Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 48-1 (2015) 115–120

Identifying Structural Changes in Austrian Social Insurance Data ⋆

Thomas Ortner * **Peter Filzmoser** * **Gottfried Endel** **

* *Department of Statistics and Probability Theory, Vienna University of Technology, Austria, (e-mail: thomas.ortner@tuwien.ac.at, p.filzmoser@tuwien.ac.at)*

** *Main Association of Austrian Social Security Institutions, Austria (e-mail: gottfried.endel@hvb.sozvers.at).*

Abstract:
Testing for structural changes is a well studied field. Classical tests for breakpoint detection utilize F-statistics which depend on independent and identically normal distributed residuals. In general, this condition is not satisfied which leads to distorted test results when the p-values of classical tests are close to the significance level. Thus, permutation tests are used to properly estimate the critical values.
Based on the accounting data of Austrian hospitals, collected by the Austrian social insurance institutions, specific observations (hospital stays) which are connected to pre-defined diseases are analysed. For those groups of observations we use characteristic factors, to test for structural changes from different perspectives. The first test analyses the temporal trend and identifies breakpoints, caused by changes in the underlying system between years. The second analysis focuses on identifying differences between hospitals. Both implemented tests ensure the often ignored aspect of a homogeneous data base for further analysis.

Keywords: Breakpoints, Structural Change, Permutation Test, Cut Off Value, Time Series

1. INTRODUCTION & RELATED WORK

Instability of parameters in models is a common phenomenon in econometrics (Bai, 1997). Naturally there exists a well developed theory for a certain class of models. For independent and identically distributed variables, multiple regression models have been studied by Bai (1994), among others. One of the first tests for structural change, defined by alterations of the parameters of a describing underlying model, was introduced by Chow (1960). The idea emerged from analysing economical relations which are often assumed to be linear. Therefore the properties of linear regression models are utilized for the proposed test which will be presented later.

The focus of this paper is the adaption of existing methods to specific characteristics of data, supplied by the Austrian social insurance. Especially economic data, recorded in the context of hospital stays are of interest. The time series of characterizing factors of those observations, which are created by aggregation methods are assumed to be depending on certain regressor variables, but also on other factors which cannot be measured. Examples for those factors are the introduction of new treatments or changes in health politics.

Whenever system-changing decisions are made at the basis of historical data it is of high importance to know, whether there has been a structural change in these data or not. One simple example would be a re-distribution of money for hospitals because of new rating systems for medical services. Such a new rating could be associated with changes in the coefficients of the underlying model and makes it easy to properly analyse and readjust the distribution of money.

1.1 Chow Test

One of the first tests for structural change and most likely the most famous one, has first been introduced by Chow (1960). It is assumed, that a multivariate time series of $\boldsymbol{x}_t$ describes a univariate time series y_t properly using a linear regression model. Let $\boldsymbol{y}$ be defined as $(y_1, \ldots, y_n)$ and $\boldsymbol{X}$ as $(\boldsymbol{x}_1', \ldots, \boldsymbol{x}_n')'$, where $'$ defines the transposed and $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})'$, for $i = 1, \ldots, n$. We further assume that the first element of each $\boldsymbol{x}_i$ is equal to 1, thus accounting for an intercept in the following regression model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \tag{1}$$

To be able to estimate the regression parameters $\boldsymbol{\beta}$ properly using the least-squares estimation we further assume, that $n > p$ holds and $\boldsymbol{X}$ is of full rank. The null-hypothesis, tested by Chow, states that two disjoint models $\boldsymbol{y}_1 = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1$ and $\boldsymbol{y}_2 = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}_2$ describe $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ at the same quality as the joint model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, using the same regression coefficients, where $\boldsymbol{y}_1 = (y_1, \ldots, y_{t_0})$,

⋆ This work has been partly funded by the K-project DEXHELPP through COMET - Competence Centers for Excellent Technologies, supported by BMVIT, BMWFI and the province Vienna. The COMET program is administrated by FFG.

10.1016/j.ifacol.2015.05.152

$\boldsymbol{y}_2 = (y_{t_0+1}, \ldots, y_n)$, $\boldsymbol{X}_1 = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t_0})$ and $\boldsymbol{X}_2 = (\boldsymbol{x}_{t_0+1}, \ldots, \boldsymbol{x}_n)$.

$$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta} \tag{2}$$

For testing this hypothesis, first the same assumptions, used for formula (1) need to be fulfilled. Thus, it is assumed, that $p < \min(t_0, n - t_0)$ and that neither $\boldsymbol{X}_1$ nor $\boldsymbol{X}_2$ are singular. Taking advantage of various F tests (Mood, 1950; Davis, 1952) a test statistic is defined by

$$F_C = \frac{||\boldsymbol{X}_1\boldsymbol{\beta}_1 - \boldsymbol{X}_1\boldsymbol{\beta}||^2 + ||\boldsymbol{X}_2\boldsymbol{\beta}_2 - \boldsymbol{X}_2\boldsymbol{\beta}||^2}{||\boldsymbol{y}_1 - \boldsymbol{X}_1\boldsymbol{\beta}_1||^2 + ||\boldsymbol{y}_2 - \boldsymbol{X}_2\boldsymbol{\beta}_2||^2} \frac{n-2p}{p}. \tag{3}$$

Under the assumptions above, the additional assumption of indipendent and identically normal distributed errors $\boldsymbol{\epsilon}$ and the null hypothesis, F_C follows a central F-distribution $F_{p,n-2p}$. An implicit assumption in this test is the knowledge of t_0. The problem with this assumption is, that the choice of t_0 directly affects the alternative hypothesis. This is pointed out and discussed in Davies (1977).

1.2 *strucchange*

Based on the work of Andrews (1993), Zeileis et al. (2001) implemented the R package *strucchange* for identifying structural changes with unknown breakpoints. There, two different approaches for identifying structural changes are implemented. Empirical fluctuation processes are computed for which the limiting processes are known and certain aggregations of F_C are introduced for testing for structural changes in intervals.

The first approach defines two different processes. The CUSUM process contains cumulative sums of standardized residuals following an approach of Brown et al. (1975). It can be shown, that the cumulative sum follows the standard Brownian motion (Borodin and Salminen, 2002). The MOSUM process analyses the moving sum of residuals instead of using cumulative sums. As shown by Chu et al. (1995), the limiting process for MOSUM processes are increments of Brownian motion. For both processes, the decision, whether there exists a structural change or not is made by testing if the process exceeds the defined boundary.

A second, in *strucchange* implemented approach, which is closer to Chow (1960) uses direct aggregations of F_C. Following the suggestions of Andrews (1993) and Andrews and Ploberger (1994), the following statistics have been implemented:

$$supF = \sup_{\underline{i} \le i \le \bar{i}} F_i, \tag{4}$$

$$aveF = \frac{1}{\bar{i} - \underline{i} + 1} \sum_{i \in \{\underline{i}, \ldots, \bar{i}\}} F_i, \tag{5}$$

$$expF = log\left(\frac{1}{\bar{i} - \underline{i} + 1} \sum_{i \in \{\underline{i}, \ldots, \bar{i}\}} exp(0.5F_i)\right), \tag{6}$$

where F_i are the Chow statistics defined by (3) for $t_0 = i$ and $\{\underline{i}, \ldots, \bar{i}\}$ defines the possible breakpoints. The aggregated statistics from (4) to (6) are obviously not F-distributed any longer but they are members of a distribution family of aggregated F-statistics. This family was analysed by Andrews (1993). Asymptotic results allow the estimation of critical values for statistics from this family. *strucchange* utilizes these results and implements the simulation results by Hansen (1997) as critical values. Note that for both, the Chow-statistic, defined by (3) and the aggregated statistics (4) to (6), the assumptions stay the same. Moreover it is necessary that all three models, used in (3) need to be of full rank to ensure (3) is well-defined.

2. DATA STRUCTURE & MODEL

The majority of tests on structural change are based on linear models, since they emerged from economic theory where mostly linear models are applied. There also exists a wide theory for structural changes in generalized linear models especially for count data (e.g. Doukhan and Kegne, 2013; Franke et al., 2012) and binary data (e.g. Hudeocova, 2013). In general, real world data cannot be properly modelled long term using linear models. While linear regressions are often useful for local approximation of time series, there usually lies a higher polynomial or exponential trend underneath. We therefore propose the use of functions similar to piecewise linear functions.

The data structure for our application is constructed from accounting data from Austrian hospitals. For each hospital stay certain information is collected, including the main diagnosis, some patient information and the performed services combined with the number of LDF points. LDF points are the Austrian version of a DRG system which determines how much money a hospital receives for their performed services (BMG, 2010). This structure allows us to define intervals of an average length of 15 days where the number of patients, number of services, sum of LDF points etc., are measured. This definition of intervals ensures 24 observations per year for the time period of 10 years from 2001 to 2010. Therefore 240 observations of all of the mentioned variables are available for distinct hospitals as well as distinct groups of patients with predefined diagnoses. Changes in the Austrian DRG system cannot be implemented during the year. Thus, it is not necessary to test for structural change for each observation but only at the end of each year.

A variety of predictors are available for describing the extracted time series. It is well known that the number of services is strongly depending on the age of the observed population. The consequence would be to model the time series, using the distribution of the population by age and gender as well as a time component. Since only yearly population data is available, the observations in between are estimated using splines (Forsythe et al., 1977), where an exact cubic is fitted through the four points at each end of the data. Even for this non-linear estimation of the age and gender related time series the data structure of the predictors is highly correlated. We used various classification intervals for the aggregation of the population according to age intervals. One representative example of the correlation structure is given in table 1. m_1 denotes the male population between 0 and 50 years, m_2 the male