EUROPEAN COMMISSION
Employment, Social Affairs and Equal Opportunities
Social Protection and Integration
Social Protection, Social Services

CAPP
Centro di Analisi delle Politiche Pubbliche

ESTIMATION AND SIMULATION OF EARNINGS USING IT-SILC

Emanuele Ciani

CAPP, University of Essex

Marcello Morciano

CAPP, University of East Anglia and ISER

Dissemination of research results

"Assessing adequacy and long term distributive effects of the Italian Pension System. A Microsimulation Approach"

under the auspices of the Community Program for Employment and Social Solidarity (PROGRESS), European Commission

University of Modena and Reggio Emilia, 26th September 2011

Aim(s)

Prediction of individual earnings

Ideally, we would like to :

1. Estimate a model of mean earnings, taking into account dynamics
2. Predict the evolution of an individual's earning, accounting for:
   - the evolution of observables
   - serial correlation in the earnings residuals
   - increases in productivity

However:

1. We have limited and long- and up-to-date panel that would also be appropriate as initial population
2. Some DMM use cross-sectional estimates. See DESTINIE (INSEE, 1999), MIDAS (Dekkers et al, 2010) and MIRTODIN (Maitino & Sciclone, 2009)
3. Other use panel data. See PenSim2 (Emmerson et al., 2004), CORSIM (Favreault & Caldweel, 1998), the Dynamod (2002) and CeRPSIM (Borella & Coda Moscarola, 2006, 2009)

Data

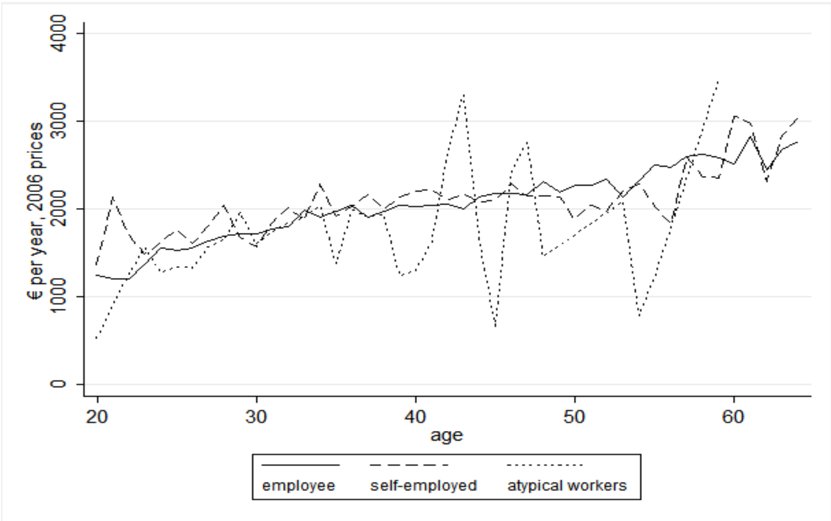
Cross-sectional and longitudinal component of IT-SILC

- Pros:
  - Large sample size, representative at the regional level
  - Link with administrative data
  - Earnings are provided both gross and net of taxes and SSC
- Cons:
  - Short rotating panel

Data

- Sample selection rules:
  - Employed individuals
  - Aged restriction [20, SPA]
  - Bottom and top-coding
- Definition of earnings:
  - Employee cash income and cash benefits or losses from self-employment
  - Gross of taxes and social contribution paid by the workers
  - (log) monthly earnings in euro 2006

Mean gross monthly earnings by age and employment status



Cross-sectional estimates, log earnings

	Men, not graduated, employees	Women, not graduated, employees	Men, graduated, employees	Women, graduated, employees	Graduated, self-employed	Men, not graduated, self-employed	Women, not graduated, self-employed
North	0.1006***	0.1187***	0.0651*	0.0595*	0.2425***	0.0646**	0.0442
	(0.023)	(0.027)	(0.039)	(0.033)	(0.057)	(0.032)	(0.050)
South	-0.1458***	-0.0961***	-0.1686***	-0.0772**	0.0023	-0.2655***	-0.3025***
	(0.027)	(0.034)	(0.046)	(0.038)	(0.070)	(0.036)	(0.059)
Private	-0.0845***	-0.1619***	-0.0694*	-0.0625*			
	(0.012)	(0.014)	(0.036)	(0.034)			
Part-time	-0.6440***	-0.5145***	-0.5826***	-0.5718***	-0.4604***	-0.2742***	-0.2830***
	(0.034)	(0.015)	(0.147)	(0.040)	(0.080)	(0.078)	(0.054)
Secondary	0.0507***	0.1402***				0.1926***	0.1990***
	(0.019)	(0.025)				(0.026)	(0.043)
Immigrant	-0.2228***	-0.2118***	-0.4556***	-0.3970***	-0.4060***	-0.0313	-0.2395**
	(0.020)	(0.032)	(0.074)	(0.081)	(0.139)	(0.079)	(0.110)
Age	0.0336***	0.0147***	0.0736***	0.0493***	0.0247***	-0.0029	0.0340**
	(0.004)	(0.005)	(0.015)	(0.012)	(0.005)	(0.010)	(0.017)
Age squared	-0.0004***	-0.0002***	-0.0006***	-0.0004***		0.0001	-0.0004**
	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
Contributions	0.0120***	0.0154***	0.0024	0.0111***	0.0396***	0.0060**	0.0082**
	(0.001)	(0.002)	(0.004)	(0.003)	(0.011)	(0.003)	(0.003)
Contributions squared					-0.0009***		
					(0.000)		
Women					-0.1621***		
					(0.054)		
Atypical					-0.0002	-0.2127***	-0.1089
					(0.062)	(0.077)	(0.070)
Observations	7478	5349	1005	1169	911	2627	1124
R^2	0.334	0.427	0.293	0.349	0.265	0.112	0.119
RESET (p-value)	0.6852	0.4375	0.5726	0.4482	0.5171	0.8156	0.6057

Note: the regressions also include a constant, SecondaryXcontributions, NorthXcontributions, SouthXcontributions

Panel estimates

- We model the earnings residuals (Lillard and Willis, 1978)

$$\ln y_{it} = z_{it}\beta + \varepsilon_{it}$$
$$\varepsilon_{it} = \mu_i + \xi_{it}$$
$$\xi_{it} = \rho\xi_{it-1} + \omega_{it},\ \omega_{it} \sim iid(0, \sigma^2_\omega),\ |\rho| < 1$$
$$\mu_i \sim iid(0, \sigma^2_\mu)$$

- We focus on the estimates for the error component
- Comparable results can be found in:
  - Ramos (2003) for the UK using BHPS
  - Lillard and Willis (1978) using the American PSID panel
  - Borella and Coda-Moscarola (2009) using INPS administrative data for Italy

Panel estimates

IT-SILC results, net log earnings

	Men, not graduated, employees	Women, not graduated employees	Men, graduated, employees	Women, graduated, employees	Graduated, self-employed	Men, not graduated, self-employed	Women, not graduated, self-employed
$\sigma^2_\mu/\sigma^2_\varepsilon$	0.627	0.543	0.728	0.642	0.550	0.539	0.538
ρ	0.322	0.346	0.249	0.282	0.123	0.171	0.237
δ(2,1)	0.747	0.701	0.795	0.743	0.605	0.618	0.647
$\sigma^2_\mu+\sigma^2_\xi$	0.092	0.108	0.138	0.127	0.373	0.295	0.339

Results in Borella and Coda Moscarola (2009), gross log earnings

	Males			Females		
	Blue collar	White collar	Self-employed	Blue collar	White collar	Self-employed
$\sigma^2_\mu/\sigma^2_\varepsilon$	0.750	0.870	0.407	0.748	0.799	0.353
ρ	0.432	0.529	0.165	0.419	0.440	0.070
δ(2,1)	0.858	0.869	0.543	0.787	0.813	0.465
$\sigma^2_\mu+\sigma^2_\xi$	0.078	0.129	0.170	0.147	0.162	0.148

Simulation

- If the individual is employed in the initial sample
  1. his/her log earnings are split into the deterministic component and a residual
  2. in the following years, the deterministic components evolves following the change in observables
  3. we simulate the evolution of the residual using the autoregressive component
- When an individual starts his/her first job
  1. the deterministic component is predicted using cross-sectional or panel estimates *(and multiply by 12)*
  2. we impute a residual assuming that the errors are normally distributed
  - Aggregate increase in productivity are assigned to all workers in each simulation period

Simulating the evolution of unobserved individual effect

- Following Pudney (1992)

$$\varepsilon_{it} = \mu_i + \xi_{it}$$

$$\xi_{it} = \rho\xi_{it-1} + \omega_{it}, \omega_{it} \sim N(0, \sigma_\omega^2), |\rho| < 1$$

$$\mu_i \sim N(0, \sigma_\mu^2)$$

- the autoregressive component implies that:

$$cov(\xi_{it}, \xi_{it-\mathrm{k}} | 1, z_i) = \rho^k \sigma_\xi^2$$

- using normality:

$$E(y_{is} | y_{it}, z_{it}, z_{is}) = z_{is}\beta + \delta(s,t)(y_{it} - z_{it}\beta)$$

- where:

$$\delta(s,t) = \frac{\sigma_\mu^2 + \rho^{|s-t|}\sigma_\xi^2}{\sigma_\mu^2 + \sigma_\xi^2}$$

The increase in productivity

- Y is multiplied by a factor $(1+\tau_s)$ allowing the individual earning in s to be linked to the medium-long term productivity growth
- But… the demographic evolution and the increase in the stock of human capital in the coming decades increase the average earning level, since age and education have a positive effect on average labour earnings
- we avoid over/under-estimations of earnings growth rates, using:

$$\tau_s = m_s - \left(\frac{E(y_s)}{E(y_{s-1})} - 1\right)$$

official productivity growth

Endogenous growth generated by the model

REFERENCES:

- Bækgaard, H. (2002). Modelling the Dynamics of the Distribution of Earned Income. NATSEM Technical Paper, no. 24.
- Baltagi, B., & Li, Q. (1991). A transformation that will circumvent the problem of autocorrelation in an error-component model. Journal of Econometrics, 48, 385-393.
- Blanchet, D., Buffeteau, S., Crenner, E., & Le Minez, S. (2010). Le modèle de microsimulation Destinie 2: principales caractéristiques et premiers résultats. Institut National de la Statistique et des Études Économiques Document de travail, G 2010 / 13.
- Borella, M. (2004). The Error Structure of Earnings: an Analysis on Italian Longitudinal Data. Rivista Internazionale di Scienze Economiche e Commerciali, 51(3), 405-422.
- Borella, M., & Coda Moscarola, F. (2006). Distributive Properties of Pensions Systems: a Simulation of the Italian Transition from Defined Benefit to Defined Contribution. Giornale degli Economisti e Annali di Economia, 65(1), 95-125.
- Borella, M., & Coda Moscarola, F. (2009). Microsimulation of pension reforms: behavioural versus nonbehavioural approach. Center for Research on Pensions and Welfare Policies Working Paper, 86/09.
- Dekkers, G., Desmet, R., & De Vil, G. (2010). The long-term adequacy of the Belgian public pension system: An analysis based on the MIDAS model. Federal Planning Bureau Working Paper, 10/10.
- Emmerson, C., Reed, H., & Shephard, A. (2004). An Assessment of PenSim2. The Institute of Fiscal Studies Working Paper, WP04/21.
- Favreault, M., & Caldweel, S. (1998). Assessing Distributional Impacts of Social Security by Cohort, Gender, Family Status, Lifetime Earnings and Race Using Dynamic Microsimulation. Paper presented at the annual meeting of the Association for Public Policy and Management, New York City.
- Giarda, E. (2007). The worsening of wage expectations in Italy: a study based on administrative data. International Journal of Manpower, 29, 64-87.
- Heckman, J., Lochner, L., & Todd, P. (2003). Fifty Years of Mincer Earnings Regressions. IZA Discussion Paper, No. 775.
- Lillard, L. A., & Willis, R. J. (1978). Dynamic Aspects of Earning Mobility. Econometrica, 46(5), 958-1012.

REFERENCES (cont.):

- Maitino, M., & Sciclone, N. (2009). Assessing the Implications of Population Ageing on Tuscan Wellbeing: a Microsimulation Approach. unpublished, available online at Mincer, J. (1974). Schooling, Experience, and Earnings. Columbia University Press.
- Pudney, S. (1992). Dynamic Simulation of Pensioner's Incomes: Methodological Issues and a Model Design for Great Britain. Dept. of Applied Economics Microsimulation Unit Discussion Paper, No. MSPMU 9201.
- Ramos, X. (2003). The Covariance Structure of Earnings in Great Britain: 1991-1999. Economica, 70(278), 353-374.
- Ramsey, J. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. Journal of the Royal Statistical Society. Series B (Methodological), 31, 350-371.
- Santos Silva, J., & Tenreyro, S. (2006). The Log of Gravity. The Review of Economics and Statistics, 88(4), 641-658.
- Wooldridge, J. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts: The MIT Press.