Walter OBERHOFER, Uwe BLIEN, Alexandros TASSINOPOULOS

Forecasting Regional Employment with a Generalised Extrapolation Method

Unpublished paper of the IAB

**Corresponding author:** Uwe Blien

Addresses:

Uwe Blien, Alexandros Tassinopoulos  
Institute for Employment Research (IAB), Regensburger Str. 104, D-90327 Nuernberg, Germany  
E-mail.: uwe.blien@iab.de; alexandros.tassinopoulos@iab.de

Walter Oberhofer  
University of Regensburg, Universitätsstraße 31, D-93040 Regensburg, Germany,  
T.: 0049-941-943-2737, Fax: 0049-911-943-4917,  
E-mail: walter.oberhofer@wiwi.uni-regensburg.de
Forecasting Regional Employment with a Generalised Extrapolation Method

**Abstract:** In the paper an innovative method is applied to generate a forecast of regional employment. In common extrapolation problems the entries of a base matrix are used to estimate the elements of an unknown matrix. In the present problem of forecasting the base matrix is given for a time period in the past and the matrix to be estimated is related to a future period. In this paper a generalisation of this approach is used since the data about the past is given for a number of points in time and not only for a single period.

The Generalized Extrapolation Method (GEM) developed to take this information into account is related to well-known methods of panel analysis based on models with fixed effects. Additionally, it is similar to the techniques of entropy optimization. The objective function to compare the past and the future is a modified chi-squared statistic, which is approximately equal to the relative entropy function often used in gravity models and in input-output analysis.

The technique is demonstrated with respect to the example of employment in all 327 (western) German districts for a time span of two years. The method chosen uses any available information extensively. Therefore, the estimates are reliable, as is shown in an ex-post forecast.

Keywords: Regional Employment, Regression analysis, Entropy Optimization, Labor Markets, Econometrics
1. Introduction

The problem the paper is related to is the forecast of employment in the 327 districts of western Germany for a time horizon of two years. The data basis for the forecast is a panel of the employment in 11 industries and 327 districts (Stadt- und Landkreise) of western Germany for 11 years (1987-1997).

The main subject of the paper is the methodology necessary for the extrapolation. In addition a brief discussion of some theoretical aspects is given to motivate single steps of the method. The econometric techniques used are new and are presented here for the first time. They connect two large fields of statistical method, since they are related to regression analysis and to entropy optimization procedures.

In the next two sections of the paper the economic background and the data basis are presented. Then, in section 4, the strategy of entropy optimisation and related methods are described. The details of the Generalized Extrapolation Method GEM used here can be found in the section 5. Then, the results obtained are described and an ex-post forecast is evaluated. Finally, a conclusion is given.

2. Economic background and strategy of the forecast

Market economies are based on decentralised decisions of the economic subjects. In principle it can be seen only ex post whether these decisions are compatible. This is one of the reasons why forecasts of economic developments are very difficult. There are many other problems relevant in forecasts of the labour market, e.g. the lack of data and inconsistency problems with existing information (cf. the papers in Hejike 1994, e.g. Lindley 1994).

There is hope, however, that it is possible to gain forecasts of the labour market which are reliable up to a certain degree. The labour market reacts to changes of basic parameters of the economy with a lag, especially compared with the development of products markets. Firms react to changes in the demand for their products first by changing their capacity utilisation. Hires and layoffs come later. This lag can be used for forecasts of the labour market. It is subject to the discussion about leading indicators for economic development (Lahiri, Moore 1991). The lag is especially large in the case of Germany, since in this country the firms have to consider legal restrictions of layoffs, which are stronger than in the Anglo-Saxon countries. Therefore, German firms can be expected to follow at least partly a policy of labour-hoarding.
The forecast can be prepared by using information about employment expectations of the firms. Because of the lag between the demand situation on products markets and the personnel policy of the firms, it can be expected that the employment expectations are a relatively solid basis for future decisions of the firms. Additionally, theoretical arguments about regional specialisation and location are important to include causal factors in the preparation of the forecast:

(i) From Krugman’s (1991, cf. Krugman, Fujita, Venables 1999) analyses and from location theory (cf. Puu 1997 for an overview) it is known that regions tend to specialise in single industries. These industries show differing patterns of business cycles. Therefore, the exploitation of the composition of employment with respect to industries is important. There is a new discussion about the impacts of structural change on employment, partly following Appelbaum, Schettkatt (1993, 1994, Schettkat 1997). In this discussion it is shown that changes in the elasticity of demand on products markets may have important employment effects (cf. Möller 1998 for an empirical analysis and Blien 1999 for a regional application).

(ii) The location of a region is important for its fate as can be seen again from location theory. The quality of a location changes slowly according to changes in transportation technology, institutional restrictions and political changes. Since the advantages or disadvantages of the location of a region are relatively stable, this property can be used for causal analyses of regional developments and for forecasts.

(iii) The type of a region is important. There are strong trends which affect all regions of the same type. Due to a tendency towards sub-urbanisation in Germany there has been a loss of the importance of large city centres compared to the periphery of agglomerations in the last twenty years.

The information about the industry structure, the location of a district within a larger region and the type of the region is used to prepare the extrapolation. The approach chosen here implies a combination of top down and bottom up methods. On the one hand, in the forecast of the employment in the 327 districts (Stadt- und Landkreise) of western Germany, autonomous developments of these regional units are relevant. On the other hand, a projection of global employment is broken down into small regional units. There are global influences affecting employment in all regions in nearly the same way, for example the interaction of the business cycle with the institutions of the Federal Republic of Germany. It can be argued that the global component of regional development is of special relevance in Germany (with the specification that it is always necessary to discriminate between the western and the eastern part of the country), since important economic processes affect the whole country. Wage
bargaining for example is rather centralised, since it takes place at the level of industries, not at the level of regions or firms. Therefore, it would be misleading to realise a forecast for the country on the basis of regional information only.

The opposite would also be misleading, since it is not possible to see regional developments as entirely determined by general influences. Large variation in the level of local unemployment rates indicates large disparities on regional labour markets. The unemployment rate (average for the country for 1997) in western Germany was 10.8 % with the extreme values of 4.4 % (district of Erding) and of 20.1 % (district of Bremerhaven). It is therefore necessary to identify the sources of this variation and to supplement the top down method by bottom up strategies.

For the top down part of the forecast especially the proportion of the industries present in the regions is used (see argument (i) above). The importance of the industrial structure is known from the bulk of literature on shift-share analyses (Brown 1969, Holden et al. 1989, Patterson 1991, Selting, Loveridge 1994). The business cycle affects different industries in differing ways. Especially manufacturing reflects slumps and peaks in the demand on the world market to a greater extent than the service sector of the economy.

The business cycle of an industry is synchronised because of spill-over effects. An especially interesting example is Germany’s automobile industry. Phases of relatively good competitiveness compared to car production in other major countries (e.g. in Japan) are followed by phases of relatively low productivity and falling shares in sales on the world market. Though the performance of the individual companies in a country differs, there is a common trend affecting all car producers. The spill-over effects between the companies are described in general terms by the new growth theory (Lucas 1988, Romer 1986) and in more detail by evolutionary economics (Nelson, Winter 1982, Dodi, Nelson 1994, Nelson 1995). The growth perspectives of industries also differ as can easily be seen from the growing share of employment which is absorbed by the service sector and is described as tertiarisation of the economy.

3. Data

The data used in the forecast is from the employment statistics of western Germany. The data basis is not a sample; it comprises a sequence of complete cross-sections of all people employed on any 30 June of the period between 1987 and 1997 (statistics based on location of
The individual records of about 270 million employment spells available in the databank of our institute form the basis of the analysis.

The employment statistics include information about the entire population of people in gainful employment and covered by the social insurance system, i.e. about 80% of all employment in Germany. Two main groups of people whose data is not included are civil servants and workers with an income lower than DM 620 a month.

The employment statistics give continuous information on employment spells, earnings, job and personal characteristics. It is based on microdata delivered by individual firms about their individual employees. For every employee a new record is generated every year. The same is done if he or she changes establishment.

Since the employment statistics contain detailed regional information, an assessment of regional employment is possible. Every employed person is registered at his or her place of work. Several additional variables are available. Here, the information about the respective industry is used. This variable has 11 categories (cf. Table 1). The classification of industries follows the work on structural change (Dietz 1988) and has the advantage of separating in particular the various components of the service sector. The evaluation of the data of the tertiary sector showed that the overall gains are not equally distributed among all types of services. Especially the financial services and the services for the society (e.g. education, health service, research) show - in contrast to other types - a relative increase of people employed in their area. This implies a crucial differentiation between the types of service in order to forecast the growing importance of this sector with more accuracy. To control for industry-specific business cycles the forecast is based on a calculation of the prospects of each industry.

Table 1 about here

To include the time dimension in the data of the employment statistics, a panel is constructed. Its individual units are repeated measurements for employment of industries in districts.

The data is ordered in a sequence of matrices $X(t)$ (with $t = 1...T$). Each $X(t)$ contains the data of a specific year $t$ about the districts in the rows and about the industries in the columns. $x_{ij}(t)$ and represent the single entries of the matrix $X(t)$, $i$ is the index of regions, i.e. districts,
j the index of industries. If \( x_i \) is total employment in region \( i \) and \( x_j \) total employment in industry \( j \), the data of a single point in time \( t \) can be represented by Table 2. The primary aim of the forecast is to estimate the future distribution of employment for all regions \( i \), that is to extrapolate the single \( x_j \) by taking into account the complete data of \( X(t) \), all available \( t \).

Table 2 about here

In the present case the data refer to eleven years (\( t = 1...11 \)), since the time span under observation is 1987 to 1997. The data matrix \( X(t) \) is of the dimension 327 to 11, since in the rows \( i \) it shows the districts of western Germany (with \( i = 1...327 \)) and its columns \( j \) represent the industries (with \( j = 1...11 \)). In addition to the sequence of matrices \( X(t) \) external information is included, which is of heterogeneous form. Normally, it is available only referring to aggregates, i.e. with respect not to a \( x_{ij} \), a specific industry of a single district, but with respect to larger units. These are either industries \( x_j \) or larger regions, represented by sums of \( x_i \) for districts which belong e.g. to the same federal state. Sources of separate forecasts for special industries are assessments based on surveys about firms’ business expectations. Usually the results of this kind of surveys are available only for the whole country, not for small regional units.

4. Methods of entropy optimization and \( \chi^2 \)-criterion

In principle it would be possible to integrate both sources of information (from the employment statistics and from external sources) by estimating first a base matrix \( X_b \) (with \( X_b = X(T + h) \) and \( h \) being the ‘horizon’ of the forecast) of employment for the target year by extrapolation within a panel model (cf. Baltagi 1995). In a second step a matrix \( Y \) (with \( Y = \hat{X}(T + h) \)) could be estimated which satisfies the requirements specified by the additional external information and which is most similar to the base matrix \( X_b \). In the following we concentrate on this single step and show later in section 5 how the estimation of a panel model and the inclusion of additional information can be integrated in one approach.
The extrapolation of X to Y task could be done by minimising a distance function between X (we drop the subscript b in the following) and Y under the constraints given by additional external information. A distance function with ‘nice’ properties is a modified relative entropy:

\[ d = E_R(Y,X) = \sum_{i,j} y_{ij} \left[ \ln \left( \frac{y_{ij}}{x_{ij}} \right) - 1 \right] + \sum x_{ij} \]  

(1)

This is a modification of the standard relative entropy:

\[ E'(Y) = \sum_{i,j} y_{ij} \ln \left( \frac{y_{ij}}{x_{ij}} \right) \]  

(1)’

A minimisation of \( E' \) without any constraints gives \( y_{ij} \) not identical to the \( x_{ij} \). They differ by the factor \( e^{-1} \). This is a disadvantage. To avoid this effect (1) instead of (1)’ is used, which is minimised subject to the following constraints:

\[ Ay = b \]  

(2)

The vector \( y \) is obtained by setting \( \text{vec}(Y) = y \), with \( \text{vec}(\cdot) \) being the operator which transforms a matrix to a vector by stacking the columns of the matrix. In this section it is assumed that the system of constraints is consistent. If \( b \) contains the row and column sums, optimising (1)’ produces the same results as optimising (1). If \( x_{ij} \) and \( y_{ij} \) do not represent absolute but relative frequencies, which could be interpreted as probabilities, (1) is obviously reduced to (1)’. In this case \( E(Y) \) is called the "information gain" (following Kullback 1968). For all \( x_{ij} = 0 \), the corresponding \( y_{ij} \) are set \( y_{ij} = 0 \). If only row and column sums are specified the matrix \( A \) contains \( r=n+m \) rows. This is the number of constraints. The rank of \( A \) is \( n + m - 1 \). Minimising the relative entropy leads to the well-known equation:

\[ y_{ij} = x_{ij} \alpha_i \beta_j, \quad 1 \leq i \leq n, 1 \leq j \leq m \]  

(3)

with \( \alpha_i \) and \( \beta_j \) as dual variables (Lagrange multipliers). These can be obtained iteratively by using the RAS-method (Deming, Stephan 1940, Bacharach 1970), which is identical with the Iterative Proportional Fitting Algorithm (IPF) used in log-linear models (cf. Bishop et al. 1975).

\[ x_{ij}^{r+1} = \frac{x_{ij}^t b_j^t}{\sum_p x_{ip}^t} \quad \text{all i and j} \]

\[ x_{ij}^{r+1} = \frac{x_{ij}^t b_j^t}{\sum_p x_{pj}^t} \quad \text{all j and i} \]
Here, the $b_i^r$ represent the row and the $b_j^c$ the column sums, $s$ is the step in the iteration process. In one step the RAS operations are carried out for all rows and then for all columns. If in (3) it is set $\alpha_i = R_i$, $\beta_j = S_j$, and $x_{ij} = A_{ij}$, it can easily be seen why the iterative procedure normally applied to find the solution is called the RAS method. The property of the solution, that all resulting $y_{ij}$ can be calculated by using constant factors specific for the respective row $i$ and column $j$ is called "double proportionality".

The RAS-method is a very simple procedure, including only linear operations, which has contributed much to the widespread use of the relative entropy as an objective function in solving extrapolation problems, especially in input-output analysis. Other applications of entropy optimization procedures are related to gravity models (cf. Wilson 1970, Nijkamp 1975, Nijkamp, Reggiani 1992).

By using the RAS method or an equivalent procedure, a matrix $Y$ is generated that is similar to the base matrix $X$, since small elements of $Y$ are small in $X$. This effect is called the property of structure conservation. Additionally, under the assumption of a multinomial distribution, the estimated matrix is the most probable one, given the specified constraints (Blien, Graef 1998). The entropy function was used in an employment forecast (Blien, Tassinopoulos 2000) with the same data base as in the present paper. For these estimations the ENTROP algorithm (Blien, Graef 1998) was adapted. Here, instead of the relative entropy the $\chi^2$-criterion is used:

$$\chi^2(Y, X) = \sum_{i,j} \frac{(y_{ij} - x_{ij})^2}{x_{ij}}$$  \quad (4)

There is only a slight difference between the two objective functions, since they are identical up to a Taylor approximation. (1) can be written as

$$E(Y) = \sum_{i,j} y_{ij} \left[ \ln \left( \frac{y_{ij} - x_{ij}}{x_{ij}} \right) + 1 \right] + \sum_{i,j} x_{ij}$$  \quad (5)

$$\approx \sum_{i,j} y_{ij} \frac{y_{ij} - x_{ij}}{x_{ij}} - \sum_{i,j} y_{ij} + \sum_{i,j} x_{ij}$$  \quad (6)

$$= \sum_{i,j} \left( y_{ij} - x_{ij} \right)^2 / x_{ij} = \chi^2$$  \quad (7)

In step two the Taylor approximation $\ln(u+1) \approx u$ is used, which is especially accurate if $|u| << 1$. Therefore, the weighted sum of squares, i.e. the $\chi^2$-criterion defined by (7), is close
to the relative entropy, if \(|y_{ij} - x_{ij}| / x_{ij} << 1\) (the proof is slightly modified after Kádas, Klafszky 1976).

It is possible to use the \(\chi^2\) criterion instead of the relative entropy to estimate a matrix \(Y\) on the basis of a matrix \(X\) subject to linear constraints. Since the relative entropy and the \(\chi^2\)-criterion are only approximately equivalent, the size of the deviations in practical applications of the two functions is important. Additionally, it is important by which function better results are generated. This can be assessed by using known matrices \(X\) and \(Y'\) and checking the deviation between \(Y'\) and the extrapolated \(Y\). This experiment was carried out with input-output tables. The results differed only slightly and the comparison with the known exact table showed that there is no clear ranking of the objective functions. In some cases the results of an extrapolation were better if the weighted sum of squares was used, in others the application of the relative entropy was superior (Oberhofer 1998).

Therefore, in many cases it is better to use the \(\chi^2\)-criterion, since its optimization leads to simple linear expressions and there is a close relationship to standard techniques of regression analysis. Both aspects will be made clearer in the following.

An application of the \(\chi^2\) criterion implies the minimisation of this Lagrange function:

\[
\sum_{i,j \in \mathbb{N}} \frac{(y_{ij} - x_{ij})^2}{x_{ij}} = 2\lambda' \left( \text{Avec}(Y) - b \right) \to \text{Min} \quad (75')
\]

To simplify the following exposition a special case is regarded, where only the row and the column sums of the matrix \(Y\) are specified. The first order conditions can be written as:

\[
y_{ij} = x_{ij}(1 + \alpha_i + \beta_j), \quad 1 \leq i \leq n, \quad 1 \leq j \leq m, \quad (8)
\]

Again \(\alpha_i\) and \(\beta_j\) are elements of the vector \(\lambda\) with the Lagrange multipliers. Comparing (8) and (3) it is obvious that (8) has a simpler structure, since it is a linear system, which can easily be solved. Formally (8) is equivalent to a two-way panel model with fixed effects, if the model is formulated in growth rates. The effects \(\alpha_i\) and \(\beta_j\) can be calculated exactly from the specification of the constraints and are not gained by an estimation.

5. The Generalised Extrapolation Method

The optimization of the \(\chi^2\)-criterion is one central property of the “General Extrapolation Method” (GEM) applied here. But this method differs from previous work in more than one respect. A special advantage of the new method is the inclusion of ‘soft’ constraints. It is possible to specify constraints which do not have to be exactly binding, but could be violated to a
certain degree. The extent to which this violation is tolerated can be specified ex-ante. An additional advantage of the new method is the fusion (in fact integration) of this extrapolation technique with regression analysis.

The application of the generalised extrapolation method includes three steps. In the first step, which is a preparatory one, the external information, which is available about the matrix to be estimated, has to be made consistent. In the second step an appropriate regression model has to be specified and in the third step the estimations are carried out. We start the outline of the method with the preparatory step.

5.1 Consistent constraints

External constraints are important in solving the problem of extrapolation. An obstacle often encountered in empirical work is that these constraints may be inconsistent, since the data basis can include errors. In this case the set of solutions of the extrapolation problem is empty.

The errors in the data can be caused simply by "noise" or because of sampling errors. An additional possibility is that the constraints may be based on expert judgements which are accurate only up to an assessment error. The constraints may be inconsistent because the underlying definitions of facts (e.g. the specifications of industries) may differ or they may refer to differing time periods. A typical source of all these inconsistencies is that the constraints may be taken over from various external sources and the design of the surveys, which is the basis of the information used, may vary.

In the following a procedure is discussed that has the purpose of transforming the constraints to make them consistent. The correction is carried out in a way that minimises changes. Only linear constraints in the form of equations (2) are taken into account. The matrix $Y$ represents the dependent variables i.e. it contains the distribution of employment (the number of employees in defined regions and industries) at the year $T + h$. In the following, it is set $h = 1$ for reasons of simplicity. $Y$ has $n$ rows and $m$ columns. $A$ is an $r \times k$ matrix of coefficients with $r < k = nm$. $r$ is the number of constraints and the vector $b = (b_f)$ - consisting of all the $f = 1...r$ elements $b_f$ - the right hand side of the equations. If the rank of $A$ is smaller than $r$, the system (2) might be inconsistent.
In many extrapolation problems only the row and the column sums are given. Then, the matrix $A$ contains only ones and zeroes and the vector $b$ represents the row and column sums. Because the row sums and column sums are not independent, the matrix $A$ is singular.

To correct inconsistent constraints it is assumed that the right hand side of (1) is stochastic:

$$b = c + u$$

The vector $u$ represents an unobserved error term with a diagonal and nonsingular variance-covariance matrix $D = (d_{ij})$. The variance $d_{ii}$ is a measure of the accuracy of $b_i$. To obtain a consistent system of constraints, the right hand side $b$ has to be substituted by an estimated $\hat{b}$ calculated by using an appropriate vector $z$ so that:

$$\hat{b} = Az$$

with $\hat{b}$ as close as possible to $b$. To find such a vector $z$ the following minimisation problem has to be solved:

$$(b - Az)'D^{-1}(b - Az) \rightarrow \text{Min}$$

The first order conditions are:

$$\begin{align*}
(A'D^{-1}A)z &= A'D^{-1}b \\
\end{align*}$$

(9)

The system of $k$ unknown variables $z_i$ has always at least one solution though the target matrix of coefficients

$$A'D^{-1}A$$

is singular. It has the same rank as $A$, i.e. in general a rank smaller than $k$. Since the right hand side

$$\hat{b} = Az$$

is independent of the special solution of (9), the Moore-Penrose inverse can be used to find a unique solution:

$$z = (A'D^{-1}A)^+ A'D^{-1}b$$

(10)

Thus, the problem of finding consistent constraints is reduced to the problem of computing a Moore-Penrose inverse to solve the system of linear equations.

5.2 Extrapolation with panel data and regression analysis

It is now assumed that data is available which has a panel structure, like the one shown in section 3. The data is interpreted as consisting of a sequence of matrices:
The aim of the extrapolation implies the calculation of a matrix $Y = X(T+1)$. Again the resulting matrix is subject to $r$ linear constraints

$$\text{Avec}(Y) = b \quad \text{(2)}$$

Additionally, it is assumed that the sequence of $X(t)$, with individual elements $x_{ij}(t)$, can be modelled in a simple regression approach:

$$x_{ij}(t) = \mu_{ij} + u_{ij}(t), \quad 1 \leq i \leq n, 1 \leq j \leq m \quad \text{(11)}$$

As usual, the $u_{ij}(t)$ are assumed as being mutually independent:

$$\text{Var}(u_{ij}(t)) = v_{ij}, \quad 1 \leq i \leq n, 1 \leq j \leq m$$

According to the regression approach (11) the $\mu_{ij}$ are estimated by $\hat{\mu}_{ij}$. A simple extrapolation is possible by setting:

$$y_{ij} = \hat{\mu}_{ij}, \quad 1 \leq i \leq n, 1 \leq j \leq m$$

To include the constraints (2) in the estimations, the results must satisfy the following condition (with $M = (\mu_{ij})$):

$$\text{Avec}(M) = b$$

Then, the Lagrange equation is

$$\sum_{i,j,t} \left( \frac{(x_{ij}(t) - \mu_{ij})^2}{v_{ij}} - 2x'(t)\text{Avec}(M) - b \right) \rightarrow \text{Min} \quad \text{(12)}$$

If $T = 1$ and $v_{ij} = x_{ij}(1)$ the results are equivalent to the extrapolation $(7)'$, as can be easily seen. The extrapolation is:

$$Y = \hat{M}$$

5.3 Discussion

The relative entropy is widely used to solve problems of extrapolating a matrix $X$ to a matrix $Y$. Since this objective function is approximately equivalent to the $\chi^2$ criterion, the approach chosen here is a fusion of extrapolation techniques related to entropy optimization with regression analysis. Two separately developed branches of econometric methods and statistics turn out to be closely related.

This interpretation is the basis of several generalisations:

1. The regression equation can be modified. Instead of using the simple equation (11) various kinds of panel models (as discussed in Baltagi 1995) with time dependent co-variables can be estimated.
2. A sequence of matrices can be used instead of only one basis matrix. Auto-regressive models are also possible.

3. The variances $v_{ij}$ can be modified to express special weightings of the data.

A relatively simple generalization of the basic model can be seen from the application of the method in the next section. Advantages of the new technique are its flexibility and the extensive use of the available information. It is possible to use the full information content of the data in one estimation process together with external information. To include theoretical considerations the models can be adapted in many ways. The extrapolation can be based on various influences represented by exogenous variables, which are regarded as important. The trend of the development can be directly taken from panel (X(1)...X(T)), which is the basis of the forecast. Additionally, all kinds of external information, expert judgements etc. can be included in the form of external constraints. Therefore, forecasts generated with the new method can be expected to be more reliable than those based on traditional techniques.

The treatment in section 5.1 shows that it is possible to include information which is affected by errors and therefore inconsistent. According to an a-priori assessment, the size of these errors can be included in a step of calculations carried out before starting the genuine estimations. This first step facilitates the inclusion of ”soft” information, which is valuable and important to be observed, but which is not very accurate.

In previous work with entropy optimising procedures, especially with the ENTROP method (cf. Blien, Graef 1998, Blien, Tassinopoulos 2000), soft information is observed by including inequalities in the estimation process. The ENTROP procedure is a generalisation of the traditional RAS method, since not only the row and the column sums can be included, but all kind of linear equations and inequalities.

To specify soft information with ENTROP, it is assumed that the exact value $b_k$ of the right hand side of a constraint $k$ lies between two bounds $b_k^l \leq b_k \leq b_k^u$. Two cases are possible for an individual constraint. If a constraint is binding, it could be substituted by an equation, e. g. by $b_k^l = b_k$. If it is not, it could be dropped without changing the result of the extrapolation. However, the researcher is often not interested to specify exact bounds, but to obtain a result which is close to $b_k$. Then $b_k$ is measured with an error $b_k = b_k + u_k$, as it is assumed in section 5.1. In this case the method introduced here is more appropriate.
6. Preparation of the forecast and results

The data basis for the forecast were the matrices \( X(t) \) of the employment statistics for the years \( t \) from 1987 to 1995. The aim of the analysis was the extrapolation to 1997. Since the actual employment matrix of 1997 is known, this is an ex post forecast and in a comparison the error of the calculated matrix could be obtained.

Additionally to the data of the employment statistics ‘external’ information of the year 1995 about employment in 1997 was available. These sources were mainly based on the Establishment Panel of the Institute for Employment Research (IAB). The IAB has built up this survey to gain information about the structures and developments of the demand side of the labour market. Major determinants of corporate employment and personnel policies can be analysed. This survey is the only existing establishment panel representative for employment of a large economy (Bellmann et al. 1996).

For the present forecast a question of the 1995 wave of the panel was used that concerned the prospects of employment in the respective establishments. Since the IAB panel is not representative for small regional units this information had to be aggregated according to industries. The results could then be included as external information related to column sums \( x_{i,j} \) of the matrix \( \hat{X}(1997) \) to be estimated.

In addition forecasts on industry specific business developments were included, which were published 1995 by the German research institutes IFO and IW. Since the available external data were contradictory, they were made consistent in the described way (cf. Section 5.1).

Then, the following simple regression model was used, instead of (11):

\[
x_{i,j}(t) = a_{i,j}x_{i,j}(t-1) + u_{i,j}(t) \quad 1987 \leq t \leq 1995, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m
\]  

(13)

The estimates were obtained by minimising the weighted mean square i.e. the \( \chi^2 \)-criterion.

\[
\sum_{i,j,t} (x_{i,j}(t)-a_{i,j}x_{i,j}(t-1))^2 / v_{i,j}
\]

with

\[
v_{i,j} = \sum_t x_{i,j}(t)
\]

(14)

This is equivalent to a feasible generalized least squares procedure.
The model used is a very simple one. Since individual coefficients $a_{ij}$ are estimated, region specific effects are included and so are effects of variables which do not change in time e.g. the type and the location of a region. More sophisticated models are possible, the one used here serves primarily as a reference to test whether the method is useful.

All $n \cdot m = 327 \cdot 11$ elements of the matrix were based on a uniform regression model. The matrix elements follow individual trends which differ much. Therefore, the extrapolation of the status quo situation

$$\hat{x}_{ij}(1997) = x_{ij}(1995)$$

is robust and – in many cases – better than uniform regression models. Consequently, a status-quo extrapolation was used as a base line for a comparison with the regression extrapolation. To assess the value of the external information, a pure regression extrapolation was carried out without any external constraints.

As a measure for the error the mean absolute deviation of the ‘real’ from the extrapolated employment in per cent $F_1$ was used:

$$F_1 = \frac{100}{nm} \sum_{i,j} \left| x_{ij}(1997) - \hat{x}_{ij}(1997) \right| / x_{ij}(1997)$$

(15)

This error measure implies the problem that the deviation of small $x_{ij}$ is weighted relatively high. To avoid this disadvantage, the $\chi^2$-criterion can be used as an error measure $F_2$:

$$F_2 = \frac{100}{nm} \sum_{i,j} \left( x_{ij}(1997) - \hat{x}_{ij}(1997) \right)^2 / x_{ij}(1997)$$

(16)

Table 3 shows a comparison of the reliability of the differing forecast methods based on the two error measures. Row one with the naive status-quo extrapolation serves as a reference.
For the time span of two years, which is relevant here, the status quo extrapolation generates a result, which is better than a pure regression extrapolation without including any external information. The forecast obtained by the generalised extrapolation method, which improves the regression by using external information, is better than both the simple regression and the status-quo forecast.

7. Conclusion

In this paper a new method is presented. The Generalized Extrapolation Method GEM is appropriate for forecasting and for other extrapolation problems. Here, the results of an application in forecasting regional employment for a two-year time span is shown.

Even by using a simple version of the regression model, which is one genuine part of the new generalised extrapolation method, results are obtained which are better than the status-quo forecast. It can be expected, that the results could be improved significantly, if more exogenous variables from the employment statistics would be included. The discussion about shift-share techniques in regional economics shows that only a smaller part of regional developments can be explained by the structure of industries, which is the most important set of variables here. The GEM used here has the advantage of taking into account extensively any available information and is therefore preferable to standard shift-share or regression techniques.

The Generalised Extrapolation Method can be used for many problems of empirical economics. Besides forecasting, the disaggregation of aggregate units is one possible application. Others are the estimation of transition or flow matrices, e. g. in Markov processes. The matrices could be from input-output analysis, from migration analysis or from demographic accounting.

Application in forecasting can help to satisfy the demand for prospective information in regional and active labour market policy. Strict tests of economic theories are also possible.
References:


Balderjahn, Ingo; Mathar, Rudolf; Schader, Martin (1998) (eds.): "Classification, Data Analysis, and Data Highways" (Proceedings of the 21st Annual Conference of the Gesellschaft für Klassifikation), Berlin etc.: Springer


Blien, Uwe; Tassinopoulos, Alexandros (1998): "Forecasting Regional Employment with the ENTROP Method", paper prepared for presentation at the EALE Conference, Blankenberge


Heijke, Hans (1994) (ed.): "Forecasting the Labour Market by Occupation and Education", Boston etc.: Kluwer


Möller, Joachim (1998): "Income and Price Elasticities in Different Sectors of the Economy. An Analysis of Structural Change for Germany, the UK and the USA", unpublished paper of the University of Regensburg

Möller, Joachim; Tassinopoulos, Alexandros (2000): "Zunehmende Spezialisierung oder Strukturkonvergenz?", in: Jahrbuch für Regionalwissenschaft 20/1: 1-38


Nijkamp, Peter; Reggiani, Aura (1992): "Interaction, Evolution and Chaos in Space", Berlin etc.: Springer


Table 1: Industries (IAB typology):

<table>
<thead>
<tr>
<th>1. Primary sector</th>
<th>7. Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Energy/mining</td>
<td>8. Distributive services</td>
</tr>
<tr>
<td>Goods-producing industry</td>
<td>9. Financial services</td>
</tr>
<tr>
<td>5. Consumer goods</td>
<td>11. Services for society</td>
</tr>
<tr>
<td>6. Manufacture of food products and beverages</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Data matrix $X(t)$

<table>
<thead>
<tr>
<th>Regions</th>
<th>Industries</th>
<th>$\Sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{11}$ $x_{12}$ $x_{13}$ $x_{14}$ $x_{15}$ ...</td>
<td>$x_{1\cdot}$ $x_{2\cdot}$ $x_{3\cdot}$ ...</td>
<td></td>
</tr>
<tr>
<td>$x_{1m}$ $x_{21}$ $x_{22}$ $x_{23}$ $x_{24}$ $x_{25}$ ...</td>
<td>$x_{1\cdot}$ $x_{2\cdot}$ $x_{3\cdot}$ ...</td>
<td></td>
</tr>
<tr>
<td>$x_{2m}$ $x_{31}$ $x_{32}$ $x_{33}$ $x_{34}$ $x_{35}$ ...</td>
<td>$x_{2\cdot}$ $x_{3\cdot}$ ...</td>
<td></td>
</tr>
<tr>
<td>$x_{3m}$ $x_{41}$ $x_{42}$ $x_{43}$ $x_{44}$ $x_{45}$ ...</td>
<td>$x_{3\cdot}$ ...</td>
<td></td>
</tr>
<tr>
<td>$x_{nm}$ $x_{n1}$ $x_{n2}$ $x_{n3}$ $x_{n4}$ $x_{n5}$ ...</td>
<td>$x_{n\cdot}$ ...</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Error of differing forecast methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean absolute error (percent)</th>
<th>$\chi^2$-criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status-quo extrapolation</td>
<td>7.99</td>
<td>40.88</td>
</tr>
<tr>
<td>Regression extrapolation without external information</td>
<td>9.05</td>
<td>47.99</td>
</tr>
<tr>
<td><strong>Generalized extrapolation method GEM</strong> (regression extrapolation with external information)</td>
<td>7.14</td>
<td>30.33</td>
</tr>
</tbody>
</table>