Dipartimento di Statistica
"Giuseppe Parenti"

WORKING PAPER 2011/10

Productivity change of Italian firms: an analysis of panel data using log linear models

Laura Grassini, Gianni Marliani

Università degli Studi di Firenze

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - www.ds.unifi.it

PRODUCTIVITY CHANGE OF ITALIAN FIRMS: AN ANALYSIS OF PANEL DATA USING LOG LINEAR MODELS

Laura Grassini, Gianni Marliani[1]
Department of Statistics, University of Florence, Italy

Abstract. *A provisional version of panel data, built in collaboration with Istat (the Italian National Statistical Institute), is used to synthetically describe the productivity profile of Italian manufacturing firms in different economic sectors (Ateco divisions). For each sector, a mobility table is built by classifying firms according to the categorized values of a labor-productivity index in two different years (1998 and 2004). The analysis is carried out using log-linear models, in order to identify specific association patterns (inertia, downgrading, upgrading) within these mobility tables. The empirical analysis reveals typical patterns in some economic sectors but, on the whole, it shows an objective difficulty in interpreting the results.*

Keywords: *Labor productivity, mobility table, log-linear models.*

1. INTRODUCTION

The Italian economy is going through a critical period. Italian firms seem to have difficulties in keeping up with competitors, mainly because their productivity (particularly labor productivity) is not increasing.

This general economic trend is not equally shared by the different economic sectors, that exhibit different performances (Grassini and Marliani, 2007).

To deepen the analysis of firm productivity in different sectors, a research group, involving researchers of the Italian Universities and of Istat, is building a panel data-set by matching individual data coming from survey responses given by the same firm for different years. Data refer to the surveys carried out annually by Istat to provide Eurostat with the Structural Business Statistics (SBS).

At the moment, a provisional version of the panel referring to firms interviewed both in 1998 and 2004 is available.

Aim of the paper is to verify the possibility of using such panel data set, to synthetically describe the productivity profile of the different economic sectors

[1] Authors' contact: grassini@ds.unifi.it, marliani@ds.unifi.it

by the use of parsimonious models.

More specifically, our study applies log-linear models to *turnover* or *mobility* tables. For each economic activity, a mobility table is built by classifying firms according to the categorized values of a performance index in two different years.

The use of this type of classified data is not new in monitoring the performance of financial instruments. In the literature, analyses are mainly carried out by *migration models* to analyze portfolio risk. For example, a transition matrix with probabilities of migration from one credit quality (rating) to another, over a given time horizon, is the key component of the CreditMetrics™ model proposed by J.P. Morgan (1997). The *migration analysis* technique is commonly used by the most important financial and rating companies and is also considered in the Basel 2 accords (Altman, 1998; Saunders, 1999).

Recently, some studies (Barry, Escalante and Ellinger, 2002; Deng, Escalante, Barry, Yu, 2007) have applied the principles of *migration analysis* on transition matrix built on firms panel data by categorizing a profitability ratio (ROA: return on assets).

Here we follow this last approach and analyse a transition matrix built by categorizing a firm-performance index but we use log-linear models to identify specific association patterns (inertia, downgrading, upgrading) within the mobility table. The firm-performance measure considered is "value added/employment". The analysis is expected to detect typical profiles of the different economic sectors.

The paper is structured as follows. Section 2 introduces the log-linear models applied in the analysis, Section 3 contains a description of the data and the construction of the mobility tables. Section 4 contains the preliminary results of the analysis. Some conclusions are drawn in section 5.

2. LOG–LINEAR MODELS FOR MOBILITY TABLES

A *mobility* table contains counts obtained by classifying the same individuals in at least two occasions. In the case of two occasions (our case), the table displays the same categories for both row and column classification. The subjects observed at time t are the same at time t+1 and the pairs of observations are called *matched pairs*. Hence, in a mobility table, row and column categories refer to the same variable, observed in two points in time.

A variety of models has been introduced to deal with features commonly associated with counts from temporal observations. In general, three main phenomena are important to be analyzed (Fingleton, 1984; Agresti, 2002): (i) loyalty (or inertia), i.e. the large presence of identical responses at different times; (ii) the dependence of transition between categories on inter-category distance; (iii) the directional balance of inter-category transitions (symmetry).

Log-linear models, that are designed to model the association structure among categorical variables, have been largely used in the analysis of mobility tables. Models that allow great flexibility typically employ many structural parameters and the substantive interpretation of such parameters is very often difficult. Therefore, we limit the analysis to a restricted numbers of log-linear models, favouring the more parsimonious ones.

Let us indicate by μ_{ij} the expected number of individuals moving from category i at time t to category j at time $t+1$, and be K the number of categories. The general log-linear model for a mobility table (i.e., for squared contingency tables) is:

$$ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \delta(i,j) \qquad i,j=1,2,...,K \qquad (1)$$

where λ is the mean effect, λ_i^R expresses the row (superscript R) effect, λ_j^C the column (superscript C) effect and $\delta(i,j)$ aims to express the association among the two variables. A special case of $\delta(i,j)$ is $\delta(i,j)=\lambda_{ij}^{RC}$ which produces the saturated log-linear model, where λ_{ij}^{RC} is the interaction term.

The term $\delta(i,j)$ determines the association structure within the table and alternative specifications of $\delta(i,j)$ produce special log linear models that have been called *hybrid* log linear models (Fingleton, 1984). More specifically, such linear models present a simplified interaction structure and are therefore unsaturated models with parameters that are not part of the conventional log-linear modelling.

A long debate on the application of log-linear models to mobility tables (often to analyze intergenerational occupational mobility) dealt with the possibility to differentiate mobility due to changes in the marginal distributions of origins and destinations (structural mobility), and mobility that arises from the ‘openness’ of the system represented in the table (circulation or pure mobility). However, the direct link between these two types of mobility and the parameters of a log-linear models are difficult to find or to interpret (Sobel et al., 1985, 1998).

Consequently, in our opinion, it is preferable to use the concept of relative mobility as it can be directly related with the parameters of a log-linear model (Cobalti, 1988). The measure of *relative mobility* compares the chance of mobility of individuals having different origins. If a mobility table has ordered rows and ordered columns, one should take this ordering into account by focusing on the 2x2 sub-tables formed by adjacent rows and adjacent columns of the full table (Goodman, 1979; Lawal, 1993); thus *relative* mobility can be measured by the *local* odds ratio:

$$odds(i,j) = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i,j+1}\pi_{i+1,j}} = \frac{\mu_{ij}\mu_{i+1,j+1}}{\mu_{i,j+1}\mu_{i+1,j}} \qquad \text{(i,j=1,…,K-1)}$$

where π_{ij} represents the joint probability associated to the cell *(i,j)*.

The local odds ratio involves four adjacent cells; a KxK matrix admits $(K-1)$x$(K-1)$ local odds ratios. The entire set of local odds ratios in a contingency table encloses all the relevant information regarding the association structure between the ordinal factors (Goodman, 1979). Since the factor levels have meaningful ordering, the local odds ratios are also individually informative (Goodman, 1979; Rudas, 1998). In the followings, we will denote the log-odds ratio by $\theta_{ij}=ln(odds(i,j))$.

The class of models here adopted for the analysis can be partitioned into three broad categories:

- independence models: independence, main diagonal, mover-stayer models;
- association models: uniform association, row association, column association, row and column association;
- minor diagonal models: minor diagonal symmetry, minor diagonal model.

For each kind of models, a specific parameterization can be derived from alternative constraints imposed on the local odds ratios that determine the association structure in the table, that is, the form of the term $\delta(i,j)$ in (1) (Goodman, 1979, 1979a; Fingleton, 1984).

2.1 INDEPENDENCE (IND), MAIN-DIAGONAL (D), MOVER-STAYER (MS) MODELS

The independence model is defined as

IND: $ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C$

This model corresponds to an ordinary null log-linear model. It assumes marginal independence between the variables in the table.

However, in a mobility table, where there is a concentration of frequencies in the main diagonal cells denoting loyalty or inertia, other models are more suitable. The main-diagonal (D) and mover-stayer (MS)[2] models are designed just to account for this recurring feature of panel data.

Specifically, in the D model, a parameter is added to allow for the inflated frequencies of the main diagonal cells. Hence:

D: $$ln\,\mu_{ij} = \begin{cases} \lambda + \lambda_i^R + \lambda_j^C + \delta & \text{for } i = j \\ \lambda + \lambda_i^R + \lambda_j^C & \text{for } i \neq j \end{cases}$$

When $\delta > 0$ permanence is higher than expected under the independence assumption.

The MS model is a refinement of the D model and attributes the inflated main diagonal frequencies to inertia effects that differ among levels:

MS: $$ln\,\mu_{ij} = \begin{cases} \lambda + \lambda_i^R + \lambda_j^C + \delta_i & \text{for } i = j \\ \lambda + \lambda_i^R + \lambda_j^C & \text{for } i \neq j \end{cases}$$

When $\delta_i > 0$ permanence is higher than expected under independence for units in level i at time t.

Notice that under D and MS models, $\theta_{ij}=0$ for all the 2x2 tables formed by cells that do not belog to the main diagonal. These models can be therefore defined as *quasi-independence models* as, apart from a selected number of cells, the variables are independent (Agresti 2002). In addition, D model, implies constant odds ratios for the main diagonal cells.

2.2 UNIFORM ASSOCIATION MODELS (UNI, UNID, UNIMS)

The following models are called *association models* and have been mostly developed by Goodman (Goodman, 1979; 1979a). We will give the general definition and derive some alternative specifications (as for the IND model) to account for the inflated frequencies in the main diagonal cells.

The UNI model accounts for the presence of ordered categories both in

[2] The term *mover-stayer* was firstly used by Fingleton (1984). It is assumed that the population is composed of *movers,* those moving from one category to another, and *stayers*, those remaining in the same category. The MS model was also called *non-uniform loyalty* model by Lawal and Upton (1990)

rows and columns, and is specified as:

UNI: $ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \gamma\, u_i\, u_j$

where u_i is the *score* associated with each category i and γ is the parameter of uniform association. The term *uniform* reflects the equality of the odds ratio measure of association when u_i=i. In this case, θ_{ij}=γ constant, for all adjacent 2x2 sub-tables. A positive value of this parameter indicates that higher row categories are associated with higher column categories.

The UNID model includes also a constant parameter δ for the main diagonal cells while UNIMS allows this parameter to vary among main diagonal cell. Both UNID and UNIMS models implies θ_{ij}=γ constant for all 2x2 tables formed by cells all falling off the main diagonal.

For all the three models above presented, specific constraints can be imposed on the parameters to model particular situations.

Firstly, equality constraints on the main effects, $\lambda_i^R = \lambda_i^C = \lambda_i$ (equality of row and column main effects) can be imposed to consider the case of symmetric contingency tables: $\mu_{ij} = \mu_{ji}$ (i.e. the symmetry property).

Assigning value i to the scores u_i of the (ordinal) row and column variables have the effect of equalizing the local odds ratios. These kinds of constraint can be adopted to reflect an assumption of equal inter-category distance.

Alternative scoring systems can be adopted (instead of u_i=i), which are sensible whenever one assumes the observed ordinal variables as the result of an underlying continuous variables. Hence, a real value is assigned to each factor category. Here, we assigned integer values maintaining the ordinal meaning of the categories. This simplification can be viewed as a positive aspect of scoring, without needing to regard these scores as indices of how apart the ordered levels truly are (Agresti, 1983).

Alternatives to the uniform association model exist in which the association is not longer uniform but it depends on row and/or columns categories. The following models (row association, column association and row-column association models) have been introduced in Goodman (1979) and adopted, among the others, by Lang and Agresti (1988), Lawal (1993), Lang and Eliason, (1997) to analyze mobility tables. In the following, we assume u_i=i.

2.3 ROW ASSOCIATION MODEL (ROW, ROWD, ROWMS)

The model

ROW: $$ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \alpha_i\, u_j$$

is called row-effect association model. The case $u_j=j$ implies $\theta_{ij}=\theta_{i.}=\alpha_{i+1}-\alpha_i$, $i=1,\ldots,K-1$ for all adjacent formed 2x2 tables (Goodman, 1979; Ishii-Kuntz, 1994). It means that the local odds ratios are constant across different combination of adjacent columns. Moreover, when $\alpha_{i+1}=\alpha_i$, rows $i+1$ and i have identical conditional distributions on the column variable; when $\alpha_{i+1}>\alpha_i$, the subjects in row $i+1$ are more likely to be found in higher categories on the ordinal scale of the column variable, compared to the subjects in row I (Ishii-Kuntz, 1994). In other terms, when $\alpha_{i+1}>\alpha_i$, the association between the variables determines higher frequencies below the main diagonal.

The ROWD and ROWMS versions of the row association model contain also the usual parameters for the main diagonal cells. In both ROWD and ROWMS models, it holds $\theta_{ij}=\theta_{i.}=\alpha_{i+1}-\alpha_i$, $i=1,\ldots,K-1$ for all formed 2x2 tables in which all cells fall off the main diagonal (Goodman, 1979).

These models are not considered adequate for mobility tables (where row and column variables are intimately related) as they treat differently the two variables: row as nominal, column as ordinal (Lawal and Upton, 1990).

2.4 COLUMN ASSOCIATION MODEL (COL, COLD, COLMS)

The COL model is similar to the ROW model and is defined as follows

COL: $$ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \beta_j\, u_i$$

The case $u_i=i$ implies $\theta_{ij}=\theta_{.j}$, where $\theta_{.j}=\beta_{j+1}-\beta_j$, $j=1,\ldots,K-1$, for all formed 2x2 tables (Goodman, 1979). This parameter describes the differences among columns with respect to their conditional distribution on the row variable.

Within a given column, a positive β_j indicates that more observations occur in rows representing high values of the row variable and fewer in rows representing lower values, compared with what we would expect under the independence model (Ishii-Kuntz, 1994).

When $\beta_{j+1}=\beta_j$, columns $j+1$ and j have identical conditional distributions on the row variable; when $\beta_{j+1}>\beta_j$, the subjects in column $j+1$ are more likely to be found in higher categories on the ordinal scale of the row variable, compared to the subjects in column j. In other terms, when $\beta_{j+1}>\beta_j$, association tends to put frequencies above the main diagonal.

The COLD and COLMS specifications contain also the usual additional parameters for the main diagonal cells. In both COLD and COLMS models, it holds $\theta_{.j}=\beta_{j+1}-\beta_j$, $j=1,\ldots,K-1$, for all formed 2x2 tables in which all cells fall off the main diagonal.

As the row effect models, also these models have been criticized when applied to mobility tables, because they treat differently the two variables: row variable as ordinal and column variable as nominal.

In the perspective of a mobility table, the interpretation of the row model seems to be more immediate in term of conditional distribution as in a transition matrix[3]. In this respect, the row-association-models can be associated with an active behavior of the firms (a sort of push effect of the row category), whereas the column-association-models with a passive behavior (a sort of pull or attraction effect of the column category).

2.5 ROW AND COLUMN ASSOCIATION MODEL (ROWCOL, ROWCOLD, ROWCOLMS)

These models include the row and column association effects without and with the main diagonal parameters, and are defined as

ROWCOL: $$ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \alpha_i u_j + \beta_j u_i$$

The case $u_i=i$ implies $\theta_{ij}=\theta_{i.}+\theta_{.j}=(\alpha_{i+1}-\alpha_i)+(\beta_{i+1}-\beta_i)$ for all formed 2x2 tables (ROWCOL) and only in those cells that fall off the main diagonal (ROWCOLD, ROWCOLMS).

On the log scale, the two association effects are additive. The interpretation is more complicated than the last two mentioned models as the local odds ratios are specific for each cell. However, the additive expression for θ_{ij} provides a simple decomposition of the association effect.

When $\lambda_i^R = \lambda_i^C = \lambda_i$ and $\alpha_i=\beta_i$, for $i=1,\ldots,K$, the models satisfy the symmetry conditions.

MINOR DIAGONAL MODELS (MINDSYM, MIND)

The MINDSYM model is specified as follows (Fingleton, 1984)

MINDSYM $$ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \eta_s \quad s = |\,i-j\,|$$

[3] A transition matrix or table is different from a mobility matrix since the cells of a transition matrix contain the conditional probabilities p_{ij} to move to (column) category j from (row) category i, given the category i at time t.

It assumes a constant local odds ratio for the cells on the same diagonal and on those ones equidistant from the main diagonal. Specifically:

$$\theta_{ji} = \theta_{ij} = 2\,\eta_{|i\text{-}j|} - \eta_{|i\text{-}j+1|} + \eta_{|i\text{-}j\text{-}1|}.$$

This formula involves *K-1* additional parameters with respect to the D model. Notice that, when $\lambda_i^R = \lambda_i^C = \lambda_i$ for $i=1,\ldots,K,$ the model still satisfies the symmetry conditions.

The minor diagonal model (MIND) is a version of MINDSYM that assumes distinct parameters for each diagonal (Fingleton, 1984; Lawal, 1993):

MIND $$ln\,\mu_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \eta_s \qquad s = i - j$$

It entails constant local odds ratios for the cells on the same diagonal and involves *2(K-1)* additional parameters with respect to the D model.

Minor diagonal models can be viewed as a refinement of the D model by replacing specific parameters for each diagonal in the table. By this perspective, diagonal models assume the presence of a sort of barrier effect that inhibits mobility but which remains constant irrespective of the origin and destination. The development of minor-diagonal models like MINDSYM and MIND arises just from this idea (Fingleton, 1984).

3. DATA, PERFORMANCE-MEASURE AND CONSTRUCTION OF THE MOBILITY TABLE

Data derive from surveys carried out by Istat to provide Eurostat with the Structural Business Statistics (SBS). SBS data are collected on a large sample of Italian firms (about 55000) and consist of a rich set of variables.

Currently, SBS cover NACE Rev.1.1 sections C to K (Industry: section C-E, Construction: F, Trade: G and Services: H-I-K), which broadly speaking refer to market activities. Financial services (sector J) are not considered.

Tab. 1 shows the economic sectors analyzed in the paper, which belong to the manufacturing activities.

The performance measure considered is labour productivity defined as “value added(at factor cost)/total employment”. The data used refer to two years: 1998 and 2004.

The construction of the mobility tables is defined in terms of quintiles of the performance measure. Quintiles are computed by sector on the individual means of the two years[4]. The values of the quintiles (thousand Euro) are reported in Tab. 1.

Tab. 1. Quintiles of labour productivity (value added/employment). Thousand Euros.

ATECO division	**P_20**	**P_40**	**P_60**	**P_80**
15 - Food	37.17	46.27	56.52	72.21
17 - Textiles	27.63	34.80	40.99	49.40
18 - Manufacture of wearing apparel¸ dressing and dyeing of fur	18.85	23.53	31.07	43.77
19 - Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	22.39	27.57	34.21	45.77
20 - Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	28.02	33.66	37.81	45.28
21 - Manufacture of paper and paper products	36.55	44.18	52.68	65.95
22 - Publishing, printing and reproduction of recorded media	36.11	42.80	51.68	71.80
24 - Manufacture of chemicals and chemical Products	46.16	58.67	71.99	91.74
25 - Manufacture of rubber and plastics products	34.29	41.30	48.31	58.42
26 - Manufacture of other non-metallic mineral Products	34.65	42.79	50.60	65.06
27 - Manufacture of basic metals	39.57	47.41	56.49	70.64
28 - Manufacture of fabricated metal products, except machinery and equipment	33.56	39.45	45.71	54.57
29 - Manufacture of machinery and equipment n.e.c.	37.52	43.90	50.28	59.34
30-31 Manufacture of office, accounting and computing machinery. Manufacture of electrical machinery and apparatus n.e.c.	33.95	39.69	46.21	58.17
32 - Manufacture of radio, television and communication equipment and apparatus	32.91	41.56	49.25	67.30
33 - Manufacture of medical, precision and optical instruments, watches and clocks	32.38	40.92	49.11	60.61
34 - Manufacture of motor vehicles, trailers and semi-trailers	32.92	39.59	47.07	58.64
35 - Manufacture of other transport equipment	30.22	36.89	45.55	59.87
36 - Manufacture of furniture; manufacturing n.e.c.	27.39	32.49	37.54	45.64

[4] Also Krueger (2005) used the mean values, as in this paper. More specifically, Krueger used the mean of the first five years and the mean for the last five years to build the transition table referred to a ten-year interval.

Some words shouol be spent on the construction of the mobility tables by categorizing continuous variables. It is well known that such a categorization can affect dynamics in important ways and the obtained table is influenced by the association (correlation) between the two variables (Mosteller, 1968).

The issue of categorizing continuous data (often used in the analysis of income distribution) has been widely discussed in the study of social mobility[5] when a transition matrix is assumed to be derived from a continuous state-space Markov model. But, it is important to stress here that these issues are beyond the aim of this study, mainly because:

1) our analysis is focused on a single mobility table and we are not, for the moment, interested in constructing a more complex model describing the evolution of productivity through time;
2) the use of fractile categorization is a suitable solution when one is interested in the relative position of individuals as, for example, in the analysis of income distributions since the relative position of the individuals is directly related to the inequality measures; but it is less suitable for the present case, where we want to investigate the decrease (or increase) of the values of labour productivity.

Instead, there is another question that is worth mentioning. In analysing the labour productivity variation by mobility tables, we do not take into account the effect that other characteristics (for example firm size) could have had on the labour productivity itself.

In order to deepen this aspect we have compared the mean firm-size, measured by the number of employees, in the two years (1998 and 2004) and computed the t-test for the mean difference. Most of the sectors does not exhibit a statistically significant mean difference (see Table 2). Moreover, in comparing size and productivity variation of each firm in the two years, the correlation between the two variables is substantially null.

---

[5] See, for example, Shorrocks, 1978; Geweke *et al.,* 1986. In some studies, the type of categorization is justified on the basis of the properties of the underlying Markov process (Singer and Spilerman, 1976). Generally, the solution adopted to satisfy some properties required by the Markov model theory is the aggregation based on fractiles that allows marginal homogeneity (Quah, 1996; Krueger, 2005; Lucas and Klaassen, 2008).

Tab. 2. Mean number of employees in 1998 and 2004. Test t (matched pairs) between the two means.

ateco 1998	N Obs	mean number employees 1998	mean number employees 2004	% variation 98-04	Mean difference 2004-1998	t value (matched pairs)	Pr > |t|
15	876	109.8	119.5	8.9%	9.8	2.5	0.014
17	986	98.9	94.1	-4.8%	-4.8	-1.1	0.280
18	468	96.3	94.6	-1.8%	-1.7	-0.6	0.535
19	506	72.1	65.4	-9.3%	-6.7	-1.7	0.084
20	314	63.6	68.2	7.3%	4.6	2.8	0.005
21	318	82.7	92.8	12.2%	10.1	3.2	0.002
22	380	96.8	88.9	-8.2%	-7.9	-1.4	0.163
24	510	210.3	193.4	-8.0%	-16.9	-0.9	0.361
25	744	93.4	98.8	5.7%	5.4	2.3	0.024
26	765	98.3	109.3	11.2%	11.0	4.4	0.000
27	409	148.9	160.1	7.5%	11.2	1.1	0.254
28	1680	65.0	70.0	7.8%	5.1	5.7	0.000
29	1784	114.0	116.9	2.5%	2.8	1.0	0.298
30-31	450	130.9	121.4	-7.3%	-9.6	-1.2	0.214
32	110	253.9	254.6	0.3%	0.7	0.0	0.988
33	245	98.4	118.6	20.5%	20.2	1.5	0.140
34	234	253.7	248.8	-1.9%	-4.9	-0.2	0.844
35	125	224.2	209.0	-6.8%	-15.2	-1.0	0.306
36	845	68.8	76.1	10.6%	7.3	3.2	0.001

4. PRELIMINARY RESULTS

Tab. 3 below summarizes some information derived from the mobility table built considering only the firms that maintain the same Ateco code in the two years[6].

As one can see, in all sectors the majority of firms has changed state (% of no-change is less than 50%).

Table 4 shows the p-values (scaled deviance) associated with the fitted models. In most cases, there are several competing models. In this case, they have been compared either on the basis of the log-likelihood ratio test, when models are nested (in the sense that one model is a reduced form of the other), or on the basis of the AIC criterion.

[6] Sectors 32-35 are not considered in the analysis because of the small frequencies (see Tab. 2)

Tab. 3. Analysis of the mobility table of labour productivity (value added/employment)

ATECO 1998	Change ATECO	Same ATECO in 1998 and 2004							
		Decrease	No change	Increase	Total	Decrease (%)	No change (%)	Increase (%)	Total (%)
15	42	216	370	290	876	24.7	42.2	33.1	100.0
17	79	432	344	210	986	43.8	34.9	21.3	100.0
18	51	115	247	105	467	24.6	52.9	22.5	100.0
19	25	174	206	126	506	34.4	40.7	24.9	100.0
20	44	66	128	120	314	21.0	40.8	38.2	100.0
21	29	104	125	89	318	32.7	39.3	28.0	100.0
22	29	96	185	99	380	25.3	48.7	26.1	100.0
24	87	172	212	126	510	33.7	41.6	24.7	100.0
25	108	223	315	206	744	30.0	42.3	27.7	100.0
26	50	209	304	252	765	27.3	39.7	32.9	100.0
27	79	146	154	109	409	35.7	37.7	26.7	100.0
28	248	599	681	400	1680	35.7	40.5	23.8	100.0
29	253	626	693	465	1784	35.1	38.8	26.1	100.0
30-31	141	132	185	133	450	29.3	41.1	29.6	100.0
32	51	31	46	32	109	28.4	42.2	29.4	100.0
33	75	54	115	76	245	22.0	46.9	31.0	100.0
34	52	68	91	75	234	29.1	38.9	32.1	100.0
35	26	52	42	31	125	41.6	33.6	24.8	100.0
36	86	292	338	215	845	34.6	40.0	25.4	100.0
Total	1555	3807	4781	3159	11747	32.4	40.7	26.9	100.0

The p-values in the framed cells in Tab. 4 identify the selected models; the shaded cells refer to p values greater than 0.05. Only sector 26 (Manufacturing of non-metallic products) did not obtain a satisfactory fit with any of the adopted models.

The results show that different association patterns can be detected for different sectors. However, it seems always necessary to use a specific parameter for the main diagonal counts. With the exception of the ROW model for sector 21 (Manufacture of paper and paper products) all the other selected models include such parameters.

The UNIMS model is selected for sectors 19 and 30-31. In both these cases, the mobility table implied by the selected model is symmetric. This feature denotes a sort of stability (main diagonal parameters) or the tendency of occupying far positions. In fact, the UNIMS models give a significant positive value of γ (respectively, 0.5174 and 0.2826). This is particularly evident for sector 19 but with movements limited to the higher categories (θ_{34}).

Tab. 4. Results: p-values (shaded area: p-value>0.05; with borders: selected model)

Model	DF	15	17	18	19	20	21	22	24	25	26	27	28	29	30-31	36
IND	16	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
D	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
MS	11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
UNI	15	0.0003	0.0045	0.0000	**0.0608**	0.0010	**0.0578**	0.0182	0.0002	0.0000	0.0000	**0.0736**	0.0000	0.0000	0.0028	0.0031
UNID	14	0.0187	0.0163	0.0023	**0.0665**	0.0019	**0.0924**	**0.2975**	0.0288	0.0002	0.0002	**0.0848**	0.0000	0.0000	**0.0766**	0.0441
UNIMS	10	0.0123	**0.0944**	0.0023	**0.2412**	0.0137	**0.0708**	**0.4069**	**0.1570**	0.0065	0.0033	**0.1666**	0.0220	0.0000	**0.5379**	0.0153
ROW	12	0.0003	**0.0991**	0.0000	**0.0634**	0.0012	**0.2554**	0.0319	0.0001	0.0000	0.0032	**0.0634**	0.0000	0.0000	0.0099	0.0007
ROWD	11	0.0239	**0.1314**	0.0010	**0.0692**	0.0014	**0.2803**	**0.2836**	0.0119	0.0000	0.0092	**0.0694**	0.0000	0.0000	**0.1537**	0.0161
ROWMS	7	**0.2479**	**0.0618**	0.0054	**0.1737**	0.0142	**0.0943**	**0.2467**	**0.0732**	0.0179	0.0072	**0.0520**	**0.0639**	0.0000	**0.3148**	0.0129
COL	12	0.0004	0.0012	0.0000	**0.1338**	0.0006	0.0326	**0.1456**	0.0000	0.0000	0.0000	**0.0730**	0.0000	0.0000	0.0171	0.0014
COLD	11	0.0153	0.0062	0.0040	**0.1450**	0.0010	**0.0708**	**0.6581**	0.0098	0.0007	0.0000	**0.0966**	0.0001	0.0000	**0.3052**	0.0184
COLMS	7	**0.0512**	**0.3885**	0.0008	**0.3376**	0.0053	**0.7235**	**0.4123**	**0.2306**	0.0252	0.0336	**0.3934**	0.0187	0.0000	**0.5154**	0.0094
ROWCOL	9	0.0017	0.0473	0.0000	0.0487	0.0003	**0.2070**	**0.0830**	0.0000	0.0001	0.0032	**0.0672**	0.0000	0.0000	0.0131	0.0002
ROWCOLD	8	**0.0824**	**0.0568**	0.0024	**0.0522**	0.0003	**0.2768**	**0.4455**	0.0028	0.0008	0.0054	**0.0855**	0.0000	0.0000	**0.2485**	0.0043
ROWCOLMS	4	**0.4068**	**0.7006**	0.0027	**0.4866**	**0.2862**	**0.5881**	**0.1452**	**0.4075**	**0.1068**	0.0208	**0.2512**	**0.7088**	**0.4708**	**0.6648**	**0.3992**
MINDSYM	12	**0.0736**	**0.0599**	0.0074	**0.0734**	0.0278	**0.0900**	**0.2320**	**0.1265**	0.0027	0.0004	**0.1891**	0.0001	**0.0791**	**0.0872**	**0.2573**
MIND	9	**0.3114**	0.0168	**0.2061**	0.0319	0.0079	0.0421	**0.2254**	**0.0573**	0.0091	0.0040	**0.2321**	0.0001	0.0370	0.0361	**0.1467**

The association is different among rows in sector 21 and it holds for every row $\alpha_{i+1}-\alpha_i>0$ (respectively: 0.5771, 0.3667, 0.0633, 0.769); rows 2 and 3 are not significantly different. It follows that, except for these rows, units in row $i+1$ are more likely to be found in higher categories of the column variable than units in row i. Hence, firms in higher category at time t seem to possess a sort of competitive advantage but the association pattern tends to put frequencies below the main diagonal.

The association is differentiated by column in sector 22, where there is also a positive main diagonal effect denoting loyalty. $\beta_{i+1}-\beta_i>0$ always holds for every i (respectively 0.4648, 0.3608, 0.5159, 0.998); hence, association tends to push frequencies above the main diagonal. However the permanence effect is predominant.

In sector 17 (Textiles), association is differentiate by column and main diagonal cells. A positive diagonal parameter is present in cells (1,1) and (5,5) denoting stability. $\beta_{i+1}-\beta_i$ equals to, respectively, –0.0301, 0.3814, 0.3956, 0.081; that is, only columns 2, 3 and 4 exhibit differences.

In sectors 20, 25 and 28, ROWCOLMS model is selected. This form is more difficult to interpret but it means that both pull and push effects are present.

Finally we considered the minor diagonal models that, according to some researchers, are more appropriate in describing mobility tables (Lawal and Upton, 1990; Lawal, 1993).

MINDSYM is selected for sectors 24, 27, 29 and 36. The model for sector 24 shows the presence of loyalty as only the main diagonal local odds ratios are significantly larger than 1. The models for sectors 27 and 29 exhibit significant (above and below the main diagonal), denoting the tendency to remain in the same position or to upgrade/downgrade by one.

MIND is selected for sector 15. The local odds ratios based on the main diagonal and on the first lower minor diagonal are significantly larger than 1, denoting the presence of loyalty and the tendency to downgrade by one position. MIND is also selected for sector 18. In this case, the main features (odds ratio statistically different from 1) are permanence and the tendency to upgrade by one position

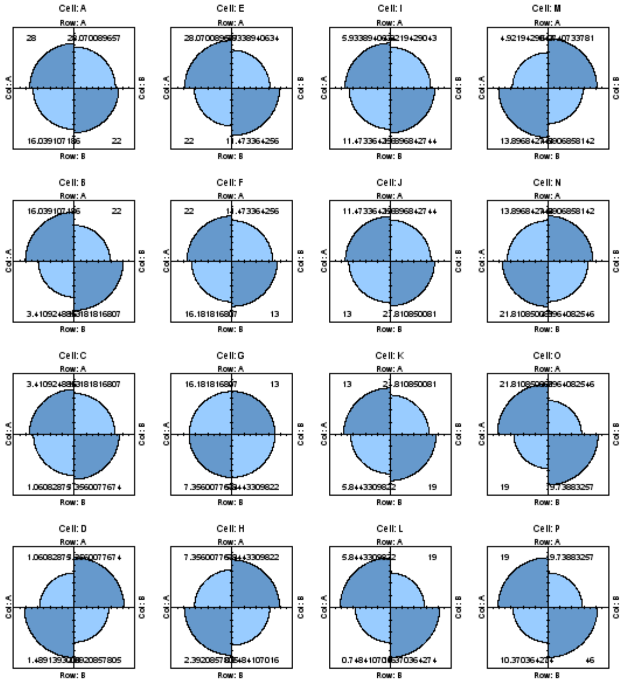
However the interpretation of data in Tab. 4 is not easy and can be facilitated through the inspection of the *fourfold plot* (Friendly, 1994), a special graphic tool that allows the visualization of the local odds ratios implied by the fitted model.

In order to illustrate the utility of this tool, we report in Fig.1 the fourfold plot referred to the UNIMS model for sector 19 and to the MIND model for sector 20.

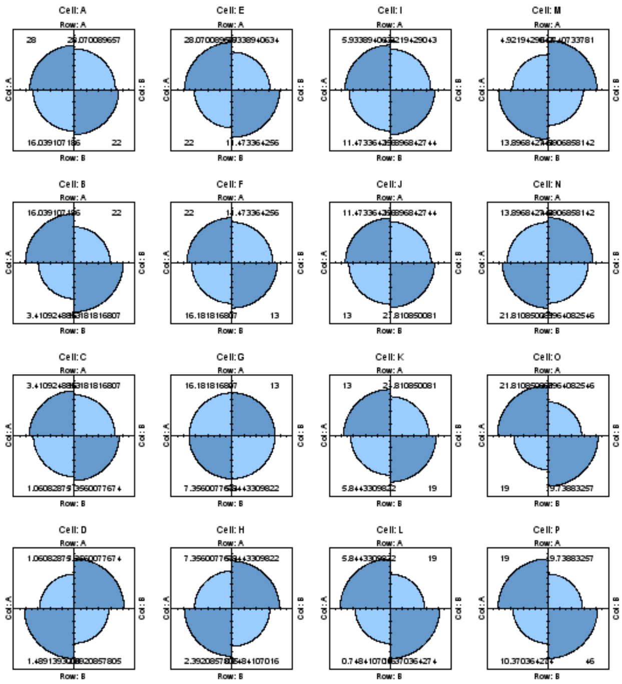
In the fourfold plot, the cell frequencies *f[i,j]* of a 2 x 2 table are shown as a quarter circle whose radius is proportional to *sqrt(f[i,j])* so that its area is proportional to the cell frequency. An association (odds ratio different from 1) between the binary row and column variables is indicated by the tendency of diagonally opposite cells in one direction to differ in size from those in the other direction. Colour is used to show this direction.

Fig. 1 – Fourfold plot: local odds-ratios of the estimated frequencies

A) Sector 19 selected models UNIMS



B) Sector 20 Selected model ROWCOLMS



In our case, cell frequencies are the ones estimated by the model and the 2 x 2 tables are composed of adjacent cells (local odds ratios). To understand graphs in Fig. 1, it must be considered that the categories are ordered left-right for columns, top-bottom for rows.

The fourfold plot for sector 15 (model UNIMS), for example, clearly shows the symmetry of the relationship. Whereas more difficult is the interpretation of the fourfold plot for sector 20 (model ROWCOLMS) that shows more articulated movements, differently oriented graphs and stability only in cell (5,5).

5. CONCLUDING REMARKS

The empirical analysis conducted in this paper detected three types of prevalent pattern.

- *Loyalty* (or stability): it is always present with the exception of sector 21.
- *Row* and *column* effects. Row effect occurs for sector 21, denoting that the position at time *t* is determinant for the position at time *t+1;* the estimated effects denote a weakness of the sector in the analyzed interval. Column effect, interpreted as a pull effect at time *t+1*, occurs especially for sector 22 . In sector 17 loyalty predominates. Both row and columns effects are present in other three sectors.
- *Barriers* effects (diagonal models) are detected for six sectors but, looking at the p-values, such forms can be accepted also for other five sectors.

This empirical analysis states an objective difficulty in interpreting the results even through the four-fold plot (actually, this difficulty is typical of the log-linear models).

It is true that column-effect and row-effect models are relatively simpler to be understood, but such forms are always present with the D or MS effect, so that the interpretation gets harder. Hence, as loyalty is always relevant, minor diagonal models seem to be more suitable to describe transitions.

Generally, diagonal models, that preserve several degrees of freedom, provide a satisfactory fit (p-values>0.05) for most of the sectors. Only for three sectors (20, 25 and 28) the only models that have a not-significant p-value are the ROWCOLMS models. Hence, it could be interesting to introduce additional parameters in the diagonal models to provide more flexibility but preserving the simplicity of interpretation. For example, starting from the quasi-symmetry form MINDSYM, a good suggestion is in the contribution in Tomizawa (1993) and some more recent studies that analyze tables with ordered categories (Miyamoto, Ohtsuka, Tomizawa, 2004; Tomizawa, Miyamoto, Ouchi, 2006).

REFERENCES

Agresti A. (1983). A Survey of Strategies for Modelling Corss-Classifications Having Ordinal variables, *Journal of the American Statistical Association*, 78:381, pp.184-198.

Agresti, A. (2002). *Categorical Data Analysis*. 2° ed. New York: Wiley.

Altman E. I. (1998). The Importance and Subtlety of Credit Rating Migration, *Journal of Banking and Finance*, 22, pp. 1231-1247.

Barry P.J., Escalante C.L., Ellinger P.N. (2002). Credit Risk Migration Analysis of Farm Businesses, Agricultural Finance Review, 62:1, pp.1-11.

Cobalti A. (1988). Alternative conceptual frameworks for the analysis of mobility tables and the log-linear models, *Quality and Quantity,* 22:1, pp. 31-47.

Deng X., Escalante C.L., Barry P.J., Yu Y. (2007). Markov Chain Models for Farm Credit Risk Migration Analysis, *Agriculture Finance Review*, 67:1, pp. 99-117.

Fingleton B. (1984). *Models of category counts,* Cambridge University Press.

Friendly, M. (1994), *A fourfold display for 2 by 2 by k tables*. Technical Report 217, York University, Psychology Department.

Geweke, J., Marshall, R.C., Zarkin, G.A. (1986). Mobility indices in continuous time Markov chains. *Econometrica,* 54, pp.1407–1423.

Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74, 537-552.

Goodman, L. A. (1979a). Multiplicative models for square contingency tables with ordered categories *Biometrika,* 66:3, pp. 413-418.

Krueger J.J. (2005). Structural change in U.S. manufacturing: Stationarity and intra-distributional changes, *Economic Letters,*87 pp. 387-392.

Ishii-Kuntz M.(1994). *Ordinal log-linear models,* SAGE Publications, n. 07/097.

J.P. Morgan (1997). *CreditMetrics. Technical document*,. Online document available on the web (for example: http://www.riskmetrics.com).

Lang J.B., Eliason S.R. (1997). Application of Association-Marginal Models to the Study of Social Mobility, *Sociological methods and Research,* 1997:26, pp.183-212.

Lawal, H.B. (1993). Association, Symmetry and Diagonal models for occupational mobility and other similar square contingency tables having ordered categorical variables. *Biometrical Journal*, 35, 193-206.

Lawal, H.B., Upton, G.J.G. (1990). Alternative interaction structures in square contingency tables having ordered classificatory variables. *Quality & Quantity*, 24, 107-127.

Lucas A., Klaassen P. (2006). Discrete versus continuous state switching models for portfolio credit risk, *Journal of Banking and Finance,* 30:1, pp.23-35.

Marliani G., Grassini L. (2007). Italian Labour Productivity Changes: an Analysis of Firm Survey Data 1998-2004, *Proceedings of the 56th Session of the International Statistical Institute,* Lisboa.

Miyamoto, N.; Ohtsuka, W., Tomizawa, S. (2004). Linear diagonals-parameter symmetry and quasi-symmetry models for cumulative probabilities in square contingency tables with ordered categories, *Biometrical Journal,* 46, pp. 664–674.

Mosteller F. (1968). Association and Estimation in Contingency Tables, *Journal of the American Statistical Association*, 63:321, pp. 1-28.

Quah D. T. (1996). Aggregate and regional Disaggregate Fluctuations, *Empirical Economics,* 21, pp.147-159.

Rudas T. (1998). *Odds ratios in the analysis of contingency tables,* SAGE Publications, n. 07/119.

Saunders A. (1999). *Credit Risk Measurement*. New York: Wiley.

Shorrocks A.F. (1978). The Measurement of Mobility, *Econometrica*, 46:5, pp. 1013-1024.

Singer B., Spielerman S. (1976). The Representation of Social Processes by Markov Models, *The American Journal of Sociology*, 82:1, pp. 1-54.

Sobel M.E., Becker M., Minick S.M. (1998). Destinations and Associations in Occupational Mobility, *The American Journal of Sociology,* 104:3, pp.687-721.

Sobel M. E:, Hout M., Duncan O.D. (1985). Exchange, Structure and Symmetry in Occupational Mobility. *The American Journal of Sociology*, 91:2, pp.359-372.

Tomizawa, S. (1993). Diagonals-parameter symmetry model for cumulative probabilities in square contingency tables with ordered categories, *Biometrics*, 49, 883–887.

Tomizawa S., Miyamoto N., Ouchi M. (2006). Decomposition of Symmetry Model into Marginal Homogeneity and Distance Subsymmetry in Square Contingency Tables with Ordered Categories, *REVSTAT - Statistical Journal,* 4:2, pp. 153-161.