The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$ 98	1-year subscription	$138
2-year subscription	$165	2-year subscription	$245
3-year subscription	$225	3-year subscription	$345
1-year student subscription	$ 75	1-year student subscription	$ 99
1-year institutional subscription	$245	1-year institutional subscription	$285
2-year institutional subscription	$445	2-year institutional subscription	$525
3-year institutional subscription	$645	3-year institutional subscription	$765
Electronic only		**Electronic only**	
1-year subscription	$ 75	1-year subscription	$ 75
2-year subscription	$125	2-year subscription	$125
3-year subscription	$165	3-year subscription	$165
1-year student subscription	$ 45	1-year student subscription	$ 45

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Stata, Stata Press, Mata, Mata, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2014)
14, Number 2, pp. 398–406

Note on Lilien and modified Lilien index

Muhammad Rashid Ansari
INSEAD Business School
Singapore
rashid.ansari@insead.edu

Chiara Mussida
Department of Economics and Social Sciences
Università Cattolica del Sacro Cuore
Piacenza, Italy
chiara.mussida@unicatt.it

Francesco Pastore
Department of Law
Seconda Università di Napoli
Santa Maria Capua Vetere, Italy
Institute for the Study of Labor
Bonn, Germany
francesco.pastore@unina2.it

Abstract. This article is a companion to the Lilien (`lilien`) and modified Lilien commands for computing relative indices in Stata. In this article, we illustrate the main features of the commands with an application to the structural determinants of regional unemployment.

Keywords: st0343, lilien, Lilien and modified Lilien index, structural change, regional unemployment

1 Lilien index

The Lilien index (LI) is an important measure of structural change in several fields of economic research. One field, which garners special attention here, is the common use of the LI as a measure of structural change in the employment composition on the determinants of structural unemployment in the research literature. Indirectly, it measures the degree to which labor demand is affected by sectoral shifts in the composition of output. Lilien (1982) developed an index that measures the standard deviation of the sectoral growth rates of employment from time period $t-1$ to time period t. For each region (or geographical area) of the country, the LI measures the variance in industry employment growth.

The LI is bounded from below at 0; that is, it takes the value 0 if there are no structural changes within one period. This index is a dispersion measure, and it considers the size (or share) of the sectors. In this article, the size is measured by the share of sectoral employment in total regional employment in a given period.

The LI is also considered a useful measure of the speed of structural changes. The higher the value of this indicator, the faster the structural changes and the bigger the reallocations of employment between branches (sectors or industries). It also indicates the ability of an economy to flexibly react and quickly adapt to changes in aggregate demand.[1]

Most of the literature on structural unemployment deals with the relationship between worker turnover[2] and unemployment. The literature on this issue is neither clear nor unambiguous. Different signs of the relations mentioned have indeed been introduced theoretically. There might be a positive relationship (Aghion and Blanchard 1994), a negative relationship (Krugman 1994), or an absence of relationship.

Many empirical works have added evidence on the nature of the relationship. For Italy, Basile et al. (2012) and Mussida and Pastore (2012) find a positive relationship between worker turnover and unemployment rates. In other words, larger turnover (high rates of layoff and hiring) is found in the higher unemployment regions of the South of Italy.

In addition, the Lilien hypothesis suggests that industrial restructuring causing sectoral shifts might explain the high level of turnover in high unemployment regions. The LI, therefore, is a useful tool to explore this theoretical hypothesis at a detailed (regional) geographical level.

The `lilien` command constructs the LI, originally proposed in Lilien (1982), to measure the variance in industry and total employment growth. The command calculates the index for two time periods (t and $t-1$) and allows users to export output in table format.

The data used to compute the LI must be in long format (see `help long`). Also an identifier variable representing the number of sectors (of employment) in each region is required. The data used might be either regional- or individual-level data.

The index is computed on the sum of employment in each sector, region, and time period (for example, year). One of the main advantages of the index is that it allows for geographically detailed investigations. The index, indeed, can be computed also at detailed geographical levels. In this article, we refer to the regional or nomenclature of units for territorial statistics as the second level of disaggregation.

1. Nevertheless, we should keep in mind that the LI only partially shows the region's ability to change, because the shift of resources that occurs within the framework of each sector cannot be captured by the LI. For this and other reasons that we will explain in section 2, other indicators of structural change, such as the modified Lilien index, have been introduced.
2. Worker turnover (WT) at time t is the number of accessions to employment from unemployment and inactivity plus the number of separations from employment to unemployment and inactivity, respectively. WT rates are computed by dividing WT by the average employment level (between $t-1$ and t).

For each region (or geographical area) of the country, the LI measures the variance in industry employment growth as follows,

$$\text{LI} = \sqrt{\sum(s_{irt}) \times \{\ln(x_{irt}/x_{irt-1}) - \ln(X_{rt}/X_{rt-1})\}^2}$$

where $s_{irt} = x_{irt}/X_{rt}$ is the sector i share in total regional employment in time period t, x_{irt} is the employment in sector i in region r, X_{rt} is the employment in the entire region, $\ln(x_{irt}/x_{irt-1})$ is the employment growth in sector i in time period t, and $\ln(X_{rt}/X_{rt-1})$ is the employment growth in the entire region in time period t.

2 Modified Lilien index

The need for additional indicators of structural change is well documented in the literature (Stamer 1997, 1999). For instance, Stamer (1999) shows that any indicator of structural change might ideally fulfill five "required" conditions:

1. The index should take the value 0 if there are no structural changes within one time period.
2. Structural change between two time periods should be independent of the time sequence.
3. Structural change in one time period should be smaller than or equal to structural change between two subperiods.
4. The index should be a dispersion measure.
5. The index should consider the weight (size) of the sectors.

Stamer (1999) demonstrates that the LI violates conditions 2 and 3. He shows that a small change is sufficient to solve this flaw and to obtain an index, the modified Lilien index (MLI), that fulfills all the required conditions.

The LI is modified by augmenting it with the weighting by the shares of the sectors in both periods. Hence, the influence or relevance of sector i is growing in proportion to its size and with respect to the value of its relative growth.

The availability of two measures of structural change (the LI and the MLI) leads to several advantages. First, the user has the opportunity to choose the index that best fits the aim of the user's analysis. Second, the user is not required to think that the choice of the indicator of structural change (either the LI or the MLI) is the only reason of the observed relationships. Finally, the choice of these indices allows for comparison of the results with other works done for Italy, Germany, and the UK.

Nonetheless, the similarities in the computation of the LI and the MLI imply a strong correlation between the two indicators of structural change.

The MLI, computed by including the option `method(MLI)` to the `lilien` command (see *Options* in section 3), measures the structural change as follows:

$$\text{MLI} = \sqrt{\sum(S_{irt}) \times \{\ln(x_{irt}/x_{irt-1}) - \ln(X_{rt}/X_{rt-1})\}^2}$$

$\ln(x_{irt}/x_{irt-1})$ is the employment growth in sector i in time period t, $\ln(X_{rt}/X_{rt-1})$ is the employment growth in the entire region in time period t, $S_{irt} = \text{avg\_}x_{irt}/X_{rt}$ is the average share of sector i in total regional employment for time periods t and $t-1$, and $\text{avg\_}x_{irt}$ is the average regional share of sector i in regional employment in time period t or the average sectoral employment in region r in time period t.

3 Command for LI and MLI

Lilien generates the LI variable for two time periods (t and $t-1$). The LI measures relative standard deviation of sector employment growth relative to overall growth in the region. There is only one restriction on the data to be used for a correct computation of the index: data cannot be negative. This does not limit the use of the command because employment data—most frequently used for LI computation—cannot be negative. The command also works on both balanced and unbalanced panel data. In the LI (and the MLI), `lilien` first computes the annual growth rates of employment (x) for each sector ($x_{it} - x_{it-1}/x_{it-1}$) and automatically excludes a sector with missing growth rates in a given time period (t). Data used might be either regional- or individual-level data.

The command computes the LI and gives an option (`method(MLI)`) to compute the MLI. The LI and the MLI are computed according to the formulas given in sections 1 and 2, respectively. Now we describe `lilien`'s options and provide some examples.

3.1 Syntax

`lilien` *varname*`, i(`*sec*`) j(`*time*`)` [`by(`*varlist*`) method(MLI) outfile replace`]

3.2 Options

`i(`*sec*`)` specifies the sector variable. `i()` is required.

`j(`*time*`)` specifies the time variable. `j()` is required.

`by(`*varlist*`)` allows groups defined by the variables listed in *varlist*, for example, `region`.

`method(MLI)` computes the MLI instead of the LI (the default).

`outfile` exports the output in `.csv` format by using *varname* as `Lilien_`*varname*`.csv`.

`replace` allows the user to overwrite the `.csv` file. `replace` should be used with `outfile`.

3.3 Examples

LI computed on the variable `x` by sector and time:

```
lilien x, i(sectors) j(time)
```

LI computed on the variable `x` by sector, time, and region:

```
lilien x, i(sectors) j(time) by(region)
```

LI computed on the variable `x` by sector, time, and region saved in a `.csv` format table (option `outfile`):

```
lilien x, i(sectors) j(time) by(region) outfile
```

The `outfile` option permits the export of the dataset in `.csv` format and saves the `.csv` file using *varname* as `Lilien_`*varname*`.csv`. In this way, if the user calls it many times in the same do-file, the results are not overwritten. To overwrite a file, use the options `outfile replace`.

MLI computed on the variable `x` by sector, time, and region:

```
lilien x, i(sectors) j(time) by(region) method(MLI)
```

4 Examples

Data for Italy suggest the presence of a positive relationship between worker turnover and the unemployment rate. Figure 1 displays that, indeed, the worker turnover rate is higher in regions with higher unemployment rates (in the South of Italy).

The map in figure 1 is obtained using the `spmap` package (Pisati 2007). By using a regional dataset on Italy that provides information on worker turnover, we plot the worker turnover by region.

Figure 1. Turnover rate by region

A higher degree of industrial change is one explanation for the high level of turnover in high unemployment regions. In fact, the Lilien hypothesis suggests that industrial restructuring causing sectoral shifts might explain the high level of turnover in high unemployment regions. To test the Lilien hypothesis, we first compute the LI by region and year by using the command described above. We obtain the `.csv` table showing the LI by year and region.

Then, by exploiting the available commands for Stata (`spmap`), we plot the regional LI to get a visual inspection of the distribution of the indicator across the Italian regions. Figure 2 displays the LI by region in Italy over the time period 2004–2010.

Figure 2. LI by region

Our data seem to confirm the Lilien hypothesis: the LI is higher in the regions where unemployment is also higher. The LI might also be used as a regressor for the worker turnover as a proxy of the relevance of the impact of structural change on worker turnover.[3]

3. Empirical evidence on LI is available in Basile et al. (2012) and Mussida and Pastore (2012).

Finally, by exploiting the available commands for Stata (`spmap`), we plot the regional MLI to get a visual inspection of the distribution of the indicator across the Italian regions. Figure 3 displays the MLI by region in Italy over the period 2004–2010.

Figure 3. MLI by region

The MLI is higher in the regions where unemployment is also higher. This is in line with expectations. The two indices for structural changes are indeed highly correlated.

5 References

Aghion, P., and O. Blanchard. 1994. On the speed of transition in Central Europe. *NBER Macroeconomics Annual* 9: 283–330.

Basile, R., A. Girardi, M. Mantuano, and F. Pastore. 2012. Sectoral shifts, diversification and regional unemployment: Evidence from local labour systems in Italy. *Empirica* 39: 525–544.

Krugman, P. 1994. Past and prospective causes of high unemployment. *Economic Review* Q IV: 23–43.

Lilien, D. M. 1982. Sectoral shifts and cyclical unemployment. *Journal of Political Economy* 90: 777–793.

Mussida, C., and F. Pastore. 2012. Is there a southern-sclerosis? Worker reallocation and regional unemployment in Italy. Discussion Paper No. 6954, Institute for the Study of Labor (IZA). http://ftp.iza.org/dp6954.pdf.

Pisati, M. 2007. spmap: Stata module to visualize spatial data. Statistical Software Components S456812, Department of Economics, Boston College. http://ideas.repec.org/c/boc/bocode/s456812.html.

Stamer, M. 1997. *Interrelations Between Subsidies, Structural Change and Economic Growth in Germany: A Vector Autoregressive Analysis*. Kiel, Germany: Kiel Institute of World Economics.

———. 1999. *Strukturwandel und wirtschaftliche Entwicklung in Deutschland, den USA und Japan*. Aachen, Germany: Shaker.

About the authors

Muhammad Rashid Ansari is an information technology research specialist at the INSEAD Business School in Singapore.

Chiara Mussida is an assistant professor at the Università Cattolica del Sacro Cuore in Piacenza, Italy.

Francesco Pastore is an assistant professor at the Seconda Università di Napoli in Santa Maria Capua Vetere, Italy.