UNIVERSITÀ DEGLI STUDI DI TRENTO

DIPARTIMENTO DI ECONOMIA

A Monte Carlo EM Algorithm for the Estimation of a Logistic Auto-logistic Model with Missing Data

Marco Bee
Giuseppe Espa

Discussion Paper No. 1, 2008

The Discussion Paper series provides a means for circulating preliminary research results by staff of or visitors to the Department. Its purpose is to stimulate discussion prior to the publication of papers.

Requests for copies of Discussion Papers and address changes should be sent to:

Dott. Luciano Andreozzi
E.mail luciano.andreozzi@economia.unitn.it
Dipartimento di Economia
Università degli Studi di Trento
Via Inama 5
38100 TRENTO ITALIA

A Monte Carlo EM Algorithm for the Estimation of a Logistic Auto-logistic Model with Missing Data

Marco Bee

Department of Economics, University of Trento

Giuseppe Espa

Department of Economics, University of Trento

Abstract

This paper proposes an algorithm for the estimation of the parameters of a Logistic Auto-logistic Model when some values of the target variable are missing at random but the auxiliary information is known for the same areas. First, we derive a Monte Carlo EM algorithm in the setup of maximum pseudo-likelihood estimation; given the analytical intractability of the conditional expectation of the complete pseudo-likelihood function, we implement the E-step by means of Monte Carlo simulation. Second, we give an example using a simulated dataset. Finally, a comparison with the standard non-missing data case shows that the algorithm gives consistent results.

JEL Classification: C13, C15, C51

Keywords: Spatial Missing Data, Monte Carlo EM Algorithm, Logistic Auto-logistic Model, Pseudo-Likelihood.

1 Introduction

The missing-data problem has a long history in statistics; since the early 1970's the literature has grown quite rapidly, mainly because of the widespread availability of cheap computing power; see Little and Rubin (2002) for a review. In the framework of spatial statistics, however, most techniques have to be modified in order to take care of the features of spatially dependent data; several tools have been developed, according to the estimation methodology, the nature of missing data and the goals of the analysis. While referring the reader to Haining (2003, sect. 4.4.1) for details concerning the various approaches, in this paper we concentrate on the missing-data problem in the setup of maximum likelihood estimation of the Logistic Auto-logistic Model.

The missing-data problem in this framework was treated by several authors: see Haining (2003, sect. 4.4.1) and the references therein. Although their approach is conceptually similar to the one developed in the present paper, it is mainly based on multivariate normality, and can possibly be extended only to cases where the likelihood function can be written in closed form. The latter requirement is not satisfied by the Logistic Auto-logistic Model (LAM) considered here, so that different technical solutions are needed. Considering that the likelihood function for the LAM is not available in closed form, in this paper we will use the Pseudo-Likelihood approach: more precisely, we will maximize the Pseudo-Likelihood function by means of the EM algorithm.

The rest of the paper is organized as follows. Section 2 reviews the Logistic Auto-logistic Model and develops the estimation procedure with missing data in the target variable; section 3 gives an example of the mechanics of the algorithm, based on simulated data, with different percentages of missing data. Section 4 concludes and outlines possible directions for future research.

2 The estimation procedure

From now on, for notational simplicity and without any loss of generality, we focus on regular grids. A spatial random process $\tilde{Y}$ on a $(k \times k)$ grid $\mathcal{D} \subset I\!\!R^2$ is naturally described by a matrix of random processes. Each element of $\tilde{\boldsymbol{Y}}$ is then characterized by two indexes: $\tilde{Y}_{ij}$ $(i, j = 1, \ldots, k)$. However, it is common to stack the columns of the random field $\tilde{\boldsymbol{Y}}$ on top of each other; in this manner, the data generating process becomes a random vector $\boldsymbol{Y} = \mathrm{vec}(\tilde{\boldsymbol{Y}})$ and a single index is sufficient to identify each dependent variable Y_i $(i = 1, \ldots, k^2)$. In the following we will adhere to this convention.

Thus, let Y_i be a Bernoulli random variable: $Y_i = 1$ if a success is observed in area i and 0 otherwise, i.e. $Y_i = 1$ with probability π_i $(i = 1, \ldots, k^2)$, where π_i is given in definition 1. The data generating process assumed in this paper is called Logistic Auto-logistic Model (LAM), defined as follows.

Definition 1 *A random field $\boldsymbol{Y}(\boldsymbol{s})$, $\boldsymbol{s} \in \mathcal{D}$, $\mathcal{D} \subset I\!\!R^2$ is called a Logistic Auto-Logistic Model (LAM) if the conditional distribution of Y_i given Y_j, $j \in C(i)$, is given by*

$$\pi_i = P(Y_i = y_i | Y_j = y_j,\ j \neq i,\ j \in C\{i\}) = \frac{\exp\{y_i(\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{x}_i + \beta \sum_{j \in C\{i\}} y_j)\}}{1 + \exp\{\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{x}_i + \beta \sum_{j \in C\{i\}} y_j\}}, \qquad (1)$$

where $y_i \in \{0, 1\}$, $\boldsymbol{x}_i = (x_1, \ldots, x_m)'$ is the vector containing the m auxiliary variables for the i-th area and $C(i)$ is the neighborhood set of cell i.

A review of the most common neighborhood criteria and of the corresponding identifications of the neighborhood set can be found in Haining (2003, pag. 80-85). Notice also that the LAM is an extension of the well known Auto-logistic model (Haining 2003, sect. 9.1.2 or Cressie 1993, pag. 423); this approach considers both the logistic (covariates) and autologistic (autocovariates) components (Arbia 2006, pag. 124-126).

Unfortunately, direct application of the maximum likelihood method of estimation to the LAM is usually impossible, because the Y_i's are dependent and their joint distribution is not computable: see Strauss (1992) for details. However, a solution which combines simplicity of implementation and good statistical properties consists in treating the observations as if they were independent; this approach allows to obtain a "likelihood-type" function as the product of the conditional densities:

$$PL = \prod P(y_i|\text{all other } y_j\text{'s}). \tag{2}$$

Formula (2), considered as a function of the parameters, is called *Pseudo-Likelihood* (PL), and is then maximized with respect to the parameters, as in the ML approach. Notice that, in this setup, (2) is just the likelihood function of a logistic model, so that maximization can be performed by means of standard techniques. This methodology, known as *Maximum Pseudo-Likelihood* (MPL), was first introduced by Besag (1975, 1977; see also Strauss and Ikeda 1990, Arnold and Strauss 1991, Strauss 1992), who has also shown that the estimators obtained enjoy all the properties of standard MLE's, with the exception of efficiency. However, inefficiency is usually negligible, and compensated by huge computational advantages.

The problem studied here consists in the fact that for some cells Y_i is missing, whereas the auxiliary information is available; it is also assumed that the observations are missing at random. In such a setup it does not make sense to discard these observations, because the auxiliary information would be discarded as well.

In cases when the maximum likelihood estimation method can be applied, the most common way of tackling the missing data problem is based on the EM algorithm; this technique, developed by Dempster *et al.* (1977) and specifically devoted to likelihood maximization with missing data, has several desirable properties and has been used for the solution of a variety of problems; referring the reader to McLachlan and Krishnan (1996) for a thorough treatment of the algorithm, here we limit ourselves to recall that the algorithm iterates until convergence two steps, called E (Expectation) and M (Maximization). The first one is given by the conditional expectation (given the observed data and the current estimated parameters values) of the so called complete likelihood function, i.e. the hypothetical likelihood function that would be available if the missing data were observed; the M-step consists in maximizing this conditional expectation.

To implement the algorithm, we would thus need to write down the complete and observed (incomplete) likelihood functions, but, as mentioned above, the observations are not independent, so that even when no observations are missing, we cannot write the likelihood function. We therefore resort to the MPL estimation methodology and work with the complete and incomplete pseudo-likelihood functions:

$$PL_c = \prod_{i=1}^{k^2} \pi_i^{y_i}(1-\pi_i)^{1-y_i}, \quad PL_{obs} = \prod_{i=1}^{k_{obs}} \pi_i^{y_i}(1-\pi_i)^{1-y_i},$$

where π_i is given by (1) and k_{obs} is the number of observed data. As can be seen, in the complete-data case estimation would be based on standard MPL methods, which implies that we know how to perform the M-step: we just have to estimate a logistic model as if the observations were independent.

Denoting with $\boldsymbol{Y}_{mis}$ the (k_{mis})-vector of the missing data ($k_{mis} = k^2 - k_{obs}$), with $\boldsymbol{Y}_{obs}$ $(k_{obs} \times 1)$ the observed data and with $\boldsymbol{Y} = (\boldsymbol{Y}'_{mis}, \boldsymbol{Y}'_{obs})'$ $(k^2 \times 1)$ the complete data, the E-step requires to compute the conditional expectation of the complete pseudo log-likelihood function $pl_c = \log(PL_c)$, given the current values of the parameters and the observed data: $E_{\boldsymbol{\pi}^{(t)}}\{pl_c(\boldsymbol{Y}|\boldsymbol{y}_{obs}, \boldsymbol{\pi}^{(t)})\}$, where $E_{\boldsymbol{\pi}^{(t)}}$ denotes expectation with respect to the current (at the t-th iteration) distribution of the complete data.

Unfortunately, pl_c is not linear in the missing data, so that its conditional expectation cannot be simply obtained by computing the conditional expectation of the missing data $E_{\boldsymbol{\pi}^{(t)}}\{\boldsymbol{Y}_{mis}|\boldsymbol{y}_{obs}, \boldsymbol{\pi}^{(t)}\}$ and plugging it into the complete pseudo log-likelihood function. In order to compute it, we start by writing down explicitly the conditional expectation:

$$\begin{aligned}
&E_{\boldsymbol{\pi}^{(t)}}\{pl_c(\boldsymbol{Y}|\boldsymbol{y}_{obs}, \boldsymbol{\pi}^{(t)})\} = \\
=&E_{\boldsymbol{\pi}^{(t)}}\left\{\sum_{i=1}^{k^2}[Y_i \log(\pi_i) + (1-Y_i)\log(1-\pi_i)] \,\Big|\, \boldsymbol{y}_{obs}, \boldsymbol{\pi}^{(t)}\right\} = \\
=&E_{\boldsymbol{\pi}^{(t)}}\left\{\sum_{i=1}^{k^2}[Y_i(\boldsymbol{\alpha}'\boldsymbol{w} + \gamma \sum_{j \in C(i)} Y_j - \log(1 + e^{\boldsymbol{\alpha}'\boldsymbol{w} + \gamma \sum_{j\in C(i)} Y_j})) - \right. \\
&\left. - (1-Y_i)\log(1 + e^{\boldsymbol{\alpha}'\boldsymbol{w} + \gamma \sum_{j\in C(i)} Y_j})]\right\}. \qquad (3)
\end{aligned}$$

It is clear that (3) is analytically intractable. In such cases, the preferred solution consists in performing the E-step by Monte Carlo simulation (Monte Carlo EM - MCEM: Wei and Tanner 1990; see also McLachlan and Krishnan 1996, sect. 6.2); in the present framework, this requires a large number (say B) of simulations of the random field $\boldsymbol{Y}$ and completes the E-step.

Using a terminology similar to Casella and Robert (2004, pag. 183), the M-step consists in maximizing the "approximate complete-data pseudo log-likelihood" pl_c. Although this is based on standard logistic regression techniques, we are now going to show that it is not completely trivial.

Example. To begin with, we illustrate by means of a toy example the mechanics of the algorithm proposed here. Suppose $k = 3$ and $C\{i\} = \{j, m\}$, with $i, j, m = 1, 2, 3$ and

$i \neq j \neq m$, that is, every area has two neighbors. Moreover, Y_1 and Y_2 are missing, Y_3 is observed.

MC E-step. In order to implement the MC E-step, we have to simulate B observations from the distributions of $(Y_1|Y_2, Y_3; \pi^{(0)})$ and $(Y_2|Y_1, Y_3; \pi^{(0)})$ and compute $pl_c(Y_1|Y_2, Y_3)$ and $pl_c(Y_2|Y_1, Y_3)$. Thus, the estimated expectation is given by:

$$
\begin{aligned}
&\hat{E}_{\boldsymbol{\pi}^{(0)}}\{pl_c(\boldsymbol{Y}|\boldsymbol{y}_{obs}, \boldsymbol{\pi})\} = \\
=&\frac{1}{B}\sum_{j=1}^{B}\left[y_{1j}\log\left(\pi_1^{(0)}\right) + (1-y_{1j})\log\left(1-\pi_1^{(0)}\right) + y_{2j}\log\left(\pi_2^{(0)}\right) + \right. \\
&\left. + (1-y_{2j})\log\left(1-\pi_2^{(0)}\right) + y_3\log\left(\pi_3^{(0)}\right) + (1-y_3)\log\left(1-\pi_3^{(0)}\right)\right],
\end{aligned}
$$

where $y_{i1}, \ldots, y_{iB}$ are the B observations simulated from $(Y_i|Y_{C(i)}, \pi^{(0)})$, $i = 1, 2$, and $Y_{C(i)}$ contains all the observations in $C(i)$. This means that we have B observations with $\sum_{j=1}^{B} y_{1j}$ successes in cell 1, B observations with $\sum_{j=1}^{B} y_{2j}$ successes in cell 2, and one observation with $y_3 \in \{0, 1\}$ successes in cell 3. When estimating π_3 we have to include in the logistic model, as auxiliary variables, the number of successes in cell 1 and 2. So the idea is to consider B observations also for cell 3, always with $y_3 \in \{0, 1\}$ successes, but with a different number of events in the neighborhood set, according to the results of the simulation. Therefore $\pi_3^{(0)}$ is given by the autologistic specification (1) with $\sum_{l\in C(3)} y_l = y_1 + y_2$. The simulation of Y_1 and Y_2, which completes the E-step, proceeds as follows:

1. simulate $y_{11}, \ldots, y_{1B} \sim \text{Bin}\left(1; \pi_1^{(0)}\right)$;
2. simulate $y_{21}, \ldots, y_{2B} \sim \text{Bin}\left(1; \pi_2^{(0)}\right)$.

M-step. In this way we get the y values needed for the first M-step, which consists in updating the estimate of $\boldsymbol{\pi}$ using the simulated values $y_{11}, \ldots, y_{1B}$ and $y_{21}, \ldots, y_{2B}$ and the observed value y_3, obtaining $\boldsymbol{\pi}^{(1)}$. This is the usual logistic regression estimation procedure, but notice that $y_{11}, \ldots, y_{1B}$ have the same auxiliary variables except the last one, i.e. the autocovariate $\sum_{l\in C(1)} y_{lj}$, which is different for each $j = 1, \ldots, B$; the same holds for y_2 and y_3.

At the t-th iteration, the steps above remain the same; only, $\boldsymbol{\pi}^{(0)}$ has to be replaced by $\boldsymbol{\pi}^{(t)}$ and the simulation in the E-step is based on the current values of the parameters, as estimated in the t-th M-step. ∎

Unfortunately, the monotonocity property of the EM algorithm does not hold for the MCEM algorithm (McLachlan and Krishnan 1996, pag. 216). This is one of the reasons why monitoring convergence is more difficult than in the standard EM setup, where the algorithm is usually stopped when the criterion $\max|\theta_i^{(t+1)} - \theta_i^{(t)}| < \epsilon$ ($i = 1, \ldots, p$, where p is the number of parameters) is satisfied; using simulated data at each iteration, it is unlikely that such a condition is met for the values of ϵ employed in the EM algorithm, unless B is huge. As pointed out by Wei and Tanner (1990), a possible solution consists in monitoring the paths of $\theta_i^{(t)}$ $(i = 1, \ldots, p)$ obtained with a small value of B; when they look reasonably stable, one may use a larger B for "fine tuning" purposes, i.e. in order to get a more precise estimate. With these premises, we are finally ready to give a formal description of the general formulation of the algorithm.

Algorithm 1 *In the LAM setup of Definition 1, assume that k_{mis} observations of the response variable Y are missing. The t-th iteration of the EM algorithm for the estimation of the parameters of (1) is given by the following two steps:*

- *E-step: for each $i \in \{1, \ldots, k_{mis}\}$, simulate $y_{i1}, \ldots, y_{iB} \sim \mathrm{Bin}\left(1; \pi_i^{(t)}\right)$;*
- *M-step: perform a standard MPL logistic estimation using the observed response variables y_i $(i = 1, \ldots, k_{obs})$, the simulated response variables y_{ij} $(i = 1, \ldots, k_{mis}$, $j = 1, \ldots, B)$, and the auxiliary variables $\boldsymbol{x}$; the last auxiliary variable, given by $\sum_{l \in C(i)} y_l$, is formed by summing the observed (when available) or the simulated values.*

Notice that it would be easier to use the quantity $\bar{y}_i = (1/B) \sum_{j=1}^{B} y_{ij}$ $(i = 1, \ldots, k_{mis})$ for logistic estimation. However, this strategy is wrong, because it corresponds to replacing the missing data with their conditional expectation (computed by means of MC simulation). Were the complete pseudo log-likelihood function linear in the missing data (which is not the case here), this would be equivalent to computing the conditional expectation of the complete pseudo log-likelihood function.

3 An example with simulated data

In this section we apply to simulated data the methodology developed above. To this aim, we simulated a 10×10 LAM; moreover, for each region we assume the existence of two auxiliary variables with coefficients $\alpha_1 = 0.2$ and $\alpha_2 = 0.3$; the parameter γ was put equal to 0.02. The auxiliary variables were simulated from an auto-normal process $\boldsymbol{X} \sim N_{k^2}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma} = (\boldsymbol{I}_{k^2} - \boldsymbol{S})^{-1}$, $\boldsymbol{S} = \rho \boldsymbol{C}$ and $\boldsymbol{C}$ the contiguity matrix (Cressie 1993, sect. 6.3.2). $\boldsymbol{C}$ is such that the neighbors of $Y_{i,j}$ are $Y_{i-1,j}$, $Y_{i+1,j}$, $Y_{i,j-1}$ and $Y_{i,j+1}$, $\boldsymbol{\mu} = \boldsymbol{0} \in \mathbb{R}^{k^2}$ and $\rho = 0.2$. After simulating the target variable, we chose randomly a percentage c of the simulated Y_i's and treated them as missing values.

Table 1 shows the results obtained with the approach proposed in this paper; for comparison purposes, we also performed a non-missing data pseudo-likelihood estimation using the actual values of the missing observations.

Table 1. Parameters estimates with complete and incomplete data

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\gamma}$
No missing data (standard LAM)	0.314	0.263	0.044
Missing data ($c = 10\%$; MCEM algorithm)	0.326	0.255	0.039
Missing data ($c = 15\%$; MCEM algorithm)	0.262	0.264	0.036
Missing data ($c = 20\%$; MCEM algorithm)	0.278	0.231	0.030

Although a measure of the asymptotic standard error of the estimators can be obtained (from standard ML estimation of the logistic model) when performing the M-step, we decided not to report it here. The reason is twofold: first, observations are not independent, and therefore the asymptotic theory of ML estimation does not hold. Second, we are using the EM algorithm, with some observations simulated at each E-step, so that the increased variability introduced by this procedure has to be taken into account. Thus, the only correct way of approximating standard errors should probably be based on simulation techniques.

Following the suggestion by Wei and Tanner (1990), we use a larger value of B for the last iterations: in particular, we noticed that the parameters estimates become stable after few iterations, so that we set the maximum number of iterations equal to 20, putting $B = 100$ for the first 15 iterations and $B = 1000$ for the last 5 iterations.

It can be seen from table 1 that, even in presence of a non-negligible percentage of missing data, the estimates are similar to those obtained when no data are missing; this is a first confirmation of the good performance of the algorithm. Of course, a more thorough comparison could be performed by means of a Monte Carlo experiment, i.e. by simulating the random field a large number of times and studying the empirical distribution of the estimates produced by the algorithm at each replication; this procedure would also allow us to get estimates of the standard errors of the estimators. However, due to space constraints, we do not perform this experiment here, but postpone it to a future investigation.

As k gets large, the computational burden increases, because the number of observations used at each EM iteration for logistic estimation in the M-step is equal to the number of replications of the Monte Carlo procedure in the E-step times the number of regions, i.e. $B \times k^2$. This is an additional reason for paying particular attention to the choice of B, and possibly increase it only in the last few iterations.

4 Conclusions

In this paper we have proposed a Monte Carlo EM algorithm for the estimation of a LAM with randomly missing observations in the target variable. The algorithm is developed in the pseudo-likelihood setup commonly employed for the LAM, and can be seen as a natural extension of the usual EM implementations for the maximization of the likelihood function. After deriving the E- and M-step, we showed that the E-step cannot be performed in closed form but requires Monte Carlo simulation. Finally we compared, by means of simulated data, the performance of the MCEM algorithm (in presence of missing data) and of the standard pseudo-likelihood approach (with no missing data); the results confirm that the algorithm works well.

Several issues remain open to further investigation. In particular, no attempt has been made here to estimate standard errors and assess other properties of the estimators; as it seems unlikely to get analytical solutions to this problem, simulation techniques will probably be necessary. Moreover, as far as we know, the (MC)EM algorithm has never been used in a maximum pseudo-likelihood setup, so that the extension of inferential results known in the standard maximum likelihood framework should not be taken for granted. In addition, it would be very important for practical applications to study the behavior of the algorithm as the percentage of missing data increases; from standard EM theory, we expect the convergence to become slower. Furthermore, our guess is that there exists a threshold such that the algorithm does not converge at all when the percentage of missing data is larger than this threshold. Finally, the methodology presented is likely to be readily extended to the Binomial Auto-binomial model (Cressie 1993, pag. 431). We already performed some numerical experiments and the results are encouraging; however, a detailed analysis of the present approach in the Binomial Auto-binomial setup will be performed in a future work.

References

Arbia, G.: Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence, New York: Springer 2006.

Arnold, B.C., Strauss, D.: Pseudo likelihood estimation: some examples, Sankhya B, **53**, 233-243 (1991).

Besag, J.: Statistical analysis of non-lattice data, The Statistician, **24**, 179-195 (1975).

Besag, J.: Efficiency of pseudo-likelihood estimators for simple Gaussian fields, Biometrika, **64**, 616-618 (1977).

Casella, G., Robert, C.P.: Monte Carlo Statistical Methods, second edition, New York: Springer 2004.

Cressie, N.: Statistics for Spatial Data, rev. ed., New York: Wiley 1993.

Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM Algorithm (with Discussion), Journal of the Royal Statistical Society Series B, **39**, 1-38 (1977).

Haining, T.: Spatial Data Analysis: Theory and Practice, Cambridge: Cambridge University Press 2003.

Little, R.J.A., Rubin, D.B.: Statistical Analysis with Missing Data, second edition, New York: Wiley 2002.

McLachlan, G.J., Krishnan, T.: The EM Algorithm and Extensions, New York: Wiley 1996.

Strauss, D.: The Many Faces of Logistic Regression, The American Statistician, **46**, 321-327 (1992).

Strauss, D., Ikeda, M.: Pseudolikelihood estimation for social networks, Journal of the American Statistical Association, **85**, 204-212 (1990).

Wei, G.C.G., Tanner, M.A.: A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithm, Journal of the American Statistical Association, **85**, 699-704 (1990).

Elenco dei papers del Dipartimento di Economia

2000.1 *A two-sector model of the effects of wage compression on unemployment and industry distribution of employment*, by Luigi Bonatti

2000.2 *From Kuwait to Kosovo: What have we learned? Reflections on globalization and peace,* by Roberto Tamborini

2000.3 *Metodo e valutazione in economia. Dall'apriorismo a Friedman* , by Matteo Motterlini

2000.4 *Under tertiarisation and unemployment.* by Maurizio Pugno

2001.1 *Growth and Monetary Rules in a Model with Competitive Labor Markets,* by Luigi Bonatti.

2001.2 *Profit Versus Non-Profit Firms in the Service Sector: an Analysis of the Employment and Welfare Implications,* by Luigi Bonatti, Carlo Borzaga and Luigi Mittone.

2001.3 *Statistical Economic Approach to Mixed Stock-Flows Dynamic Models in Macroeconomics,* by Bernardo Maggi and Giuseppe Espa.

2001.4 *The monetary transmission mechanism in Italy: The credit channel and a missing ring,* by Riccardo Fiorentini and Roberto Tamborini.

2001.5 *Vat evasion: an experimental approach,* by Luigi Mittone

2001.6 *Decomposability and Modularity of Economic Interactions,* by Luigi Marengo, Corrado Pasquali and Marco Valente.

2001.7 *Unbalanced Growth and Women's Homework,* by Maurizio Pugno

2002.1 *The Underground Economy and the Underdevelopment Trap,* by Maria Rosaria Carillo and Maurizio Pugno.

2002.2 *Interregional Income Redistribution and Convergence in a Model with Perfect Capital Mobility and Unionized Labor Markets,* by Luigi Bonatti.

2002.3 *Firms' bankruptcy and turnover in a macroeconomy,* by Marco Bee, Giuseppe Espa and Roberto Tamborini.

2002.4 *One "monetary giant" with many "fiscal dwarfs": the efficiency of macroeconomic stabilization policies in the European Monetary Union,* by Roberto Tamborini.

2002.5 *The Boom that never was? Latin American Loans in London 1822-1825,* by Giorgio Fodor.

2002.6 *L'economia senza banditore di Axel Leijonhufvud: le 'forze oscure del tempo e dell'ignoranza' e la complessità del coordinamento,* by Elisabetta De Antoni.

2002.7 *Why is Trade between the European Union and the Transition Economies Vertical?,* by Hubert Gabrisch and Maria Luigia Segnana.

2003.1 *The service paradox and endogenous economic gorwth,* by Maurizio Pugno.

2003.2 *Mappe di probabilità di sito archeologico: un passo avanti,* di Giuseppe Espa, Roberto Benedetti, Anna De Meo e Salvatore Espa.
(*Probability maps of archaeological site location: one step beyond,* by Giuseppe Espa, Roberto Benedetti, Anna De Meo and Salvatore Espa).

2003.3 *The Long Swings in Economic Understianding,* by Axel Leijonhufvud.

2003.4 *Dinamica strutturale e occupazione nei servizi,* di Giulia Felice.

2003.5 *The Desirable Organizational Structure for Evolutionary Firms in Static Landscapes,* by Nicolás Garrido.

2003.6 *The Financial Markets and Wealth Effects on Consumption* An Experimental Analysis, by Matteo Ploner.

2003.7 *Essays on Computable Economics, Methodology and the Philosophy of Science,* by Kumaraswamy Velupillai.

2003.8 *Economics and the Complexity Vision: Chimerical Partners or Elysian Adventurers?*, by Kumaraswamy Velupillai.

2003.9 *Contratto d'area cooperativo contro il rischio sistemico di produzione in agricoltura,* di Luciano Pilati e Vasco Boatto.

2003.10 *Il contratto della docenza universitaria. Un problema multi-tasking*, di Roberto Tamborini.

2004.1 *Razionalità e motivazioni affettive: nuove idee dalla neurobiologia e psichiatria per la teoria economica?* di Maurizio Pugno.
(*Rationality and affective motivations: new ideas from neurobiology and psychiatry for economic theory?* by Maurizio Pugno.

2004.2 *The economic consequences of Mr. G. W. Bush's foreign policy. Can th US afford it?* by Roberto Tamborini

2004.3 *Fighting Poverty as a Worldwide Goal* by Rubens Ricupero

2004.4 *Commodity Prices and Debt Sustainability* by Christopher L. Gilbert and Alexandra Tabova

2004.5 *A Primer on the Tools and Concepts of Computable Economics* by K. Vela Velupillai

2004.6 *The Unreasonable Ineffectiveness of Mathematics in Economics* by Vela K. Velupillai

2004.7 *Hicksian Visions and Vignettes on (Non-Linear) Trade Cycle Theories* by Vela K. Velupillai

2004.8 *Trade, inequality and pro-poor growth: Two perspectives, one message?* By Gabriella Berloffa and Maria Luigia Segnana

2004.9 *Worker involvement in entrepreneurial nonprofit organizations. Toward a new assessment of workers? Perceived satisfaction and fairness* by Carlo Borzaga and Ermanno Tortia.

2004.10 *A Social Contract Account for CSR as Extended Model of Corporate Governance (Part I): Rational Bargaining and Justification* by Lorenzo Sacconi

2004.11 *A Social Contract Account for CSR as Extended Model of Corporate Governance (Part II): Compliance, Reputation and Reciprocity* by Lorenzo Sacconi

2004.12 *A Fuzzy Logic and Default Reasoning Model of Social Norm and Equilibrium Selection in Games under Unforeseen Contingencies* by Lorenzo Sacconi and Stefano Moretti

2004.13 *The Constitution of the Not-For-Profit Organisation: Reciprocal Conformity to Morality* by Gianluca Grimalda and Lorenzo Sacconi

2005.1 *The happiness paradox: a formal explanation from psycho-economics* by Maurizio Pugno

2005.2 *Euro Bonds: in Search of Financial Spillovers* by Stefano Schiavo

2005.3 *On Maximum Likelihood Estimation of Operational Loss Distributions* by Marco Bee

2005.4 *An enclave-led model growth: the structural problem of informality persistence in Latin America* by Mario Cimoli, Annalisa Primi and Maurizio Pugno

2005.5 *A tree-based approach to forming strata in multipurpose business surveys,* Roberto Benedetti, Giuseppe Espa and Giovanni Lafratta.

2005.6 *Price Discovery in the Aluminium Market* by Isabel Figuerola-Ferretti and Christopher L. Gilbert.

2005.7 *How is Futures Trading Affected by the Move to a Computerized Trading System? Lessons from the LIFFE FTSE 100 Contract* by Christopher L. Gilbert and Herbert A. Rijken.

2005.8 *Can We Link Concessional Debt Service to Commodity Prices?* By Christopher L. Gilbert and Alexandra Tabova

2005.9 *On the feasibility and desirability of GDP-indexed concessional lending* by Alexandra Tabova.

2005.10 *Un modello finanziario di breve periodo per il settore statale italiano: l'analisi relativa al contesto pre-unione monetaria* by Bernardo Maggi e Giuseppe Espa.

2005.11 *Why does money matter? A structural analysis of monetary policy, credit and aggregate supply effects in Italy*, Giuliana Passamani and Roberto Tamborini.

2005.12 *Conformity and Reciprocity in the "Exclusion Game": an Experimental Investigation* by Lorenzo Sacconi and Marco Faillo.

2005.13 *The Foundations of Computable General Equilibrium Theory*, by K. Vela Velupillai.

2005.14 *The Impossibility of an Effective Theory of Policy in a Complex Economy*, by K. Vela Velupillai.

2005.15 *Morishima's Nonlinear Model of the Cycle: Simplifications and Generalizations,* by K. Vela Velupillai.

2005.16 Using and Producing *Ideas* in Computable Endogenous Growth, by K. Vela Velupillai.

2005.17 *From Planning to Mature: on the Determinants of Open Source Take Off* by Stefano Comino, Fabio M. Manenti and Maria Laura Parisi.

2005.18 *Capabilities, the self, and well-being: a research in psycho-economics,* by Maurizio Pugno.

2005.19 Fiscal and monetary policy, unfortunate events, and the SGP arithmetics. Evidence from a *growth-gap* model, by Edoardo Gaffeo, Giuliana Passamani and Roberto Tamborini

2005.20 *Semiparametric Evidence on the Long-Run Effects of Inflation on Growth,* by Andrea Vaona and Stefano Schiavo.

2006.1 *On the role of public policies supporting Free/Open Source Software. An European perspective,* by Stefano Comino, Fabio M. Manenti and Alessandro Rossi.

2006.2 *Back to Wicksell? In search of the foundations of practical monetary policy,* by Roberto Tamborini

2006.3 *The uses of the past,* by Axel Leijonhufvud

2006.4 *Worker Satisfaction and Perceived Fairness: Result of a Survey in Public, and Non-profit Organizations,* by Ermanno Tortia

2006.5 *Value Chain Analysis and Market Power in Commodity Processing with Application to the Cocoa and Coffee Sectors,* by Christopher L. Gilbert

2006.6 *Macroeconomic Fluctuations and the Firms' Rate of Growth Distribution: Evidence from UK and US Quoted Companies,* by Emiliano Santoro

2006.7 *Heterogeneity and Learning in Inflation Expectation Formation: An Empirical Assessment,* by Damjan Pfajfar and Emiliano Santoro

2006.8 Good *Law & Economics* needs suitable microeconomic models: the case against the application of standard agency models: the case against the application of standard agency models to the professions, by Lorenzo Sacconi

2006.9 *Monetary policy through the "credit-cost channel". Italy and Germany,* by Giuliana Passamani and Roberto Tamborini

2007.1 *The Asymptotic Loss Distribution in a Fat-Tailed Factor Model of Portfolio Credit Risk,* by Marco Bee

2007.2 *Sraffa?s Mathematical Economics – A Constructive Interpretation,* by Kumaraswamy Velupillai

2007.3 *Variations on the Theme of Conning in Mathematical Economics,* by Kumaraswamy Velupillai

2007.4 *Norm Compliance: the Contribution of Behavioral Economics Models,* by Marco Faillo and Lorenzo Sacconi

2007.5 *A class of spatial econometric methods in the empirical analysis of clusters of firms in the space,* by Giuseppe Arbia, Giuseppe Espa e Danny Quah.

2007.6 *Rescuing the LM (and the money market) in a modern Macro course,* by Roberto Tamborini.

2007.7 *Family, Partnerships, and Network: Reflections on the Strategies of the Salvadori Firm of Trento,* by Cinzia Lorandini.

2007.8 *I Verleger serici trentino-tirolesi nei rapporti tra Nord e Sud: un approccio prosopografico,* by Cinzia Lorandini.

2007.9 *A Framework for Cut-off Sampling in Business Survey Design,* by Marco Bee, Roberto Benedetti e Giuseppe Espa

2007.10 *Spatial Models for Flood Risk Assessment,* by Marco Bee, Roberto Benedetti e Giuseppe Espa

2007.11 *Inequality across cohorts of households:evidence from Italy,* by Gabriella Berloffa and Paola Villa

2007.12 *Cultural Relativism and Ideological Policy Makers in a Dynamic Model with Endogenous Preferences,* by Luigi Bonatti

2007.13 *Optimal Public Policy and Endogenous Preferences: an Application to an Economy with For-Profit and Non-Profit,* by Luigi Bonatti

2007.14 *Breaking the Stability Pact: Was it Predictable?,* by Luigi Bonatti and Annalisa Cristini.

2007.15 *Home Production, Labor Taxation and Trade Account,* by Luigi Bonatti.

2007.16 *The Interaction Between the Central Bank and a Monopoly Union Revisited: Does Greater Uncertainty about Monetary Policy Reduce Average Inflation?,* by Luigi Bonatti.

2007.17 *Complementary Research Strategies, First-Mover Advantage and the Inefficiency of Patents,* by Luigi Bonatti.

2007.18 *DualLicensing in Open Source Markets*, by Stefano Comino and Fabio M. Manenti.

2007.19 *Evolution of Preferences and Cross-Country Differences in Time Devoted to Market Work,* by Luigi Bonatti.

2007.20 *Aggregation of Regional Economic Time Series with Different Spatial Correlation Structures,* by Giuseppe Arbia, Marco Bee and Giuseppe Espa.

2007.21 *The Sustainable Enterprise. The multi-fiduciary perspective to the EU Sustainability Strategy,* by Giuseppe Danese.

2007.22 *Taming the Incomputable, Reconstructing the Nonconstructive and Deciding the Undecidable in Mathematical Economics,* by K. Vela Velupillai.

2007.23 *A Computable Economist's Perspective on Computational Complexity,* by K. Vela Velupillai.

2007.24 *Models for Non-Exclusive Multinomial Choice, with Application to Indonesian Rural Households,* by Christopher L. Gilbert and Francesca Modena.

2007.25 *Have we been Mugged? Market Power in the World Coffee Industry,* by Christopher L. Gilbert.

2007.26 *A Stochastic Complexity Perspective of Induction in Economics and Inference in Dynamics,* by K. Vela Velupillai.

2007.27 *Local Credit ad Territorial Development: General Aspects and the Italian Experience,* by Silvio Goglio.

2007.28 *Importance Sampling for Sums of Lognormal Distributions, with Applications to Operational Risk,* by Marco Bee.

2007.29 *Re-reading Jevons's Principles of Science. Induction Redux,* by K. Vela Velupillai.

2007.30 *Taking stock: global imbalances. Where do we stand and where are we aiming to?* by Andrea Fracasso.

2007.31 *Rediscovering Fiscal Policy Through Minskyan Eyes,* by Philip Arestis and Elisabetta De Antoni.

2008.1 *A Monte Carlo EM Algorithm for the Estimation of a Logistic Auto-logistic Model with Missing Data,* by Marco Bee and Giuseppe Espa.

PUBBLICAZIONE REGISTRATA PRESSO IL TRIBUNALE DI TRENTO