

An estimate of female labour force participation with endogenous switching ⁽¹⁾

Stima dell'attività lavorativa femminile con 'endogenous switching'

Gustavo De Santis

DESMS, V. T. Cannizzaro 298, 98122 ME, gdesantis@unime.it

Antonino Di Pino

DESMS, V. T. Cannizzaro 298, 98122 ME, dipino@unime.it

Riassunto: In assenza di dati longitudinali, la stima dell'offerta di lavoro femminile, che interagisce con i processi demografici (matrimonio, fecondità, ecc.), richiede che si cerchi almeno di separare l'influenza inversa, dell'attività lavorativa sulla sfera demografica, ricorrendo a variabili strumentali e, per il loro calcolo su sub-campioni selezionati, alla correzione di Heckman. Ma si può tentare un passo in più, per cogliere anche gli effetti di "endogenous switching", ossia la circostanza che trovarsi, in modo forse inatteso, in una data condizione (es. di divorziata, o di madre di molti figli) può influenzare l'atteggiamento verso il mercato del lavoro, come sembra emergere dai nostri dati.

Keywords: Labour force participation; Heckman correction; endogenous switching.

1. Introduction

We address a substantial problem, the solution of which also requires that we face a methodological obstacle. The substantial problem is that of the estimation of female labour force participation, taking into account the effect of simultaneous, and interfering, demographic processes, and notably fertility, and family formation and dissolution - a situation where it proves difficult to find truly exogenous variables, especially without longitudinal data (Browning, 1992; for a recent application to Italy, Del Boca, 2002).

And it is precisely because of the intrinsic endogeneity of our variables that the methodological issue arises: participation in the labour market L is influenced by, but also influences, what happens in the demographic sphere D , and this biases the parameters of a direct estimation of the type $L=f(D; \varepsilon)$, where ε is an *I.I.D.* error term.

In order to circumvent this difficulty, we follow two paths:

a) firstly, we use instrumental variables and, when necessary, Heckman correction factors. This part is not novel, but necessary: in short, we first estimate the reduced form $\hat{D} = g(E)$ (E =exogenous variables), and then we get back to the original formulation, but now with non-endogenous regressors $\hat{L} = f(\hat{D}) = f[g(E)]$;

b) secondly, we also deal with endogenous switching. In several cases, especially for discrete variables, the fact of having or not having experienced an event (e.g., marriage), and therefore being or not being in a specific state (e.g., currently with a spouse) does affect attitudes towards the labour market. In order to account for this, in a few cases, we

⁽¹⁾ Financial support from the Italian MIUR and from the University of Messina is acknowledged.

include in the regression: i) the instrumental variable $\hat{D}$, "explained" at a first stage by a complete set of exogenous variables, and ii) the residual, or gap, $\tilde{D} = D - \hat{D}$, i.e. the part of the demographic behaviour that our model cannot explain, and that we assume to be a stochastic component, uncorrelated with the disturbances ε . This gap, that may depend, among other things, on pure hazard (e.g., having or not having found the "right" partner), may trigger peculiar behaviours in the labour market. In short, our model occasionally becomes of the type $L = f(\hat{D}, \tilde{D}, \varepsilon)$, and we are now in a position to disentangle the two effects of the demographic situation: that expected on the basis of exogenous variables $\hat{D}$ ("expected effect"), and that determined by the a-priori unexpected situation the woman has actually found herself into $\tilde{D}$ ("residual effect"). Notice that this type of reasoning is in line with Terza's (1998) ideas on how to integrate in the analysis empirical categorical data (D , in our example) and estimated residuals ($\tilde{D}$, in this case), but it develops them in an original way.

2. The model

Our model is too complex for a detailed exposition in these few pages, but its seven equations can be summarised as in table 1 (to read by columns).

Table 1. Equation system to estimate female labour participation, in eq. 7 (Italy 2002).

	Eq. #	1	2	3	4	5	6	7
	Type	Probit	Probit	Ols	Ols	Poisson	Probit	Probit
	Sample	All	Women	All	Employed	Women	Ever married women	Women
Indep. variables	Exogenous	Age Edu Geo(d)	Age Edu Siblings Edu parents	Edu Geo(d)	Edu Geo	Age Edu Siblings	Age Geo(d) Edu parents	Age Edu Geo
	Expected effects			Sex(d) Prev. exp.	Sex(d) White collar (d)	Age diff. G.prnts av.(d)		
	instrument					Partner [2]		Seniority [3] Wage [4] CEB [5]
	Residual effect instrument				Work(d)(H)[1]	Partner [2]	Partner (H)[2]	Partner [2] CEB [5] Divorced [6]
Dep. var.		Work(d)	Partner(d)	Seniority	Wage	CEB	Divorced(d)	Work(d)

Legend. **CEB**=children ever born; **Edu**=own education in # of years of schooling; **Edu parents**=parents' education (# of years of schooling of the most educated); **G.prnts av.**= grand parent(s) available for help with childcare; **Prev. exp**=previous work experience; **Seniority**=log(seniority/age). **(d)**=dichotomous; **(H)**=Heckman correction factor; **[#]**=Eq.# from which the info is taken. Note that several variables are more complex than the scheme shows: e.g. "Age" is also taken as squared; "Geo" may mean South (vs. North-Center) or average regional added value per worker; "Edu" is in difference from average for that sex and age (to control for the spread of education over the generations); "Wage" is in log; etc.

Keeping into account the nature of the dependent variable (dichotomous, count, or continuous) we envisage three type of regressions (Probit, Poisson, OLS), while the poten-

tial bias of estimates based on samples of varying composition (everybody, employed, women, ever married women), is corrected with a Heckman factor. Indeed, there are three types of independent variables: exogenous, “expected effect” instruments (i.e. estimates derived from previous equations), and “residual effect” instruments (i.e. differences between the observed variable and its estimate).

Although we are also interested in intermediate passages - that is, in the goodness of fit and in the value of the coefficients in each of the equations - our aim is eq. 7, where female labour force participation - essentially, an estimate of the supply side of the female labour market - is assumed to depend on a number of individual and, partly, contextual factors.

3. Data

For our empirical application, we use the data of the Bank of Italy Survey of Household Income and Wealth (SHIW), in its latest edition, of year 2002, which is freely available on the internet (<http://www.bancaditalia.it/statistiche/ibf>), along with all the related documentation. We restrict our sample to 10408 adults, women aged 18 to 60 and their husbands (all ages) plus all other men aged 18 to 65. This leaves us with a global sample of slightly more than 10 thousand individuals, of whom 4526 are women, who shrink further to a minimum of 4221 ever married women analysed in eq. 6.

4. Results

We cannot show all of our results here, but we made sure that all our equations pass the standard tests, and that the exclusion restriction applies (more exogenous variables excluded from eq. 7 than instrumental variables included in it). Table 2 shows an example of our outcomes, where it is worth noticing that all the retained parameters are either highly significant, or highlight the difference between the North-Centre and the South of Italy.

Female labour participation follows a typical inverted-U curve in age, with a maximum around 40 years. Participation is higher for those who are more educated and live in richer regions. Moving to instrumental variables, we note that own and partner's wage do not seem to impact too much and, oddly enough, the strongest effect (for women, in North-Centre) is not in the expected direction: this probably reflects the fact that female labour force participation is more widespread in the North-Centre, even in relatively low-wage occupations. More seniority (i.e. years worked in the past, relative to age) increase the chances that one still holds a job, while fertility.

Finally, residuals indicate that having a husband, beyond what would be theoretically expected on the basis of eq. 2 (not shown here), lowers the chances of being a worker, especially in the South. Similarly, having had children beyond the number theoretically expected on the basis of eq. 5, reduces the odds of labour activity, especially in the North-Centre. Finally, being currently divorced, beyond what is expected on the basis of eq. 6, greatly enhances the chances of being at work in the North-Centre, while there are too few cases in the South to obtain reliable estimates

Table 2. *Parameter estimates of eq. 7 (probit, works=1; doesn't work=0; on 4526 women, Italy 2002). Source: own elaborations on SHIW data.*

		Parameter	Std Err	Wald Stat.	p
exogenous	Constant	-5.16319	0.525651	96.48077	0.000
	Age	0.27546	0.021729	160.71014	0.000
	Age2	-0.00342	0.000244	196.89844	0.000
	Edu	0.12565	0.012291	104.49459	0.000
	Geo (y/worker)	0.00003	0.000009	11.84450	0.001
instruments	Wage [^] , M, South	0.00001	0.00001	2.41968	0.120
	Wage [^] , F, CN	-0.00015	0.000022	50.41650	0.000
	Wage [^] , F, South	0.00003	0.000025	1.14734	0.284
	CEB [^] ,CN	-0.42480	0.072453	34.37524	0.000
	Seniority [^] ,CN	4.15580	0.667074	38.81148	0.000
residuals	Divorced ^v , CN	0.56761	0.166060	11.68331	0.001
	CEB ^v , CN	-0.20334	0.027458	54.84142	0.000
	CEB ^v , South	-0.07813	0.034683	5.07534	0.024
	Partner ^v ,CN	-0.39823	0.083255	22.87895	0.000
	Partner ^v , South	-0.67503	0.111915	36.38065	0.000

5. Conclusions

The investigation of the true "causes" of female labour force participation would be hard to carry out even with ad hoc longitudinal information, and proves virtually impossible with cross-sectional data. What one can try to do, in these cases, is to (a) separate the mutual influences of the dependent and the independent variables (using instruments), (b) keep under control the selection bias (Heckman), and (c) consider that once a woman finds herself in a given state (e.g. divorced, or with several children) she is likely to change her perspectives as to whether or not she should join the labour market (endogenous switching). This paper is an attempt to keep all three issues into account: the value of our parameters (shown only for eq. 7), and the general diagnostics of the model (not shown here), seem to suggest that we are moving in the right direction

References

- Browning M. (1992) Children and Household Economic Behaviour, *Journal of Economic Literature*, 30 (3), pp. 1434-1475
- Del Boca, D. (2002) The effect of child care and part time opportunities on participation and fertility decision in Italy, *Journal of Population Economics*, Vol. 15, No 3, pp. 549 - 573
- Terza J.V. (1998) Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects, *Journal of Econometrics*, 84 pp. 129-154.