4044**, **c4549**, **c5054** and **c5559**. The first cohort is taken as reference group.

Place of birth. Four regional geographic dummies for place of birth are defined: **MexCity** (base group), **North**, **Centre** and **South**.[2] There are important differences in the features of the data across the four geographical zones. Mean value and standard deviation of the dependent variable vary significantly from one region to the other, the South being the zone where the highest mean count is registered. Moreover, Mexican Indians are clearly concentrated in the South and Centre of the county. Important variations of education at age 12 are also detected across the different geographic zones (see Table 1b).

4. Econometric issues

As was discussed earlier in the text, Mexican completed fertility data exhibit some characteristic features: an excess of zeros, a recognizable proportion of women choosing a completed fertility between one and three children, and a characteristic excess of large counts contributed by women that seem to move from low to high order parities without taking measures for limiting their fertility. Clearly, successful modelling should therefore consider that the various values of the dependent variable might be generated by different mechanisms. Otherwise results are difficult to interpret and important bias might be present.

4.1 A double-hurdle model

Let individual's i-th completed fertility be y_i. The objective is to estimate a model for the probability that a fertility count j would be observed for the i-th individual from a random sample Y={$y_1,\ldots,y_n$}. The model is formulated as follows. First a standard Poisson Hurdle model (Mullahy 1986) is considered,

$$\Pr(y_i = j) = \begin{cases} \exp(-\mu_{0,i}), & j = 0 \\ \left[1 - \exp(-\mu_{0,i})\right]\Pr(y_i \mid y_i > 0), & j = 1,2,3... \end{cases} \qquad (1)$$

where the parameter $\mu_{0,i}$ maintains a deterministic log-linear relationship with a kx1 vector $x_{i,0}$ of explanatory variables (including the constant term),

$$\mu_{0,i} = \exp\left(x_{0,i}'\beta_0\right), \qquad (2)$$

β_0 is its kx1 vector of associated coefficients, and Pr(y_i | y_i > 0) represents the probability distribution function of y_i given that a positive count has been observed. Notice that, unlike most Hurdle models reported in the literature, equation (1) uses an Extreme Value (EV) distribution for modelling the probability of observing a zero count. Specifying EV rather than the commonly selected Normal or Logistic distributions has two advantages in the present context. First, in contrast to Normal and Logistic, Extreme Value delivers a non-symmetric distribution for the binary outcome model in equation 1 (see Arulampalam and Booth 2001). Second, since EV and Poisson predict the same Pr(y_i = 0), for practical proposes the hurdle in equation (1) can be seen as governed by a standard Poisson model.

Equation (1) represents a standard Hurdle Model. The model stresses the fact that the decision of entering parenthood is qualitatively different from the decision on the actual number of children, given that a strictly positive count is desired. To put it in other words, the Hurdle stresses the fact that zero and strictly positive counts may be generated by two different mechanisms. In order to allow for a second hurdle modifications are introduced in Pr($y_i \mid y_i > 0$),

$$
\Pr(y_i = j \mid y_i > 0) = \begin{cases} \left[1-\exp(-\mu_{1,i})\right]^{-1} \dfrac{\exp(-\mu_{1,i})\mu_{1,i}{}^{j}}{j!}, & j = 1,2,3 \\ \left[1 - \sum_{k=1}^{3} \left[1-\exp(-\mu_{1,i})\right]^{-1} \dfrac{\exp(-\mu_{1,i})\mu_{1,i}{}^{k}}{k!}\right] \Pr(y_i \mid y_i \geq 4), & j = 4,5,6... \end{cases} \tag{3}
$$

with,

$$
\mu_{1,i} = \exp\left(x_{1,i}'\beta_1\right). \tag{4}
$$

A standard Hurdle specifies Pr($y_i \mid y_i > 0$) as a zero-truncated Poisson distribution. In contrast, equation (3) considers the case where counts in the [1,3] and [4,∞) intervals are drawn from two different data generating processes. For the [1,3] interval a zero-truncated Poisson distribution is written as usual. However, for counts larger than three, a new distribution Pr($y_i \mid y_i \geq 4$) is introduced. Clearly Pr($y_i \mid y_i \geq 4$) will be truncated at three and, to guarantee a well behaved probabilistic model, it should be re-scaled so that Pr($y_i \mid y_i > 0$) sums up to one. Since equation (3) is similar to equation (1) in its philosophy, one could interpret the count process for the [1,3] interval as a second hurdle. From this perspective the probability of crossing such a barrier is given by

$$\Pr(y_i > 3 \mid y_i > 0) = \left[1 - \sum_{k=1}^{3} \left[1 - \exp(-\mu_{1,i})\right]^{-1} \frac{\exp(-\mu_{1,i})\mu_{1,i}{}^{k}}{k!}\right].$$

To close the model a functional form for Pr(y_i | $y_i \geq 4$) must be specified. For convenience a Poisson distribution is, once again, selected:

$$\Pr(y_i = j \mid y_i \geq 4) = \left[1 - \sum_{h=0}^{3} \frac{\exp(-\mu_{2,i})\mu_{2,i}{}^{h}}{h!}\right]^{-1} \frac{\exp(-\mu_{2,i})\mu_{2,i}{}^{j}}{j!}, \qquad j = 4,5,6... \quad (5)$$

As usual,

$$\mu_{2,i} = \exp\left(x_{2,i}'\beta_2\right). \qquad (6)$$

The model is identified as long as vectors $x_{0,i}$, $x_{1,i}$ and $x_{2,i}$ are of full rank. In principle $x_{0,i}$, $x_{1,i}$ and $x_{2,i}$ may contain some (or all) common elements and no exclusion restrictions are required to achieve identification. Similarly, the vector of parameters β_0, β_1 and β_2 are estimated without constraints. Notice that if $\beta_1 = \beta_2$ the Double-Hurdle model (DHM) collapses to a standard Poisson Hurdle model. Moreover, if $\beta_0 = \beta_1 = \beta_2$ a simple Poisson model is obtained. Hence, the advantages of DHM over standard Poisson Hurdle and Poisson models may be assessed by testing for the equality of β_0, β_1 and β_2. Parameters are estimated by maximum likelihood. The contribution of the i-th individual to the overall likelihood is simply

$$
\begin{aligned}
L_i = &\prod_{y_i=0} \exp(-\mu_{0,i}) \prod_{y_i>0} \left[1-\exp(-\mu_{0,i})\right] \prod_{1\le y_i \le 3} \left[1-\exp(-\mu_{1,i})\right]^{-1} \frac{\exp(-\mu_{1,i})\mu_{1,i}{}^{y_i}}{y_i!} \\
&\prod_{y_i \ge 4} \left[1-\sum_{k=1}^{3}\left[1-\exp(-\mu_{1,i})\right]^{-1} \frac{\exp(-\mu_{1,i})\mu_{1,i}{}^{k}}{k!}\right] \\
&\prod_{y_i \ge 4} \left[1-\sum_{h=0}^{3} \frac{\exp(-\mu_{2,i})\mu_{2,i}{}^{h}}{h!}\right]^{-1} \frac{\exp(-\mu_{2,i})\mu_{2,i}{}^{y_i}}{y_i!}.
\end{aligned}
\tag{7}
$$

At convergence minus the inverse of the Hessian matrix $-H^{-1}$ estimates the covariance matrix. Usual asymptotic hypothesis testing is valid. The likelihood function is separable. Therefore, estimates can be obtained by maximizing separately three different likelihood functions. First, a binary outcome model (the first two terms of equation 7) can report consistent and efficient estimates for β_0. Then, a model for a left truncated and right censored Poisson variable can properly estimate β_1 (third and fourth terms of equation 7: for further details see Terza 1985). Finally, a model for a left truncated Poisson (the fifth term of equation 7) can estimate β_2. Separating the likelihood function into three independent elements is possible because selection into zero, one-to-three, and larger-than-three fertility groups is exogenous.

To summarize, notice that Double-Hurdle models are composed of three parts: (i) an Extreme Value distribution governing the likelihood that a woman will remain childless for her entire lifetime, (ii) conditional on having a strictly positive outcome, a Poisson distribution governing the likelihood of observing any particular count in the [1,3] interval, and finally (iii) conditional on having more than three children, a Poisson distribution governing the likelihood of observing any count larger than or equal to four. The model has a Double Hurdle interpretation because in order to observe an outcome equal or larger than four it is necessary first to register a strictly positive count (i.e., to cross the first hurdle) and then to move to parities higher than three (i.e., to cross the second hurdle). The structure of the model is graphically represented in Figure 2.

Selection among different specifications will be based on an Akaike information criterion (AIC) statistic. For completeness, selection on the basis of a consistent Akaike information criterion (CIAC) statistic will be also performed,

$$\begin{aligned} AIC \;\; &= -2\ln(L) + 2k \\ CIAC &= -2\ln(L) + k\{\ln(n)+1\}, \end{aligned} \tag{8}$$

where k represents the number of parameters to be estimated. A best fitting model achieves the minimum AIC and CIAC among all its potential competitors.

In the count data literature competing models are also assessed by means of a goodness-of-fit χ^2 statistic. To calculate such a statistic the analyst must first predict, for each individual, the probability of observing r = 0,1,2,… children on the basis of the estimated model. The resulting probabilities are thus summed over individuals to obtain the predicted number of women with r children, $\hat{n}_r$. Finally the statistic is calculated as,

$$\chi^2 = \sum_{r=0}^{R} \frac{(n_r - \hat{n}_r)^2}{\hat{n}_r}, \tag{9}$$

where n_r represents the actual number of women with r children in the sample. The statistic has a χ^2 distribution with *R-1* degrees of freedom (Melkersson and Rooth

2000, Heckman and Walker 1990). A low value χ^2 is evidence of good fit and the best preferred model should have minimum χ^2 among all potential alternatives.

4.2 Unobserved heterogeneity

The model is easily extended to allow for unobserved individual heterogeneity. A general strategy would consider the inclusion of a random term in each section of the Double Hurdle,

$$\mu_{k,i} = \exp\lfloor x_{k,i}'\beta_k + v_{k,i}\rfloor, \quad k = 0,1,2 \tag{8}$$

Next, some assumptions about the distribution of $v_{0,i}$, $v_{1,i}$, and $v_{2,i}$ will be required to fully specify the model. Joint Normality is a natural choice.

This general approach has, however, two important drawbacks. First, various levels of numerical integration are needed so that estimation will be computing-intensive - particularly in the most interesting case where $v_{0,i}$, $v_{1,i}$, and $v_{2,i}$ are not orthogonal. Clearly, in many applications the computing cost may become large or even prohibitive. Second, and more substantially, there are no theoretical reasons to believe that selection into each fertility group is dependent on different unobservables. Tastes towards children, for instance, are likely to enter every single part of the Double-Hurdle model. To avoid the aforementioned problems one could rewrite equation (8) as

$$\mu_{k,i} = \exp\lfloor x_{k,i}'\beta_k + \theta_k v_i\rfloor; \quad \theta_2 = 1,\ k = 0,1,2. \tag{9}$$

Under the new specification there is conceptually only one unobserved random factor but its impact varies in each part of the Double-Hurdle via the inclusion of three factor loadings θ_0, θ_1, and θ_2. Since only two factor loads are identified θ_2 will be standardised to one. If σ^2 represents the variance of the random effect v, one could show that

$$\begin{aligned}
&\operatorname{var}[\log(\mu_2)] = \sigma^2 \\
&\operatorname{var}[\log(\mu_k)] = \theta_k^2\sigma^2, \quad k = 0,1
\end{aligned}$$

and,

$$\begin{aligned}
&\operatorname{cov}[\log(\mu_0), \log(\mu_1)] = \theta_0\theta_1\sigma^2 \\
&\operatorname{cov}[\log(\mu_2), \log(\mu_k)] = \theta_k\sigma^2, \quad k = 0,1.
\end{aligned}$$

Hence, over-dispersion is allowed in any component of the Double-Hurdle and correlation of any sign between the μ's may be accommodated. In a few words, the simplification does not impose serious loss of flexibility.

Once unobserved heterogeneity is included the likelihood function is no longer separable. Therefore, from this perspective selection into zero, one-to-three, and larger-than-three fertility groups is now endogenous and all parameters $\{\beta_0, \beta_1, \beta_2, \theta_0, \theta_1, \sigma^2\}$ must be estimated in a simultaneous fashion (other models with endogenous selectivity have been suggested by Greene 1997, Terza 1998, Winkelmann 1998). Notice, however, that given v_i all sections of the conditional likelihood function remain independent. Consequently, the unconditional likelihood function is simply written as

$$L_i = \int_{v_i} L_i(v_i) g(v_i) dv_i, \qquad (10)$$

where $L_i(v_i)$ represents the conditional likelihood function. The model is closed once a distribution for the unobserved heterogeneity term, $g(v_i)$, is specified. Here a Normal distribution will be used. Since the integral in equation (10) does not accept a closed solution Gauss-Hermite quadrature may be used to approximate it. As usual, the model is estimated by maximum likelihood and at convergence $-H^{-1}$ estimates the covariance matrix.

Tests for the significance of θ_0, θ_1, and σ^2 may be used to assess the adequacy of the specification for the unobservables in the Double-Hurdle model. If the null $\theta_0=0$ cannot be rejected, then unobserved heterogeneity does not enter the first hurdle (i.e., the count process that determines the probability of remaining childless for a entire lifetime). Similarly, if $\theta_1=0$ then there is no unobserved heterogeneity in the second hurdle. Finally, if $\sigma^2=0$ unobserved heterogeneity will be absent in the overall model. Clearly, testing $\sigma^2=0$ requires a boundary-value likelihood ratio test. Given that the admissible range of θ_0 and θ_1 is the whole real line, testing for $\theta_0=0$ and $\theta_1=0$ may be performed on the basis of standard likelihood tests.

4.3 Relation to the literature

To the knowledge of the author no previous study has used a Double-Hurdle count data model similar to the one introduced in the present paper. There are, however, two main previous efforts to control explicitly for the special characteristics that completed fertility data exhibit. On one hand, Melkersson and Rooth (2000) point out that, due to social norms, completed fertility data from developed countries commonly exhibit an excess of zero and two counts. In such a context Melkersson and Rooth suggest the use of a zero and two inflated count model. On the other hand, Santos Silva and Covas (2000) argue that social norms discourage individuals in developed societies from having an only child. Thus, if for instance a woman enters motherhood, the chances of observing an only child at the end of her fertile life are lower than predicted by standard count models. To control for this tendency to avoid an only child, Santos Silva and Covas develop a modified hurdle model that deflates the probability of observing such an outcome.

Double Hurdle models are widely used in the econometrics literature in various application fields. Existing models, however, are based on the modified Tobit-like model of Cragg (1971) and have a different philosophy from the Double-Hurdle model presented here. In particular, previous work has considered the case where the variable of interest must cross two different hurdles to achieve a strictly positive value. In the case of tobacco (alcohol) consumption, for instance, it is argued that a zero outcome might be equally reported for individuals who never smoke (drink) during their life - or up to the date of data collection - and for individuals who have smoke (have drunk) once but have quit the habit in the past (Yen and Jensen 1996, Blaylock and Blisard 1993, Jones 1989, Labeaga 1999). Clearly, at-least-once and current participation in the smoking (drinking) activity are potentially two different decisions. Thus, observing a strictly positive level of consumption implies that two hurdles have been crossed. Yen, Tan and Su (2001) offer a count data model with similar characteristics to the Tobit-like Double-Hurdle of Cragg (1971). Unlike previous work, the Double-Hurdle presented in this paper considers the case where the second hurdle occurs in a strictly positive value (interval) of the variable of interest. Hence, the approach is essentially different.

5. Empirical results

In this section the empirical results of a study on the socio-economic determinants of completed fertility in Mexico are presented. Special emphasis is given to enquiring how socio-economic factors such as religion and ethnic group affect the likelihood of transition from low to high parities.

5.1 Insights from standard hurdle models

Table 5 contains empirical results from standard Poisson hurdle models. For comparison purposes the hurdle at zero is modelled with an EV binary variable model in place of the usual Probit or Logit specification. Two cases are considered. Column (1) reports estimates from a hurdle model with no added unobserved heterogeneity, while column (2) reports estimates from a model where Normal unobserved heterogeneity is allowed in the post hurdle count process – i.e., for counts larger than zero. Model (2) is an important extension of model (1) as it relaxes the restrictive equi-dispersion assumption of the Poisson distribution.

To start with, notice that, though v_i is detected to have small variance, the presence of unobserved heterogeneity is strongly supported by the data via a significant positive estimate for σ^2 (see column 2 of Table 5). In fact, a boundary-value likelihood ratio test for H_0: $\sigma^2 = 0$ rejects the null at any conventional significance level with a $\chi^2(01)$ of 296. These results are consistent with the previously discussed observation that unconditional variance (7.5) is larger than unconditional mean (4.43).

According to Table 5 the likelihood of remaining permanently childless is significantly affected only by the education of the index woman – see the top panel of Table 5. In fact, a likelihood ratio test for the exclusion of **catholic**, **indspker**, **c4549** trough **c5559**, and **north** through **south** is not rejected with a χ^2 (8) = 14.6 and p-value = 0.067. The coefficient on **edu12** is reported to be negative, implying that women with a higher level of education at age 12 are more likely to remain permanently childless than women with a lower level of education at age 12. These findings conform economic theory in the sense that individuals with a higher level of education are expected to have a large opportunity cost of bearing children in relation to the cost paid by individuals with a lower level of education (Willis 1973).

Regarding strictly positive outcomes, a negative and significant coefficient on **Catholic** in models (1) and (2) indicates that Catholic individuals have fewer children than individuals with other religious backgrounds – see the bottom panel of Table 5. This is an interesting finding given the widespread opposition of the Catholic Church to the use of contraceptives as a way of limiting family size, an attitude that is traditionally thought to be a barrier to fertility reduction. The result is better understood if one considers that despite its formal opposition, the Catholic Church in Mexico has in practice been tolerant towards the adoption of contraceptives as a way of limiting family size. In fact, beyond some insignificant negative campaigns implemented by radical catholic associations – not directly related to the Catholic Church – no efforts to fight against the use of contraceptives have been undertaken in Mexico (Cabrera 1994). Under these circumstances other group-specific characteristics of the Catholic community may induce a negative coefficient on **Catholic**, say, its opposition towards out-of-wedlock sex. Other factors may also be at work. For instance, the existence of a large base of contraception users within the Catholic community may imply that a Catholic individual receives better information about the advantages of family planning relative to a non-Catholic individual.

The proxy for broad ethnic group **Indspker** has a positive coefficient attached, though it is significant only at a 5% significance level. Besides differences in culture, it is likely that the coefficient on **Indspker** may reflect differences in standards of living between indigenous and non-indigenous individuals in Mexico. As is well known, most indigenous individuals in Mexico live in small rural communities (particularly in the south) that are far from the main industrial centres. In such localities health and education services are very limited and most individuals live with a high degree of marginality (CONAPO 2001b).

According to the results in Table 5, education at age 12 has a negative and significant effect on completed fertility. This finding clearly supports theory suggesting that investment in human capital increases the opportunity cost of children (Willis 1973). A negative coefficient on **Edu12** is also consistent with recent literature stressing the idea that education might increase the bargaining power of women within the household (see for instance Klawon and Tiefenthaler 2001, Eswaran 2002, Hindin 2000).

All coefficients on cohort-of-age dummies are negative and significant (base group 1940-1944.) These results are clearly in line with the general trend that Mexican period fertility rates, including the total fertility rate TFR, have showed in the last forty years. Pair-wise tests for the equality of the coefficients on **c4549**, **c5054** and **c5559** reject the null at any conventional confidence level. More importantly, results indicate that younger cohorts of women have larger coefficients attached to their age-specific dummy. Hence, there is strong evidence that younger cohorts of Mexican women are reducing their lifetime fertility in comparison to the experience of older cohorts.

5.2 Results from double-hurdle models

5.2.1 Model selection

Table 6 presents the empirical results. For comparison proposes various specifications are reported. Column (1) contains estimates for a Double Hurdle model that does not control for the presence of unobserved individual heterogeneity. Similarly, Column (2) through (4) contain estimates for Double Hurdle models with Normal unobserved heterogeneity and three different assumptions about factor loadings. Namely, these are (a) $\theta_0= \theta_1=0$, (b) $\theta_0=\theta_1=1$, and (c) θ_0 and θ_1 free. Notice that θ_2 has been standardized to one in all cases. Case (a) corresponds to a model where unobserved heterogeneity enters exclusively in the count process (iii). In addition, selection among regimes is exogenous in the sense that the log-likelihood function can be factored into three independent components. Case (b) removes the assumption of exogenous selection but constrains unobserved heterogeneity to have a symmetric effect in all (i), (ii) and (iii). Finally, case (c) removes all restrictions on the unobservables so that for each regime a different random effect is estimated. Correlation (of either sign) among random effects is explicitly allowed. Hence, the log-likelihood cannot be factored into three independent components. In other words, there is endogenous regime selection.

A significant positive estimate for σ^2 is detected in all the alternative models with heterogeneity (column 2 through 4). In fact, a boundary-value likelihood ratio test for $\sigma^2 = 0$ rejects the null at any conventional significance level with a $\chi^2(01)$ of 78.53 for model (2), 48.62 for model (3), and 78.52 for model (4). Further, pair-wise selection performed on the basis of Akaike and Consistent Akaike information criteria strongly favours (2), (3) or (4) over (1). In a few words, unobserved heterogeneity is present and significant.

Table 7 presents a series of likelihood ratio tests that help discrimination among the different models. The first row of the top panel considers a test on the overall significance of θ_0 taking $\sigma^2 \neq 0$ as a premise and imposing no constraints on θ_1. Clearly, this is a test for H_0: $var(\log(\mu_0)) = \theta_0\sigma^2 = 0$ against H_1: $var(\log(\mu_0)) \neq 0$. Table 6 reports a $\chi^2(1)$ statistic of 0.016 for this test. Hence, the null hypothesis cannot be rejected at any conventional significance level. A similar LRT (see second row of table 7) fails to reject H_0: $var(\log(\mu_1)) = 0$ against H_1: $var(\log(\mu_1)) \neq 0$. But if H_0: $\sigma^2 = 0$ is tested against H_1: $\sigma^2 \neq 0$ a $\chi^2(01) = 78.53$ [p-val = 0.000] is obtained, indicating that unobserved heterogeneity cannot be ignored overall. These results support, then, a model where unobserved heterogeneity enters exclusively in the process that governs the realisation of large outcomes. That is, in the truncated-at-three Poisson distribution (iii). The bottom panel of Table 7 reports further evidence that $\theta_0 = \theta_1 = 0$ and $\sigma^2 \neq 0$ is the correct specification. Selection on the basis of Akaike and Consistent Akaike information criteria supports the same conclusion (see bottom of Table 6).

Before moving to discuss how explanatory variables affect fertility behaviour, it is worth pointing out that alternative assumptions about the distribution of unobservables have a limited, almost negligible, impact on the estimates. Thus results seem to be robust to various assumptions about unobservables.

5.2.2 Test for the joint equality of the coefficients

The following discussion reports findings from a model where unobserved heterogeneity enters exclusively in the Poisson process that governs the realisation of large outcomes (i.e., θ_0 and θ_1 are set to zero). As discussed in the previous section, this is the specification that fits best the ENADID data. The results are reported in Table 6. From now on the vector of parameters that enter count process (i) of the Double Hurdle model will be referred to as β_0. Similarly, parameters that enter count process (ii) and (iii) are referred to as β_1 and β_2.

Table 8 contains a formal likelihood ratio test for the joint equality of the coefficients β_1 and β_2. The reported $\chi^2(10)$ statistic takes a value of 164.27, which is enough evidence to reject the null at a 1% significance level. Similar tests strongly reject $\beta_0 = \beta_1$ with a $\chi^2(10) = 1610.30$ [p-val=0.000], and $\beta_0 = \beta_1 = \beta_2$ with a

$\chi^2(20)$ = 2339.49 [p-val=0.000]. In a few words, neither Poisson nor hurdle at zero Poisson are supported by the data (notice that in either case unobserved individual heterogeneity is being controlled for). The Double-Hurdle model is therefore preferred.

Comparing the elements of vector β_1 and β_2 various interesting observations can be made. Education at age 12, religion and ethnic group have a larger effect in the transition from low to high parities – i.e., the likelihood of crossing the 1-3 hurdle – than in determining fertility once the second hurdle has been crossed. This observation is supported by the fact that the coefficients on **Catholic**, **Indspker** and **Edu12** are larger in absolute value in vector β_1 than in vector β_2. However, pair-wise tests for (Coefficient on variable j in β_1) = (Coefficient on variable j in β_2) reject the null hypothesis exclusively in the case of **Edu12** with a t-stat = -2.27 [p-val=0.0115]. A similar exercise reveals that there are significant pair-wise differences in the coefficients on **c4549** (t-stat = 1.61, pval = 0.053), **c5054** (t-stat = 2.55, pval = 0.0054), c5559 (t-stat = 4.89, pval = 0.0000), **centre** (t-stat = -1.70, pval = 0.0444) and **south** (t-stat = -3512, pval = 0.0000). Hence, differences in the likelihood of crossing the one-to-three children and the likelihood of observing any particular count larger than three are mainly driven by education, cohort of age and place of birth. It is important to underline here that cohort of age and birthplace dummies have larger coefficients in β_2 than in β_1, implying that the impact of these socio-economic characteristics on family size is stronger once the second hurdle has been crossed.

5.2.3 Advantages of the Double-Hurdle model

Table 9 contains a detailed comparison of predicted sample distributions generated on the basis of standard Hurdle and Double-Hurdle models. Only predicted probabilities from a best fitting Double-Hurdle are reported (i.e, a model with $\theta_0 = \theta_1 = 0$). To obtain the figures presented in Table 9 the likelihood of observing any particular count, from zero to eighteen, must be estimated for each individual using the relevant model and conditioning on their observed characteristics. Individual-specific predicted probabilities should then be averaged over all individuals (cell by cell) and the results collected for tabulation. In the bottom section of Table 9 a goodness-of-fit chi-square statistic is reported for each competing model along with Akaike and Consistent Akaike information criterion statistics.

If models that do not control for unobserved heterogeneity are compared, goodness-of-fit chi-square statistics for standard Hurdle and Double-Hurdle are, respectively, 371 and 150. Even controlling for unobserved heterogeneity Double-Hurdle (chi-square = 150) does better than standard Hurdle (chi-square = 213). Therefore, empirical evidence suggests that Double-Hurdle models fit noticeably better the data than the standard Hurdle – similar conclusions may be obtained on the basis of Akaike and Consistent Akaike information criteria. It must be stressed here that even the best fitting Double-hurdle with Normal unobserved individual heterogeneity does not offer a complete description of the data, as is witnessed by its relative large goodness of fit chi-square.

Inspecting in detail Table 9, the reader can conclude that a standard hurdle with no heterogeneity under-predicts 2 and 3 counts, and over-predicts 4,5,6 counts. Clearly, a Double-hurdle model with no heterogeneity fits better 2,3,5, and 6 counts but does marginally worse predicting 1 and 4 outcomes. Accounting for unobserved heterogeneity improves the fit of both models. In particular, standard Hurdle reduces its degree of under-prediction of 2 and 3 counts. Counts 4,5 and 6 are still over-predicted but not to the same degree as in the case where unobserved individual heterogeneity is completely neglected. Similarly, controlling for unobserved heterogeneity causes the Double-Hurdle model to improve its prediction power of 4, 5, and 6 counts and to do better in predicting 2 outcomes. It seems that the relative ability to predict well 4,5, and 6 counts is what causes the Double-Hurdle model to perform better than a standard Hurdle model.

5.2.4 Effect of explanatory variables

Estimates from various specifications of a Double Hurdle Poisson model are reported in Table 6. The present section discusses results for a model in which $\theta_0 = \theta_1 = 0$. This is the best fitting specification (see column 2 of Table 6). Additionally, Table 11 contains predicted probabilities for various representative individuals. Since most Mexicans are Catholic and non-indigenous language speakers, let a Catholic and non-indigenous language speaker who was born in Mexico City between 1940 and 1944 be the benchmark case (see row 2). Set as well **Edu12** to its mean value of four years of schooling. This individual, referred as individual II for the rest of the discussion, has a likelihood of remaining childless for her whole lifetime of approximately seven per cent. Moreover, if a non-negative count has been observed individual II is expected to have a family of one, two or three children 47 out of a hundred times. To put it in other words, conditional on observing a positive count, individual II will move to parities higher than three with probability $\{1-\Pr[1 < j \leq 3 \mid j > 0]\} = 0.5316$. Finally, once a fourth child is observed Individual II will have a family larger than six with $\Pr[j > 6 \mid j > 3] = 0.3947$.

5.2.4.1 Probability of a zero count

Let the discussion start by assessing the effect of explanatory variables on the likelihood that a woman will remain childless for her entire lifespan. The chances of observing such an event are determined by an Extreme Value distribution that is dependent on a vector of coefficients β_0. The most interesting observation that one may draw from the results in Table 6 is that except for constant and **Edu12** all the elements of β_0 are insignificant. In fact, a likelihood ratio test for the exclusion of **Catholic**, **Indspker**, **c4549** through **c5559**, and **North** through **South** is not rejected with a χ^2 (8) = 14.6 and a p-value = 0.067. Thus, it seems that education is the only variable that affects the probability of observing a zero count. As expected, the coefficient on **Edu12** is negative. Further, from Table 11 the reader may learn that, ceteris paribus, a woman who had no formal education at age 12 is 3.33% more likely of remaining childless for her entire life than a woman who had six years of education at age 12. Hence, though statistically significant, the effect of **Edu12** on $\Pr[j=0]$ seems to be rather small.

5.2.4.2 Transition from low to high parities given a positive count

Conditional on having at least one child, the probability of observing any particular count in the interval [1,3] is determined by a truncated-at-zero Poisson distribution that depends on the vector of parameters β_1. Notice then that, since Pr(j > 3 | j > 0) is a function of β_1, the probability of crossing the second hurdle - or say, getting out of the [1,3] interval - is also a function of β_1.

Using this interpretation for the elements of vector β_1 the reader can conclude from the estimates in Table 6 that Catholic individuals are less likely to cross the second hurdle than non-Catholic individuals. In order to assess the relevance of such an effect Table 11 contains predicted probabilities for a non-Catholic woman (individual I) who is otherwise identical to the benchmark woman II. There the reader can learn that individual I scores a $\Pr[1 < j \leq 3 \mid j > 0] = 0.4302$ while individual II scores a $\Pr[1 < j \leq 3 \mid j > 0] = 0.4684$. That is, Catholicism reduces the chances of transition from low to high order parities by as many as 3.8 percentage points.

Various factors may be behind the negative and significant coefficient on **Catholic** in the middle panel of Table 6. Among the most significant reasons there is a rather weak opposition of the Catholic Church towards the diffusion and adoption of contraceptives among the Catholic community in Mexico. A conjecture then would argue that this lack of opposition and the wide heterogeneity of the Catholic community – which represents the far majority of Mexicans – has allowed the establishment of a large and diverse base of active users of modern contraceptives among Catholic individuals. As a consequence, relative to individuals with other religious backgrounds, Catholics receive more and better information (and stronger social pressure) about family planning and the desirability of a relatively low fertility.

Coming back to Table 6, it seems that being an indigenous language speaker increases the chances of crossing the second hurdle, as the coefficient on **Indspker** is estimated to be positive – though the coefficient is different from zero only at 5%. The finding is intuitive because, as was discussed earlier in the text, indigenous individuals in Mexico have in general a lower economic status than non-indigenous individuals. Row 3 of Table 11 reports predicted probabilities for an indigenous language speaker individual who is otherwise identical to the benchmark individual II. Comparing figures in row 2 and 3 of Table 11 it is easy to conclude that the marginal effect of **Indspker** on $\Pr[1 < j \leq 3 \mid j > 0]$ is around -.0306. In other words, holding other things constant, an indigenous language speaker has a 3% higher chance of having a family larger than three than a non-indigenous language speaker.

A negative coefficient on **Edu12** in vector β_1 of Table 6 suggests that an extra year of education at age 12 increases the likelihood that a woman will remain with less than four children during her entire lifespan. The finding confirms general economic intuition. More importantly, the effect of **Edu12** on the probability of observing such an event is estimated to be rather large. For instance, according to Table 11 increasing **Edu12** from five to six years will lead to an increment in $\Pr[1 < j \leq 3 \mid j > 0]$ of 5.93 points, other things being constant. Further, a rise of schooling at age 12 from zero to six years implies that the odds of crossing the second hurdle would shrink by as much as 36.48 percentage points.

Vector β_1 in Table 6 contains sequentially more negative coefficients on **c4549** through **c5559**. Hence, the evidence is that young generations have lower chances of crossing the second hurdle. In fact, a woman born between 1945 and 1949 who is in other aspects similar to the benchmark woman II is estimated to bear 4% lower chances of ending her fertile life with more than three children in relation to the reference individual. Such a reduced risk becomes 10% and 13% for women in cohort 1950-54 and 1955-1959 respectively (see row 4 through 6 of Table 11).

As expected, being born in a region other than Mexico City implies increments in the odds of crossing the one-to-three hurdle. For instance, an individual who was born in the North of the Country will cross the second hurdle 18.6 out of a hundred more times than individual II, other things being equal. Similarly, marginal effects of **Centre** and **South** on $\{1\text{-Pr}[1 < j \leq 3 \mid j > 0]\}$ are respectively 0.1931 and 0.1197. Thus, being born in different geographical areas of the country leads to wide variations in the likelihood of a large family.

5.2.4.3 Probability of Counts Larger than Six given that the second Hurdle has been crossed.

Conditional on having more than three children, a truncated-at-three Poisson distribution governs the likelihood of observing any particular count equal or higher than four. This last distribution depends on a vector of coefficients β_2.

Notice first from table 6 that conditional on observing a count larger than three the coefficient on **Indspker** is insignificant at all conventional levels. In other words, ethnic group seems to have no influence on completed fertility once the second hurdle has been crossed. In other issues, the negative coefficient on **Catholic** is different from zero at 5% but not 1% significance level. Such a negative coefficient on **Catholic** implies that, conditional on crossing the second hurdle, the Catholic reference individual II of Table 11 will end her fertile life with more than six children with probability 0.3947 while her non-Catholic equivalent individual I will register the same event with probability 0.4165. That is, Catholicism is associated with a reduction of 0.02181 units in Pr[j > 6 | j > 3]. Since the previous discussion has already offered some intuition for explaining this result no further comment on the issue will be made here.

Cohort of age affects significantly Pr[j > 6 | j > 3] as well. Namely, a woman born in the 1945-1949 cohort – i.e., individual IV of Table 11 – that has crossed the second hurdle is estimated to end fertile life with a family size larger than six with probability 0.3416. In comparison, woman II scores a Pr[j > 6 | j > 3] of 0.3947. Hence, ceteris paribus, a woman in the cohort 1945-1949 bears a reduced risk of 5.31 per cent of registering a large count in relation to a woman in the control group. Younger generations have even lower odds of a large completed fertility. In fact, marginal effects of **c5054** and **c5559** on Pr[j > 6 | j > 3] are -0.1105 and -0.1547 respectively.

Marginal effects for **North**, **Centre** and **South** on Pr[j > 6 | j > 3] might be obtained on the basis of row 2, and 7 through 9 of Table 11. Marginal effects are positive and large: 0.1873, 0.2414 and 0.1822 respectively.

5.2.5 Regional Results

Table 10 presents regression results for a Double-Hurdle model fitted to various sub-samples of the data constructed according to women's birthplace. Four Regions are considered: Mexico City, North, Centre, and South. In each region various specifications were estimated and Table 10 reports exclusively the resulting best fitting model. Model selection was performed on the basis of the strategy followed at the National level. With the exception of the Centre, unobserved individual heterogeneity was detected exclusively in the post second hurdle count process (that is, evidence suggested $\theta_0 = \theta_1 = 0$). In the case of the Centre, θ_1 is reported to be significantly different from zero. Except for the North, likelihood ratio tests for the joint equality of the coefficients β_1 and β_2 easily reject the null (see Table 8). In the case of the North a standard Hurdle model is supported by the data. In all cases $\beta_0 = \beta_1$ and $\beta_0 = \beta_1 = \beta_2$ are rejected at least at 5% of significance. Interpretation of the coefficients remains the same and marginal effects might be calculated on the basis of Table 11.

Some differences in the coefficients on explanatory variables across the various regions are detected. In the first place, the evidence suggests that the likelihood of observing a zero count is independent of all the explanatory variables in Mexico City and the South. And education at age 12 affects significantly Pr[j=0] only in the North and Centre of the country.

Regarding the probability of crossing the one-to-three hurdle, Pr[j >3 | j>0], empirical evidence indicates that religious background is irrelevant in Mexico City and the

Centre, while relevant in the North and South of the Country. Similarly, with the exception of the South, ethnic group seems not to affect the odds of crossing the second hurdle. Finally, education at age 12 is found to reduce the likelihood of having a large family in all cases. There are, however, some differences in the size of its effect. In particular, **Edu12** seems to have a far larger effect in Mexico City than in any other geographical region of the country.

Conditional on observing a count larger than three, Catholic individuals are expected to have a significantly lower fertility than non-Catholics only in the South. A similar observation is valid for ethnic group. That is, being an indigenous language speaker is associated significantly with increases in $Pr[\ j \mid j > 3]$ exclusively in the South of the country. Education at age 12 reduces significantly $Pr[\ j \mid j > 3]$ in all the geographic regions of the country.

In conclusion, the evidence suggests that the effect of explanatory variables on completed fertility varies across the different regional areas of the country. In some areas religion and ethnic background have significant impact on fertility behavior while in other regions such characteristics are largely irrelevant. Education at age 12 is a relevant factor across the whole country.

6. Conclusions

The present paper reports a study on the socio-economic determinants of completed fertility in Mexico. Special attention is given to how socio-economic factors such as religion and ethnic group affect the likelihood of transition from low to high parities. An innovative Poisson Double-Hurdle count model is developed for the analysis. This methodological approach allows low and high order parities to be determined by two different data generating mechanisms, and explicitly accounts for potential endogenous switching between both regimes. Unobserved heterogeneity is properly controlled.

Catholicism is found to be associated with reductions in the likelihood of transition from low to high parities. This result may be associated with the relatively weak opposition of the Catholic Church to the diffusion of contraceptives in Mexico, and its much stronger opposition to the initiation of sexual life before marriage. Other factors may be at work. For instance, the existence of a large base of contraception users within the Catholic community may imply that a Catholic individual receives better information about the advantages of family planning relative to a non-Catholic individual.

Empirical evidence suggests that being an indigenous language speaker increases the likelihood of transition from low to high parities, especially in the South and Centre of the country. Further, as suggested by economic intuition, education at age 12 is found to reduce women's odds of having a large family.

Conditional on observing a count larger than three, Catholic individuals are expected to have a significantly lower fertility than non-Catholics only in the south of the country. A similar observation is valid for ethnic group. That is, being an indigenous language speaker is associated significantly with increases in completed fertility exclusively in the South.

Appendix

Table 1. Descriptive Statistics

Variable	Description	Mean	Std. Dev.	Min	Max
Age	age in years	45.93	4.21	40	54
Children	number of children ever born alive	4.43	2.75	0	18
Edu12	Completed years of schooling at age 12	4.01	2.33	0	6
Religion and Ethnic group					
Catholic	=1 if Catholic; 0 otherwise	0.90	-	-	-
indspker	=1 if indian language speaker; 0 otherwise	0.09	-	-	-
Cohort			-	-	-
c4044 (base group)	=1 if born within 1940-1944; 0 otherwise	0.10	-	-	-
c4549	=1 if born within 1951-1955; 0 otherwise	0.29	-	-	-
c5054	=1 if born within 1956-1960; 0 otherwise	0.36	-	-	-
c5559	=1 if born within 1961-1965; 0 otherwise	0.25	-	-	-
Birth Place					
MexCity (base group)	=1 if born in Mex City; 0 otherwise	0.05	-	-	-
North	=1 if born in North; 0 otherwise	0.23	-	-	-
Centre	=1 if born in Cebtre; 0 otherwise	0.54	-	-	-
South	=1 if born in South; 0 otherwise	0.18	-	-	-
Number of observations					19,477

Table 2. Descriptive Statistics -- Region (split according to birthplace dummies)

	Mean	Std. Dev.	Min	Max		Mean	Std. Dev.	Min	Max
Mexico City					**Centre**				
Age	45.40	4.10	40	54	Age	46.01	4.24	40	54
Children	2.91	1.78	0	12	Children	4.68	2.87	0	18
Edu12	5.69	1.12	0	6	Edu12	3.75	2.39	0	6
Catholic	0.90	-	-	-	Catholic	0.93	-	-	-
indspker	0.01	-	-	-	indspker	0.07	-	-	-
c4044	0.07	-	-	-	c4044	0.10	-	-	-
c4549	0.26	-	-	-	c4549	0.30	-	-	-
c5054	0.39	-	-	-	c5054	0.36	-	-	-
c5559	0.29	-	-	-	c5559	0.25	-	-	-
N. obs				967	N. obs				10537
North					**South**				
Age	45.90	4.20	40	54	Age	45.91	4.16	40	54
Children	4.11	2.48	0	16	Children	4.51	2.78	0	16
Edu12	4.80	1.85	0	6	Edu12	3.29	2.45	0	6
Catholic	0.89				Catholic	0.81	-	-	-
indspker	0.02	-	-	-	indspker	0.29	-	-	-
c4044	0.09	-	-	-	c4044	0.09	-	-	-
c4549	0.29	-	-	-	c4549	0.28	-	-	-
c5054	0.37	-	-	-	c5054	0.38	-	-	-
c5559	0.25	-	-	-	c5559	0.25	-	-	-
N. obs				4532	N. obs				3441

Table 3. Empirical distribution of Children and a Poisson distribution with mean of 4.4

Count	Obs.	Share	Poisson
0	1,211	0.0622	0.012
1	1,134	0.0582	0.054
2	2,504	0.1286	0.119
3	3,383	0.1737	0.174
4	2,905	0.1492	0.192
5	2,349	0.1206	0.169
6	1,818	0.0933	0.124
7	1,390	0.0714	0.078
8	1,036	0.0532	0.043
9	746	0.0383	0.021
10	474	0.0243	0.009
11	241	0.0124	0.004
12-18	286	0.0147	0.002
Total	19,477	1.000	1.000

Table 4. Likelihood of high parities given y > 3

Count	4	5	6	7-18	Total
No. obs.	2,905	2,349	1,818	4,173	11,245
Pr(count | y>3)	0.26	0.21	0.16	0.37	1.00

Figure 1. Empirical distribution of Children and a theoretical Poisson with mean 4.4

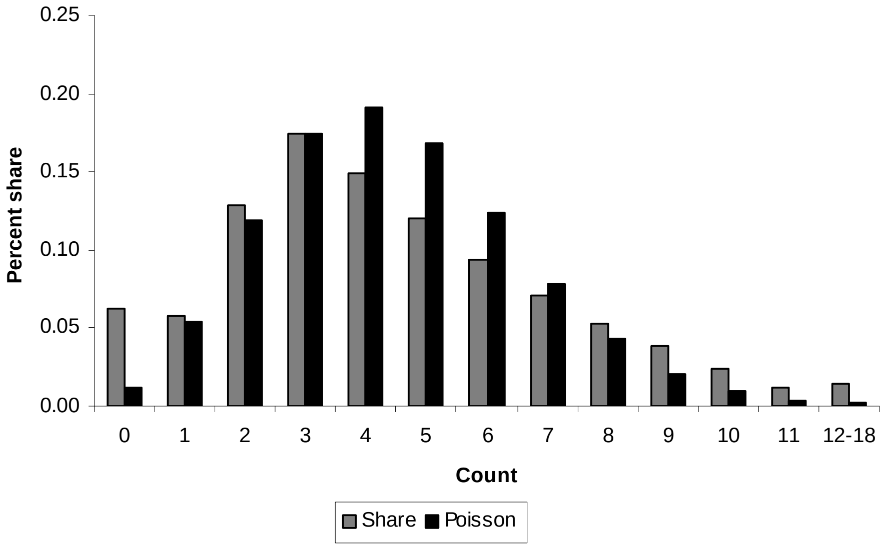

Figure 2. Double-Hurdle Model Structure

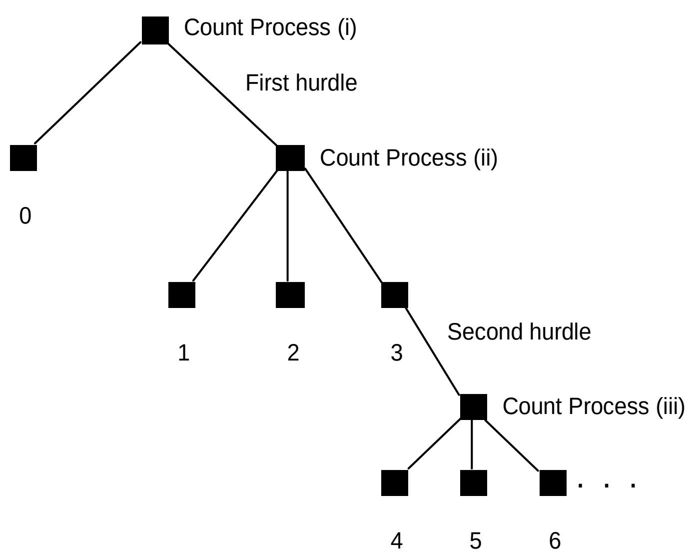

Table 5. Standard Hurdle Model -- National Data
Coefficient [Std. Err.]

Count Process	(1) No Het.	(2) Normal Het.
At Zero		
Constant	1.1547 [0.0675]**	1.1547 [0.0675]**
Education, Religion and Ethnic group		
Catholic	-0.0525 [0.0342]	-0.0525 [0.0342]
Indspker	-0.0728 [0.0381]	-0.0728 [0.0381]
Edu12	-0.0314 [0.0047]**	-0.0314 [0.0047]**
Cohort (base 1940-1944)		
c4549	0.0230 [0.0382]	0.0230 [0.0382]
c5054	0.0494 [0.0374]	0.0494 [0.0374]
c5559	0.0225 [0.0390]	0.0225 [0.0390]
Birthplace (base Mexico City)		
North	0.0558 [0.0487]	0.0558 [0.0487]
Centre	0.0001 [0.0465]	0.0001 [0.0465]
South	0.0460 [0.0519]	0.0460 [0.0519]
Larger than zero		
Constant	1.7903 [0.0260]**	1.7740 [0.0280]**
Education, Religion and Ethnic group		
Catholic	-0.0475 [0.0112]**	-0.0482 [0.0124]**
Indspker	0.0289 [0.0120]*	0.0321 [0.0133]*
Edu12	-0.0878 [0.0015]**	-0.0891 [0.0017]**
Cohort (base 1940-1944)		
c4549	-0.0836 [0.0120]**	-0.0848 [0.0134]**
c5054	-0.1868 [0.0120]**	-0.1895 [0.0133]**
c5559	-0.2563 [0.0129]**	-0.2588 [0.0143]**
Birthplace (base Mexico City)		
North	0.2669 [0.0220]**	0.2676 [0.0233]**
Centre	0.3053 [0.0214]**	0.3060 [0.0227]**
South	0.2057 [0.0228]**	0.2036 [0.0243]**
σ^2	-	0.0411 [0.0027]**
Log-likelihood	-44144.42	-43996.48
AIC	88,328.84	88,034.95
CIAC	88,506.38	88,221.37
Number of observations	19,477	19,477

Note: ** significant at 1% ; * significant at 5%.

Table 6. Poisson Double Hurdle Model -- National Data

Coefficient [Std. Err.]

	No Het.	Normal Het.		
Count Process	(1)	(2) $\theta_0=\theta_1=0$	(3) $\theta_0=\theta_1=1$	(4) θ_0 , θ_1 free
At Zero [vector β_0] -- Process (i)				
Constant	1.1547 [0.0675]**	1.1547 [0.0675]**	1.1800 [0.0698]**	1.1567 [0.0755]
Education, Religion and Ethnic group				
Catholic	-0.0525 [0.0342]	-0.0525 [0.0342]	-0.0543 [0.0353]	-0.0527 [0.0344]
Indspker	-0.0728 [0.0381]	-0.0728 [0.0381]	-0.0753 [0.0393]	-0.0731 [0.0383]
Edu12	-0.0314 [0.0047]**	-0.0314 [0.0047]**	-0.0324 [0.0049]**	-0.0314 [0.0049]**
Cohort (base 1940-1944)				
c4549	0.0230 [0.0382]	0.0230 [0.0382]	0.0237 [0.0394]	0.0230 [0.0383]
c5054	0.0494 [0.0374]	0.0494 [0.0374]	0.0513 [0.0386]	0.0496 [0.0376]
c5559	0.0225 [0.0390]	0.0225 [0.0390]	0.0235 [0.0402]	0.0226 [0.0391]
Birthplace (base Mexico City)				
North	0.0558 [0.0487]	0.0558 [0.0487]	0.0575 [0.0502]	0.0559 [0.0489]
Centre	0.0001 [0.0465]	0.0001 [0.0465]	0.0001 [0.0480]	0.0001 [0.0467]
South	0.0460 [0.0519]	0.0460 [0.0519]	0.0475 [0.0535]	0.0462 [0.0521]
At one-to-three [vector β_1] -- Process (ii)				
(vector β_1)				
Constant	1.7142 [0.0328]**	1.7142 [0.0328]**	1.7370 [0.0344]**	1.7142 [0.0328]**
Education, Religion and Ethnic group				
Catholic	-0.0509 [0.0157]**	-0.0509 [0.0157]**	-0.0535 [0.0165]**	-0.0509 [0.0157]**
Indspker	0.0408 [0.0181]*	0.0408 [0.0181]*	0.0430 [0.0191]*	0.0408 [0.0181]*
Edu12	-0.0842 [0.0022]**	-0.0842 [0.0022]**	-0.0888 [0.0024]**	-0.0842 [0.0022]**
Cohort (base 1940-1944)				
c4549	-0.0535 [0.0184]**	-0.0535 [0.0184]**	-0.0564 [0.0194]**	-0.0535 [0.0184]**
c5054	-0.1326 [0.0179]**	-0.1326 [0.0179]**	-0.1391 [0.0190]**	-0.1326 [0.0179]**
c5559	-0.1770 [0.0187]**	-0.1770 [0.0187]**	-0.1853 [0.0198]**	-0.1770 [0.0187]**
Birthplace (base Mexico City)				
North	0.2523 [0.0248]**	0.2523 [0.0248]**	0.2605 [0.0256]**	0.2523 [0.0248]**
Centre	0.2616 [0.0239]**	0.2616 [0.0239]**	0.2702 [0.0248]**	0.2616 [0.0239]**
South	0.1597 [0.0262]**	0.1597 [0.0262]**	0.1638 [0.0271]**	0.1597 [0.0262]**
Larger than three [vector β_2] -- Process (iii)				
Constant	1.7752 [0.0522]**	1.7564 [0.0542]**	1.7429 [0.0537]**	1.7554 [0.0550]**
Education, Religion and Ethnic group				
Catholic	-0.0348 [0.0156]*	-0.0359 [0.0168]*	-0.0379 [0.0164]*	-0.0361 [0.0168]*
Indspker	0.0129 [0.0156]	0.0163 [0.0169]	0.0161 [0.0165]	0.0160 [0.0169]
Edu12	-0.0753 [0.0023]**	-0.0768 [0.0024]**	-0.0798 [0.0025]**	-0.0769 [0.0027]**
Cohort (base 1940-1944)				
c4549	-0.0911 [0.0153]**	-0.0934 [0.0166]**	-0.0944 [0.0162]**	-0.0933 [0.0167]**
c5054	-0.2025 [0.0156]**	-0.2075 [0.0170]**	-0.2103 [0.0166]**	-0.2073 [0.0171]**
c5559	-0.3030 [0.0180]**	-0.3086 [0.0193]**	-0.3130 [0.0190]**	-0.3084 [0.0195]**
Birthplace (base Mexico City)				
North	0.2831 [0.0494]**	0.2810 [0.0509]**	0.2913 [0.0504]**	0.2811 [0.0510]**
Centre	0.3570 [0.0486]**	0.3559 [0.0500]**	0.3657 [0.0496]**	0.3557 [0.0501]**
South	0.2787 [0.0499]**	0.2740 [0.0515]**	0.2816 [0.0510]**	0.2740 [0.0516]**
σ^2	-	0.0340 [0.0042]**	0.0239 [0.0038]**	0.2745 [1.6784]**
θ_0	-	set to zero	set to one	-0.0110 [0.2300]
θ_1	-	set to zero	set to one	0.0341 [0.0042]
Log-likelihood	-43,980.42	-43,941.15	-43,956.11	-43,941.16
AIC	88,020.84	87,944.30	87,974.22	87,946.32
CIAC	88,287.15	88,219.49	88,249.41	88,230.38
Number of observations	19,477	19,477	19,477	19,477

Note: ** significant at 1% ; * significant at 5%.

Table 7. Model Selection
Poisson Double-Hurdle with Normal Heterogeneity -- National Data

Case	H_0	H_1	Test type	$\chi 2$ [p-val]	Inference
1	$\theta_0 = 0, \sigma^2 \neq 0$	$\theta_0 \neq 0, \sigma^2 \neq 0$	LRT	0.016 [0.8993]	Do not reject H_0
2	$\theta_1 = 0, \sigma^2 \neq 0$	$\theta_1 \neq 0, \sigma^2 \neq 0$	LRT	0.018 [0.8933]	Do not reject H_0
3	$\sigma^2 = 0$	$\sigma^2 \neq 0$	BVLRT	78.53 [0.0000]	Reject H_0
4	$\theta_0 = \theta_1 = 0, \sigma^2 \neq 0$	$\theta_0 \neq 0, \theta_1 = 0, \sigma^2 \neq 0$	LRT	0.032 [0.858]	Do not reject H_0
5	$\theta_0 = \theta_1 = 0, \sigma^2 \neq 0$	$\theta_0 = 0, \theta_1 \neq 0, \sigma^2 \neq 0$	LRT	0.002 [0.9643]	Do not reject H_0
6	$\theta_0 = \theta_1 = 1, \sigma^2 \neq 0$	$\theta_0 \neq \theta_1 \neq 1, \sigma^2 \neq 0$	LRT	29.90 [0.0000]	Reject H0

Note: Boundary-value likelihood ratio test is abbreviated as BVLRT. Likelihood ratio test is abbreviated as LRT.

Table 8. Likelihood Ratio Tests

	LR	**P-val**	**Inference**
H_0: $\beta_0=\beta_1$ vs. H_1: $\beta_0\neq\beta_1$			
National	1610.30	0.0000	Reject H_0
Mex City	22.64	0.0122	Reject H_0
North	269.44	0.0000	Reject H_0
Centre	1295.47	0.0000	Reject H_0
South	251.18	0.0000	Reject H_0
H_0: $\beta_1=\beta_2$ vs. H_1: $\beta_1\neq\beta_2$			
National	164.27	0.0000	Reject H_0
Mex City	20.36	0.0260	Reject H_0
North	12.58	0.2483	Do no reject H_0
Centre	255.82	0.0000	Reject H_0
South	35.92	0.0001	Reject H_0
H_0: $\beta_0=\beta_1=\beta_2$ vs. H_1: $\beta_0\neq\beta_1\neq\beta_2$			
National	2339.49	0.0000	Reject H_0
Mex City	35.41	0.0180	Reject H_0
North	308.92	0.0000	Reject H_0
Centre	1584.50	0.0000	Reject H_0
South	345.58	0.0000	Reject H_0

Note: Tests based on best fitting Double-Hurdle Models.

Table 9 Observed and predicted sample distribution -- National data

Count	Obs.	Standard Hurdle No Het.	Standard Hurdle Normal Het.	Double-Hurdle (best fit) No Het.	Double-Hurdle (best fit) Normal Het. (θ_0=θ_1=0)
0	0.062	0.062	0.062	0.062	0.062
1	0.058	0.058	0.070	0.066	0.066
2	0.129	0.113	0.122	0.125	0.125
3	0.174	0.152	0.152	0.163	0.163
4	0.149	0.160	0.153	0.136	0.145
5	0.121	0.142	0.132	0.128	0.128
6	0.093	0.111	0.103	0.106	0.102
7	0.071	0.079	0.074	0.079	0.074
8	0.053	0.052	0.050	0.054	0.051
9	0.038	0.032	0.033	0.035	0.033
10	0.024	0.018	0.020	0.021	0.021
11	0.012	0.010	0.012	0.012	0.013
12-18	0.015	0.010	0.016	0.013	0.016
chi-square		371	213	150	116
Pr > chi-square		0.0000	0.0000	0.0000	0.0000
logL		-44144	-43996	-43980	-43941
AIC		88,329	88,035	88,021	87,944
CIAC		88,506	88,221	88,287	88,219

Note: Sample size is 19,477.

Table 10. Poisson Double Hurdle Model -- Regional Results (Best fitting model)
Coefficient [Std. Err.]

	(1) **Mex City**	(2) **North**	(3) **Centre**	(4) **South**
Count Process				
At Zero [vector β_0] -- Process (i)				
Constant	1.4155 [0.3385]**	1.2455 [0.1087]**	1.1121 [0.0692]**	1.2510 [0.1080]**
Education, Religion and Ethnic group				
Catholic	-0.1366 [0.1575]	0.0046 [0.0682]	-0.0489 [0.0532]	-0.0968 [0.0646]
Indspker	-0.2638 [0.4871]	-0.2505 [0.1596]	-0.0482 [0.0562]	-0.0555 [0.0571]
Edu12	0.0024 [0.0398]	-0.0434 [0.0121]**	-0.0331 [0.0060]**	-0.0193 [0.0105]
Cohort (base 1940-1944)				
c4549	-0.2458 [0.2161]	-0.0443 [0.0819]	0.0733 [0.0499]	-0.0201 [0.0969]
c5054	-0.3898 [0.2084]	0.0480 [0.0809]	0.0839 [0.0489]	0.0250 [0.0944]
c5559	-0.3265 [0.2129]	0.0127 [0.0840]	0.0914 [0.0513]	-0.1216 [0.0973]
At one-to-three [vector β_1] -- Process (ii)				
Constant	2.2446 [0.1622]**	2.0979 [0.0495]**	1.9182 [0.0370]**	1.8954 [0.0483]**
Education, Religion and Ethnic group				
Catholic	0.0681 [0.0790]	-0.0958 [0.0313]**	-0.0019 [0.0257]	-0.1010 [0.0287]**
Indspker	0.0454 [0.3040]	-0.0379 [0.0806]	0.0012 [0.0281]	0.1015 [0.0260]**
Edu12	-0.1770 [0.0220]**	-0.0940 [0.0054]**	-0.0859 [0.0036]**	-0.0770 [0.0049]**
Cohort (base 1940-1944)				
c4549	-0.1037 [0.0896]	-0.0748 [0.0385]	-0.0354 [0,0255]	-0.0801 [0.0447]
c5054	-0.2653 [0.0870]**	-0.1887 [0.0376]**	-0.0889 [0.0251]**	-0.1748 [0.0432]**
c5559	-0.3413 [0.0904]**	-0.2384 [0.0392]**	-0.1469 [0.0262]**	-0.1676 [0.0454]**
Larger than three [vector β_2] -- Process (iii)				
Constant	1.8688 [0.2595]**	2.1311 [0.0516]**	2.0352 [0.0304]**	2.0693 [0.0413]**
Education, Religion and Ethnic group				
Catholic	-0.0919 [0.1845]	-0.0493 [0.0396]	0.0120 [0.0250]	-0.1071 [0.0282]**
Indspker	-0.3206 [0.6569]	-0.0621 [0.0971]	-0.0361 [0.0234]	0.0938 [0.0257]**
Edu12	-0.0733 [0.0279]**	-0.0867 [0.0060]**	-0.0761 [0.0032]**	-0.0843 [0.0061]**
Cohort (base 1940-1944)				
c4549	-0.2200 [0.1592]	-0.1274 [0.0380]**	-0.0671 [0.0212]**	-0.1314 [0.0390]**
c5054	-0.4044 [0.1644]**	-0.3055 [0.0455]**	-0.1804 [0.0216]**	-0.1810 [0.0390]**
c5559	-0.8182 [0.2170]**	-0.4044 [0.0455]**	-0.2785 [0.0245]**	-0.2823 [0.0446]**
σ^2	0.1520 [0.0507]**	0.0507 [0.0111]**	0.0277 [0.2688]**	0.0304 [0.0096]**
θ_0	set to zero	set to zero	set to zero	set to zero
θ_1	set to zero	set to zero	0.7686 [0.2688]**	set to zero
Log-likelihood	-1,793.98	-9,839.05	-24,332.7	-7,799.3
AIC	3,649.96	19,740.10	48,729.40	15,660.60
CIAC	3,832.06	19,970.09	48,993.80	15,882.05
Number of observations	967	4,532	10,537	3,441

Note: ** significant at 1% ; * significant at 5%.

Table 11. Predicted Probabilities -- Double Hurdle Poisson Model

	Caracteristics	Pr(j = 0)	Pr(1 < j ≤ 3 | j > 0)	Pr(j > 6 | j > 3)
National				
(1)	edu12=mean, all dummies set to zero	0.0609	0.4302	0.4165
(2)	edu12=mean,catholic=1, other dummies set to zero	0.0703	0.4684**	0.3947*
(3)	edu12=mean,catholic=1,indspker=1, other dummies set to zero	0.0847	0.4378*	0.4044
(4)	edu12=mean,catholic=1, c4549=1, other dummies set to zero	0.0661	0.5081**	0.3416**
(5)	edu12=mean,catholic=1, c5054=1, other dummies set to zero	0.0615	0.5648**	0.2842**
(6)	edu12=mean,catholic=1, c5559=1, other dummies set to zero	0.0662	0.5955**	0.24**
(7)	edu12=mean,catholic=1, north=1, other dummies set to zero	0.0604	0.2818**	0.582**
(8)	edu12=mean,catholic=1,centre=1, other dummies set to zero	0.0703	0.2753**	0.6361**
(9)	edu12=mean,catholic=1, south=1, other dummies set to zero	0.0620	0.3487**	0.5769**
(10)	edu12=0,catholic=1, other dummies set to zero	0.0493**	0.2243**	0.6015**
(11)	edu12=5,catholic=1, other dummies set to zero	0.0762**	0.5298**	0.3511**
(12)	edu12=6,catholic=1, other dummies set to zero	0.0826**	0.5891**	0.3107**
Mex City				
(1)	edu12=mean, all dummies set to zero	0.0169	0.3126	0.4987
(2)	edu12=mean,catholic=1, other dummies set to zero	0.0285	0.2646	0.4381
(3)	edu12=mean,catholic=1,indspker=1, other dummies set to zero	0.0650	0.2342	0.2648
(4)	edu12=mean,catholic=1, c4549=1, other dummies set to zero	0.0619	0.3386	0.3124**
(5)	edu12=mean,catholic=1, c5054=1, other dummies set to zero	0.0899	0.4596**	0.2299**
(6)	edu12=mean,catholic=1, c5559=1, other dummies set to zero	0.0768	0.5159**	0.11**
(10)	edu12=0,catholic=1, other dummies set to zero	0.0275	0.0096**	0.6423**
(11)	edu12=5,catholic=1, other dummies set to zero	0.0287	0.3919**	0.3934**
(12)	edu12=6,catholic=1, other dummies set to zero	0.0290	0.5239**	0.3514**
North				
(1)	edu12=mean, all dummies set to zero	0.0539	0.1886	0.6471
(2)	edu12=mean,catholic=1, other dummies set to zero	0.0532	0.2493**	0.6113
(3)	edu12=mean,catholic=1,indspker=1, other dummies set to zero	0.1020	0.2752	0.5666
(4)	edu12=mean,catholic=1, c4549=1, other dummies set to zero	0.0604	0.3012	0.5207**
(5)	edu12=mean,catholic=1, c5054=1, other dummies set to zero	0.0461	0.3851**	0.4044**
(6)	edu12=mean,catholic=1, c5559=1, other dummies set to zero	0.0513	0.4224**	0.3474**
(10)	edu12=0,catholic=1, other dummies set to zero	0.0305**	0.0624**	0.8459**
(11)	edu12=5,catholic=1, other dummies set to zero	0.0602**	0.3143**	0.5498**
(12)	edu12=6,catholic=1, other dummies set to zero	0.0678**	0.3838**	0.4897**
Centre				
(1)	edu12=mean, all dummies set to zero	0.0698	0.2849	0.6084
(2)	edu12=mean,catholic=1, other dummies set to zero	0.0792	0.2862	0.6171
(3)	edu12=mean,catholic=1,indspker=1, other dummies set to zero	0.0892	0.2853	0.5911
(4)	edu12=mean,catholic=1, c4549=1, other dummies set to zero	0.0653	0.3115	0.5687**
(5)	edu12=mean,catholic=1, c5054=1, other dummies set to zero	0.0634	0.3506**	0.4898**
(6)	edu12=mean,catholic=1, c5559=1, other dummies set to zero	0.0621	0.3939**	0.4259**
(10)	edu12=0,catholic=1, other dummies set to zero	0.0553**	0.092**	0.8259**
(11)	edu12=5,catholic=1, other dummies set to zero	0.0859**	0.3478**	0.5629**
(12)	edu12=6,catholic=1, other dummies set to zero	0.0931**	0.412**	0.5096**
South				
(1)	edu12=mean, all dummies set to zero	0.0394	0.2759	0.6092
(2)	edu12=mean,catholic=1, other dummies set to zero	0.0531	0.3486**	0.5328**
(3)	edu12=mean,catholic=1,indspker=1, other dummies set to zero	0.0622	0.2755**	0.5996**
(4)	edu12=mean,catholic=1, c4549=1, other dummies set to zero	0.0563	0.4086	0.4444**
(5)	edu12=mean,catholic=1, c5054=1, other dummies set to zero	0.0493	0.4796**	0.4132**
(6)	edu12=mean,catholic=1, c5559=1, other dummies set to zero	0.0743	0.4742**	0.3541**
(10)	edu12=0,catholic=1, other dummies set to zero	0.0419	0.1477**	0.7724**
(11)	edu12=5,catholic=1, other dummies set to zero	0.0561	0.4056**	0.4757**
(12)	edu12=6,catholic=1, other dummies set to zero	0.0593	0.4636**	0.4214**

Note: ** (*) indicates that the relevant coefficient in Table 5 and 7 is significant at 1% (5%) of significance.

Endnotes

[1] Notice that under these assumptions consistent estimators are obtained. The argument might be outlined as follows. Suppose that **indspker** contains a measurement error so that **indspker** = **indspker*** + u, with **indspker*** representing the variable without error and u representing its measurement error. Suppose further that E(u)=0 and that u is uncorrelated with any observed and unobserved explanatory variable considered in the fertility equation **-**including the observed proxy for Indian language speaker itself, **indspker**. Represent unexplained heterogeneity by the random term v, which is assumed to have a zero mean and to be uncorrelated with all explanatory variables. In particular, suppose that v is uncorrelated with both **indspker*** and **indspker**. Clearly, using **indspker** in place of **indspker*** in a simple OLS fertility equation will shift unexplained heterogeneity term (i.e., the error term) from v to w = (v-βu), where β represents the OLS coefficient on **indspker**. Under these set of assumptions, however, Cov(indspker,u)=0 and Cov(v,u)=0. Hence, the shifted heterogeneity term, w, has mean zero and is uncorrelated with all explanatory variables, including the proxy **indspker**. **Indspker** possesses then all the properties of an instrumental variable and therefore it might be used in place of **indspker*** to produce consistent estimators (for more details on this issue see Wooldridge 2002, ch. 4).

[2] **North** is integrated by Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nuevo León, Sinaloa, Sonora and Tamaulipas. **Centre** is integrated by

Aguascalientes, Colima, Guanajuato, Guerrero, Hidalgo, Jalisco, Estado de México, Michoacán, Morelos, Nayarit, Puebla, Querétaro, San Luis Potosí, Tlaxcala, Veracruz, and Zacatecas. Finally, Campeche, Chiapas, Oaxaca, Quintana Roo, Tabasco, and Yucatan integrate the **South**.

References

Arulampalam W, Booth A. 2001. Learning and Earning: Do Multiple Training Events Pay? A Decade of Evidence from a Cohort of Young British Men. *Economics* **68** 379-400.

Bergstrom T. 1989. A Fresh Look at the Rotten Kid Theorem--and Other Household Mysteries. *The Journal of Political Economy* **97** (5): 1138-1159.

Blaylock J, Blisard W. 1993. Women And The Demand For Alcohol: Estimating Participation And Consumption. *The Journal of Consumer Affairs* **27** (2): 319-334.

Cabrera G. 1994. Demographic Dynamics and Development: The role of Population Policy in Mexico. *Population and Development Review* **20**:105-120.

Cameron A, Trivedi P. 1986. Econometric Models Based On Count Data: Comparison And Applications Of Some Estimators and Test. *Journal of Applied Econometrics* **1** (1): 29-53.

CONAPO. 2001a. El Comportamiento Reproductivo De La Población Indígena.

CONAPO. 2001b. Marginality Index 2000. Document downloadable at www.conapo.gob.mx.

CONAPO. 2002. National Population Development Plan 2001-2006.

Cragg J. 1971. Some Statistical Models for Limited dependent Variables with Application to the Demand for Durable Goods. *Econometrica* **39** (5): 829-844.

Eswaran M. 2002. The Employment Of Women, Fertility, And Child Mortality: Towards A Theoretical Analysis. *Journal of Population Economics* **15**:433-454.

Faddy M. 1997. Extended Poisson Process Modelling And Analysis Of Count Data. *Biometric Journal* **39**:431-440.

Faddy M, Bosch R. 2001. Likelihood-based Modelling And Analysis Of Data Underdispersed Relative To The Poisson Distribution. *Biometrics* **57**:620-624.

Gomez J. 1996. Fecundidad y Anticoncepcion . *Demos* **9**:8-10.

Greene W. 1997. FIML Estimation of Sample Selection Models for Count Data. Manuscript.

Heckman J, Walker J. 1990. Estimating Fecudability From Dara On Waiting Times To First Conception. *Journal of The American Statistical Association* **85** (410): 283-294.

Hindin M. 2000. Women's Autonomy, Women's Status and Fertility-Related Behavior in Zimbabwe. *Population Research and Policy Review* **19**:255-282.

INEGI. 1999. National Survey of Demographic Dynamics 1997.

INEGI. 2000. Estadisticas Historicas de Mexico.

INEGI. 2001a. Censo de Población 2000.

INEGI. 2001b. Sistema De Indicadores Para El Seguimiento De La Situacion De La Mujer En Mexico.

Jones A. 1989. A Double-Hurdle Model for Cigarette Consumption. *Journal of Applied Econometrics* **4** (1): 23-39.

Klawon E, Tiefenthaler J. 2001. Bargaining over family size: The determinants of fertility in Brazil. *Population Research and Policy Review* **20**:423-440.

Labeaga J. 1999. A Double-Hurdle Rational Addiction Model with Heterogeneity: Estimating the Demand for Tobacco. *Journal of Econometrics* **93**: 49-72.

Lindstrom D. 1998. The Role of Contraceptive Supply and Demand in Mexican Fertility Decline: Evidence From a Microdemographic Study. *Population Studies* **52**:255-274.

Melkersson M, Rooth D. 2000. Modeling Female Fertility Using Inflated Count Data Models. *Journal of Population Economics* **13** (2): 189--203.

Mullahy J. 1986. Specification And Testing Of Some Modified Count Data Models. *Journal of Econometrics* **33**:341-365.

Podlich H, Faddy M, Smyth G. 2004. Likelihood Computations For Extended Poisson Process Models. Manuscript.

Santos Silva J, Covas F. 2000. A Modified Hurdle Model for Completed Fertility. *Journal of Population Economics* **13** (2): 173--88.

Terza J. 1998. Estimating Count Data Models with Endogenous Switching: Sample selection and endogenous treatment effects. *Journal of Econometrics* **84**:129-154.

Terza J. 1985. A Tobit-Type Estimator For The Censored Poisson Regression Model. *Economics Letters* **18:** 361-365.

Willis R. 1973. A New Approach to the Economic Theory of Fertility Behavior. *Journal of Political Economy* **81** (2): s14-s63.

Winkelmann R. 1995. Recent Developments in Count Data Modelling: Theory and Application. *Journal of Economic Surveys* **9** (1): 1-24.

Winkelmann R. 1998. Count Data Models with Selectivity. *Econometric Reviews* **17**: 339-359.

Winkelmann R. 2000. Econometric Analysis of Count Data, 3rd rev. and enl. ed. Springer, Berlin Heidelberg, New York.

Wooldridge J. 2002. Econometric Analysis Of Cross Section And Panel Data, The MIT Press, Cambridge, Massachusetts.

World Bank. 2001. World Development Indicators.

Yen S, Jensen H. 1996. Determinants of Household Expenditures on Alcohol. *Journal of Consumer Affairs* **30** (1): 48-67.

Yen S, Tang C, Su S. 2001. Demand for Traditional Medicine in Taiwan: A Mixed Gaussian-Poisson Model Approach. *Health Economics* **10**:221-232.

Zavala de Cosio E . 1990. Politicas de Poblacion en Mexico. *Revista Mexicana de Sociologia* **52** (1): 15-32.