Space and time in comparative political research: pooled time-series cross-section analysis and multilevel designs compared

Stadelmann-Steffen, Isabelle; Bühlmann, Marc

Space and Time in Comparative Political Research

Pooled Time-series Cross-section Analysis and Multilevel Designs Compared

Raum und Zeit in der vergleichenden Politikwissenschaft

Die gepoolte Zeitreihenanalyse und die Mehrebenenanalyse im Vergleich

Isabelle Stadelmann-Steffen und Marc Bühlmann

Abstract
The combination of cross-section and time dimension is a central issue in current comparative political research. The state-of-the-art procedure in this context is pooled time-series cross-section analysis (PTSCS), which is en vogue in today's relevant literature but not uncontested. An interesting option are multilevel designs, which allow the combination of time and space by considering observations over time nested within country-specific contexts. The purpose of this paper is to illustrate the advantages of multilevel designs in comparative political research, which mainly concern the modeling of time-invariant variables, the possible distinction between cross-sectional and time related variance in the data, and the possibility to model heterogeneity instead of just correcting it. Using the example of an analysis of public education expenditure in the 26 Swiss cantons between 1978 and 2003, it can be shown that multilevel analysis – mainly due to its statistical and conceptual advantages – is indeed a promising alternative to PTSCS.

Zusammenfassung
Die Kombination von Längs- und Querschnittvergleich ist ein zentrales Thema in der vergleichenden Politikwissenschaft. In aller Regel wird ein gepooltes Zeitreihen-Design angewandt, um verschiedene politische Einheiten über die Zeit und miteinander zu vergleichen. Diese Methode, obwohl sehr in Mode, ist allerdings nicht unumstritten. Eine interessante Alternative stellt die Mehrebenenanalyse dar, welche ebenfalls die Kombination von Zeit und Raum zulässt, indem sie Beobachtungen in der Zeitachse quasi als Eigenschaften verschiedener länderspezifischer Kontexte betrachtet. Das Ziel dieses Beitrags ist es, die Vorteile eines Mehrebenen-Designs für die vergleichende Politikforschung zu illustrieren. Diese ergeben sich v. a. in Bezug auf die Modellierung zeit-invarianter Variablen, die Unterscheidung zwischen querschnitt- und zeitbedingter Varianz sowie die Möglichkeit, Heterogenität zu modellieren statt lediglich zu korrigieren. Als praktisches Anwendungsbeispiel werden die öffentlichen Bildungsausgaben in den 26 Schweizer Kantonen zwischen 1978 und 2003 analysiert. Es zeigt sich, dass die Mehrebenenanalyse sowohl aus statistischen Gründen, aber auch aufgrund ihrer konzeptuellen Vorteile tatsächlich eine viel versprechende Alternative zu gepoolten Zeitreihen-Designs darstellt.

1 Introduction[1]

Combining cross-sectional and time-dimensional data is one of the central issues in the present methodological discussion in comparative political research. Since data sets in social science are often limited to about 20 to 30 cases, and as dynamics over time are an important aspect of typical political and societal phenomena, the simultaneous examination of time and space seems to be indispensable in quantitative research. Pooled time-series cross-sectional analysis (PTSCS) is the currently prevailing method for analysing this type of data and is correspondingly en vogue in the relevant literature.[2]

However, together with the design's growing popularity a discussion on its methodological pitfalls has arisen. By pooling the time and cross-sectional dimensions the data obtains a specific structure, which violates important condition for ordinary least squares regression (OLS). Although the techniques to deal with these problems are quite sophisticated by now, concerns and criticism on the one hand and the search for better alternatives on the other remain.

An interesting option in this discussion is multilevel analysis (MLA), which allows for the combination of different levels of analysis. Prevalent in educational science for a long time, multilevel applications have more recently been used in political science, too. Although in most cases the method is used for the simultaneous analysis of individual and contextual data, it also offers the possibility of combining time and space dimensions in one single model: time and space are considered as two levels of analysis, whereas observations over time are nested within country-specific contexts.

While there is an extensive debate on the strengths and limits of pooled time-series cross-section analysis, multilevel designs as a possible alternative have been much less discussed. An exception is Western (1998: 1234) who claims that a multilevel specification of pooled time-series models has both substantial and statistical advantages. He stresses that MLA allows for greater causal complexity which provides a closer fit between typical comparative theory and model specifi-

1 We thank two anonymous referees for helpful remarks on an earlier version of this paper, and K. Gilland for linguistic assistance.

2 If we refer to the PTSCS method in the following we mean the "de facto Beck-Katz standard" which has become the "accepted econometric technique in comparative political economy" (Plümper et al. 2005: 327), even though PTSCS actually only describes the type of data structure, which would actually allow for a wide variety of model specifications.

cation. Hence, multilevel models not only control for, but model heterogeneity in the data. However, only recently multilevel models for time-series cross-section data have been subject to a more systematic evaluation. Shor et al. (2007) employing Monte Carlo simulations show that multilevel models indeed perform better than OLS estimators. They stress that MLA is more flexible offering the researcher additional advantages that can be summarized as follows:

1. MLA performs better than other common estimators in Monte Carlo Simulations. When the number of observations is low (either at the time- or country-level) MLA does better, since each "estimated parameter has the potential to borrow strength from other parameters in the model" (Shor et al. 2007: 5).
2. As a consequence, MLA generally provides a better model fit (Shor et al. 2007: 4f.). While complete pooling ignores important parts of contextual variance, no pooling – i.e. separate intercepts for single time series – is affected by effects of outliers and does not allow the sharing of important information from across units. Multilevel analysis which can be seen as a process of "partial pooling" (Shor et al. 2007: 5) offers an advantageous middle way. Multilevel models thereby allow distinguishing and analyzing the variances at both the cross-sectional and the time level.
3. Another advantage of MLA is the handling of slowly moving or completely time-invariant variables: While such variables often produce estimation problems in PTSCS, they are not a problem in multilevel models: Partial pooling as done in MLA "allows estimates of units to borrow strength from the whole sample and shrink toward a common mean" (Shor et al. 2007: 4). Moreover, theoretically and conceptionally it makes more sense to view such time-invariant variables as contextual effects which influence all observations of a given unit.
4. The basic random intercept multilevel model can easily be expanded by also including varying slopes, i.e. different effects of a predictor variable for different units.

Starting from these considerations the purpose of this paper is to demonstrate the differences between pooled time-series cross-sectional models and multilevel analysis. Its main aim is to illustrate the advantages of multilevel models in practical applications that are typical in comparative political research. First and foremost, we focus on the differences between the two methods in terms of the modeling of time-invariant variables and model fit. Time-invariant or slowly moving variables such as institutions or cultural settings are crucial elements of theories explaining differences between political entities. The simultaneous modeling of time-invariant variables and dynamic determinants within comparative politics therefore is an important task. Drawing on Shor et al. (2007) we suggest that the estimation of the influence of time-invariant variables is more appropriately done by MLA than by PTSCS. Furthermore, it can be expected that due to its conceptual advantages MLA provides a better model fit than PTSCS.

As an example of a practical application we use a comparative analysis of public education expenditure in the 26 Swiss cantons between 1978 and 2003 (in so doing we particularly refer to Freitag/Bühlmann 2003). Following the theories of welfare state development (Schmidt 2000) we distinguish institutional, socio-economic, party-political, and historic-cultural variables, and derive respective hypotheses to explain the differences in cantonal education expenditure.

The paper unfolds as follows. First the theoretical background of our illustrative example is shortly demonstrated. Afterwards the main part of this contribution consists of two analyses of cantonal education expenditure, one using pooled time series and the other applying a multilevel design. The analyses are followed by a discussion of the results and the practical implications of this methodological comparison. The paper closes with concluding remarks.

2 Pooled Time-series Cross-section Analysis versus Multilevel Design – Two Comparative Analyses of the Public Education Expenditure in the Swiss Cantons

The aim of this contribution is to compare a pooled times-series cross-section design and multilevel analysis by means of an empirical example. We will therefore present two analyses of cantonal public education expenditure, one applying pooled times-series cross-sectional analysis and the other using a multilevel design. Both analyses will start with a short description of the method and the state-of-the-art procedures used in practical applications. Afterwards, the hypotheses on the differences in cantonal education expenditure, briefly outlined in the following section, will be empirically tested by applying the two methods. After having illustrated both methods, we will compare the results and try to demonstrate the advantages of a multilevel design.

2.1 Public education expenditure in the Swiss cantons – theory and hypotheses

In the Swiss federal system there are various areas in which the cantons have substantial autonomy in policy-making. This is especially true for the education system, which has been the responsibility of the cantons and the communes since the beginning of the Swiss nation-state. As a consequence, the education systems and the financial investments of public authorities in human resources substantially vary among the cantons. In the following we will therefore analyze the determinants of cantonal differences in public education expenditure.

Even though this paper focuses on the methodological implications rather than on substantial results, we will shortly lay out the theoretical background of our analysis in this paragraph.

Most of the quantitative studies done in this field (e.g. Freitag/Bühlmann 2003; Boix 1997; Castles 1999; Cameron/Hofferbert 1974) refer to theories of comparative welfare state research: education policy is seen as a part of a comprehensive welfare state concept (Castles 1999: 10). Theories of welfare state development (Schmidt 2000) can thus help to explain the variances in public education expenditure. Following this literature we distinguish four different approaches: socio-economic determinants, the partisan theory of public policy, institutional explanations, and historical and cultural legacies.

The first approach hypothesizes that public policy is substantially shaped by existing *socio-economic challenges* (Schmidt 2000: 23f.). In this sense public education expenditure depends on the demand for public education and the possibility to provide a corresponding supply. First, it can be hypothesized that public education expenditure is high, if the share of children of school age and thus the demand for investments in public education is considerable (Schmidt 2002: 10; Hega 1999: 62; Poterba 1999; Fernandez/Rogerson 1997). Similarly, a substantial share of foreigners in schools is expected to raise the need for public investments in education in order to finance necessary integration measures (Freitag/Bühlmann 2003: 146). However, theoretically, only those parts of the population support an extensive education system, which in some way can profit from the educational supply. All others will try to restrict public investment in education. Boix (1997: 837) finds that especially high employment shares in the agricultural sector tend to reduce public education expenditure since this group has only very limited interest in the public education system. In contrast, it can be expected that a large service sector demands a well-educated labor force and therefore has a positive effect on public education expenditure. Similarly and according to Castles (1999: 180), the demand for human resources is higher in urban areas. On the other hand it can nevertheless be argued that the provision and organization of public education is cheaper in densely populated regions, which in turn leads to lower expenditure.

A second explanation for differences in public education expenditure is provided by the *partisan theory* of public policy (Schmidt 2000: 26f.; Hibbs 1977). It develops the idea that the party-political composition of government influences public investments in public policy in general and in education in particular. According to the findings in the literature a substantial share of leftist parties in government lead to higher (education) expenditure, as these parties attribute the responsibility for education to the state rather than to the private sector (Schmidt 2002: 13f.; Hega 1999: 63f.; Castles 1999: 53, 1989: 438; Boix 1997: 815). This effect should be even more pronounced, if strong leftist parties are supported by powerful unions (Schmidt 2002: 12). Conversely, Christian democratic parties disapprove of government interference in the area of education for two reasons: On the one hand education had been in the hands of the church for a long time, on the other hand public intervention in education policy goes against the principle of subsidiarity, which is an important element of Christian democratic politics (Schmidt 2002: 14; Wilensky 1981).

Thirdly, in the analysis of public policy provision *institutional explanations* are of particular importance. Formal or informal rules build the framework in which state activities take place. Institutional factors therefore constrain but also allow for governmental action (Schmidt 2000: 28f.). In Switzerland the three typical elements of the political system are of overriding interest: direct democracy, the amount of decentralization, and the character of decision making (i.e. whether decision making is rather consensual or follows the majority rule) (Vatter/ Freitag 2002). With respect to direct democracy previous studies have found that the extensive use of direct-democratic rights have a restricting effect on public expenditure (Vatter 2002; Vatter/Freitag 2002; Hega 1999: 67). Generally, voters tend to avoid additional taxes (Hega 1999: 66). Direct democracy provides them with a strong instrument to bring these preferences into politics and, thus, to prevent further investment in public policy (Schaltegger 2001: 6).

A similar effect can be attributed to decentralized decision-making structures and federalism. Federalist power-sharing and decentralized fiscal power in particular narrows the scope for state intervention (Vatter/Freitag 2002). For reasons of possible migration, decentralized structures additionally lead to competition for the most attractive location – not least in terms of taxes – among the decentralized entities (Freitag/Bühlmann 2003: 149). In sum, these tendencies lead to the hypothesis that decentralized decision-making structures (in our case: high local autonomy) result in lower public education expenditure (Cameron/Hofferbert 1974).

An inverse effect is generally attributed to consensus democracy, which is expected to foster welfare state expansion (Lijphart 1999). The integration of various political, social or cultural minorities into the decision-making process tends to produce cost-intensive barter transactions and thereby a higher degree of state intervention. It can therefore be assumed that a high degree of consensual alignment in the government leads to more substantial public education expenditure (Freitag/Bühlmann 2003: 150).

Finally, *historical and cultural legacies* can help explaining variances in public education expenditure. As mentioned above, the relationship between church and state has played an important role in the development of educational systems. In societies with strong conservative-catholic background state intervention in education has been limited, since catholic forces tried to maintain their traditional competence in this field. In contrast, dominant protestant forces facilitated co-operation between church and state as Protestantism stood for a fast education expansion lead by the state. It can actually be shown that the power balance between catholic forces and the state at the end of the 19th century is an important factor for the path chosen in education policy (Archer 1979), and also for the degree to which the state engages in education even today. In Schmidt's (2000: 31) argumentation present education policy is thus path-dependent, which means that it is a result of earlier decisions and their intended or unintended consequences.

Cultural differences exist in Switzerland not only for religious but also for linguistic reasons. While the French- and Italian-speaking parts of Switzerland share a lot of institutional and cultural elements with the neighboring countries France and Italy, the German-speaking part is much more oriented towards Germany and Austria. This is also reflected in different education systems. Contrary to the French- and Italian-speaking parts of Switzerland, in the German-speaking cantons we find a strong tradition of vocational education like in Germany and Austria, and accordingly higher education and thus financial investments in public schooling are lower (Schmidt 2002: 9; Castles 1999: 175; Hega 1999: 63).

In the following, these hypotheses will be tested in a comparative quantitative analysis. For the operationalization of the variables see the appendix.[3]

The four approaches provide us with different variables. While some variables (mainly the socio-economic and the party political determinants) vary over time, this is not the case for others. Hence, the historic-cultural factors as well as most of the institutional variables do not or only slightly change over time. In what follows we will see that the two methods differently handle and interpret these two groups of explanatory variables.

2.2 The dependent variable and methodological challenges

The dependent variable is the public education expenditure in the Swiss cantons between 1978 and 2003. This period of investigation has been chosen mainly for data related reasons: The year 1978 is the first year (after the foundation of the 26th Swiss canton, Jura) for which data for all 26 cantons is available. Starting in 1978 thus provides us with enough data for applying a multilevel analysis (Bullen et al. 1994; Jones 1997). Since the cantons vary substantially in size and population we analyze the education expenditure per capita.

The cantonal public education expenditure has been steadily growing during the last 26 years, which lets one assume that the data follows a non-stationary development. The statistical tests for unit-root (Hadri 2000; Im et al. 1997; Levin/Lin 1993) indeed all conclude that public education expenditure is non-stationary.[4] In the methodological discussion of PTSCS this phenomenon is an important issue. If the data follows a permanent growing development over time the mean of the residuals is not constant over time either, which violates the Gauss-Markov-assumption for the execution of OLS. As a consequence OLS-results are wrong. Additionally, there often result so-called "spurious regressions", which display relationships only resting on a common development of variables over time (Kittel 1999: 249). Generally, the best way to handle non-stationary data is the examination of first differences. However, in this case conceptual characteristics of the model change: It is the short-term changes that are now modeled, while the differences in the level between units and long-term developments have been

3 To allow for a comparison of the results all data were standardised (mean = 0, standard deviation = 1).

4 Further analyses not presented here show that this finding also applies for the logarithmized education expenditure per capita.

eliminated from the model. Thus, such a model is not suited to sensibly evaluate the effect of the time-invariant variables. To sum up, the analysis of first differences may solve the statistical problem of non-stationarity, but it is not a valuable alternative for most applications in political science where differences between countries and effects of (almost) time-invariant variables (in our case: institutions and historical-cultural characteristics) are of interest.

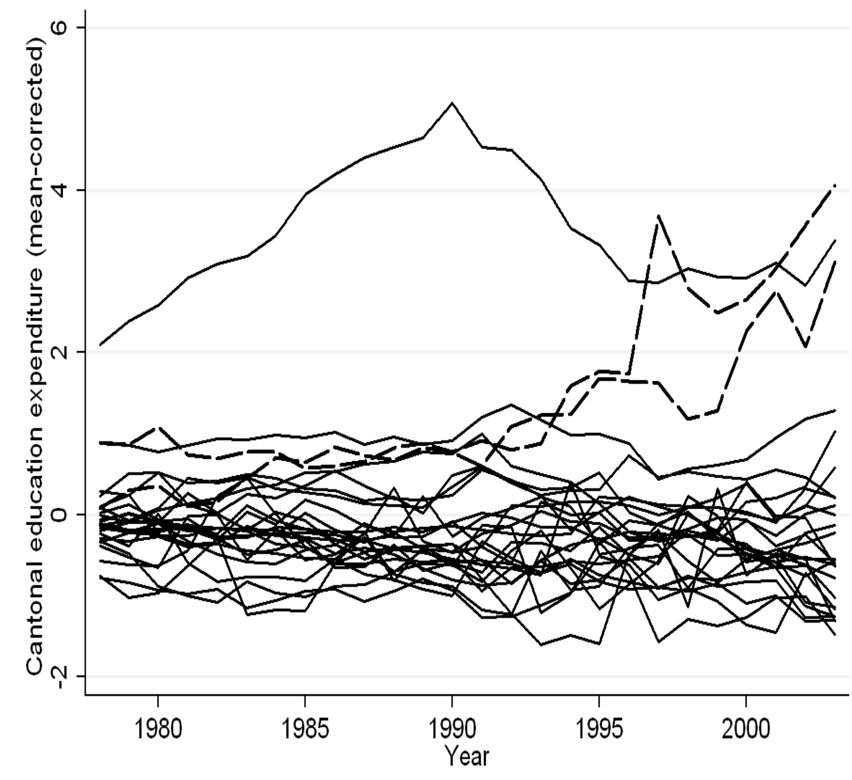
Therefore we proceed in another way that in some cases can also eliminate the problem of non-stationarity. We transform the education expenditure variable by sorting out the annual average values, which means that we remove the steadily increasing development in the time dimension.[5] Using this transformation a big part of the development over time cannot be analyzed either. However, even though the model indeed focuses on cross-sectional variation, it also contains a substantial part of the time dimension. That is, the relative differences in the development of education expenditure between cantons can still be analyzed.

As shown in figure 1, this procedure largely solves the problem. Two cantons, however, possibly still follow a non-stationary development. For the purpose of this paper these two cantons – Zug and Basle Town – are excluded from the analysis. In so doing both the graphical examination of the data (figure 1) as well as the statistical tests indicate that the time-series under investigation are stationary.[6] Thus, for the following analysis we use the mean-corrected public education expenditure for 24 cantons.

5 In the language of PTSCS this is actually equivalent to a model with fixed time effects.

6 The Levin-Lin-Chu test (lag[1]) tests the null hypothesis of Non-stationarity. In the present case the test produces a coefficient of -0.28 (p=0.02), thus implying that the time-series are stationary.

Figure 1 Standardized education expenditure in the Swiss cantons (mean-corrected)



Notes: The lines for Zug and Basle-Town (dashed, bold lines) indicate a non-stationary development.

2.3 Pooled time-series cross-section analysis (PTSCS)

In comparative political research the prevailing way of simultaneously analyzing variances over time and across countries is to pool cross-sectional data for different points in time:

$$Y_{ij} = \beta_0 + \beta X_{ij} + \varepsilon_{ij} \quad (1)$$

In our example Y_{ij} specifies the predicted educational investments of year *i* in canton *j*. This predicted value is explained by the overall mean (β_0) and the explanatory variables (the X variables and their respective β) measured for each year and unit separately (subscripts *i* [year] and *j* [canton] respectively). β has no subscript since one coefficient per variable for all years and cantons is estimated.

However, this produces a data structure which practically by definition violates important conditions for OLS regression. In particular, PTSCS models suffer from temporally and spatially correlated errors as well as from panel heteroscedasticity (Beck/Katz 1995: 636; Kittel 1999: 228ff.; Podestà 2002: 10; Plümper et al. 2005). The consequence of these data characteristics is that the standard errors are systematically biased. Hence, conclusions concerning the statistical significance of the estimations become meaningless (Gujarati 1995: 366). In order to successfully use a PTSCS design it is therefore necessary to correct for these problems. To this end, the following procedures have won recognition, and represent the state of the art in this field. First, for the correction of heteroscedasticity and the spatial autocorrelation so-called panel-corrected standard errors (PCSE) are applied. This procedure draws on the fact that OLS still produces unbiased, but inefficient estimates. The OLS residuals allow for a reliable estimation of the covariances, which then are used to correct the standard errors (Beck/Katz 1995: 638). The PCSE procedure requires that any serial correlation has been eliminated before calculating the corrected standard errors (Beck/Katz 1995: 638). This can be done either by introducing a lagged dependent variable or by estimating the autocorrelation coefficient and correct for the correlation in the residuals (Kittel 1999: 230). The two proceedings represent different approaches: On the one hand, by modeling one or more lagged dependent variables autocorrelation is not seen as a problem, but is like other explanatory factors used for the explanation of the dependent variable. This procedure corresponds to the theoretical proposition that the economy is a self-regulating system which is only marginally influenced by political intervention (Kittel 1999: 230). The so-called AR1-approach, on the other hand, treats autocorrelation as a problem which has to be eliminated (Kittel 1999: 231). The corresponding correction of the standard errors is based on the proposition that the residuals follow a first order autoregressive process:

(2) $$\varepsilon_{ij} = \rho\varepsilon_{i-1j} + \upsilon_{ij}$$

In this function the autocorrelation coefficient ρ (Rho), which takes on a value between 0 and 1, is estimated and is then used to correct the standard errors by applying a so-called Prais-Winsten-transformation (Kittel/Winner 2002: 16). This second approach is often preferred in comparative political research (Plümper et al. 2005: 342f.).[7] Thus, the model takes the following form:

(3) $$Y_{ij} = \beta_0 + \rho\gamma_{i-1j} + \beta X_{ij} + \varepsilon_{ij}$$

A further lively debate has evolved on whether unit fixed effects (FE) should be controlled for in order to account for unobserved unit-specific developments (equation 4). Shor et al. (2007: 4) call this option 'no pooling' and the 'opposite extreme' of complete pooling, since a separate intercept is calculated for each unit. The main disadvantage of this procedure from a political science point of view is, however, that such unit dummies "completely absorb differences in the level of independent variables across units" (Plümper et al. 2007: 331) and thus the time-invariant variables cannot be analyzed anymore. As a FE model is the 'opposite' of complete pooling[8] and since a FE is conceptionally even more closely related to a multilevel design than a completely pooled model (Shor et al. 2007: 4) we will also present a FE model based on a "fixed effects vector composition", which should be a better way to estimate the effect of timeinvariant variables in a FE-context (Plümper/Troeger 2007).

(4) $$Y_{ij} = \beta_{0j} + \rho\gamma_{i-1j} + \beta X_{ij} + \varepsilon_{ij}$$

In what follows, these procedures are applied to the analysis of the cantonal public education expenditure.

7 A discussion of the disadvantages of the lagged-variable approach can be found in Plümper et al. (2005).

8 A multilevel model would lie in between these two options (see section 2.4 and Shor et al. 2007: 4).

Table 1 TSCS-models for the explanation of public education expenditure (mean-corrected)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6 (xtfevd)
Pupils per cantonal population	.10*** (.04)				.10*** (.03)	.02 (.09)
Proportion of foreign pupils	.15*** (.05)				.17*** (.05)	.00 (.03)
Sector 1	-.08 (.07)					
Urbanization	.29*** (.08)				.20*** (.07)	.24*** (.02)
Leftists in government		.09** (.05)			-.01 (.04)	.02 (.03)
CVP in government		-.07 (.05)				
Union		.07 (.05)				
Direct democracy			-.15*** (.05)		-.08 (.05)	-.21*** (.02)
Governmental concordance			-.03 (.03)			
Local autonomy			-.44*** (.06)		-.39*** (.06)	-.41*** (.02)
Strength of Catholics 19th century				-.25*** (.06)	.00 (.05)	-.07*** (.02)
Efforts in education policy 1880				.06 (.05)		
language				-.32*** (.08)	.08* (.06)	.13*** (.02)
eta						.93*** (.04)
constant	-.09** (.05)	-.12*** (.03)	-.11*** (.03)	-.13*** (.04)	-.08** (.04)	-.08*** (.01)
N	624	624	624	624	624	600
R2	.13	.03	.13	.05	.26	.87
AIC	189.5	322.3	206.2	179.6	232.5	415.4
Rho	.83	.79	.82	.86	.76	.67

*Notes: Non-standardized OLS-coefficients; standard errors in brackets. Heteroscedasticity and autocorrelation is treated as follows: To correct for heteroscedasticity and spatial autocorrelation panel corrected standards errors (PCSE) are calculated as suggested by Beck/Katz (1995: 638). Beforehand, serial autocorrelation was eliminated by using a Prais-Winsten-transformation to correct the standard errors (in model 6 the Cochran-Orcutt method is used for the correction of the standard errors which reduces the number of observations by one in each panel). Thereby it is assumed that the residuals follow an AR1-process (Kittel/Winner 2002: 16). *** significant at the 1%-level; ** significant at the 5%-level; * significant at the 10%-level. In model 6 the following variables are used to explain unit fixed effects: urbanization, direct democracy, autonomy of the municipalities, strength of Catholics, language. The language variable and the strength of Catholics are typical examples for variables that per definition do not change over time, while urbanization is time-invariant due to the limited period of investigation and data availability. The same applies to local autonomy that can be seen as a constant institutional feature during our period of investigation. In contrast direct democracy is an example of a slowly-moving variable, since we have three different measures of the direct democracy index indicating only minor changes over time.*

Table 1 reports the results for the multivariate models with the mean-corrected public education expenditure as dependent variable. In models 1 to 4 the explanatory power of the four theoretical approaches (socio-economic, party political, institutional, and historic-cultural approach) is tested separately. Afterwards a fifth model is estimated which incorporates the most important (i.e., significant) variables of the four approaches. Finally, a FE model is calculated on the basis of model 5 allowing for separate cantonal intercepts.

It can be shown that socio-economic determinants on the one hand and institutional factors on the other hand are the most important variables in the model. However, only for two variables a consistent and significant influence on cantonal education expenditure can be demonstrated: While a high degree of urbanization increases public investments in education, highly decentralized structures within the Swiss cantons (i.e., high autonomy of municipalities) decrease cantonal education expenditure. Moreover, held everything constant, the German speaking cantons spend more on public education, which is contrary to our hypothesis. This result can be seen against the background that the dependent variable incorporates expenditure for vocational training, but not for higher education. The latter is particularly important in the Latin cantons, and thus their total investments in education are underestimated by the variable chosen here. In contrast, the size of the agricultural sector, the share of Christian-democratic parties in government, the union strength as well as our indicator for consensus democracy seem not to influence public education expenditure. For some other variables, i.e. the number of (foreign) pupils, direct democracy, and the historical strength of the Catholic Church the PTSCS models are inconclusive.[9]

The Akaike Information Criterion (AIC)[10] shows that neither model 5 nor model 6 are satisfactory in terms of model fit. Actually, even though the two models incorporate more (significant) variables they are not better in explaining

9 In what concerns direct democracy these results actually reflect a standard finding in Swiss political science in that this variable is highly correlated with the language region and therefore is highly sensitive to the introduction of the latter. This is because cantonal differences with regard to direct democracy strongly coincide with the language regions. It is mainly the German speaking cantons that show very extensive direct-democratic procedures (Linder 2005: 272, 1994; Vatter 2002: 401ff.).

10 The Akaike Information Criterion tests the model fit by penalizing for the addition of parameters, and thus selects a model that fits well but has a minimum number of parameters. By means of the AIC models based on different estimation methods and specifications can be compared, the sample however needs to be identical (Burnham/Anderson 2004: 267f.). Therefore, the AIC of model 6 cannot strictly be compared with the others.

the data than model 4 which includes the cultural variables only. Finally, even though we use the mean-corrected data, the high values of Rho show that autocorrelation in the model is still substantial. The investigation of the residuals however indicates that the first order autocorrelated error terms (AR1) in the model do satisfactorily account for these processes.

2.4 Multilevel analysis

"Multilevel analysis (MLA) is a methodology for the analysis of data with complex patterns of variability, with a focus on nested sources of variability" (Snjiders/Bosker 1999: 1). In most of the empirical analyses with MLA, individuals are nested in contexts. Mainly applied in education science, a typical multilevel design consists of data on pupils nested within classes and/or schools, whereby it is argued that their school achievements, for instance, not only depend on individual capacities but also on class- and school-related factors.

In political science, the use of MLA is increasing. While the idea that context influences individual behavior has been crucial at least since the suicide-analysis by Durkheim (1897), corresponding quantitative analyses in political science in the past century have lacked the methodological base to properly analyze contextual effects. With MLA the simultaneous modeling of individual and contextual determinants to explain individual behavior is now possible (Bühlmann 2006).[11]

The underlying principle of multilevel modeling is quite simple. Intercepts of common OLS regression analysis are allowed to vary:

(5) $Y_{ij} = \beta_{0j} + \beta X_{ij} + \varepsilon_{ij}$, where

(6) $\beta_{0j} = \beta_0 + \mu_{0j}$ (μ_{0j} stands for the residuals at level 2).

11 With pure individual or pure ecological analyses, we cannot explain contextual influence on individual behaviour. For more detailed discussion see Bühlmann 2006; Goldstein 1987: 19; Hox/Kreft 1994: 285; Rasbash et al. 2002; Snijders/Bosker 1999; Steenbergen/Jones 2002; Stoker/Bowers 2002; Teachman/Crowder 2002: 284.

If individual behaviour is modelled as it is most often the case in MLA, such a model implies that individual behaviour can vary between contexts. In other words, unlike standard regression analysis this model does not assume that the constant (β_0) is the same in all units; but rather it varies from context to context. In our case we adapt the common multilevel model as follows: instead of modeling individuals within contexts, we apply the model to multiple observations nested within a unit. More precisely: the data for different points in time is seen as nested within cantons. Thereby, we suppose that the development of educational expenditure over time differs from canton to canton.[12] In so doing, we are able to analyze the effects of time-dependent developments (level 1) and of time-invariant cantonal characteristics (i.e., institutions or historical cultural characteristics) (level 2) separately. While, the number of pupils and of foreign pupils, or the strength of parties in government change over time and are, thus, level-1 variables, most of the institutional and historical cultural variables (e.g., local autonomy, the strength of the Catholic church at the end of the 19th century, or the language region) are constant over time. They are now explicitly modeled as context factors which differently influence all observations of a given canton over time. Furthermore, we can distinguish between the cross-sectional variance (between the cantons) on the one hand, and the variance over time within cantons on the other. To sum up, such a model enables us to model simultaneously different effects due to time (time-dependent variables) and due to cantonal characteristics (time-independent cantonal properties) on cantonal education expenditure.[13] Our standard model, hence, takes the following form.

$$(7) \qquad Y_{ij} = \beta_{0j} + \beta X_{ij} + \alpha W_j + \mu_{0j} + \varepsilon_{ij}$$

Education expenditure in canton j at time point i can be explained by an overall mean (β_0), time-dependent variables (the X variables and their respective β), time-independent cantonal properties (the W variables and their respective α), cantonal variation (μ_{0j} with an assumed mean of 0 and a total between-canton

12 It would also be possible to change the two levels, i.e., to consider the cantons as level 1 and the time dimension as level 2.

13 We do not go further in this article. It is however worth mentioning that MLA makes it possible to also model random slopes or even cross-level interactions. In other words: One can model canton-specific effects on the strength of time effects on education expenditure.

variance of σ_{μ}^{2}), and time variation (ε_{ij} with an assumed mean of 0 and a total within-canton variance of σ^{2}). Hence, the overall variation ($\sigma_{\mu}^{2}+\sigma^{2}$) is divided into differences at the time level (level 1 variance), which shall be explained by time-dependent variables, and differences between contexts (level 2 variance), which shall be explained by cantonal, cross-sectional factors.[14]

However, as with PTSCS the issue of autocorrelation has to be discussed: In the multilevel model, too, correlated errors have to be expected. This is first of all true for level 1, which consists of 26 subsequent and thus related observations for the cantons. Just as in the PTSCS model, in the multilevel model, too, we therefore model the level-1 error term to follow a first-order autocorrelated process ($\rho_i\varepsilon_{ij}$).[15]

(8) $$Y_{ij} = \beta_{0j} + \beta X_{ij} + \alpha W_j + \mu_{0j} + \rho_i\varepsilon_{ij}$$

As we did in the PTSCS, we use this multilevel model for the analysis of the cantonal public education expenditure.

Table 2 presents the results of the multilevel models with mean-corrected public education expenditure as dependent variable. As in the PTSCS analysis, we first estimate four models for the four theoretical approaches. Again, we calculate a joint fifth model by including the significant variables from the previous analyses. To further compare the MLA and the PTSCS design, we estimate a sixth model containing the variables of the PTSCS-model 5 from table 1.

Before calculating the multivariate models we estimate empty models without any specifications in order to assess the variance components. With MLA we can distinguish between the effects of cantonal characteristics and time-related changes. In the empty model we estimate the variance of the education expenditure which can be separated in a level 1 part (differences between time points) and a level 2 part (differences between cantons). The estimation shows

14 We refer to the relevant literature on MLA for a more thorough discussion of the method and expansions of the basic model presented here (Bullen et al. 1994; Ditton 1998; Goldstein 1987, 1995; Hox 1995; Jones 1997; Jones/Duncan 1996; Snijders/Bosker 1999; Teachman/Crowder 2002; for a critic see Rohwer 1998).

15 We use MlwiN to estimate our multilevel models (Rasbash et al. 2002). In this software program it is possible to model autocorrelated errors terms by applying time-series macros (cp. Yang et al. 2004).

that both, differences between the cantons and differences between time points within a canton contribute to the model's total variance: 78% of total variance in the empty model is found at level 2, while 22% can be attributed to level 1. This can be interpreted to mean that differences between countries are more pronounced than differences over time (which may be also due to fact that the data is mean corrected).

It can be seen from table 2 that the same procedure of model building does not result in the same final model as in the PTSCS context (model 5, table 2). Surely, like in the PTSCS model the most crucial variables are again the degree of urbanization and local autonomy. Compared to PTSCS, however, both the language region and the share of (foreign) pupils (complete pooled model 5, table 1), as well as direct democracy and the strength of Catholics at the end of the 19th century (FE model 6, table 1) are not significantly associated with public education expenditure. This is also the case, if the final model identified by the PTSCS method (model 5 and 6 respectively in table 1) is calculated (model 6, table 2). Hence, the results of different model specification are highly consistent in the multilevel context. Finally, and similar to PTSCS, we find high autocorrelation in the model, which is corrected for in the error terms at level 1. As for the model fit, the joint model 5 is now best in parsimoniously explaining cantonal education expenditure.

2.5 Comparison of the results and methodological implications

Formally, the PTSCS- and the multilevel model are quite similar (cf. equations 3 and 4 respectively, and equation 8). Actually, the main difference between the models can be found in the structure of the error term. First and foremost, the multilevel model incorporates an additional residual component which is μ_{0j} and accounts for country-specific random intercepts. This means that compared to PTSCS the multilevel model allows time-series coefficients "an extra stochastic component, an additional uncertainty which influences inferences" (Western 1998: 1234). Hence, as the basic idea of multilevel analysis suggests, the difference between PTSCS and MLA is that the latter allows for cantonal heterogeneity, which does not exist in the PTSCS setup.

Table 2 Multilevel models for the explanation of public education expenditure (mean-corrected)

	Empty Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6 (PTSCS)
Pupils per cantonal population		.06 (.07)					.06 (.07)
Proportion of foreign pupils		.05 (.04)					.05 (.04)
Sector 1		-.04 (.05)					
Urbanization		.40*** (.14)				.32** (.15)	.26* (.15)
Leftists in government			.04 (.03)				.03 (.03)
CVP in government			.01 (.06)				
Union			.13** (.05)			.11** (.05)	
Direct democracy				.02 (.09)			.04 (.09)
Governmental concordance				.03 (.02)			
Local autonomy				-.59*** (.17)		-.43** (.17)	-.42** (.17)
Strength of Catholics 19th cent.					-.26* (.15)	-.04 (.14)	-.05 (.13)
Efforts in education policy 1880					.05 (.14)		
Language					-.32** (.15)	.04 (.15)	-.03 (.16)
constant	-.12 (.17)	-.08 (.13)	-.12 (.16)	-.10 (.13)	-.12 (.15)	-.07 (.12)	-.08 (.11)
Between-variance of cantons (σ^2_μ)	*.64*** (.20)*	*.37*** (.14)*	*.62*** (.19)*	*.40*** (.13)*	*.48*** (.15)*	*.31** (.10)*	*.26** (.09)*
Between-variance of between time-points within cantons (σ^2)	*.18*** (.03)*	*.18*** (.03)*	*.17*** (.03)*	*.17*** (.03)*	*.18*** (.03)*	*.17*** (.03)*	*.18*** (.03)*
N	*624*	*624*	*624*	*624*	*624*	*624*	*624*
Autocorrelation level 1	*.798*	*.798*	*.798*	*.798*	*.798*	*.798*	*.798*
-2loglikelihood	*148.1*	*136.3*	*140.00*	*135.7*	*141.6*	*126.6*	*127.5*
AIC	*154.1*	*150.3*	*152.0*	*147.7*	*153.6*	*142.6*	*149.5*

Notes: Non-standardized coefficients; standard-errors in brackets; time-independent variables, i.e. variables explaining contextual level variance, are highlighted (see Notes Table 1). Estimation with iterative generalized least squares (Goldsmith 1995). A first-order autocorrelation process was modeled in order to correct for autocorrelation at level 1 (MlwiN TS macro).

From a substantial point of view both methods show that structural and institutional characteristics are most important for the explanation of public education expenditure in the cantons (see tables 1 and 2). Particularly, cantonal education expenditure is high, where an urbanized structure demands high investments in education. In contrast, a high degree of decentralized decision making structures (e.g. substantial local autonomy in education) hinders extensive public education expenditure. Other factors such as the number of (foreign) pupils, direct democracy, the strength of Catholicism at the end of the 19th century, and the linguistic region tend to somehow influence cantonal education expenditure, the results are however not consistent and obviously depend on the model specification and estimation method. These inconsistent results may be due to fact that the two approaches handle heterogeneity differently. While in PTSCS heteroscedasticity and spatial autocorrelation is corrected for by calculating panel-corrected standard errors, in multilevel analysis heteroscedasticity is not corrected for, but modelled by distinguishing two levels of analysis. In other words: It is seen as "substance" and not as "nuisance" (Jones 1997).

Moreover, our results indicate that the multilevel design provides some clear advantages. The first aspect concerns the cross-sectional and time-related variance respectively. Most of our explanatory variables do not change much over time or are even constants, as it is typical for institutional and cultural variables in political research. This means that, logically, our model should be better in explaining cross-sectional variance than the development over time. Empirically, this can be best seen in the multilevel models: While the final models explain a substantial part of the variance between the cantons, namely about 60 percent, the variance within the cantons, i.e. the development over time, is not reduced at all. This means that the model is actually not able to explain variations in public education expenditure over time. Against the background that our model includes a quite typical set of explanatory variables for comparative political research these findings imply that one of the main advantages of pooling, i.e. the modeling of "*dynamics* inherent in the panel" (Plümper et al. 2005: 334), must be indeed questioned. Second, our results confirm the claim made by Shor et al. (2007) that multilevel models generally provide a better model fit. A glance at the AIC for the different models confirms that the multilevel models do better in explaining the differences in cantonal education expenditure and should therefore be preferred. This is also the case if we model fixed unit effects in the PTSCS context. Moreover, different to PTSCS MLA produces a final model which exhibits the best model fit, which again speaks for the higher sensibility and consistency of multilevel models.[16]

3 Conclusion

The aim of this paper was to compare multilevel analysis and pooled time-series cross-section designs in the analysis of time and space in comparative political research. It has to be mentioned that by means of our analyses we are not able to actually *prove* whether MLA performs better than PTSCS, since such a comparison would require that the real error structures of the data are known which is normally not the case in empirical data. However, while such a systematic evaluation has recently been provided by Shor et al. (2007), this was not the aim of this paper. Rather, our analyses allow for some implications in terms of practical application of PTSCS and MLA respectively. From the discussion of the two approaches and the empirical examples we can make the following remarks that may be important for empirical analyses of time-serial cross-section data in comparative political research:

Generally, a simple two-level model with repeated measures (level 1) nested within units (level 2) and correlated error terms roughly corresponds to a basic pooled time-series cross-sectional model. The main difference is that the multilevel model allows for random intercepts. This can be seen as "partial pooling where the estimated parameter has the potential to borrow strength from other parameters in the model." (Shor et al. 2007: 4f.). As a result, both methods produce quite similar results for our data, implying that urbanized structures and high local autonomy are crucial for explaining cantonal education expenditure. For the rest, multilevel analysis is more conservative, but also more consistent in that it finds no further effects where some PTSCS-specifications do, thus, providing inconclusive results. Goodness of fit measures (AIC) thereby indicate that multilevel models better explain the variance in the data and should therefore be preferred.

16 It has to be mentioned that in this paper very simple models have been chosen. Both approaches would offer some more sophisticated procedures that would probably lead to better model fits (but probably also to more problems).

Comparing the two designs, MLA moreover offers some advantages particularly with regard to conceptual and theoretical aspects.

Firstly, a multilevel design is theoretically better suitable for modeling the effect of time-invariant variables (e.g., institutions) which often take center stage in comparative political research. If (almost) time-invariant variables are integrated into a PTSCS design the corresponding hypothesis is that a particular (not or very slowly changing) variable parameter at time *t* (or t-x if a time-lag is incorporated) influences the dependent variable at time *t*. However, such a theoretical link between (time-invariant) explanatory factors and the dependent variables is often questionable. This is most pronounced in the presence of serial autocorrelation, when such a model implies that specific time-invariant characteristics explain some short-term changes in the dependent variable. In such a context a multilevel design is theoretically more sensible because it is based on the assumption that time-invariant factors build a unit-specific context that influences all observations of this unit in a particular way.

Secondly, a multilevel design allows for a clear distinction between cross-sectional effects and time effects.[17] It is possible to analyze to what extent cross-sectional differences and developments over time each account for the overall variance in the data. Concerning substantial interpretation of the estimation results this feature is very useful: In our case it shows that the models do quite well in explaining cross-sectional variance, while they are not able to explain developments over time.

Thirdly, heterogeneity in the data is not just corrected for as in the pooled time-series cross-section analysis, but is modeled. Conceptually this is a more elegant way to handle the statistical problem, because heterogeneity is not just seen as a problem that has to be eliminated, but as an important (realistic) characteristic of the data that should be adequately modeled.

A further advantage is that while in this paper we applied a very simple model in order to make comparison easier, MLA would allow for various model extensions, be it the introduction of an additional level (e.g., an intervening municipality level), the estimation of random slopes or even cross-level interactions (e.g., the effect of urbanization on the strength of relationship between for-

17 In the case presented in this paper time effects corespond to the relative changes in the cantons' position over time.

eign pupils and education expenditure), or the calculation of growth curves in order to model particular developments over time (Bryk/Raudenbush 1987; Cnaan et al. 1997; Hox 2000; Huggins/Loesch 1998; Luke 2004; Yang/Goldstein 1996). In a multilevel design more sophisticated models are relatively easy to apply, which could be especially helpful when working with data reflecting changes over time.

However, multilevel analysis also has its shortcomings in practical applications. First and foremost, the necessity of large and particularly structured data sets has to be mentioned. Indeed, in order to use MLA, data has not only to be measured at different levels, but additionally each group at all levels should consist of at least 25 observations in order to get valid estimations for multivariate analyses.[18] Obviously, in comparative political research such data is still quite rare. Furthermore, the analysis and handling of crucial statistical pitfalls is much better developed in the methodological discussion of PTSCS than it is in the multilevel context. Mainly in relation to serial autocorrelation, which is a crucial aspect in case of repeated measurements, procedures to correct for the correlated errors are much better elaborated and standardized in pooled time-series cross-section designs than in multilevel analysis. When carrying out a MLA with data that includes developments over time, it is useful to consider the methodological debate on non-stationarity and serial autocorrelation in the PTSCS-context.

To conclude, we would propose the following procedure for the analysis of time and space in comparative political research: If a large data set satisfying the requirements for multilevel analysis is available, the implementation of a multilevel design is recommended due to its statistical, conceptual and theoretical strengths discussed above. In smaller data samples a PTSCS design might generally be more appropriate. In both situations, however, we think a triangulation of methods worthwhile. In the multilevel case additional analyses from the PTSCS context help clarifying and understanding problems of autocorrelation and non-stationarity. In smaller data samples the calculation of empty multilevel models allows at least for an approximate distinction between longitudinal and cross-sectional variances in the data. Hence, the combination of pooled time-series cross-

18 If one is only interested in an analysis of variances at different levels the case-restriction is less problematic. If we have only few cases, but many time points within these cases it makes sense to control for the effects of the level 2-variance with MLA. However, in this case we would not specify level 2 with independent variables.

section analysis and multilevel analysis offers the possibility of profiting from the advantages of both methods, but also to compensate for their weaknesses.

References

Année politique Suisse, various years: Schweizerische Politik. Bern: Institut für Politikwissenschaft.

Archer, M. S., 1979: Social Origins of Educational Systems. London/Beverly Hills: Sage.

Bassand, M., 1988: Les enjeux de l'urbanisation. Bern/Frankfurt a. M./Paris: Lang.

Beck, N. and J. Katz, 1995: What to do (and not to do) with time-series cross-section data. American Political Science Review 89(3): 634-647.

Boix, C., 1997: Political Parties and the Supply Side of the Economy. The Provision of Physical and Human Capital in Advances Economies 1960-90. American Journal of Political Science 41(3): 814-845.

Bryk, A. S. and S. W. Raudenbush, 1987: Application of Hierarchical Linear Models to Assesing Change. Psychological Bulletin 101(1): 147-158.

Bühlmann, M., 2006: Politische Partizipation im kommunalen Kontext. Der Einfluss lokaler Kontexteigenschaften auf individuelles politisches Partizipationsverhalten. Bern: Haupt.

Bullen, N., K. Jones and C. Duncan, 1994: Modelling complexity-analysing between-individual and between-place variation. A multilevel tutorial. Mimeo. Portsmouth: University of Portsmouth, Department of Geography.

Burnham, K. P. and D. R. Anderson, 2004: Multimodel Inference. Understanding AIC and BIC in Model Selection. Sociological Methods Research 33: 261-304.

Cameron, D. R. and R. I. Hofferbert, 1974: The Impact of Federalism on Education Finance. A Comparative Analysis. European Journal of Political Research 2: 225-258.

Castles, F. G., 1989: Explaining public education expenditure in OECD nations. European Journal of Political Research 17: 431-448.

Castles, F. G., 1999: Comparative Public Policy. Patterns of Post-war Transformation. Cheltenham/Northampton: Edward Elgar Publishing Limited.

Cnaan, A., N. M. Laird and P. Slasor, 1997: Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. Statistics in Medicine 16: 2349-2380.

Ditton, H., 1998: Mehrebenenanalyse. Grundlagen und Anwendungen des Hierarchisch Linearen Modells. Weinheim/München: Juventa.

Eidgenössische Finanzverwaltung (EFV), various years: Öffentliche Finanzen der Schweiz. Neuchâtel: Bundesamt für Statistik.

Fernandez, R. and R. Rogerson, 1997: The Determinants of Public Education Expenditure. Evidence from the States 1950-1990. NBER-Working Paper 5995. Cambridge: National Bureau of Economic Research.

Freitag, M. and M. Bühlmann, 2003: Die Bildungsfinanzen der Schweizer Kantone. Der Einfluss sozioökonomischer Bedingungen, organisierter Interessen und politischer Institutionen auf die Bildungsausgaben im kantonalen Vergleich. Swiss Political Science Review 9(1): 139-168.

Goldstein, H., 1987: Multilevel Models in Educational and Social Research. London/New York: Oxford University Press.

Goldstein, H., 1995: Multilevel Statistical Models. New York: Halstead Press.

Gujarati, D. N., 1995: Basic Econometrics. Third Edition. New York: McGraw-Hill International Editions.
Hadri, K., 2000: Testing for stationarity in heterogeneous panel data. Econometrics Journal 3: 148-161.
Hega, G. M., 1999: Consensus Democracy? Swiss Education Policy between Federalism and Subsidiarity. New York: Peter Lang.
Heidenheimer, A. J., 1999: External and Domestic Determinants of Education Expansion. How Germany, Japan, and Switzerland Have Varied. Governance 6(2): 194-219.
Hibbs, D., 1977: Political Parties and Macroeconomic Policy. American Political Science Review 71: 1467-1487.
Hox, J. J., 1995: Applied Multilevel Analysis. Den Haag: CIP-Gegevens Koninklijke Bibliotheek.
Hox, J. J., 2000: Multilevel Analyses of Grouped and Longitudinal Data. Pp. 15-32 in: T. D. Little, K. U. Schnabel and J. Baumert (Eds.): Modeling longitudinal and multilevel data. Mahwah. NJ: Lawrence Erlbaum Associates.
Hox, J. J., and I. G. Kreft, 1994: Multilevel Analysis Methods. Sociological Methods & Research 22(3): 283-299.
Hsiao, C., 2003: Analysis of Panel Data. 2. Edition. Cambridge: Cambridge University Press. http://assets.cambridge.org/0521818559/sample/0521818559WS.pdf (21.8.2003).
Huggins, R. M. and D. Z. Loesch, 1998: On the analysis of mixed longitudinal growth data. Biometrics 54: 583-95.
Hunziker, O., 1882: Geschichte der Schweizerischen Volksschule in gedrängter Darstellung. Vol. 3. Zürich: Friedrich Schulthess.
Im, K. S., H. M. Pesaran and Y. Shin (Eds.): Testing for unit roots in heterogeneous panels. Journal of Econometrics 115(1): 53-74.
Jones, K., 1997: Multilevel Approaches to Modelling Contextuality. From Nuisance to Substance in the Analysis of Voting Behaviour. Pp. 19-43 in: G. P. Westert and R. N. Verhoeff (Eds.): Places and people. Multilevel modelling in geographical research. Utrecht: The Royal Dutch Geographical Society.
Jones, K. and C. Duncan, 1996: People and places. The multilevel model as a general framework for the quantitative analysis of geographical data. Pp. 79-104 in: P. Longley and M. Batty (Eds.): Spatial Analysis: Modelling in a GIS Environment. New York: John Wiley and Sons.
Jones, K. and C. Duncan, 1996: People and places. The multilevel model as a general framework for the quantitative analysis of geographical data. In: P. Longley and M. Batty (Eds.): Spatial Analysis: Modelling in a GIS Environment. New York: John Wiley and Sons.
Kittel, B., 1999: Sense and Sensitiveness in Pooled Analysis of Political Data. European Journal of Political Research 35(2), 225-253.
Kittel, B. and H. Winner, 2002: How Reliable is Pooled Analysis in Political Economy? The Globalization-Welfare State Nexus Revisited. Max-Planck-Institut für Gesellschafts-forschung: Discussion Paper 2(3).
Ladner, A., 1994: Finanzkompetenzen der Gemeinden – ein Überblick über die Praxis. Pp. 64-85 in: F. Eng, A. Glatthard and B. H. Koenig (Eds.): Finanzföderalismus. Zürich: Emissionszentrale der Schweizer Gemeinden.
Levin, A. T. and C.-F. Lin, 1993: Unit Root Tests in Panel Data. New Results. University of California at San Diego: San Diego Discussion Paper (93-56).
Linder, W., 1994: Swiss Democracy. Possible Solutions to Conflict in Multicultural Societies. New York: St. Martin's Press.

Linder, W., 2005: Schweizerische Demokratie. Institutionen, Prozesse, Perspektiven. 2nd Edition. Bern: Paul Haupt.

Lijphart, A., 1999: Patterns of Democracy. Government forms and Performance in Thirty-Six Countries. New Haven/London: Yale University Press.

Luke, D. A., 2004: Multilevel Modeling. Thousand Oaks: Sage.

Pennings, P., H. Keman and J. Kleinnijenhuis, 1999: Doing Research in Political Science. An Introduction to Comparative Methods and Statistics. London/Thousand Oaks/New Delhi: Sage Publications.

Plümper, T. and V. E. Troeger, 2007: Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects. Political Analysis 15: 124-139.

Plümper, T., V. E. Troeger and P. Manow, 2005: Panel data analysis in comparative politics. Linking method to theory. European Journal of Political Research 44: 327-354.

Podestà, F., 2002: Recent Developments in Quantitative Comparative Methodology. The Case of Pooled Time Series Cross-Section Analysis. Universität Brescia: Discussion Paper (3-02). http://www.eco.unibs.it/segdss/paper/pode202.pdf (21.8.2003).

Popper, K. R., 1992: Die offene Gesellschaft und ihre Feinde. Vol. 2. Tübingen: J.C.B Mohr.

Poterba, J. M., 1999: Demographic Structure and the Political Economy of Public Education. Cambridge: Working Paper 5677 of National Bureau of Economic Research.

Rasbash, J., W. Browne, H. Goldstein, M. Yang, I. Plewis, M. Healy, G. Woodhouse, D. Draper, I. Langford and T. Lewis, 2002: A user's guide to MLwiN. Version 2d for use with MLwiN 1.10. London: Centre for Multilevel Modelling, Institute of Education, University of London.

Rohwer, G., 1998: Mehrebenenmodelle und das Mikro-Makro-Problem in der empirischen Sozialforschung. Bochum: Ruhr Universität, Fakultät für Sozialwissenschaften.

Schaltegger, C., 2001: Ist der Schweizer Föderalismus kleinräumig? Swiss Political Science Review 7(1): 1-18.

Schmidt, M. G., 2000: Die sozialpolitischen Nachzüglerstaaten und die Theorien der vergleichenden Staatstätigkeitsforschung, Pp. 22-36 in: H. Obinger and U. Wagschal (Eds.): Der gezügelte Wohlfahrtsstaat. Sozialpolitik in reichen Industrienationen. Frankfurt am Main/New York: Campus.

Schmidt, M. G., 2002: Warum Mittelmaß? Deutschlands Bildungsausgaben im internationalen Vergleich. Politische Vierteljahresschrift 43 (1): 3-19.

Shor B., J. Bafumi, L. Keele and D. Park, 2007: A Bayesian Multilevel Modeling Approach to Time Series Cross-Sectional Data. Political Analysis 15: 165-181.

Steenbergen, M. R. and B. S. Jones, 2002: Modelling Multilevel Data Structures. American Journal of Political Science 46(1): 218–237.

Stoker, L. and J. Bowers, 2002: Designing multilevel studies. Sampling voters and electoral contexts. Pp. 77-109 in: Mark N. Franklin and C. Wlezien (Eds.): The Future of Election Studies. Amsterdam: Pergamon.

Stutzer, A., 1999: Demokratieindizes für die Kantone der Schweiz. Zürich: Working Paper 23 of Institutes for Empirical Research in Economic.

Teachman, J. and K. Crowder, 2002: Multilevel Models in Family Research. Some Conceptual and Methodological Issues. Journal of Marriage and Family 64(2): 280-294.

Vatter, A., 2002: Kantonale Demokratien im Vergleich. Entstehungsgründe, Interaktionen und Wirkungen politischer Institutionen in den Schweizer Kantonen. Opladen: Leske & Budrich.

Vatter, A. and M. Freitag, 2002: Die Janusköpfigkeit von Verhandlungsdemokratien. Zur Wirkung von Konkordanz, direkter Demokratie und dezentralen Entscheidungsstrukturen auf den öffentlichen Sektor der Schweizer Kantone. Swiss Political Science Review 8(2): 53-80.

Vatter, A. et al., 2004: Politische, soziale und ökonomische Daten zu den Schweizer Kantonen, 1983-2000. Bern: Institut für Politikwissenschaft.

Western, B., 1998: Causal Heterogeneity in Comparative Research. A Bayesian Hierarchical Modelling Approach. American Journal of Political Science 42 (4): 1233-1259.

Wilensky, H., 1981: Leftism, Catholicism, and Democratic Corporatism. The Role of Political Parties in Recent Welfare State Development. Pp. 345-382 in: P. Flora and A. J. Heidenheimer (Eds.): The Development of Welfare State in Europe and America. New Brunswick: Transaction.

Yang, M., and H. Goldstein, 1996. Multilevel Models for longitudinal data. Pp. 191-220 in: U. Engel and J. Reinecke (Eds.): Analysis of Change, Advanced Techniques in Panel data Analysis. Berlin: Walter de Gruyter.

Contact Addresses:

Isabelle Stadelmann-Steffen
University of Konstanz
Department of Politics and Management
P.O. Box D84
D-78457 Konstanz
Isabelle.Stadelmann-Steffen@uni-konstanz.de

Marc Bühlmann
Universität Zürich
Zentrum für Demokratie
Küttigerstraße 21
CH-5000 Aarau
Marc.buehlmann@zda.uzh.ch

Appendix Variables, Hypotheses, and Operationalization

Variable	Hypothesis (ceteris paribus)	Expected Relationship	Operationalization / Source
Dependent variable			
Public education expenditure per capita, mean corrected			Cantonal education expenditure in CHF per capita (for elementary school and vocational education, less federal contributions), according to the functional classification of the expenditure of cantons and communes, annual mean corrected. Source: EFV (various years).
Socio-economic variables			
Pupils per cantonal population	The more pupils a canton has, the higher is cantonal public education expenditure.	+	Number of pupils in elementary school per cantonal population. Source: Statistisches Jahrbuch (various years).
Proportion of foreign pupils	The higher the share of foreigners in school, the higher is expenditure for educational integration.	+	Share of foreigners in elementary schools, in percent. Source: Statistisches Jahrbuch (various years).
Employment in the third sector	The higher the employment in the service sector, the higher is cantonal education expenditure.	+	Share of employed in the third sector according to the federal census of enterprises 1991, in percent. Source: Swiss Federal Statistical Office.
Urbanization	The higher urbanization, the higher is cantonal public education expenditure.	+	Share of residents living in urban areas, in percent. Source: Own calculation on the basis of Bassand (1988).
Party-political variables			
Strength of leftist parties in the government	The higher the share of leftist parties in the government, the higher is public education expenditure in a canton.	+	Share of seats in cantonal government held by SPS, DSP and the Green parties. Source: Année Politique Suisse (various years).
Strength of Christian-democratic parties in the government	The higher the share of Christian democratic parties in the cantonal government, the lower the public education expenditure turn out to be.	-	Share of seats in cantonal government held by Catholic parties (CVP and CSP). Source: Année Politique Suisse (various years).
Strength of unions	The better unions in a canton are organized, the more this canton spends for public education.	+	Union-membership in percent of total employment. Source: Own calculation on the basis of the Swiss federation of trade unions and BAK Basel.

Variable	Hypothesis (ceteris paribus)	Expected Relationship	Operationalization / Source
Institutional variables			
Local autonomy	The more decentralized the decision making structures in a canton are, the lower is public education expenditure.	-	Degree of communal autonomy in relation to the Federation and the Canton (index) according to Ladner (1994: 81) (0 = no autonomy at all; 10 = very substantial autonomy).
Direct democracy	The more extensive direct democratic rights in a canton are, the lower is public education expenditure.	-	Index of direct democracy according to Stutzer (1999); 1978-1991: value for 1970; 1992-1995: value for 1992; 1996-2003: values for 1996.
Consensus democracy	The higher the consensual alignment of the government is, the more a canton spends for public education.	+	Summarized percentages of votes of the governing parties (if the majority rule is used for parliamentary elections (Grisons, Uri, Appenzell Inner- and Outer-Rhodes) the summarized share of parliamentary seats of the governing parties is calculated instead). Source: Année Politique Suisse (various years).
Historical and cultural variables			
Strength of the Catholic Church	The stronger the influence of the Catholic Church on education was at the beginning of the education expansion (end 19th century), the lower is public education expenditure today.	-	Estimation of the influence of the catholic church when the cantonal educational system was built up. 1 = weak influence; 2 = medium influence; 3 = strong influence. Source: Hunziker (1882: 304-367).
Efforts in education policy 1880	The higher the efforts of the government in education were around 1880, the higher are investments in public education still today.	+	Efforts in public education of the cantonal governments around 1880. Index of ranking points based on five indicators: expenditure per capita, pupils per capita, pupil-teacher-quota, the introduction of an education law, number of school weeks. 1=little effort; 2=medium effort; 3=strong effort. Source: Hunziker (1882: 304-367).
Language	German-speaking cantons spend less for public education than the Latin cantons.	-	Dummy-variable. 1= majority of German-speaking inhabitants, otherwise =0. Source: Swiss federal statistical office, Statistisches Jahrbuch der Schweiz.

Note: The data used is based on the data set of Vatter et al. (2004).