Department of Quantitative Social Science

A double-hurdle count model for completed fertility data from the developing world

Alfonso Miranda

DoQSS Working Paper No. 10-01
June 2010

Leading education and social research
Institute of Education University of London

DISCLAIMER

Any opinions expressed here are those of the author(s) and not those of the Institute of Education. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

DoQSS Workings Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Department of Quantitative Social Science. Institute of Education, University of London. 20 Bedford way, London WC1H 0AL, UK.

A double-hurdle count model for completed fertility data from the developing world

Alfonso Miranda[*][†]

Abstract. This paper reports a study on the socio-economic determinants of completed fertility in Mexico. 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 regimes. Unobserved heterogeneity is properly controlled. 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. 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.

JEL classification: J13, J15, C25.

Keywords: Completed fertility, count data models, double-hurdle model.

[*]Department of Quantitative Social Science, Institute of Education, University of London. 20 Bedford Way, London WC1H 0AL, UK. E-mail: (A.Miranda@ioe.ac.uk).

[†]**Acknowledgements**. I am grateful to Wiji Arulampalam, Mark Stewart, John Ermisch, Jeremy Smith, and João Santos Silva for useful comments. I am also grateful to the National Council for Science and Technology (CONACYT) for its financial support.

1 Introduction

Nowadays it is widely recognised that social norms induce special features in most known completed fertility data sets. In the case of the developed world, for instance, Melkersson and Rooth (2000) suggest that it is social norms what induces the relative excess of zero and two counts reported for Swedish fertility data. Similarly, Santos Silva and Covas (2000) argue that, among other reasons, social norms may be behind a relatively rare only child outcome in data from Portugal. Various count data strategies have been developed to account for the special features depicted by fertility data collected in developed countries, including hurdle and inflated count models.

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).

There are various potential explanations for the special features of completed fertility data from the developing world. Among other potential explanations, this 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, this type of problem will be more prevalent in the developing world where women still face important labour market discrimination after taking a maternity leave and, as a consequence, find it difficult to recover from a career break. Other explanation for the features of the data may be that most developing countries are in the middle of their demographic transition and, as a consequence, women live in an environment where various competing fertility norms coexist Kholer (2000). Innovative individuals (pioneers) adopt a low fertility norm and limit the number of their children at a relatively low parity using modern contraceptives. Non-innovators, in contrast, stick to a traditional high fertility norm and tend to transit from low to high order parities without taking any action for limiting the number of their descendants.

Clearly, the relative *exess of large counts* that exhibit fertility data from the developing world needs explicit modelling to allow marginal increments in family size at low and high parities to be created by different data generating processes. Part of the modelling strategy, from the point of view of the author, should consider that women move from a *low fertility* regime to a *high fertility* regime when their fertility crosses certain pre-established thresholds that are largely determined by social fertility norms — In the case of México, for instance, data suggest that the relevant thresholds are 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 is easily extended to account for the problem that unobserved individual heterogeneity is expected to influence a woman's propensity to cross the socially pre-established fertility thresholds. That is, switching among fertility regimes is explicitly allowed to be endogenous. The paper places special emphasis in learning how socio-economic characteristics such as religion and ethnic group affect the probability of transition from low to high order parities in Mexico.

2 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 (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 modelled here, the present work adopts the convention in applied and work and completed fertility will be defined as number of children ever born to a woman during her lifetime. Children ever born is then the dependent variable. According to the descriptive statistics (see Table 3) children has mean 4.43 and variance 7.56. The data is therefore over-dispersed.

Table 1 presents 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, 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 Table 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 2, 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 labour 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). 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.

$Edu12$. 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*. 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 3).

3 Econometric issues

3.1 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,

$$Pr\left(y_i = j\right) = \begin{cases} \exp\left(-\mu_{0i}\right) & \text{if } j = 0 \\ \left[1 - \exp\left(-\mu_{0i}\right)\right] P\left(y_i | y_i > 0\right) & \text{otherwise,} \end{cases} \tag{1}$$

where the parameter μ_{0i} maintains a deterministic log-linear relationship with a $K \times 1$ vector, $\mathbf{x_{0i}}$, of explanatory variables (including the constant term),

$$\mu_{0i} = \exp\left(\mathbf{x'_{0i}}\boldsymbol{\beta_0}\right),$$

$\boldsymbol{\beta_0}$ is its $K \times 1$ vector of associated coefficients, and $Pr\left(y_i|y_i > 0\right)$ 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 $Pr\left(y_i = 0\right)$, 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

$Pr\left(y_i|y_i>0\right)$,

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

with,

$$\mu_{1i}=\exp\left(\mathbf{x}'_{1i}\boldsymbol{\beta}_1\right).$$

A standard Hurdle specifies $Pr(y_i|y_i>0)$ as a zero-truncated Poisson distribution. In contrast, equation (2) considers the case where counts in the $[1,3]$ and $[4,\infty)$ 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 $Pr(y_i|y_i\geq 4)$ is introduced. Clearly $Pr(y_i|y_i\geq 4)$ will be truncated at three and, to guarantee a well behaved probabilistic model, it should be re-scaled so that $Pr(y_i|y_i>0)$ sums up to one. Since equation (2) 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

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

To close the model a functional form for $Pr(yi|yi\geq 4)$ must be specified.

For convenience a Poisson distribution is, once again, selected:

$$Pr\left(y_i|y_i \geq 4\right) = \left[1 - \sum_{k=1}^{3}\left[1-\exp\left(-\mu_{1i}\right)\right]^{-1} \cdot \frac{\exp\left(-\mu_{1i}\right)\mu_{1i}^{k}}{k!}\right]^{-1} \frac{\exp\left(-\mu_{2i}\right)\mu_{2i}^{j}}{j!} \quad \text{if } j = 4,5,6,\ldots \tag{3}$$

As usual,

$$\mu_{2i} = \exp\left(\mathbf{x}'_{\mathbf{2i}}\boldsymbol{\beta}_{\mathbf{2}}\right).$$

In principle $\mathbf{x_{0i}}$, $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ may contain some (or all) common elements as no exclusion restrictions are required to achieve identification. Similarly, the vector of parameters$\boldsymbol{\beta}_{\mathbf{0}}$, $\boldsymbol{\beta}_{\mathbf{1}}$ and $\boldsymbol{\beta}_{\mathbf{2}}$ are estimated without constraints. Notice that if $\boldsymbol{\beta}_{\mathbf{1}} = \boldsymbol{\beta}_{\mathbf{2}}$ the Double-Hurdle model (DHM) collapses to a standard Poisson Hurdle model. Moreover, if $\boldsymbol{\beta}_{\mathbf{1}} = \boldsymbol{\beta}_{\mathbf{1}} = \boldsymbol{\beta}_{\mathbf{2}}$ 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 $\boldsymbol{\beta}_{\mathbf{0}}$, $\boldsymbol{\beta}_{\mathbf{1}}$ and $\boldsymbol{\beta}_{\mathbf{2}}$. Parameters are estimated by maximum likelihood. The contribution of the $i-th$ individual to the overall likelihood is simply,

$$\begin{aligned} L_i = & \prod_{y_i=0}\exp\left(-\mu_{0i}\right)\prod_{y_i>0}\left[1-\exp\left(-\mu_{0i}\right)\right] \\ & \cdot \prod_{1\leq y_i\leq 3}\left[1-\exp\left(-\mu_{1i}\right)\right]^{-1}\frac{\exp\left(-\mu_{1i}\right)\mu_{1i}^{y_i}}{y_i!} \\ & \cdot \prod_{y_i\geq 4}\left[1-\sum_{k=1}^{3}\left[1-\exp\left(-\mu_{1i}\right)\right]^{-1}\frac{\exp\left(-\mu_{1i}\right)\mu_{1i}^{k}}{k!}\right] \\ & \cdot \prod_{y_i\geq 4}\left[1-\sum_{k=1}^{3}\left[1-\exp\left(-\mu_{1i}\right)\right]^{-1}\frac{\exp\left(-\mu_{1i}\right)\mu_{1i}^{k}}{k!}\right]\frac{\exp\left(-\mu_{2i}\right)\mu_{2i}^{y_i}}{y_i!} \end{aligned} \tag{4}$$

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 maximising separately three different likelihood functions. First, a binary outcome model (the first two terms of equation 4) can report consistent and efficient estimators for $\boldsymbol{\beta_0}$. Then, a model for a left truncated and right censored Poisson variable can properly estimate $\boldsymbol{\beta_1}$ — third and fourth terms of equation 4; for further details see Terza (1985). Finally, a model for a left truncated Poisson (the fifth term of equation 4) can estimate $\boldsymbol{\beta_2}$. 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 summarise, 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 1.

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 (CAIC) statistic will be R also performed,

$$AIC = -2\ln(L) + 2k$$
$$CAIC = -2\ln(L) + k[\ln(n) + 1]$$

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 $r = \{0, 1, 2, \ldots\}$ 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, Finally the statistic is calculated as,

$$\chi^2 = \sum_{r=0}^{R} \frac{(n_r - \widehat{n}_r)^2}{\widehat{n}_r}$$

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

3.2 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,

$$\mu_{ki} = \exp\left[\mathbf{x}'_{\mathbf{ki}}\boldsymbol{\beta_k} + v_{ki}\right], \quad k = 0, 1, 2. \tag{5}$$

Next, some assumptions about the distribution of v_{0i}, v_{1i}, and v_{2i} 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_{0i}, v_{1i}, and v_{2i} 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 (5) as

$$\mu_{ki} = \exp\left[\mathbf{x}'_{\mathbf{ki}}\boldsymbol{\beta_k} + \theta_k v_i\right], \quad \theta_2 = 1; \quad k = 0, 1, 2. \tag{6}$$

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 is normalised to one. If σ^2 represents the variance of the random effect v, one could show that

$$\begin{aligned} \text{Var}\left[\ln\left(\mu_2\right)\right] &= \sigma^2 \\ \text{Var}\left[\ln\left(\mu_k\right)\right] &= \theta_k\sigma^2, \quad k=0,1 \end{aligned}$$

and,

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

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 $\{\boldsymbol{\beta_0}, \ \boldsymbol{\beta_0}, \ \boldsymbol{\beta_0}, \theta_0, \theta_1, \sigma^2\}$ 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

$$L_i = \int_{v_i} L_i\left(v_i\right) g\left(v_i\right) dv_i, \tag{7}$$

where $L_i\left(v_i\right)$ represents the conditional likelihood function. The model is closed once a distribution for the unobserved heterogeneity term, $g\left(vi\right)$, is specified. Here a Normal distribution will be used. Since the integral in equation (7) 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 $\theta_2 = 0$ 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.

3.3 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 work. There are, however, two main previous efforts to control explicitly for the special characteristics that completed fertility data exhibit.

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. In a similar vain, 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 (see, for instance, Bratti and Miranda 2009). Thus, observing a strictly positive level of consumption implies that two hurdles have been crossed. Yen et al. (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 chapter 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.

4 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.

4.1 Insights from standard hurdle models

Table 4 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 H0: $\sigma^2 = 0$ 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 4 the likelihood of remaining permanently childless is significantly affected only by the education of the index woman — see the top panel of Table 4. In fact, a likelihood ratio test for the exclusion of catholic, indspker, c4549 trough c5559, and north through south is not rejected with a $\chi^2(8) = 14.6$ 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 4. 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 (2001).

According to the results in Table 4, 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.

4.2 Results from double-hurdle models

Table 5 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 6 presents a series of likelihood ratio tests that help discriminating 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 H0: Var($log(\mu_0)$) $= \theta_0\sigma^2 = 0$ against H1: Var($log(\mu_0)$) $\neq 0$. Table 7 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 H0: var(log(m1)) = 0 against H1: Var($log(\mu_1)$) $\neq 0$. But if H0: $\sigma^2 = 0$ is tested against H1: $\sigma^2 \neq 0$ 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 5 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 5).

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.

4.2.1 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 7. From now on the vector of parameters that enter count process (i) of the Double Hurdle model will be referred to as $\boldsymbol{\beta}_0$. Similarly, parameters that enter count process (ii) and (iii) are referred to as $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$.

Table 7 contains a formal likelihood ratio test for the joint equality of the coefficients $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. 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 $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_1$ with a $\chi^2(10) = 1610.30$ [p-val = 0.000], and $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$ 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 $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ 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 $\boldsymbol{\beta}_1$ than in vector $\boldsymbol{\beta}_2$. However, pair-wise equality tests (Coefficient on the j-th variable in $\boldsymbol{\beta}_1$) = (Coefficient on the j-th variable in $\boldsymbol{\beta}_2$) reject the null hypothesis exclusively in the case of Edu12 with a t-stat = -2.27 [p-val = 0.011]. A similar exercise reveals that there are significant pair-wise differences in the coefficients on c4549 (t-stat = 1.61, p-val = 0.053), c5054 (t-stat = 2.55, p-val = 0.005), c5559 (t-stat = 4.89, p-val = 0.000), centre (t-stat = -1.70, pval = 0.044) and south (t-stat = -3512, p-val = 0.000). 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 $\boldsymbol{\beta}_2$ than in $\boldsymbol{\beta}_1$, implying that the impact of these socio-economic characteristics on family size is stronger once the second hurdle has been crossed.

4.2.2 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 after controlling for unobserved heterogeneity Double-Hurdle ($\chi^2 = 150$) does better than standard Hurdle ($\chi^2 = 213$). 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.

4.2.3 Effect of explanatory variables

One of the most interesting observations that one may draw from the results in Table 5 is that, except for constant and Edu12, all the elements of b_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 $\chi^2(8) = 14.6$ and a p-val = 0.067. Thus, education is the only variable that affects the probability of observing a zero count. As expected, the coefficient on Edu12 is negative.

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 b1. Notice then that, since $\Pr(j > 3|j > 0)$ is a function of $\boldsymbol{\beta}_1$, the probability of crossing the second hurdle — or say, getting out of the [1,3] interval — is also a function of $\boldsymbol{\beta}_1$. Using this interpretation for the elements of vector $\boldsymbol{\beta}_1$ the reader can conclude from the estimates in Table 5 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 9 contains predicted probabilities for a non-Catholic woman (individual I) who is otherwise identical to a benchmark Catholic woman (individual II).[1]. From this Table the reader can learn that individual I scores a $\Pr(1 < j < 3|j > 0) = 0.43$ while the benchmark individual II scores a $\Pr[1 < j < 3|j > 0) = 0.47$. Hence, the marginal effect of Catholic is around -3.8 percentage points.

[1]The benchmark individual is a Catholic, non-indigenous language speaker who was born in Mexico City between 1940 and 1944 and has the average education in the study sample

Various factors may be behind the negative and significant effect of Catholic in the probability of crossing the second hurdle. The point of view of the author is that this effect is a consequence of the rather weak opposition of the Catholic Church towards the diffusion and adoption of contraceptives in Mexico. The conjecture is 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 the Catholic community. 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.

A negative coefficient on Edu12 in vector $\boldsymbol{\beta}_1$ in Table 5 suggests that an extra year of education at age 12 increases the likelihood that a woman will remain with three or less 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 9 increasing Edu12 from five to six years will lead to an increment in $\Pr(1 < j < 3|j > 0]$ 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. Other effects on the probability of crossing the second hurdle are interesting. For instance, the effect of the cohort dummies are increasingly negative. This obviously implies that younger generation of Mexican women are becoming less and less likely to have more than three children. Clearly, this finding is consistent with the fact that the total fertility rate in Mexico dropped rapidly in the last few decades. Finally, being an indigenous language speaker increases the chances of crossing the second hurdle. 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.

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 the vector of coefficients $\boldsymbol{\beta}_2$. Notice first from table 5 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 at 1% of significance. The effect of Catholic is important. In fact, Table 9 shows that a Catholic will end her fertile life with more than six children with probability 0.39 while her non-Catholic equivalent will register the same event with probability 0.42. In other words, the marginal effect on Catholic on $\Pr(j > 6|j > 3)$is around -3 percentage points. 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. As expected, younger generations of women are found to be less likely to have a large family than their predecessors.

5 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.

References

Arulampalam, W. and 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. and Blisard, W. (1993). Women and the demand for alcohol: Estimating participation and consumption. *The Journal of Consumer Affairs*, 27(2):319–334.

Bratti, M. and Miranda, A. (2009). Non-pecuniary returns to higher education: The effect on smoking intensity in the UK. *Health Economics.*, Published Online: Jul 13 2009 12:08PM (Early View). DOI: 10.1002/hec.1529.

Cabrera, G. (1994). Demographic dynamics and development: The role of population policy in mexico. *Population and Development Review*, 20:105–120.

Cameron, A. and 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 (2001). Marginality index 2000. Document downloadable at www.conapo.gob.mx.

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.

Gomez, J. (1996). La fecundidad y el crecimiento de la descendencia. *Demos*, 9:8–10.

Greene, W. (1997). Fiml estimation of sample selection models for count data. Stern School of Business, New York University. Manuscript.

Heckman, J. (1990). Varieties of selection bias. *The American Economic Review*, 80(2):313–318.

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.

Jones, A. (1989). A double-hurdle model for cigarette consumption. *Journal of Applied Econometrics*, 4(1):23–39.

Kholer, H. (2000). Fertility decline as a coordination problem. *Journal of Development Economics*, 63(2):231–63.

Klawon, E. and 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.

Melkersson, M. and 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.

Santos Silva, J. and Covas, F. (2000). A modified hurdle model for completed fertility. *Journal of Population Economics*, 13(2):173–88.

Terza, J. (1985). A tobit-type estimator for the censored poisson regression model. *Economics Letters*, 18:361–365.

Terza, J. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics*, 84:129–154.

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*. Springer, Berlin Heidelberg, New York, 3rd, rev. and enl. edition.

Yen, S. and Jensen, H. (1996). Determinants of household expenditures on alcohol. *Journal of Consumer Affairs*, 30(1):48–67.

Yen, S., Tang, C., and Su, S. (2001). Demand for traditional medicine in taiwan: A mixed gaussian-poisson model approach. *Health Economics*, 10:221–232.

Appendix

Table 1. Empirical distribution of Children and Poisson with mean 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	1.0000	1.000

Table 2. Likelihood of high parities given children > 3

Count	4	5	6	7-18	Total
N. obs.	2,905	2,349	1,818	4,173	11,245
$\Pr(\text{children}\mid\text{children} > 3)$	0.26	0.21	0.16	0.37	1.00

Table 3. 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
Catholic	= 1 if Catholic; 0 otherwise	0.90	-	-	-
Indspker	= 1 if indigenous language speaker; 0 otherwise	0.09	-	-	-
C4044	= 1 if born within 1940-1944; 0 otherwise	0.10	-	-	-
C4549	= 1 if born within 1945-1949; 0 otherwise	0.29	-	-	-
C5054	= 1 if born within 1950-1954; 0 otherwise	0.36	-	-	-
C5559	= 1 if born within 1955-1959; 0 otherwise	0.25	-	-	-
MexCity	= 1 if born in Mexico City; 0 otherwise (base group)	0.05	-	-	-
North	= 1 if born in the North; 0 otherwise	0.23	-	-	-
Centre	= 1 if born in the Centre; 0 otherwise	0.54	-	-	-
South	= 1 if born in Mexico the South; 0 otherwise	0.18	-	-	-
Number of observations					19,477

Table 4. Standard Hurdle Model

Count Process	(1) No Het.	(2) Normal Het.
At zero		
Constant	1.1547 [0.0675]**	1.1547 [0.0675]**
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]**
C4549	0.0230 [0.0382]	0.0230 [0.0382]
C5054	0.494 [0.0374]	0.0494 [0.0374]
C5559	0.0225 [0.0390]	0.0225 [0.0390]
North	0.0558 [0.0487]	0.0558 [0.0487]
Centre	0.0460 [0.0519]	0.0460 [0.0519]
Larger than zero		
Constant	1.7903 [0.0260]**	1.7740 [0.0280]**
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]**
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]**
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	-44,144.42	-43996.48
AIC	88,328.84	88,034.95
CAIC	88,506.38	88,221.37
N. observations	19,477	19,477

Note: Standard errors in brackets. ** significant at 1%; significant 5%.

Table 5. Poisson Double-Hurdle Model

	No Heterogeneity	Normal Heterogeneity		
	(1)	(2)	(3)	(4)
Count Process		$\theta_0 = \theta_1 = 0$	$\theta_0 = \theta_1 = 1$	θ_0, θ_1 free
At zero − (i)				
Constant	1.1547 [0.0675]**	1.1547 [0.0675]**	1.1800 [0.0698]**	1.2119 [0.4945]
Catholic	-0.0525 [0.0342]	-0.0525 [0.0342]	-0.0543 [0.0353]	-0.0562 [0.0481]
Indspker	-0.0728 [0.0381]	-0.0728 [0.0381]	-0.0753 [0.0393]	-0.0780 [0.0598]
Edu12	-0.0314 [0.0047]**	-0.0314 [0.0047]**	-0.0324 [0.0049]**	-0.0337 [0.0201]**
C4549	0.0230 [0.0382]	0.0230 [0.0382]	0.0237 [0.0394]	0.0246 [0.0431]
C5054	0.0494 [0.0374]	0.0494 [0.0374]	0.0513 [0.0386]	0.0531 [0.0512]
C5559	0.0225 [0.0390]	0.0225 [0.0390]	0.0235 [0.0402]	0.0244 [0.0449]
North	0.0558 [0.0487]	0.0558 [0.0487]	0.0575 [0.0502]	0.0244 [0.0449]
Centre	0.0001 [0.0465]	0.0001 [0.0465]	0.0001 [0.0480]	-0.0002 [0.0498]
South	0.0460 [0.0519]	0.0460 [0.0519]	0.0460 [0.0535]	0.0489 [0.0610]
At one-to-three − (ii)				
Constant	1.7142 [0.0328]**	1.7142 [0.0328]**	1.7370 [0.0344]**	1.7142 [0.0328]**
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]**
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]**
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 [0.0519]**	0.1638 [0.0271]**	0.1597 [0.0262]**

Note: Standard errors in brackets. ** significant at 1%; * significant 5%.

Table 5. Poisson Double-Hurdle Model (continued)

	No Heterogeneity	Normal Heterogeneity		
	(1)	(2)	(3)	(4)
Count Process		$\theta_0 = \theta_1 = 0$	$\theta_0 = \theta_1 = 1$	θ_0, θ_1 free
Larger than three − (iii)				
Constant	1.7752 [0.0522]**	1.7564 [0.0542]**	1.7429 [0.0537]**	1.7554 [0.0550]**
Catholic	-0.0348 [0.0156]*	-0.0359 [0.0168]*	-0.0379 [0.0164]*	-0.0351 [0.0171]*
Indspker	0.0129 [0.0156]	0.0163 [0.0169]	0.0161 [0.0165]	0.0173 [0.0175]
Edu12	-0.0753 [0.0023]**	-0.0768 [0.0024]**	-0.0798 [0.0025]**	-0.0763 [0.0033]**
C4549	-0.0911 [0.0153]**	-0.0934 [0.0166]**	-0.0944 [0.0162]**	-0.0937 [0.0167]**
C5054	-0.2025 [0.0156]**	-0.2075 [0.0170]**	-0.2103 [0.0166]**	-0.2082 [0.0174]**
C5559	-0.3030 [0.0180]**	-0.3086 [0.0193]**	-0.3130 [0.0190]**	-0.3089 [0.0195]**
North	0.2831 [0.0494]**	0.2810 [0.0509]**	0.2913 [0.0504]**	0.2801 [0.0511]**
Centre	0.3570 [0.0486]**	0.3559 [0.0500]**	0.3657 [0.0496]**	0.3558 [0.0501]**
South	0.2787 [0.0499]**	0.2740 [0.0515]**	0.2816 [0.0510]**	0.2732 [0.0517]**
σ^2		0.0340 [0.0042]**	0.0239 [0.0038]**	0.0346 [0.0065]**
θ_0		set to zero	set to one	-1.2450 [5.3790]
θ_1		set to zero	set to one	-0.0080 [0.2314]
log-likelihood	-43,980.42	-43,941.15	-43,956.11	-43,941.13
AIC	88,020.84	87,944.30	87,974.22	87,948.26
CAIC	88,287.15	88,219.49	88,249.41	88,241.20
N. of observations	19,477	19,477	19,477	19,477

Note: Standard errors in brackets. ** significant at 1%; significant 5%.

Table 6. Model Selection

Case	H_0	H_1	Test type	χ^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	$\theta_0 = \theta_1 = 1$, $\sigma^2 \neq 0$	$\theta_0 \neq 1, \theta_1 \neq 1$, $\sigma^2 \neq 0$	LRT	29.90 [0.0000]	Reject H_0

Boundary-value likelihood test is abbreviated as BVLRT; Likelihood ration test is abbreviated as LRT.

Table 7. Likelihood ratio test

H_0	H_1	LR	P-val	Inference
$\boldsymbol{\beta}_0 = \boldsymbol{\beta}_1$	$\boldsymbol{\beta}_0 \neq \boldsymbol{\beta}_1$	1610.30	0.000	Reject H_0
$\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$	$\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$	164.27	0.000	Reject H_0
$\boldsymbol{\beta}_0 = \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$	$\boldsymbol{\beta}_0 \neq \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$	2339.49	0.000	Reject H_0

Table 8. Observed and predicted sample distribution

		Standard Hurdle		Double hurdle (best fit)	
Count	Obs	No. Het.	Normal Het.	No Het.	Normal Het. ($\theta_0 = \theta_1 = 0$)
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
χ^2		371	213	150	116
Pr> χ^2		0.0000	0.0000	0.0000	0.0000
LogL		-44,144	-43,996	-43,980	-43,941
AIC		88,329	88,035	88,021	87,944
CAIC		88,506	88,221	88,287	88,219

Note: Sample size is 19,477.

Table 9. Predicted probabilities — Double-Hurdle Poisson

Case	Characteristics	$\Pr(j = 0)$	$\Pr(1 < j \leq 3 \mid j > 0)$	$\Pr(j > 6 \mid j > 3)$
1	edu12=mean, all dummies zero	0.0609	0.4302	0.4165
2	edu12=mean, catholic=1, other dummies zero	0.0703	0.4684**	0.3947*
3	edu12=mean, catholic=1, indspker=1, other dummies zero	0.0847	0.4378	0.4044
4	edu12=mean, catholic=1, C4549=1, other dummies zero	0.0661	0.5081**	0.3416**
5	edu12=mean, catholic=1, C5054=1, other dummies zero	0.0615	0.5648**	0.2842**
6	edu12=mean, catholic=1, C5559=1, other dummies zero	0.0662	0.5955**	0.24**
7	edu12=mean, catholic=1, north=1, other dummies zero	0.0604	0.2818**	0.582**
8	edu12=mean, catholic=1, centre=1, other dummies zero	0.0703	0.2753**	0.6361**
9	edu12=mean, catholic=1, south=1, other dummies zero	0.0620	0.3487**	0.5769**
10	edu12=0, catholic=1, other dummies zero	0.0493**	0.2243**	0.6015**
11	edu12=5, catholic=1, other dummies zero	0.0762**	0.5298**	0.3511**
12	edu12=6, catholic=1, other dummies zero	0.0826**	0.5891**	0.3107**

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

Figure 1. Double-Hurdle Model Structure.