WIDER TECHNICAL NOTE | 3/2022

SOUTHMOD – simulating tax and benefit policies for development

Discrete choice modelling of labour supply and informal employment using ECUAMOD

H. Xavier Jara[1] and Pia Rattenhuber[2]

May 2022

wider.unu.edu

Abstract: This technical note presents one of the modelling approaches used in Jara and Rattenhuber (2022) to estimate formal employment elasticities, namely the estimation of a discrete choice model of labour supply with informal employment. The approach makes use of ECUAMOD, the tax-benefit model for Ecuador, to calculate the disposable income at each discrete alternative defined in terms of hours of work and sector of employment (formal or informal employment). The ECUAMOD outputs are then used as input in the estimation of a conditional logit model of labour supply. The note first describes the data preparations to simulate disposable income at each discrete alternative using ECUAMOD. Then, it presents the modelling set-up for the estimation of the discrete choice model of labour supply with informal employment. Finally, it presents formal employment elasticities across population subgroups based on our estimations.

Key words: labour supply, informal employment, tax-benefit system, Ecuador

JEL classification: J23, J42

Acknowledgements: The work presented in this note is based on ECUAMOD v2.0. ECUAMOD is developed, maintained, and managed by UNU-WIDER in collaboration with the EUROMOD team at ISER (University of Essex), SASPRI (Southern African Social Policy Research Insights), and local partners in selected developing countries (Ethiopia, Ghana, Mozambique, Tanzania, Uganda, Zambia, Ecuador, and Viet Nam) in the scope of the SOUTHMOD project. The local partner for ECUAMOD is Facultad Latinoamericana de Ciencias Sociales (FLACSO) Sede Ecuador. We are indebted to the many people who have contributed to the development of SOUTHMOD and ECUAMOD.

Related publications:

Jara, H.X., and P. Rattenhuber (2022). 'Female Labour Supply and Informal Employment in Ecuador'. WIDER Working Paper 47/2022. Helsinki: UNU-WIDER.

---

[1] Institute for Social and Economic Research, University of Essex, Colchester, UK, corresponding author: hxjara@essex.ac.uk; [2] UNU-WIDER, Helsinki, Finland

This note is published within the UNU-WIDER project SOUTHMOD – simulating tax and benefit policies for development Phase 2, which is part of the Domestic Revenue Mobilization programme. The programme is financed through specific contributions by the Norwegian Agency for Development Cooperation (Norad).

Information and requests: publications@wider.unu.edu

https://doi.org/10.35188/UNU-WIDER/WTN/2022-3

United Nations University World Institute for Development Economics Research

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

1 Introduction

Reforms to tax-benefit policies do not only affect household incomes but they might also influence behavioural reactions of individuals in the labour market. In the context of low- and middle-income countries, reforms to tax-benefit systems might influence individuals' choice between formal and informal employment. Thus, it seems appropriate to complement static tax-benefit microsimulation models with behavioural components allowing to assess the effect of potential reforms on employment.

This technical note describes the use of ECUAMOD, the tax-benefit microsimulation model for Ecuador, to generate input for the estimation of a discrete choice model of labour supply with informal employment. The note has been produced alongside a WIDER Working Paper (Jara and Rattenhuber 2022) analysing female labour supply with informal employment in Ecuador. The paper and the technical note are based on ECUAMOD version v2.0. For the analysis, the model makes use of four waves of the National Survey of Employment, Unemployment and Underemployment of Urban and Rural Households (ENEMDU).

The note is structured as follows. Section 2 describes the data preparation with ECUAMOD for the modelling of labour supply and informality. Section 3 describes the set-up of the discrete choice model of labour supply including hour alternatives in formal and informal employment. Section 4 presents the formal employment elasticities estimated with the discrete choice model.

2 Data preparation for the modelling of labour supply and informality

The estimation of labour supply elasticities with ECUAMOD is based on ENEMDU, a nationally representative survey conducted on a quarterly basis, which represents the main data source to track labour market changes and the evolution of poverty and inequality in Ecuador. ENEMDU contains information on employment, labour and non-labour income, public pensions, cash transfers, private transfers, as well as personal and household characteristics for the simulation of tax-benefit policies. It also contains information on affiliation to social security, which we use to define formal employment. Informal employment in our analysis is therefore defined as non-affiliation to social security. For our analysis, we use four waves of ENEMDU which correspond to the December rounds of years 2011, 2013, 2017, and 2019. Pooling cross-sectional waves for the estimation of discrete choice labour supply models has two advantages. First, it allows having sufficiently large data to apply the sample restrictions imposed for the estimations. Second, time variation in tax-benefit rules represents an additional source for identification of the estimated parameters

For the purpose of the labour supply estimations, households where only one labour supply unit is present are selected for the analysis. A labour supply unit is defined as single men, single women, or couples with or without dependent children or dependent elderly. Dependent children are defined as children aged 18 or below who are in education and have no earnings, and dependent elderly are defined as parents and parents-in-law aged 60 or above who are retired and have no earnings. Extended households, where more than one labour supply unit cohabit, are therefore excluded.

Due to the flexibility offered by ECUAMOD to create different 'tax units', these labour supply units were created in ECUAMOD under the TUDef_ec policy using the DefTu function of the software, as shows in Figure 1. Then, the ECUAMOD output file containing identifiers of each labour supply unit was used to select the households kept for the labour supply estimation.

Figure 1: Definition of labour supply assessment unit

DefTu		on	Labour Supply Tax Unit
Name		tu_family_ls_ec	
Type		SUBGROUP	unit is a subgroup of the household
Members		Partner & OwnDepChild & DepParent	members: head, partner, dependent children and dependent parents
DepChildCond		dag<5 | (dag<=18 & dec>=1 & dec<4 & ils_earns#1=0)	dependent children: those aged below 5 or aged below 18 in education, and with zero earnings
DepParentC...		Default & dag>60 & ils_earns#1=0	dependent parents: those aged above 60 and with zero earnings
#_Level	1	tu_individual_ec	
LoneParentC...		Default & !IsMarried	

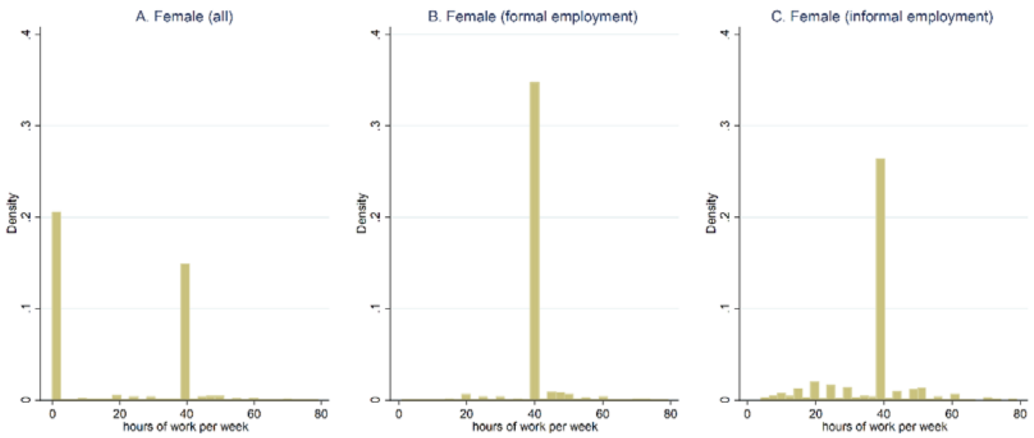
Source: screenshot from ECUAMOD v2.0, EUROMOD software.

Following the labour supply literature, the sample of analysis was further restricted to households where the head of the household and their partner are of working age (aged between 19 and 59), available for the labour market (not disabled, in education, or retired) and excluding those in self-employment and those with more than one job. Finally, we restricted our analysis to females, but the same model can be estimated for males.

Our selected sample includes 12,722 households: 3,065 single women and 9,657 women in couples, which represents 18 per cent of all households with female members in our data. The sample restrictions described above should be considered with care prior to the estimation of labour supply models in other SOUTHMOD countries, depending on the structure of households in each country.

To assess the number of discrete hour alternatives to be used in the estimation of the structural labour supply model, we then look at the distribution of hours of work in our sample of analysis. Figure 2 presents the distribution of actual worked hours for all women in our selected sample (Panel A). For those in work in each of these categories, Figure 2 further shows the distribution of hours of work, distinguishing between those in formal (Panel B) and informal (Panel C) employment.

Figure 2: Distribution of weekly hours of work (pooled selected sample)

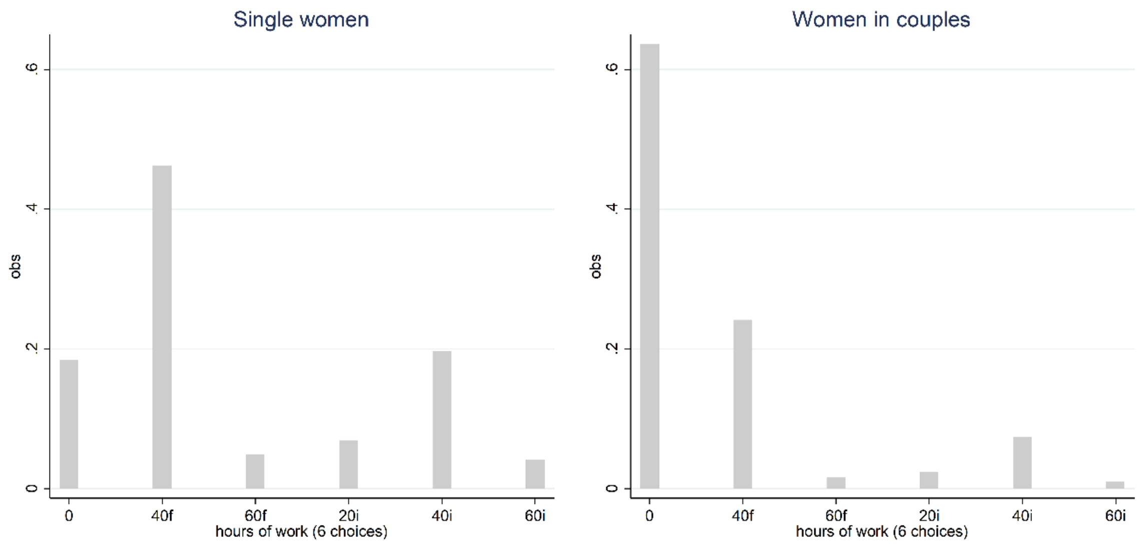


Source: authors' elaboration based on ENEMDU and ECUAMOD v2.0.

We observe two marked peaks at non-participation and full-time work (Panel A in Figure 2). The accompanying WIDER Working Paper (Jara and Rattenhuber 2022) shows that the distribution varies across single women and those in couples. For single women, the largest concentration is observed around full-time work, whereas the majority of women in couples do not participate. The distribution of female hours of work in formal employment (Panel B in Figure 2) has a very strong concentration around full-time work, whereas more variation is observed in the distribution of hours of work of women in informal employment (Panel C in Figure 2), with a somewhat larger concentration around part-time work.

Based on the observed distributions, we define six alternatives for our discrete choice model, characterizing non-participation, full-time, and overtime for formal employment, and part-time, full-time, and overtime for informal employment. The discretized set for each individual is, therefore, given by $h = \{0, 40f, 60f, 20i, 40i, 60i\}$, where f stands for formal employment and i for informal employment. Figure 3 shows the distribution of discrete hours of work for single women and women in couples in our analysis.

Figure 3: Distribution of discrete hours of work alternatives (pooled selected sample)



Source: authors' elaboration based on ENEMDU and ECUAMOD v2.0.

The final step of the data preparation for the estimation of the discrete choice labour supply model consists of saving an ECUAMOD input data file containing only the households in our selected sample. ECUAMOD is then run a number of times equal to the number of discrete hour alternatives defined for the analysis. In our case, ECUAMOD is run six times. More precisely, for each working-age individual (i.e. aged between 19 and 59) in our selected households, we calculate gross hourly wages and multiply them by each discrete alternative of hours of work to run ECUAMOD and obtain household disposable income under each alternative.[1] The six different ECUAMOD outputs are then merged together for the estimation of the discrete choice model.

[1] Gross hourly wages are predicted for non-workers using a two-stage Heckman selection model separately for formal and informal employment.

3 Estimating a discrete choice model of labour supply with informality

Our model specification represents an extension of the unitary discrete choice model of household labour supply of van Soest (1995). The approach is derived under the random utility maximization framework and consists of defining a finite set of working hour alternatives and to explicitly specify a utility function to evaluate individuals' utility at each discrete alternative. In our case, we allow individuals to choose between discrete alternatives representing non-participation and hours of work in formal or informal employment. Based on the observed distribution of hours of work across sectors shown in Figure 2, our discrete choice model considers six alternatives: (i) inactivity, (ii) full-time in formal employment, (iii) overtime in formal employment, (iv) part-time in informal employment, (v) full-time in informal employment, (vi) overtime in informal employment.

Let U_{ij} represent individual's i utility at alternative $j = 1, \dots, J$, where the utility function can be decomposed into a deterministic and a stochastic component: $U_{ij} = V_{ij} + \varepsilon_{ij}$. Assuming that individuals maximize utility (i.e. individual i chooses alternative j if and only if $U_{ij} > U_{ik}, \forall k \neq j$) and the stochastic component is independent and identically distributed over alternatives and follows a type-one extreme value distribution, then the probability that an alternative j is chosen follows is expressed as (McFadden 1974):

$$P_{ij} = \frac{e^{V_{ij}}}{\sum_{k=1}^{J} e^{V_{ik}}}$$

In our setting, individuals choose among a finite number of alternatives to maximize their utility, defined over net income and hours of work in each sector (formal or informal employment). More formally, let h_i be the number of hours worked and s_i the sector of employment of individual i. We define J discrete alternatives so that h_{ij} represents the number of hours worked by individual i under alternative j and s_{ij} the sector of employment of individual i in alternative j, with $j = 1, \dots, J$.

Let y_{ij} be individual i's household disposable income given the hours h_{ij} and sector s_{ij} choice, and z_i a vector of individual characteristics. Household disposable income y_{ij}, when $h_i = h_{ij}$ and $s_i = s_{ij}$ are chosen, is defined as:

$$y_{ij} = w_{ij}h_{ij} + \mu_i + G\left(w_{ij}, h_{ij}, \mu_i, z_i\right),$$

where w_{ij} are gross hourly wage rates in sector s_{ij} with gross hourly wages varying across sectors but fixed across hour alternatives within each sector. μ_i is non-labour income, and the function $G\left(w_{ij}, h_{ij}, \mu_i, z_i\right)$ represents the tax-benefit function which depends on gross wages, hours of work, non-labour income, and individual characteristics. y_{ij} is calculated using ECUAMOD for each discrete alternative. In the tax-benefit simulations, workers in informal employment are assumed not to pay social insurance contributions and personal income tax but might still receive non-contributory social assistance benefits depending on their household characteristics.

Several specifications of the deterministic part of the utility function have been used in the literature. In our analysis, we follow Keane and Moffitt (1998), Brewer et al. (2006), and Kabátek et al. (2014) and specify a quadratic utility function given by:

$$V(y_{ij}, h_{ij}, s_{ij}, z_i) = \alpha_{yy} y_{ij}^2 + \alpha_{hh} h_{ij}^2 + \alpha_y y_{ij} + \alpha_h h_{ij} + \alpha_s s_{ij} + \alpha_{yh} y_{ij} h_{ij} + \alpha_{ys} y_{ij} s_{ij}$$

$$+\alpha_{hs} h_{ij} s_{ij},$$

where s_{ij} is a dummy equal to 1 for informal employment alternatives, zero otherwise. Interactions between income, leisure, and the informal employment dummy are included to consider their differentiated effect across sectors.

The coefficients of hours of work and informal employment are allowed to vary with personal characteristics z_i, specified as:

$$\alpha_h = \alpha_{h0} + \alpha'_{hx} z_i$$

$$\alpha_s = \alpha_{s0} + \alpha'_{sx} z_i$$

The conditional logit model is estimated by maximum likelihood. The model is estimated separately for single women and women in couples. For women in couples, the labour supply of their partners is fixed at their observed number of hours of work and sector of employment.

4 Formal employment elasticities

Elasticities in discrete choice labour supply models are calculated numerically using the estimated parameters of the utility function. Formal employment elasticities are obtained based on a 10 per cent increase in formal gross wages (wage elasticity) or non-labour income (income elasticity). More precisely, first, we increase gross hourly wages (income) in formal employment by 10 per cent and compute the new disposable income for each alternative using ECUAMOD. Then, with the estimated coefficients from the estimation of the conditional logit model, we calculate the average probability of being at each alternative for the new and baseline value of disposable income. Finally, we compare the new and the baseline predicted frequencies to assess the change in the predicted frequencies of formal employment. Table 1 presents formal employment elasticities obtained from our simulations by education and income quintiles.

Table 1: Formal employment elasticities by education and income

	Single women		Women in couples	
	Wage elasticity	Income elasticity	Wage elasticity	Income elasticity
Low education	0.10	0.14	0.10	0.34
Middle education	0.08	0.12	0.13	0.44
High education	0.05	0.06	0.12	0.29
Q1	0.08	0.12	0.11	0.24
Q2	0.09	0.13	0.12	0.32
Q3	0.09	0.13	0.13	0.38
Q4	0.08	0.11	0.14	0.45
Q5	0.05	0.07	0.11	0.33
All	0.07	0.09	0.12	0.36

Source: authors' elaboration based on ECUAMOD v2.0.

References

Brewer, M., A. Duncan, and A. Shephard (2006), 'Did Working Families' Tax Credit Work? The Impact of In-Work Support on Labour Supply in Great Britain'. *Labour Economics*, 13(6): 699–720. https://doi.org/10.1016/j.labeco.2005.11.002

Jara, H.X., and P. Rattenhuber (2022). 'Female Labour Supply and Informal Employment in Ecuador'. WIDER Working Paper 47/2022. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2022/178-5

Kabátek, J., A. van Soest, and E. Stancanelli (2014). 'Income Taxation, Labour Supply and Housework: A Discrete Choice Model for French Couples'. *Labour Economics*, 27. 30–43. https://doi.org/10.1016/j.labeco.2014.01.004

Keane, M., and R. Moffitt (1998). 'A Structural Model of Multiple Welfare Program Participation and Labor Supply'. *International Economic Review*, 39 (3), 553–89. https://doi.org/2527390

McFadden, D. (1974). 'Conditional Logit Analysis of Qualitative Choice Behaviour'. In P. Zarembka (ed.), *Frontiers in Econometrics*, pp. 105–42. New York: Academic Press.

van Soest, A. (1995). 'Structural Models of Family Labor Supply: a Discrete Choice Approach'. *Journal of Human Resources*, 30(1): 63–88. https://doi.org/10.2307/146191