Nr. 1, 2006

Cliometrics of Academic Careers and the Impact of Infrequent Large Shocks in Germany before 1945

Olivier DARNÉ, Claude DIEBOLT
Abstract: Using the outliers methodology, the focus of this paper is to present the time series dynamics of the German higher education system before 1945. The outline of the paper is as follows. In Section 2, we define the outliers and describe the outliers identification procedure. We apply this procedure to the annual series of the students enrolled in the protestant theology, law, medicine and literature faculties and analyse the results in Section 3. Section 4 concludes.

Keywords: Academic Careers, Overeducation, Outliers, Cliometrics, Germany.

JEL-classification: C22, N3.

This version: 25/01/2006.
1. INTRODUCTION

The starting point of this article was to discuss, as Titze in his seminal work (1990) and Müller-Benedict later on (1991, 2000), the erroneous representation of the glutting crisis in German universities from 1885-1890.

In the late 19th century, the hypothesis was put forward, that this glutting was manipulated and augmented by the conservative Prussian government to prevent members of the lower social classes from studying and to restore the social exclusivity of the Prussian Universities. This crisis had a large impact on the public discussion surrounding academic education and culminated in references to a growing “academic proletariat”. During a great debate in parliament the Prussian minister of culture —Gustav von Goßler— asked, with the acclamation of the conservatives: How many students are necessary to maintain the size of the governing classes? A secret prognosis of student enrolments, which had been made by Wilhelm Lexis was used to successfully push through political measures designed to diminish student populations. But all these efforts were unsuccessful and after a short cyclical recession, student populations grew at a much faster rate than they had beforehand. It became rapidly clear that the dynamic of the process that led to glutting was probably independent of political influences. Further, the hypothesis that the glutting crisis was manipulated by particular interest groups, could not be sustained. Therefore new investigations into the nature of the generated dynamics present in the time series were necessary (Müller-Benedict, 2000, p. 37).

First of all, it was possible to publish a complete stock of data on the number of students for all subjects at all universities. This data, together with other information such as age, sex, and social origin of the students, has been published in three books (Titze et al., 1987, 1995, Diebolt, 1997).
Although this large historical database enables a preliminary approach to reality, it is not sufficient to account for the whole complexity of the phenomenon.

Using the outliers methodology (Darné & Diebolt, 2004) and following previous researches on education and labour markets (based essentially on the cyclical phases of saturation and shortage in the numbers of students enrolled at universities over a century and a half)3 this study on the growth of the numbers of students enrolled at German universities across 120 years should make it possible to extend and control the previous investigations.

Our basic assumption is to say that the regular shocks we observe for the evolution of the German higher education system are superposed by irregular shocks which appear rarely (infrequent large shocks). This includes the question whether the long-term development of higher education in Germany is caused (or not) by such extraordinary shocks such as wars, political measures and institutional changes. If this was the case, educational development could probably not be explained as a systematic endogenous process but would have to be traced back to specific historical events.

The outline of the paper is as follows. In Section 2, we define the outliers and describe the outliers identification procedure. We apply this procedure to the annual series of the students enrolled in the protestant theology, law, medicine and literature faculties4 and analyse the results in Section 3. Section 4 concludes.

---


4 Our reference population for the whole of the observation period forms 84% of all persons enrolled at universities (ranging from 100% in 1830 to 71% in 1941) and 69% of those enrolled in German higher education (ranging from 99% in 1830 to 55% in 1941). In 1830, 41% of the persons attending German universities were enrolled in Prussian universities. The figure was 52% in 1941 In short, Prussia is not Germany but it has made Germany.
2. OUTLIERS METHODOLOGY

We introduce here a new econometric technique for shock analysis in historical economics: the outliers methodology: «Outliers represent sudden temporary or permanent shifts in the level of a time series. There are several methods for the detection of outliers based on intervention analysis as originally proposed by Box and Tiao (1975). An often used procedure is that of Tsay (1988). This method was also used by Balke and Fomby (1994), although with some modifications. Here we will use an improved algorithm by Chen and Liu (1993), which is readily available, with slight modifications, in the computer program TRAMO developed by Gómez and Maravall (1997, 2001).»

Consider a univariate time series \( y_t \) which can be described by the ARIMA\((p, d, q)\) model:

\[
\alpha(B)\phi(B)y_t = \theta(B)a_t \tag{1}
\]

where \( B \) is the lag operator, \( a_t \) is a white noise process, \( \alpha(B), \phi(B), \theta(B) \) are the lagged polynomials with orders \( d, p, q \), respectively. The outliers can be modelled by regression polynomials as follows:

\[
y_t = y_t^* + \sum_i \omega_i \nu_i(B) I_t(\tau) \tag{2}
\]

where \( y_t^* \) is an ARIMA process, \( \nu_i(B) \) is the polynomial characterizing the outlier occurring at time \( t = \tau \), \( \omega_i \) represents its impact on the series and \( I_t(\tau) \) is an indicator function with the value 1 at time \( t = \tau \) and 0 otherwise.

\[\text{\textsuperscript{5}}\text{For the reader interested in other cliometric applications and a complete mathematical and statistical presentation of the outlier methodology, please cf. Darné and Diebolt, 2004, 2006, Metz, 2002, 2004.}\]

\[\text{\textsuperscript{6}}\text{Darné and Diebolt, 2004, p. 1452.}\]
In this paper, four main outliers are classified as:

– Additive Outliers (AO) that affect only a single observation at some points in time series and not its future values. In terms of regression polynomials, this type can be modelled by setting: \( \nu_i(B) = 1 \).

– Innovational Outliers (IO) that affect temporarily the time series with the same dynamics as an innovation. The polynomial is then \( \nu_i(B) = \theta(B) / \phi(B) \).

– Level Shifts (LS) that increase or decrease all the observations from a certain time point onward by some constant amount. In this case, the polynomial: \( \nu_i(B) = 1/(1 - B) \).

– Temporary Changes (TC) that allow an abrupt increase or decrease in the level of a series which then returns to its previous level exponentially rapidly. Their speeds of decay depend on the parameter \( \nu_i(B) = 1/(1 - \delta B) \), where \( 0 < \delta < 1 \).

It is considered that AOs and IOs are outliers which are related to an exogenous and endogenous change in the series, respectively, and that TCs and LSs are more in the nature of structural changes. TCs represent ephemeral shifts in a series whereas LSs are more the reflection of permanent shocks. However, IOs will have a relatively persistent effect on the level of the series.

An ARIMA model is fitted to \( y_i^* \) in (1) and the residuals are obtained:

\[
\hat{a}_i = \pi(B)Y_i, \quad (3)
\]

where \( \pi(B) = \frac{\alpha(B) \phi(B)}{\theta(B)} = 1 - \pi_1 B - \pi_2 B^2 - ... \).
For the three types of outliers in (2), the equation in (3) becomes:

AO: \[ \hat{a}_i = a_i + \omega_t \pi(B) I_i(\tau) \]

IO: \[ \hat{a}_i = a_i + \omega_t I_i(\tau) \]

LS: \[ \hat{a}_i = a_i + \omega_t \left[ \frac{\pi(B)}{(1-B)} \right] I_i(\tau) \]

TC: \[ \hat{a}_i = a_i + \omega_t \left[ \frac{\pi(B)}{(1-\delta B)} \right] I_i(\tau) \]

These expressions can then be viewed as a regression model for \( \hat{a}_i \), i.e.,

\[ \hat{a}_t = \omega_t x_{i,t} + a_i \]

With:

for all \( i \) and \( t < \tau \):
\[ x_{i,t} = 0 \]

for all \( i \) and \( t = \tau \):
\[ x_{i,t} = 1 \]

\[ x_{1,t+k} = -\pi_k \] \hspace{1cm} (AO);
\[ x_{2,t+k} = 0 \] \hspace{1cm} (IO);

for \( t > \tau \) and \( k \geq 1 \):
\[ x_{3,t+k} = 1 - \sum_{j=1}^{k-1} \pi_j \] \hspace{1cm} (LS);
\[ x_{4,t+k} = \delta^k - \sum_{j=1}^{k-1} \delta^{k-j} \pi_j - \pi_k \] \hspace{1cm} (TC).

The test statistics for the types of outliers are given by:

AO: \[ \hat{e}_1(\tau) = \left[ \hat{\omega}_1(\tau) / \hat{\sigma}_a \right] / \left[ \sum_{i=1}^{\infty} x_{1,i}^2 \right]^{1/2} \]

IO: \[ \hat{e}_2(\tau) = \hat{\omega}_2(\tau) / \hat{\sigma}_a \]
An outlier is identified at time $t = \tau$ when the test statistics $\hat{t}_i(\tau)$ exceeds a critical value. In TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) the critical value is determined by the number of observations in the series based on simulation experiments. The different test statistics at time $t = \tau$ are compared in order to identify the type of outlier. The one chosen has the greatest significance such as $\hat{t}_{\text{max}} = \max |\hat{t}_i(\tau)|$.

When an outlier is detected, we can adjust the observation $Y_i$ at time $t = \tau$ to obtain the corrected $Y_i^*$ via (2) using the $\hat{\omega}_i$, i.e. $Y_i^* = Y_i - \hat{\omega}_iY_i(\tau)$. Finally, the procedure is repeated until no outlier is detected. A multiple regression on $Y_i^*$ is performed on the various outliers detected to identify spurious outliers.

where $\hat{\omega}_i(\tau)(i = 1 - 4)$ denotes the estimation of the outlier impact at time $t = \tau$, and $\hat{\sigma}_a$ is an estimate of the variance of the residual process.
3. CLIOMETRIC APPLICATION

We now examine the nature of the shocks on the annual series of the students enrolled in the protestant theology, law, medicine and literature faculties in Germany. The period studied runs from 1830 to 1941.

Tables 1 and 2 display the results of outlier identification. All detected outliers are given by series, with their timing, type, value and \( t \)-statistic. The last column presents one possible economic, financial, political etc. explanation of outlier occurring. In a general manner, outliers are detected in all the series.

**Table 1: Outlier detection**

<table>
<thead>
<tr>
<th>Faculty</th>
<th>Date</th>
<th>Type</th>
<th>Value</th>
<th>( t )-stat</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1870</td>
<td>AO</td>
<td>-0.879</td>
<td>-5.43</td>
<td>Franco-Prussian War</td>
</tr>
<tr>
<td></td>
<td>1924</td>
<td>TC</td>
<td>-0.101</td>
<td>-3.81</td>
<td>Hyperinflation</td>
</tr>
<tr>
<td></td>
<td>1936</td>
<td>LS</td>
<td>-0.170</td>
<td>-4.71</td>
<td>Nazi Regime</td>
</tr>
<tr>
<td></td>
<td>1939</td>
<td>TC</td>
<td>-2.050</td>
<td>-53.58</td>
<td>World War 2</td>
</tr>
<tr>
<td></td>
<td>1941</td>
<td>AO</td>
<td>-0.294</td>
<td>-5.16</td>
<td>World War 2</td>
</tr>
<tr>
<td>Law</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1849</td>
<td>TC</td>
<td>0.812</td>
<td>7.36</td>
<td>End of the Revolutions in the German States</td>
</tr>
<tr>
<td></td>
<td>1870</td>
<td>AO</td>
<td>-0.186</td>
<td>-25.11</td>
<td>Franco-Prussian War</td>
</tr>
<tr>
<td></td>
<td>1883</td>
<td>TC</td>
<td>-0.878</td>
<td>-7.45</td>
<td>Crisis and Economic Depression</td>
</tr>
<tr>
<td></td>
<td>1914</td>
<td>TC</td>
<td>-0.158</td>
<td>-12.00</td>
<td>World War 1</td>
</tr>
<tr>
<td></td>
<td>1916</td>
<td>IO</td>
<td>0.159</td>
<td>3.70</td>
<td>World War 1</td>
</tr>
<tr>
<td></td>
<td>1924</td>
<td>TC</td>
<td>-3.421</td>
<td>-22.12</td>
<td>Hyperinflation</td>
</tr>
<tr>
<td></td>
<td>1933</td>
<td>IO</td>
<td>-0.207</td>
<td>-4.41</td>
<td>Nazi Regime</td>
</tr>
<tr>
<td></td>
<td>1935</td>
<td>IO</td>
<td>-0.155</td>
<td>-3.63</td>
<td>Nazi Regime</td>
</tr>
<tr>
<td></td>
<td>1936</td>
<td>IO</td>
<td>-0.189</td>
<td>-4.41</td>
<td>Nazi Regime</td>
</tr>
<tr>
<td></td>
<td>1939</td>
<td>LS</td>
<td>-0.660</td>
<td>-24.83</td>
<td>World War 2</td>
</tr>
</tbody>
</table>
Table 2: Outlier detection (continued)

<table>
<thead>
<tr>
<th>Faculty</th>
<th>Date</th>
<th>Type</th>
<th>Value</th>
<th>t-stat</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>1866</td>
<td>LS</td>
<td>0.197</td>
<td>4.34</td>
<td>Austro-Prussian War</td>
</tr>
<tr>
<td></td>
<td>1870</td>
<td>AO</td>
<td>-0.238</td>
<td>-9.52</td>
<td>Franco-Prussian War</td>
</tr>
<tr>
<td></td>
<td>1919</td>
<td>TC</td>
<td>-0.159</td>
<td>4.06</td>
<td>End of World War 1</td>
</tr>
<tr>
<td></td>
<td>1924</td>
<td>TC</td>
<td>-0.158</td>
<td>-4.03</td>
<td>Hyperinflation</td>
</tr>
<tr>
<td>Literature</td>
<td>1866</td>
<td>LS</td>
<td>-0.513</td>
<td>-6.47</td>
<td>Austro-Prussian War</td>
</tr>
<tr>
<td></td>
<td>1939</td>
<td>AO</td>
<td>-0.596</td>
<td>-12.54</td>
<td>World War 2</td>
</tr>
</tbody>
</table>

The ‘Value’ represents the estimated outlier effects by which the series of the first differences can be adjusted. The identified AO in 1870, 1939 and 1941 for theology, law, medicine and literature causes a change of the series only in the year of their appearance. In contrast to this, the IO, LS and TC for all the series in various years do not only change the adjusted series in the year of their appearance, but also in the following years.

Statistically, figures 1, 3, 5, 7 display (see Appendix) the original series (in full line) and the outlier corrected series (in dotted line) for the students enrolled in the protestant theology, law, medicine and literature faculties in Germany. They clearly show the presence of outliers in each original series and their correction from the outlier identification procedure.

A comparison of the first differences with the adjusted series demonstrates the extent of the adjustment. In other words, the outlier effects can also be demonstrated by subtracting the adjusted series from the unadjusted series. The main advantage of such a representation (see figures 2, 4, 6, 8 in Appendix) is that it is directly open to scientific interpretation.
So far, with the outliers methodology it became clear that the dynamic of the process that led to higher education development was strongly dependent on economic and political events. War periods are characterized by irregularities. Especially the Austro-Prussian, the Franco-German and the two World Wars had an exceptional influence on the evolution of students in German faculties.

As was put forward early by Von Ciriacy-Wantrup (1936), the greatest changes seem to be those caused by wars and revolutions through the economic, institutional, legal and population changes that they necessarily cause. More recently, Goldstein (1988) established a link between the turning points in long cycles of the economy and the periodic outbreak of wars. This seems to be an interesting approach. Wars are special moments in the metamorphosis of economic and social structures. They take the form of stages during which the socio-economic sphere has to change under the weight of political issues. However, wars deform part of reality. They prevent the apprehending of the overall range of the phenomenon. This being so, they are not just milestones marking the turning points of history. They exist as facts of considerable importance for the socio-economic growth of a country.

On the other hand, wars and revolutions do not come out of a clear sky, and they are not caused by arbitrary acts of individual personalities. They originate from real, especially economic, circumstances. The assumptions that wars acting from the outside evokes the question as to why they themselves follow each other with regularity and mostly during the upswing of the economic cycles? Much more probable is the assumption that wars originate in the acceleration of the pace and the increased tension of socioeconomic life, in the heightened economic struggle for new markets, and that the social shocks happen most easily under the pressure of new economic and political forces.
Wars therefore can also be fitted into the rhythm of the economic growth process and do not prove to be the forces from which these movements originate, but rather to be one of their symptoms. But once they have occurred, they naturally exercise a potent influence on the pace and direction of socioeconomic dynamics.

4. CONCLUSION

The nature and magnitude of the shocks that have affected the German higher education system were analysed over a long period using the outlier approach.

We found strong proof of infrequent large permanent and transitory shocks resulting essentially from the major economic and political events formed by the two World Wars in the twentieth century, the end of the Revolutions in the German States in 1849, the struggle for supremacy in Germany, i.e. the Austro-Prussian (1866) and the Franco-Prussian (1870-1871) wars, crisis and economic depression (1883) in the nineteenth century, the recession in the 1920s and the Nazi regime after 1933.

Finally, our cliometric results are also mainly different from previous conclusions developed especially by Titze (1987, 1990) on the autonomous process, i.e. the internal dynamic and social mechanisms of academic status recruitment in Germany before World War 2.

5. BIBLIOGRAPHY


Diebolt, C. Ed. [2003], Analyse cliométrique de la relation éducation-croissance en Europe aux 19ème et 20ème siècles, Rapport de recherche: Aides à Projets Nouveaux, Centre National de la Recherche Scientifique (CNRS), juillet.


6. APPENDIX

Figure 1

Figure 2
Figure 7

Figure 8
WP2006-1  Olivier DARNÉ, Claude DIEBOLT
"Cliometrics of Academic Careers and the Impact of Infrequent Large Shocks in Germany before 1945"

WP2006-2  Claude DIEBOLT, Catherine KYRTSOU
"Non-Linear Perspectives for Population and Output Dynamics: New Evidence for Cliometrics"

WP2006-3  Claude DIEBOLT, Karine PELLIER
"L’intérêt des systèmes de gestion de bases de données relationnels en cliométrie"

WP2006-4  Claude DIEBOLT, Jean-Pascal GUIRONNET
"The Dynamics of Education Returns"

WP2006-5  Claude DIEBOLT, Jean-Pascal GUIRONNET
"Vers une théorie économique de la suréducation ?"

WP2006-6  Claude DIEBOLT, Mishra TAPAS
"Cliometrics of the Abiding Nexus between Demographic Components and Economic Development"

WP2006-7  Magali JAOUUL
"Cliométrie de l’engorgement en France. Evaluation théorique et empirique"

WP2006-8  Claude DIEBOLT, Antoine PARENT
"Were the Anomalies in the Sterling-Franc Exchange Rate Regulation during the Mid-19th Century?"

WP2006-9  Claude DIEBOLT
"Progrès technique et cycles économiques dans la pensée allemande de l’entre-deux-guerres : l’apport d’Emil Lederer"

WP2006-10  Claude DIEBOLT
"Croissance et éducation"

WP2006-11  Mohamed CHIKHI, Claude DIEBOLT
"Nonparametric Analysis of Financial Time Series by the Kernel Methodology"

WP2006-12  Claude DIEBOLT, Antoine PARENT
"A Note on Juglar, Bonnet and the Intuition of the Interest Parity Relation"