Federal Governance and Public Production of Education: A Political Economy Perspective

Carolin Amann [a] Bernd Süssmuth [a] Robert K. von Weizsäcker [a b c d]

[a] *Munich University of Technology, Germany*[*]

[b] *CEPR London* [c] *CESifo Munich* [d] *IZA Bonn*

This version: February 28, 2007

ABSTRACT. This paper investigates the optimal design of federal governance for the public production of education. In a simple yardstick competition model of vote-seeking incumbents, we address the question of the ideal organization of education in a federal system. In our empirical exercise, we rely on student achievement data pooled by test subjects. They are drawn from the national extensions of the PISA 2000 and 2003 studies for Germany. We use them to analyze two central aspects of the model: First, the announcement effect of achievement test results on the popularity function of state governments, and, second, production inefficiencies by test subjects at the state level. Our method of choice is an event study analysis and a stochastic frontier approach, respectively. In both cases, our estimates support the model.

Keywords: Secondary schooling, federal governments, production efficiency
JEL classification: I21, H77

---

[*] Correspondence: TUM, Department of Business and Economics, Arcisstr. 21, D-80333 Munich E-mail: `suessmuth@wi.tum.de`. We would like to thank Joachim Behnke and Christian Feilcke for helpful comments and discussion. Bernd Süssmuth gratefully acknowledges the hospitality of the economics department at UC Santa Barbara, where parts of this paper were finalized.

1. Introduction and Motivation

The impending reform of federal structures in Germany will almost certainly lead to a less standardized organization of the German education system. Major parts of decision-making authorities will be relocated from the Federal Government (*Bund*) to the lower-tier governments of Federal States (*Länder*). For public schooling the reform implies an imminent heterogeneity with regard to structural cornerstones such as tracking norms, the classification of education levels, examination standards, etc. The opposite applies, for example, to Switzerland,[1] where the advisory board of education directors in the 26 cantons[2] opts for a reform designed to "harmonize" and standardize existing cantonal education practices. The "HarmoS" reform program of the EDK takes the form of a legally binding state treaty proposal. Among others, it defines superordinate targets for compulsory schooling, sets out guidelines of quality assurance, and decrees mandatory educational standards across cantons.[3] It was put on the vote in May 2006. After its ratification, the HarmoS Concordat will be fully enacted in 2009. These observations on the development of federal governance of the education system of two neighboring democracies in opposite directions raise the fundamental question of the optimal allocation of public education responsibilities to the two levels of government in a federal system.

Proponents of a reform targeted toward more autonomy at the Federal State level argue along the following lines: A nationwide agreeing on and setting of binding educational norms, for example in the form of minimum average student achievements, would necessarily lead to a substandard decree. They attribute this to an externality of democratic decision-making inherent to the federal system. The argument is simple and can best be illustrated by the following scenario: Suppose that for rea-

[1]Similarly, in the early 1970s, a California Supreme Court ruling transferred responsibilities for public schools from localities to the state. The ruling led to similar rulings in other states, although no state reformed its structure quite as radically as California (Sonstelie, 2003).

[2]*Schweizerische Konferenz der kantonalen Erziehungsdirektoren*, henceforth EDK.

[3]For detail see `http://www.edk.ch`.

sons of an unspecified x-inefficient inertia each Federal State wants to stick to its realized (average) level of student performance and proposes the respective figure as the new superordinate target. In the case of German *Länder*, for example, the highest educational standard capable of winning a majority in the Federal Council, i.e. the upper house of German parliament (*Bundesrat*), would then correspond to the proposal made by the council members of Schleswig-Holstein (SH). This acceptable standard, however, would lie below the average student performance of all states. Table 1 makes the point.

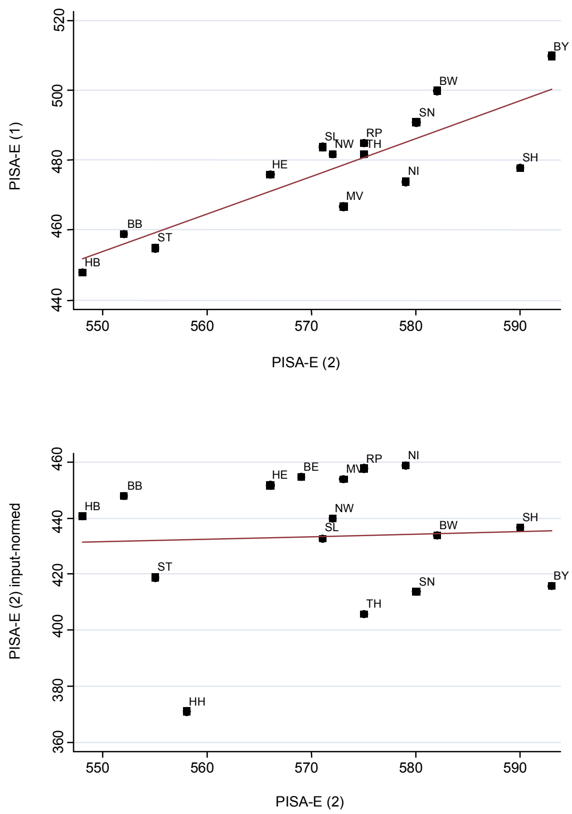
Table 1. PISA-E 2000 scores and votes in the Federal Council: German Länder

Federal State	PISA-E (1)	PISA-E (2)	PISA-E (2) per input units	Votes	Votes cumulated
Bavaria (BY)	510 [1]	593 [1]	416 [13]	6	6
Baden-Württemb. (BW)	500 [2]	582 [3]	434 [10]	6	12
Saxony (SN)	491 [3]	580 [4]	414 [14]	4	16
Rhinel.-Palatinate (RP)	485 [4]	575 [6]	458 [2]	4	20
Saarland (SL)	484 [5]	571 [10]	433 [11]	3	23
N. Rhine-Westph. (NW)	482 [6]	572 [9]	449 [8]	6	29
Thuringia (TH)	482 [7]	575 [7]	406 [15]	4	33
Schleswig-Holstein (SH)	478 [8]	590 [2]	437 [9]	4	37
Hesse (H)	476 [9]	566 [12]	452 [5]	5	42
Lower Saxony (NI)	474 [10]	579 [5]	459 [1]	6	48
M.-W. Pomerania (MV)	467 [11]	573 [8]	454 [4]	3	51
Brandenburg (BB)	459 [12]	552 [15]	448 [6]	4	55
Saxony-Anhalt (ST)	455 [13]	555 [14]	419 [12]	4	59
Bremen (HB)	448 [14]	548 [16]	441 [7]	3	62
Hamburg (HH)		558 [13]	371 [16]	3	65
Berlin (BE)		569 [11]	455 [3]	4	69

Note: Respective ranks given in square brackets.

The last column displays the cumulated votes of *Länder* in the Federal Council ranked in decreasing order of student achievements according to the national extension of the OECD Program for International Student Assessment 2000 study (PISA-E) reported in the second column.[4]

Figure 1. PISA-E 2000 average student test scores: German Länder



Note: (1) – all school tracks, (2) – Gymnasium (in per input units)

[4]PISA-E was conducted simultaneously to the PISA 2000 test. It covers 15 year old students and ninth graders. It represents a supplement to the international test in the form of national test items. Its sample size is several times the one of the international test (5,000) comprising two overlapping samples of 33,809 15 year old students and 33,744 ninth graders. The overlap is 47 percent.

As can be seen from the entries of the last column in Table 1, the vote of SH is the pivotal one as it represents the median vote (37 of 69 possible votes). Notably, the SH score slightly falls below the average score of all *Länder*. This result suggests that only substandard norms are capable of winning a majority if they have to be voted on and cannot be set and decided upon autonomously. In the course of the present paper we will demonstrate that this line of reasoning represents a fatal fallacy. A straightforward objection is the reverse side of the educational production process – or in other words the resources spent by the state to generate the respective student performance – that this kind of argumentation obviously neglects. The third column of Table 1 displays the average student performances of students enrolled in high schools and comprehensive schools (*Gymnasien*) across states tested in the course of PISA-E 2000.[5] The fourth column reports these test outputs in per input units.[6] The second scatter plot in Figure 1 shows how markedly these relative figures differ from the absolute test scores. However, we should not fail to mention that this type of average products merely represent more than the most crude measure of productivity (Bonaccorsi and Daraio, 2004). Hence, a central aspect of our study is to estimate and discuss more meaningful measures of the efficiency of public educational production against the background of the recent policy issues of accountability, value-for-money, and cost-effectiveness (Worthington, 2001).

Given a parsimonious theoretical model and adequate econometric techniques to test its central implications, our study seeks to answer the following questions: What

[5]We concentrate on this school type as only for the *Gymnasien* a representative sample containing all *Länder* results (in particular, results of the city states Hamburg and Berlin) is available. However, as can be seen from the first diagram of Figure 1, they proxy the average scores for all examined school types across states very well.

[6]We simply divided the average scores by the cumulated nominal lessons taught (aggregated over the tested fields) plus the cumulated material resources in the respective Federal State. The cumulation was done from first to ninth grade covering a period from 1991 to 1999. The per input units have been multiplied by a factor of 10^4. Curiousy, students' "exposure" to particular subjects in terms of cumulated hours is in contrast to class size a widely neglected measure of educational resources hitherto.

is the optimal allocation of educational competencies among upper-tier and lower tier governments of a federal system? Do perceived discrepancies in student test results among states affect the re-election probability of state governments such that educational output really interacts with political objective functions? Does more or less standardization of education practices across provinces enhance the efficiency in educational production? As we deal with the assignment of policy responsibilities to provinces on the one hand and to the federal state on the other as well as with collective decision-making at the federal level (Table 1), we address the two central path-independent issues of (federal) constitutional choice according to Inman and Rubinfeld (1997).

The rest of the paper is organized as follows: Section 2 outlines our stylized theoretical model of political yardstick competition taking into account the double-tiered government structure of federal systems. Evidence of central implications of this model based on German *Länder* data is presented in Section 3. It consists of three parts: First, we analyze voting intentions and PISA test score announcements in an event study framework. Second, we briefly outline the stochastic frontier techniqes that we rely on in order to assess and quantify the educational production efficiency of German province governments. In the course of our efficiency estimates, we investigate central hypotheses established in the literature. Finally, we discuss our results. Section 4 concludes.

2. Model

The basic yardstick competion model of vote-seeking incumbents that will be outlined in the present section resembles the theoretical framework of the seminal contributions by Salmon (1987) and Besley and Case (1995) in the following points:[7] Voters are assumed to make comparisons between jurisdictions. However, in our model they are largely modelled as exogenous to the two agents game played by the Federal

[7]This type of political yardstick competition modelling is sometimes also referred to as 'Salmon Mechanisms.' A recent and comprehensive survey of the related theoretical and empirical literature is given in Bodenstein and Ursprung (2005).

Government and the lower-tier (province) governments. Nevertheless, as in Besley and Case (1995) voting is the main incentive mechanism to discipline incumbents with regard to a best practice education policy. Voters are able to appraise incumbents' *relative* performance. From the media or other sources, voters can gain access to information about what other incumbents are doing and the corresponding (average) achievements of the respective state's students, which serves as a benchmark for their own jurisdiction. These premises force incumbents into a yardstick competition in which they care about what other incumbents are doing. In the following model, however, we adhere to Bodenstein and Ursprung (2005) in explicitly considering the double-tiered governance in federal systems. The degree of competition the province governments are exposed to is up to the upper-tier government. More standardization of education practices fosters yardstick competition. More autonomy at the state level reduces it.

The Federal States

Let us suppose the following most simple (linearized) educational production process

$$a_i = e_i + \theta_i + \varepsilon_i, \tag{1}$$

where a_i denotes the observed average students' achievement level in state i, e_i (deterministic) effort in terms of gubernatorial resources spent by governor i, θ_i some idiosyncratic efficiency to transform effort into output, and ε_i an i.i.d. disturbance term.

Analogously, for a group of reference states we may write $a = e + \theta + \varepsilon$. Since average test scores are stochastic varibales, we suppose for the production efficiencies

$$E\theta = E\theta_i = \mu,\ \mathrm{var}\,(\theta_i) = \sigma^2,\ \mathrm{cov}\,(\theta, \theta_i) = r\sigma^2 \tag{2}$$

and for the erratic shocks $E\varepsilon_i = E\varepsilon = 0$, $\mathrm{var}(\varepsilon_i) = \sigma_\varepsilon^2$, $\mathrm{cov}(\varepsilon, \varepsilon_i) = 0$.

Following Besley and Case (1995), we allow voters to be able to appraise incumbents' *relative* performance, i.e. $a_i - a$. From the media or other sources, they can gain access to information about what other incumbents are doing and the corresponding (average) achievements of the respective state's students, which serves as a benchmark for their own jurisdiction. We, therefore, model the re-election probabilty R as a positive function of the lead in terms of student achievements with regard to the benchmark group of states:[8]

$$R_i(a_i, a) = \overline{R} + \alpha a_i - \delta a = \overline{R} + \alpha\delta\left(\frac{1}{\delta}a_i - \frac{1}{\alpha}a\right), \tag{3}$$

where $\alpha, \delta \geq 0$ are part of the policy space set by the upper-tier government (the Federation) and $\overline{R}$ is constant subsuming all factors unrelated to education policies. As simple as equation (3) may seem, it captures a variety of policy solutions. For $\alpha \to 0$ ($\delta \to 0$), for example, we face the two corner solutions, for which decreed standards by the Federal Government based on absolute (relative) measures tend to zero.[9]

With regard to the gubernatorial utility function we plausibly suppose it to be positively dependent on the respective re-election probability R_i (and, therefore, on the reputational lead in terms of education policy). It is also assumed to negatively depend on invested public resources which are ultimately payed by the voters that is the state's taxpayers. Since politicians are risk averse agents, uncertainty of the election outcome enters the utility function with a negative sign. Consequently, we write the expected utility function of incumbents at the state level as

$$E(U_i) = E\left(R_i - \frac{1}{2}e_i^2\right) - \frac{1}{2}\gamma \text{var}(R_i). \tag{4}$$

From (1), (2), and (3), we can rewrite the two central components of (4) as

$$\begin{aligned} E\left(R_i - \frac{1}{2}e_i^2\right) &= \overline{R} + (\alpha - \delta)\mu + \alpha e_i - \delta e - \frac{1}{2}e_i^2 \text{ and} \\ \text{var}(R_i) &= \left(\alpha^2 + \delta^2\right)\left(\sigma^2 + \sigma_\varepsilon^2\right) - 2\alpha\delta\sigma^2 r. \end{aligned} \tag{5}$$

[8] We may also interpret this lead as some kind of reputational advantage as part of the location factor of a state i.

[9] As α and δ are set by the upper-tier government they can be interpreted as assignments of policy responsibilities to the Federal Government. While the case $\alpha = 0$ ($\delta = 0$) would clearly charecterize a confederation as the extreme version of a federation, $\alpha \to \infty$ ($\delta \to \infty$) would not necessarily imply a unitary state.

The optimal level of inputs to be spent by a state government i for education purposes is given as a result of maximizing $E(U_i)$ with respect to e_i. From the first order condition $e_i = a$ (that *ex ante* has to be met as well by the effort of the benchmark group of states), we obtain

$$E\left(R_i - \frac{1}{2}e_i^2\right) = \overline{R} + (\alpha - \delta)\mu + \frac{1}{2}\alpha^2 - \delta e. \qquad (5')$$

Let us now turn to the second player of our set-up.

The Federal Government

We suppose the upper-tier government acts like a social planner. It seeks to parallely maintain a high national students' attainment standard on the one hand and to ensure the highest possible nation-wide level of equity, in the sense of equal living conditions, on the other.[10] In the context of education policy, the latter corresponds to minimizing the reputational gap between a benchmark group of provinces and the other states of the Federation due to perceived quality differences as measured by average students' test achievements. Assuming additivity in these two (potentially conflicting) targets, the upper-tier government maximizes the welfare function W:

$$\max_{\overline{R},\alpha,\delta} W = E(a_i - R_i) = \mu + \alpha - \overline{R} - (\alpha - \delta)\mu - \alpha^2 + \delta e \qquad (6)$$

subject to $E(U_i) = \text{const}$.

The corresponding Lagrangian is given by

$$\begin{aligned}\Lambda \;=\; & \mu + \alpha - \left[\overline{R} + (\alpha - \delta)\mu + \alpha^2 - \delta\alpha\right] + \\ & \lambda\left\{\overline{R} + (\alpha - \delta)\mu + \frac{1}{2}\alpha^2 - \delta\alpha - \frac{1}{2}\gamma\left[\left(\alpha^2 - \delta^2\right)\left(\sigma^2 + \sigma_\varepsilon^2\right) - 2\alpha\delta\sigma^2 r\right]\right\}. \qquad (7)\end{aligned}$$

Using $\frac{\partial\Lambda}{\partial\overline{R}} = -1 + \lambda = 0$, we can re-write this as

$$\Lambda = \mu + \alpha - \frac{1}{2}\alpha^2 - \frac{1}{2}\gamma\left[\left(\alpha^2 - \delta^2\right)\left(\sigma^2 + \sigma_\varepsilon^2\right) - 2\alpha\delta\sigma^2 r\right]. \qquad (7')$$

[10] In the German Constitution the latter maxim is manifested in Art. 72 II GG.

From (7') we obtain as optimal values of α and δ, in the sense of a Nash Equilibrium of the federal government's and the i lower-tier governments' objectives,

$$\alpha^* = \frac{1}{1 + \gamma\left(\sigma^2 + \sigma_\varepsilon^2\right) - r^2 \frac{\sigma^4}{\sigma^2 + \sigma_\varepsilon^2}} \tag{8}$$

and

$$\delta^* = \alpha r \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}. \tag{9}$$

These results suggest the following interpretation: While it is in any case optimal to let the Federal Government decree educational standards based on absolute measures, i.e. $\alpha^* > 0$, pursuing education policies based on relative measures, i.e. $\delta^* > 0$, is only optimal if and only if $r > 0$. The intuition behind this finding is the following. Only if there exist discernable different shades of efficiency in the public production of education across provinces, our assumption $\text{cov}(\theta, \theta_i) = r\sigma^2$ holds such that $\text{cov}(\theta, \theta_i) = r\sigma^2 \Leftrightarrow r > 0$.

In sum, there are two central testable premises and corresponding implications of our model:

1. Absolute and relative educational attainment[11] has an impact on the popularity functions of incumbent province governments. It disciplines incumbents.

2. Idiosyncratic production efficiencies at the Federal State level exist. They show a positive covariance with the efficiency vector of a reference group of provinces leading in terms of student test scores. This circumstance suggests a subordinate governance based on relative measures ($\delta > 0$). A democratically voted on nationwide standard will not (unequivocally) generate a subaverage efficiency norm.

[11] As made visible to voters through the publication of average test scores in international student assessment exercises like the OECD PISA studies and their national extensions.

3. Evidence

3.1. Event study: Voting intention and PISA test score announcements

In the empirical investigations of this section we rely on opinion polls data to gauge voting intentions or the popularity of ruling parties in Germany in order, to ultimately capture the re-election probability of incumbent governments. This is a frequent practice of the empirical strand of the political economy literature (see Kirchgässner, 1985, for German federal elections, Carlsen, 1997, for the US, and Wolfers and Leigh, 2002, for federal elections in Australia). Wolfers and Leigh (2002) comparing opinion polls outcomes with projections from economic models and betting market data find that opinion polls do a good job in accurately assessing both the popularity and re-election probabilities of incumbents. This holds in particular over short-run, i.e. close to election, time horizons which are the relevant ones with regard to our event study analysis.

Announcement effects: Federal Government

As our period of investigation at the level of the Federal Government we choose the period from the month following the penultimate election to the Lower House of German Parliament (*Bundestag*) in fall 1998 to the month of the latest early *Bundestag* election in September 2005. The dotted vertical lines in Figure 2 mark the months corresponding to the two elections covered by our sample, while the vertical solid lines give the months in which the PISA test scores have been published.[12] With regard to our measure of the re-election probabilities, we rely on opinion poll (*Sonntagsfrage*) data collected by one of the major independent psephological institutes in Germany. The same source of opinion polls data (now: Infratest-dimap − then: Infratest) has been used by Kirchgässner (1985) to gauge voting intentions of the German voters and the popularity functions of parties for the Federal Republic of Germany from 1971 to 1982. The underlying sample is representative and the opinion poll has been

[12]Notably, these dates do not correspond to the announcements of the scores in the national extensions of the respective international studies.

conducted at least once a month over our observation period. If there were more than one opinion poll a month, we simply averaged the figures.

In order to assess and quantify the announcement effects of the two PISA study results on voting intention, we follow the usual event study practice of a two stage least squares (2SLS) strategy. In the first stage of our 2SLS regression we estimate a stylized Fair-type model (cf. Fair, 1996) of the incumbents' percentage of votes[13]

$$V_t = \overline{V} + \beta_1 UER_t + \beta_2 IR_t + \beta_3 \Delta \ln PROD_t + s_t, \tag{10}$$

where UER denotes the unemployment rate, IR the inflation rate computed from growth rates of the consumer price index (CPI), and $\Delta \ln PROD_t$ the growth rate in total production, respectively. Corresponding time series in monthly frequency were obtained from the Federal Employment Agency (*Bundesagentur für Arbeit*)[14] and the Federal Statistics Office (*Statistisches Bundesamt*). For the considered period, V_t corresponds to the share of voters intending to vote the social democrats (SPD) and/or the Green Party (*Bündnis 90/Die Grünen*) that formed the incumbent Red-Green coalition.

For an event study analysis the choice of the time slot for investigation is crucial. In our case, it is straightforward to concentrate on the time slot from the month of the respective announcement to the month of the next *Bundestag* election in order to assess the studies' popularity impact. In Figure 2 this corresponds to the time span between the respective solid and dotted vertical lines. Therefore, the second step of our 2SLS model consists of testing the obtained residual vector $\widehat{s}_t$ for a structural break in the form of a permanent shock beginning with the announcement of the PISA test score to the month of the next federal election. Practically, we split $\widehat{s}_t$

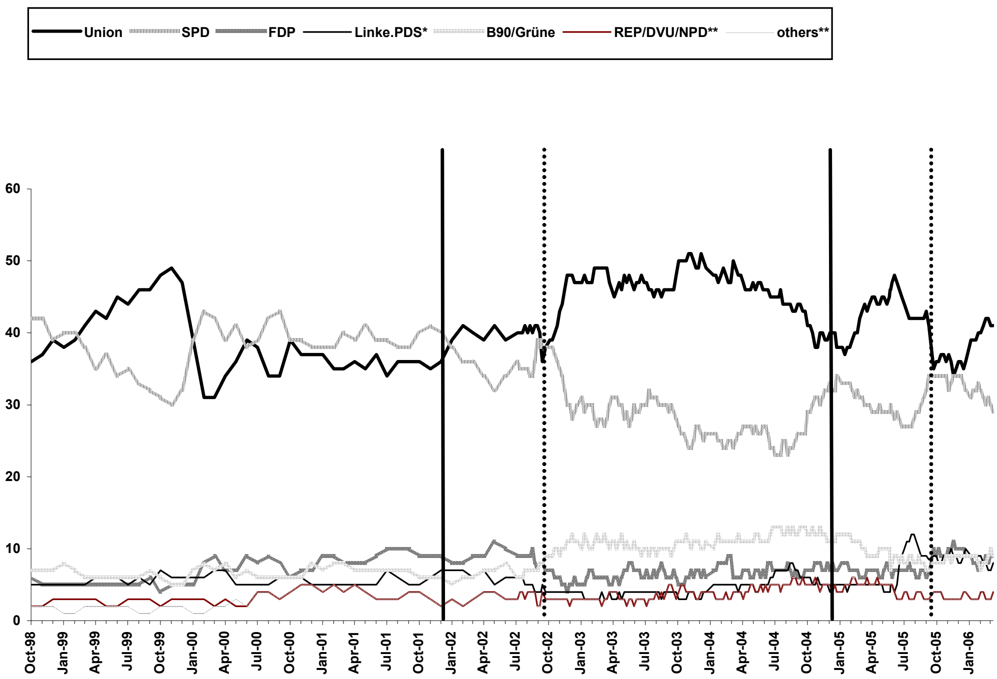
[13]Of course, this specification disregards several central determinants of popularity discussed in voting theory, including campaigns, issues, candidates, quality of challengers, and party identification (PID). However, given scarcity of aggregate data and the notoriously good performance of Fair-type models in predicting election outcomes, we are confident to rely on a "best second best" practice.

[14]As of January 2004 the UER figures exclude people participating in aptitude tests and training measures. End-of-month values were ascribed to the respective proceeding month.

into two subsamples and regress it on a respective dummy variable: ANN_EL02 and ANN_EL05 identify the period from the announcement of PISA 2000 (December 2001) and PISA 2003 (December 2004) to the following federal elections of September 2002 and September 2005, respectively.

The results are summarized in the column labelled "Sample (1)" in Table 2.

Figure 2. Bundestag election projections and PISA test result announcements
Source: Infratest-dimap (`http://www.infratest-dimap.de/`)



The different signs of coefficient estimates for ANN_EL02 and ANN_EL05 suggest the interpretation of a negative impact of the international PISA test results announcements for the subperiod when the incumbent coalition was leading in terms of popularity and a positive effect when the incumbents were lagging or catching up.

In the latter case, the opposition, i.e. the sum of voting shares of all non-incumbent parties, actually suffered in terms of popularity from the announcement. As can be seen from Figure 2 the incumbent Red-Green coalition led in terms of popularity up to the month of the PISA 2000 results announcement. The months from the 2002 federal election to the early 2005 election the SPD throughout fell below the popularity of the conservative CDU/CSU (Union) parties and was catching up the months before the PISA 2003 test score announcement in December 2004.

Table 2. 2SLS Estimates of PISA test results announcements

	Sample (1): all observations	Sample (2): w/o outliers
First Stage		
UER	-1.70^{***} (-2.69)	-1.63^{**} (-2.38)
IR	-1.96 (-0.83)	-0.82 (-0.64)
$\Delta \ln PROD$	0.04 (1.00)	0.04 (0.74)
adj. R^2	0.11	0.08
N obs.	84	78
Second Stage		
ANN_EL02	-2.54^{*} (-1.97)	-2.96^{**} (-2.14)
adj. R^2	0.08	0.07
N obs.	48	44
ANN_EL05	3.37^{*} (1.98)	3.47^{**} (2.07)
adj. R^2	0.23	0.25
N obs.	36	34

Note: Consistent Student t-values in parantheses (Newey and West, 1987)

*, **, *** denotes significance at 10, 5, 1% level of significance.

One may be concerned that some serious omitted variable bias plagues these estimates as central domestic and foreign-policy events and shocks were not controlled for. In particular, there were three major influential events associated with the months September/October 2001, August/September 2002, and March/April 2003 that require a special treatment. The political momentousness of the 09/11 attacks and the shoulder-to-shoulder stance of Chancellor Schröder with President Bush should have contributed exceptionally to the popularity of the incumbent government. In the summer before the 2002 federal election Germany wittnessed a hundred year flood, where due to the flooding of the Elbe River mainly East-German regions were concerned. During August and September 2002, 30,000 people got evacuated and more than 20 died. The incumbent government promised transfers and reconstruction funds amounting to 10 bn. Euro. It officially delayed the planned tax reform for 2003 due to this exceptional event. The gain in popularity for the incumbents through this taking measures is common knowledge today. Finally, the second Persian Gulf War started in March 2003. After a massive air strike coalition ground forces invaded Iraq. By mid-April, Saddam Hussein's army and government had collapsed. The German incumbents' corporate position against the invasion of Iraq also most probably affected its popularity. In event study analyses, a straightforward practice to treat exceptional and influential events simply consists of dropping these observations off the sample; see, for example, Bernanke and Kuttner (2005). Leaving out our monthly data associated with 09/11, the Elbe Flood, and the Persian Gulf War decreases our sample from 84 to 78 events. However, as can be seen from column "Sample (2)" in Table 2 our results remain qualitatively as well as quantitatively the same.

Announcement effects: Federal States

Because opinion polls data for German *Länder* governments are available in a discontinuous monthly frequency only, the resultant voting intention series at the province level seldomly total more than 70 observations for our period of investigation.[15] Therefore, they are throughout constituted of in-equidistant events. Hence, following the 2SLS strategy as in the federal government's case seems inadequate. An event study practice that has been successfully applied in widely different contexts is a one stage specification with a lower order autoregressive term and a dummy variable introducing events like announcements, accidents as the Chernobyl nuclear accident, etc. For our *Länder* PISA-E study it appears to be the appropriate one. It is evidenced, for example, in Kalra, Henderson, and Raines (1997), Berman, Brooks, and Davidson (2000), and Veraros, Kasimati, and Dawson (2004).

In our baseline regression, which does not take care of the actual discrepancies in average student test results among provinces, we estimate the following specification

$$V_{i,\tau} = \overline{V}_i + \beta_{1,i} V_{i,\tau-1} + \beta_{2,i} \varphi_{i,\tau} + \epsilon_\tau \qquad (11)$$

for all $i = 1, ...16$ *Länder* over the differently discontinuous points of observation τ. Again, $V_{i,\tau}$ represents the voting intentions in terms of vote percentages for the incumbent coalition or single ruling party.[16] Finally, $\varphi_{i,\tau}$ denotes a binary variable that equals zero until the respective opinion poll observation which corresponds to the nearest neighbor month of, or in other words the closest following event to, the PISA-E 2000 announcement (July 2002) and equals one thereafter. Theoretically we expect a negative sign for $\beta_{2,i}$ if the average student performance of *Bundesland* i fell below the average or median *Länder* score and vice versa.

In event study analyses interaction terms as, for instance, the direction (cut vs. increase) or reversals of interest rate target changes by the central bank (cf. Bernanke and Kuttner, 2005) usually are of particular interest. In our case this concerns the

[15] We draw our series from the wahlrecht.de database which comprises opinion polls data for the 16 *Länder* from different sources, i.e. from various psephological institutes. For detail see `http://wahlrecht.de/umfragen/landtage`.

[16] Note, we do account for in-sample changes in the *Länder* governments by simply adjusting the figures to the respective newly elected party or newly formed coalition.

discrepancies in average student test scores across the *Länder*. From a theoretical point of view it is reflected in equation (3) of our model. Our modified specification, therefore, reads

$$\begin{aligned} V_{i,\tau} &= \overline{V}_i + b_{1,i} V_{i,\tau-1} + \\ & b_{2,i} \left[\left(f_{i,\tau+1} - \widetilde{f}_{i,\tau} \right) \left(a_i^{00} - \overline{a}^{00} \right) + \widetilde{f}_{i,\tau} \left(a_i^{02} - \overline{a}^{02} \right) \right] + \epsilon_\tau, \end{aligned} \tag{12}$$

where $f_{i,\tau}$ is constructed in analogy to the binary variable $\varphi_{i,\tau}$ and $\widetilde{f}_{i,\tau}$ denotes a binary variable that equals zero until the respective opinion poll observation which represents the closest one following the PISA-E 2003 announcement (July 2005). Variables a_i^{00} and a_i^{02} denote the average students' test scores in PISA-E 2000 and PISA-E 2003, respectively, while $\overline{a}^{00}$ and $\overline{a}^{02}$ correspond to the respective mean result for all *Länder*. With regard to the sign of coefficients $b_{2,i}$ theory consistent estimates would now in any case require a positive value.

Table 3. Significant PISA-E test results announcement effects

	Berlin (BE)	Saarland (SL)	Baden-W. (BW)	Bremen (HB)
β_1	0.80 (19.2)	0.70** (2.65)	0.45*** (3.22)	1.01*** (11.9)
β_2	−0.02** (−2.32)	0.01 (0.80)	0.02 (2.40)	−0.02** (−2.32)
adj. R^2	0.80	0.50	0.57	0.82
b_1	0.88*** (20.3)	0.73*** (6.13)	0.49*** (3.22)	1.01*** (11.9)
b_2	0.00 (−0.03)	0.01*** (2.92)	0.01** (2.36)	0.01** (3.55)
adj. R^2	0.80	0.50	0.56	0.82
N obs.	118	22	32	10

Note: HAC Student t-values in parantheses (Newey and West, 1987)

*, **, *** denotes significance at 10, 5, 1% level of significance.

Table 3 reports four out of the 16 *Länder* to show a theory consistent reaction to the absolute and relative PISA-E results.[17] For all other provinces our estimates produced insignificant results.[18] Given the scarcity and shortfall of data, this 25 percentage share of theory consistent estimates seems remarkable and should be interpreted as a lower bound.

3.2. Production efficiency: Hypotheses and methodology

Stochastic Production Frontier Analysis

In the following, we briefly describe our methods used to quantify educational production efficiencies. The central idea of stochastic frontier analysis (SFA) is to distinguish between shifts in the technological frontier and production efficiency, i.e. a movement towards or away from the technological frontier. In contrast to an OLS regression, the frontier "envelops" the data points. The output, say educational performance measures, of two sample entities, j and k, can be expressed as a function of input H. SFA overcomes the crucial drawback of estimating the isolated Solow residual that fails to indicate whether a position below the frontier is solely due to noise or to an inefficient use of inputs or a combination of both. SFA allows this important distinction.

[17] In the case of the city state Berlin, this only holds for absolute scores. However, it should be noted that for PISA-E 2000 Berlin did not show up in the rankings which was based on student scores for all school types. Hence, for voters in Berlin an assessment of the relative educational attainment was merely possible as the representative achievements for the *Gymnasien* were not published in the mass media.

[18] There was one exception given by the estimates for Rhineland-Palatina (RP) that were inconsistent with our theoretical considerations. We attribute this to two basic facts: First, for our period of observation RP has been ruled throughout by an SPD/FDP (liberals) coalition. Both of these parties were the central losers in the 2002 *Bundestag* Election. In particular, the FDP severely lost ground due to a miserable campaign (the "18+X Prozent" Campaign) and a national scandal known as the "Jürgen Möllemann Anti-Semitic Flyer Scandal" in the final stage of the 2002 election campaign. Second, somehow unexpectedly students in RP obtained only slightly above average test scores in the two PISA-E studies.

A general model in the class of stochastic production frontier specifications is given by Battese and Coelli (1995). For our analysis, the specification of a cross section production frontier seems appropriate. Idiosyncratic effects are assumed to be distributed as exponential, half normal or truncated normal random variables (see Greene, 2003, ch. 17.6.3). These effects may be modelled by some regressors. Such models are called technical efficiency (henceforth TE) effects models

$$\begin{aligned} y_i &= \overrightarrow{\beta}\mathbf{x}_i + (v_i - u_i) \qquad \text{and} \\ m_i &= \overrightarrow{\delta}\mathbf{z}_i, \text{ where} \end{aligned} \tag{13}$$

- y_i is performance of students in the i-th state;
- $\mathbf{x}_i$ denotes the $k \times 1$ input vectors;
- $\overrightarrow{\beta}$ is a vector of unknown coefficients, over which the likelihood will be maximized;
- v_i denote random variables which are assumed to be i.i.d. $N\left(0, \sigma_v^2\right)$ and independent of the
- u_i non-negative random variables which are assumed (i) to account for technical inefficiency in production and (ii) to be independently distributed as truncations at zero of the $N\left(m_i, \sigma_u^2\right)$ distribution; where m_i is defined above and
- $\mathbf{z}_i$ is a $p \times 1$ vector of variables which may influence the educational efficiency in a certain state; and
- $\overrightarrow{\delta}$ is a $1 \times p$ vector of unknown coefficients, over which the likelihood will be maximized, and for which $\delta_0 = \mu$.

Depending on the choice of specifying $\overrightarrow{\delta}$ as zero-vector or not, we suppose an error decomposition (ED) frontier or a TE-effects model, respectively. Imposing for the former $\mu = m_i = m = 0$, we obtain a half-normal specification of the distribution of inefficiencies u_i. With no assumption on μ, the truncated normal specification of

Stevenson (1980) results. For $\mu = 0$, it is implicitly assumed that the inefficiency effects u_i are distributed half-normal (HN), whereas the specification of Stevenson (1980) is more general, inasmuch it only supposes a truncated normal (TN) distribution without restriction on its first moment. Educational efficiency relates to the estimated inefficiency as follows[19]

$$\xi_i = \exp\left(-u_i\right). \tag{14}$$

Independently of assumptions regarding the distribution of efficiencies, estimating an (ED- or TE-) efficiency model requires the statistical rejection of OLS estimates of the production process. Crucial for the rejection of OLS, and at the same time the justification of an efficiency analysis of the observed data, is the symmetry property, i.e. the third moment of the OLS residuals. This can be illustrated as follows: Given a vector of technical production inefficiencies, from the above equation, $u_i = 0$ for all $i = 1, ..., N$ observations. This would certainly be the case *iff* $\varepsilon_i = v_i$, that is the residuals ε_i of a corresponding OLS estimate showing the presupposed symmetry. Analogously, an $N \times 1$ vector $\mathbf{u} > 0$ along with a resulting negative skewness of the $\varepsilon_i = v_i - u_i$ OLS residuals would render an efficiency model the appropriate candidate. The applicable test statistic proposed by Coelli (1995) follows an asymptotically standard normal distribution[20]:

$$z = \frac{m_3}{\left(\frac{6m_2^3}{N}\right)^{\frac{1}{2}}} \overset{a}{\sim} N\left(0, 1\right), \tag{15}$$

where m_2 and m_3 denote the second and third moments of the residual distribution of the OLS estimate. Because of $m_3 < 0$ and thus $z < 0$, one-sided p-values are considered. For the single input AK-specification of the production function that will be considered in the following, z amounts to -164.6. The null of symmetric OLS

[19] Cf. Battese and Coelli (1995).

[20] Zero skewness of the residuals is under the null.

residuals is thus clearly rejected and a necessary pre-condition for OLS is violated. Results of the appropriate alternative model specifications, i.e. the ED- and TE-effects analysis, are outlined in the following section 3.3. However, before presenting our detailed specification and findings, let us briefly survey the central determinats of educational production efficiency arising from the literature.

Path-dependent and institutional variables

Peer group effects. One of the main determinants of student achievement to be found in the literature is the effect of peer group effects (Hanushek, 2002, p. 2078-2080). In our context, the realization of positive peer group effects would ceteris paribus imply a smaller variance of students' test scores in the respective province.

Tracking. Overall, there seems to be a negative relationship between the relative share of ninth graders attending the *Gymnasium* and the average test score of this group of students (Baumert *et al.*, 2002, pp. 92, 124, 141). A rigidly selected "elite" of students (those attending *Gymnasium*) may be the reason for this effect. Plausibly, a smaller elite can be assumed to reach better learning outcomes than the masses. A recent and comprehensive study on tracking based on comparisons of international student achievements and educational systems is given by Hanushek und Wössmann (2005). At the national level, a negative nexus of the proportion of students attending high school and PISA test scores might indicate future academic elites being educated and promoted better in smaller groups. This would, of course, suggest more rigorous tracking practices. Successful tracking, on the other hand, requires a strictly performance-based criterion for sorting students into different school tracks. Socially undesirable side effects can be social disparities and segregation (Hoxby, 2000).

Class size effects. It is often said that class size plays a decisive role in student achievement. Again, the hypothesis is that better learning outcome is easier to achieve in smaller groups. Recent international studies, however, do not to support the class size effect hypothesis. Positive effects seem to be rather the exception than the rule (Wössmann and West, 2006).

Central exit examination (CEE). The achievement standard of students in provinces

with CEE (*Abitur*) is considered higher than the one of students in states without CEE. This observation is usually ascribed to the exclusion of teaching to the test strategies. A recent survey is given in Jürges, Richter, and Schneider (2005).
Family background and path dependency. The historical time of abolishment of secondary school fees is a path-dependent determinant of student achievements, which is immanent to the respective schooling system (Riphahn, 2004). According to the estimates of Riphahn (2004) based on German *Länder* data, the abolishment of school fees has increased secondary school attendance by a rough 6 percentage points. In Rhineland-Palatinate this was, for example, the case in the early 1960s. The positive enrollment effect is found to have been particularly pronounced for female students. This finding suggests two lines of reasoning: First, families with a lower social status were able to send their children to secondary school after the abolishment of fees. The ninth-graders in the year 2000 may accordingly have parents or grandparents who were able to attain a high school degree after the abolishment. A reference generation of parents or grandparents from another province, however, may not have had this chance due to the fees and, therefore may not have started a tradition of higher education. Second, the awareness of the costs related to higher education wittnessed by the last birth cohort who paid school fees might be important as it might be passed on to today's students.

Besides the above listed determinants of educational attainment, we consider the number of school years till *Abitur*, which varies across *Länder*. In 2000, in some provinces students attended school for twelve years. However, in the majority of *Länder* it were thirteen years.

Political economy factors

Traditionally the political economy literature distingushes two stereotyped governments: conservative and liberal. The former is supposed to follow less redistributive strategies. We employ a dummy to control for potential effects from the political background of incumbents. In coalition governments comprising a smaller party, it is frequently the case that the secretary of education belongs to this smaller party. As it is usually their only way to make their mark, these parties not seldomly outperform incumbents of traditional parties in terms of innovation, successful reforms, etc. Conversely, due to these parties' lack of experience a negative impact on attainment might as well be resultant. Again, our strategy is to control for these potential effects by a dummy. Finally, we consider two variables depicting the general cost effectiveness of some federal state government's activity, is the degree of indebtedness and total public expenditure in p.c. terms.

Economic and demographic framework

Economic preconditions may indirectly influence student achievement. To take this into account, we include the unemployment rates, employment rates of women, and p.c. disposable income as additional regressors. Some strand of literature in applied public finance, suggests a positive relationship between population density and public expenditures. For German provinces, for example, Büttner, Schwager, and Stegarescu (2004, Table 2, S. 508) find a significant positive impact of the population density on the general support of education in the federal state. In this context, population density can be viewed as a proxy of a pro-education environment with regard to public expenditures. As a final demographic variable, we consider the foreigners share in the state's population. Assuming a high share of non-native speakers among students, this might be another channel through which educational attainment is affected.

3.3. Production efficiency: Results

As argued in the introduction, we are able to obtain data separately for the three (four) tested fields in 2000 (2003), and for the 16 Länder. We are, therefore, able to rely on pooled data of 112 observations. An ideal variable to instrument educational efforts is the number of lessons, decided upon and provided by the respective federal state.[21] Since cumulated material resources as additional inputs generated insignificant estimates throughout our baseline regressions, we refrained from considering in

[21] We use accumulated lessons attended by the PISA-E student cohorts from first to ninth grade that we adjusted for the respective curricula and holiday breaks.

the further analysis. Ultimately, we suppose some fairly basic AK production process for students' abilities

$$\ln\left(\widetilde{a}_{i,j}\right) = \beta_0 + \beta_1 \ln\left(H_{i,j}\right) + \sum_{c=1}^{C} \gamma_c Q_{c,i,j} + \phi_{i,j}, \tag{16}$$

where $H_{i,j}$ denotes lessons as inputs $(e_{i,j})$ of federal government i with regard to field j. Q represents a matrix of $j = 1, ..., C$ control variables,[22] and $\phi_{i,j}$ summarizes the residual components.

We sequentially introduce our control variables using a forward selection precedure. In each step, all variables from the categories *Institutional and path-dependent conditions*, *political economy determinants, economic conditions* and *demographic variables* are separately included in the model. Using partial F-tests, the control producing the best model fit is selected. These variables are included in the following regressions. Dummies to control for different technologies by subject, i.e. for reading and mathematical literacy, throughout showed no significant impact.

Table A.1 in the Annex reports our baseline regression results. The accumulated number of lessons taught in the respective subjects is highly significant for student test scores. This result reflects the classic human capital theory, stating that education increases productivity.[23] The constant in the regression equation can be interpreted as level of production technology in the education sector, which is independent of the explicitly modelled inputs. The negative dependence of economical public budgeting of federal states, measured by public indebtedness, and the quality of educational institutions' equipment and infrastructure seems to be an indisputable fact.[24] The share of conservatives in the government coalition, as well as government participation of a smaller party, have significantly positive effects on test scores. Additionally, we

[22]These are determinants that do not directly alter the eductational production process. However, in the simplest case, they might have an impact on the minimum level of students' performance.

[23]If schooling represented only an (effective) screening mechanism for identifying high-potentials, an effect of the number of lessons on achievements would not be reasonable.

[24]Descriptive statistics of variables is given in Table A.5 in the Appendix. Abbreviations are to be found in Table A.6.

identify an influence of private cost awareness on student achievement: The later school fees for higher secondary schooling were abolished, the better the average test scores of today's students in the respective *Bundesland.* This might be interpreted as an increase in learning productivity of the offspring through the cost awareness of their parents. We control for differnt educational length (12 or 13 years) to graduation, and for centralized exit examinations, using dummies. They are estimated highly significant, stating a negative effect of school years on student performance. The existance of central exit examinations, on the contrary, has a positive effect. We find a positive effects of population density and per capita disposable income, and a negativ one of the foreign population share on PISA-E student achievements. All other considered determinants were estimated to have no statistically significant effects. For analyzing educational efficiency of the federal states we firstly estimated a two-step Error Decomposition model. We used the alternative assumptions of a half-normally and exponentially distributed inefficiency term u_i. Both specifications have qualitatively similar results. Analogously to the procedure in the OLS case, the accumulated number of lessons is the only input factor. It is significant on the 1%, or at least on the 5% level, throughout all the specifications. Here, too, there are no subject specific differences within the federal states. The efficiency ranking of the federal states resulting from the first step of the model is given in Table 4 in the column Efficiency (ED Model). Additionally, column Efficiency (TE Model) of Table 2 shows the results from TE analysis. The according ranks given in brackets make the robustness of the efficiency ranking evident, when the choice of model is concerned.

Based on the ranks of the federal states' education production efficiency, we resume the voting scenario in *Bundesrat.* Considering the absolute average test scores of PISA test scores, a centrally targetted achievement level led to an average result. On the basis of efficiently used public ressources, now a completely different picture shows: the column cumulated votes makes is clear, that the majority of votes in *Bundesrat* could be reached by the first seven, most efficient federal states alone. The production efficiency of the median voter, which is the Saarland, is above the state-wide average. Setting "best practice" of educational production as the joint target can obviously lead to a common above average level. This was not the case when using an absolute value as a target.

The second step of the ED model helps to explain the differences in efficiency between the states. Applying the same specification design as above, we control the efficiency scores for determinants sorted blockwise. The results are shown in table A.2 in the appendix. In the TE model, the parameters denoted $\delta_0, ..., \delta_p$ in section 3.2, are estimated simultaneously with the inefficiencies.[25] Table A.3 shows the results and the parameter values for σ^2 and ρ: Following the parameterization of Battese and Coelli (1995), we use the replacement of σ_v^2 and σ_u^2 with $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\rho = \frac{\sigma_u^2}{\sigma_v^2+\sigma_u^2}$. As a result parameter ρ lies between 0 and 1. This range can be searched to provide a starting value for use in an iterative maximization process like the Davidon-Fletcher-Powell (DFP) algorithm.

Both models show clearly a negative dependence between public indebtedness (DEBT) and education efficiency. Considering high indebtedness as an indicator for "uneconomical" budgeting of the public sector, this also includes inefficient resource use in the *Gymnasien.* The influence of late abolition of secondary school fees (FEE) is highly significant, and positive. Stronger cost awareness obviously leads to more efficient input use. This effect overweighs the contrary effect of the assumed higher educational class of the parent generation. A longer duration of secondary education (DUR) has a negative effect on efficiency, whereas *Länder* with central exit examinations (CEE) are more efficient. Another positive determinant is a higher share of a smaller party in federal government (NTP). This can on the one hand be explained by the (rare) chance for the smaller party to make their mark. On the other hand, more heterogenous governments may control government spending more rigorously.

Higher share of foreigners (FOR) in the population and lower population density (DENS) are determinants for lower efficiency, see table A.4.More non-native speakers

[25]The TE model is a one-step model. Contrarily, in the second step of the ED model the effects on efficiency (not inefficiency) scores are estimated.

attending the lessons may make fast progress through the curriculum harder. It may be plausible that more teaching input, measured in lessons, is be necessary to achieve the same knowledge level. Population density instruments the quantity of cultural institutions.

In the TE specification, a more efficient environment for educational production is created by more restrictive segregation (SEG), which is measured by the share of students attending *Gymnasium.* This means that a more homogenous group of students can reach educational goals more efficiently. Two other variables have positive effects on educational environment: The share of a smaller party in the federal government (NTP), and that of the conservative party (CONS). Educational output is effected positively by the participation rate of women in the labor market (FEM) and by the migration balance (MIG). A positive migration balance has a positive effect on student performance.

4. Conclusion

This paper investigated the optimal design of federal governance for the public production of education. In a simple yardstick competition model of vote-seeking incumbents, we addressed the question of the ideal organization of education in a federal system. We relied on student achievement data drawn from the national extensions of the PISA 2000 and 2003 studies for German *Länder* to analyze two central aspects of the model: First, we estimated the announcement effect of the PISA and PISA-E test results on the popularity function of federal government and state governments in an event study approach based on opinion polls data. Second, we quantified production inefficiencies at the state level using stochastic frontier techniques. In both cases, our estimates support the model. As robust determinants of efficiency we identified the popular awareness of the costs of public education, the level of public indebtedness, and coalition governments consisting of more than two parties in the respective *Bundesland*.

Table 4a. Educational efficiency and votes in the Federal Council: German Länder 2000

Federal State	Efficiency (ED model)	Efficiency (TE model)	Votes	Votes cumulated
Bavaria (BY)	0.991 + [1]	0.980 + [3]	6	6
Schleswig-Holstein (SH)	0.991 + [2]	0.992 + [1]	4	10
Baden-Württemb. (BW)	0.989 + [3]	0.966 + [8]	6	16
Saxony (SN)	0.989 + [4]	0.974 + [5]	4	20
Lower Saxony (NI)	0.989 + [5]	0.975 + [4]	6	26
M.-W. Pomerania (MV)	0.988 + [6]	0.972 + [7]	3	29
Thuringia (TH)	0.987 + [7]	0.961 + [11]	4	33
Rhinel.-Palatinate (RP)	0.987 + [8]	0.964 + [9]	4	37
Saarland (SL)	0.986 + [9]	0.987 + [2]	3	40
N. Rhine-Westph. (NW)	0.986 + [10]	0.974 + [6]	6	46
Hesse (H)	0.985 - [11]	0.944 - [12]	5	51
Berlin (BE)	0.985 - [12]	0.961 + [10]	4	55
Hamburg (HB)	0.982 - [13]	0.904 - [16]	3	58
Saxony-Anhalt (ST)	0.979 - [14]	0.928 - [13]	4	62
Brandenburg (BB)	0.979 - [15]	0.914 - [15]	4	66
Bremen (HB)	0.976 - [16]	0.920 - [14]	3	69

Note: Average efficiencies are: 0.986 (ED) and 0.957 (TE). ED estimates are displayed for normal/half-normal assumption; results are robust for exponential assumption.

Table 4b. Educational efficiency and votes in the Federal Council: German Länder 2003

Federal State	Efficiency (ED model)	Efficiency (TE model)	Votes	Votes cumulated
Bavaria (BY)	0.993 + [1]	0.991 + [2]	6	6
Saxony (SN)	0.992 + [2]	0.985 + [3]	4	10
Baden-Württemb. (BW)	0.992 + [3]	0.977 + [6]	6	16
Schleswig-Holstein (SH)	0.991 + [4]	0.973 + [9]	4	20
Thuringia (TH)	0.990 + [5]	0.962 - [11]	4	26
Rhinel.-Palatinate (RP)	0.990 + [6]	0.981 + [4]	4	29
Lower Saxony (NI)	0.990 + [7]	0.995 + [1]	6	33
M.-W. Pomerania (MV)	0.989 + [8]	0.975 + [8]	3	37
Saxony-Anhalt (ST)	0.989 + [9]	0.972 + [10]	4	40
Saarland (SL)	0.989 - [10]	0.977 + [5]	3	46
N. Rhine-Westph. (NW)	0.988 - [11]	0.976 + [7]	6	51
Hesse (H)	0.988 - [12]	0.950 - [12]	5	55
Hamburg (HB)	0.988 - [13]	0.927 - [15]	3	58
Berlin (BE)	0.986 - [14]	0.948 - [13]	4	62
Bremen (HB)	0.985 - [15]	0.935 - [14]	3	66
Brandenburg (BB)	0.985 - [16]	0.926 - [16]	4	69

Note: Average efficiencies are: 0.989 (ED) and 0.966 (TE). ED estimates are displayed for normal/half-normal assumption; results are robust for exponential assumption.

References

Battese, G. E. and T. J. Coelli (1995), A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data, Empirical Economics 20, 325-332.

Baumert *et al.* (2002), PISA 2000: Die Länder der Bundesrepublik Deutschland im Vergleich, K.-J. Tillmann & M. Weiss (ed.), Leske + Budrich, Opladen.

Berman, B., R. Brooks, and S. Davidson (2000), The Sydney Olympic Games Announcement and Australian Stock Market Reaction, Applied Economics Letters 7 (12), 781-784.

Bernanke, B. S. and K. N. Kuttner (2003), What Explains the Stock Market's Reaction to Federal Reserve Policy?, Journal of Finance 60, 1221-1257.

Besley, T. and A. Case (1995), Incumbent Behavior: Vote-Seeking, Tex-Setting, and Yardstick Competition, American Economic Review 85, 25-45.

Bodenstein, M. and H. W. Ursprung (2005), Political Yardstick Competition, Economic Integration, and Constitutional Choice in Federation, Public Choice 124, 329-352.

Bonaccorsi, S. und C. Daraio (2004), Econometric Approaches to the Analysis of Productivity of R&D Systems, in H. F. Moed, W. Glänzel, and U. Schmoch (eds.), Handbook of Quantitative Science and Technology Research, Springer: Berlin, Heidelberg, New York, 51-74.

Brunner, E. and J. Sonstelie (2003), School Finance Reform and Voluntary Fiscal Federalism, Journal of Public Economics 87, 2157-2185.

Büttner, T., Schwager, R. and M. Stegarescu (2004), Agglomeration, Population Size, and the Cost of Providing Public Services: an Empiric Analysis for German States, ZEW Discussion Paper, No 04-18

Battese, G. E. and T. J. Coelli (1995), A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data, Empirical Economics 20, 325-332.

Carlsen, F. (1997), Opinion Polls and Political Business Cycles: Theory and Evidence for the United States," Public Choice 92, 387-406.

Fair, R. C. (1996), Econometrics and Presidential Elections, Journal of Economic Perspectives 10, 89-102.

Greene, W.H. (2003), Econometric Analysis, Pentice Hall, Upper Saddle River, New Jersey

Hanushek, E. A. (2002), Publicly Provided Education, NBER Working Paper No. 8799

Hanushek, E. A. and D. D. Kimko (2000), Schooling, Labor-Force Quality, and the Growth of Nations, American Economic Review 90, 1184-1208.

Hanushek, E. and L. Wössmann (2005), Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries, CESifo Working Paper, No.

Hoxby, C. M. (2000), Does Competition Among Public Schools Benefit Students and Taxpayers?, American Economic Review 90, 1209-1238.

Inman, R. and D. Rubinfeld (1997), Rethinking Federalism, Journal of Economic Perspectives 11, 43-64.

Jürges, H., Richter, W. F., and K. Schneider (2005), Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality, FinanzArchiv 6, 298-326

Jürges, H. and K. Schneider (2004), International Differences in Student Achievement: An Economic Perspective, German Economic Review 5, 357-380.

Kalra, R., Henderson, G.V. and Raines G.A. (1997), Effects of the Chernobyl nuslear accident on utility share prices, Quarterly Journal of Business and Economics 32, 52-77.

Kirchgässner, G. (1985), Causality testing of the popularity function: An empirical investigation for the Federal Republic of Germany, 1971-1982, Public Choice 45, 155-173.

Riphahn, R. T. (2004), Immigrant Participation in Social Assistance Programs, IZA Working Paper No. 15

Salmon, P. (1987), Decentralization as an incentive scheme, Oxford Review of Economic Policy 3, 24-43.

Stevenson, R. E. (1980), Likelihood Functions for Generalised Stochastic Frontier Estimation, Journal of Econometrics 13, 57-66.

Veraros, N., E. Kasimati, and P. Dawson (2004), The 2004 Olympic Games Announcement and its Effect on the Athens and Milan Stock Exchanges, Applied Economics Letters 11, 749-753.

Wolfers, J. and A. Leigh (2002), Three Tools for Forecasting Federal Elections: Lessons from 2001, Australian Journal of Political Science 37, 223-240.

Worthington, A. C. (2001), An Empirical Survey of Frontier Efficiency Measurement Techniques in Education, Education Economics 9, 245-268.

Wössmann, L. and M. West (2006), Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS, European Economic Review 50, 695-736.

Table A.1: Institutional, political and socio-economic determinants of educational output (OLS-Model)

	1	2	3	4	5	6	7	8
HOUR[1)]	19.050*** (3.57)	8.933* (1.93)	8.612* (1.93)	8.840** (2.12)	8.438** (2.15)	9.330** (2.28)	10.170** (2.38)	7.939* (1.82)
DEBT[2)]		-3.546*** (-11.71)	-3.787*** (-12.90)	-3.630*** (-13.28)	-3.610*** (-12.16)	-3.300*** (-11.43)	-2.987*** (-9.17)	-3.849*** (-10.19)
FEE[3)]			6.243*** (3.13)	10.390*** (4.36)	11.320*** (5.41)	12.920*** (6.68)	19.600*** (6.02)	23.300*** (7.09)
DUR[4)]				-2.092*** (-3.61)	-2.576*** (-4.56)	-2.337*** (-4.10)	-2.457*** (-4.20)	-1.950*** (-3.38)
NTP[2)]					5.244** (2.44)	5.166** (2.49)	5.811*** (2.73)	3.586 (1.64)
CEE[5)]						8.878*** (2.65)	9.479*** (2.79)	11.182*** (3.37)
FOR[5)]							-1.584** (-2.54)	-3.247*** (-4.70)
DENS[3)]								8.978*** (3.69)
Adj. R^2	0.040	0.498	0.546	0.598	0.613	0.630	0.648	0.710
Log L	259.520	296.384	302.475	309.782	312.426	315.567	318.877	326.026

Annotation: t-values are given in brackets.
Legend: * p < 0.1; ** p < 0.05; *** p < 0.01
1) Values multiplied by 10^7
2) Values multiplied by 10^4
3) Values multiplied by 10^6
4) Values multiplied by 10^2
5) Values multiplied by 10^3

Table A.2a: Institutional, political and socio-economic determinants of educational output (ED-Model, normal/half-normal)

	1	2	3	4	5	6	7	8	9	10
DEBT[1)]	-5.699** (-10.30)	-6.116** (-11.51)	-5.863** (-11.69)	-5.808** (-12.09)	-5.347** (-10.69)	-6.650** (-11.03)	-6.252** (-10.51)	-6.052** (-10.51)	-7.406** (-10.72)	-7.383** (-10.77)
FEE[2)]		1.095** (3.92)	1.774** (5.70)	1.982** (6.50)	3.048** (6.05)	3.540** (7.09)	3.951** (7.92)	4.190** (8.62)	242.070** (3.30)	248.700** (3.41)
DUR[2)]			-342.684** (-4.03)	-450.727** (-5.14)	-472.815** (-5.51)	-405.689** (-4.84)	-357.255** (-4.34)	-364.957** (-4.61)	-345.183** (-4.54)	-224.370* (-2.11)
NTP[1)]				11.685** (3.29)	12.760** (3.67)	9.493** (2.76)	9.055** (2.74)	12.343** (3.68)	14.322** (4.38)	15.806** (4.69)
FOR[1)]					-25.679* (-2.62)	-49.822** (-4.30)	-54.503** (-4.84)	-85.978** (-5.76)	-106.153** (-6.82)	-111.012** (-7.05)
DENS[2)]						1.305** (3.51)	1.463** (4.05)	1.890** (5.05)	2.604** (6.20)	2.830** (6.43)
CEE[1)]							155.144** (3.05)	183.395** (3.68)	168.431** (3.52)	132.893* (2.54)
DISP[1)]								5.509** (3.07)	8.615** (4.38)	9.230** (4.64)
FEE2[1)]									-23.773** (-3.25)	-24.474** (-3.36)
CONS[1)]										5.843 (1.61)
Adj. R^2	0.486	0.545	0.601	0.635	0.653	0.687	0.710	0.732	0.754	0.758
Log L	501.400	508.771	516.613	522.024	525.531	531.756	536.549	541.447	546.952	548.366

Annotation: t-values are given in brackets.
Legend: * p < 0.05; ** p < 0.01
1) Values multiplied by 10^5
2) Values multiplied by 10^6

Table A.2b: Institutional, political and socio-economic determinants of educational output (ED-Model, normal/exponential)

	1	2	3	4	5	6	7	8	9
DEBT[1)]	-1.620*** (-10.33)	-1.738*** (-11.55)	-1.667*** (-11.72)	-1.651*** (-12.15)	-1.528*** (-10.77)	-1.913*** (-11.24)	-1.804*** (-10.71)	-1.753*** (-10.66)	-2.137*** (-10.82)
FEE[2)]		3.108*** (3.92)	5.007*** (5.67)	5.616*** (6.52)	8.455** (5.91)	9.906*** (7.03)	11.040*** (7.81)	11.650*** (8.39)	686.000*** (3.28)
DUR[1)]			-95.889*** (-3.97)	-127.501*** (-5.14)	-133.382*** (-5.47)	-113.560*** (-4.80)	-100.271*** (-4.31)	-102.241*** (-4.52)	-96.636*** (-4.45)
NTP[1)]				3.419*** (3.41)	3.705*** (3.75)	2.740*** (2.83)	2.620*** (2.80)	3.462*** (3.61)	4.023*** (4.31)
FOR[1)]					-6.837** (-2.46)	-13.966*** (-4.27)	-15.251*** (-4.79)	-23.304*** (-5.47)	-29.024*** (-6.53)
DENS[2)]						3.853*** (3.67)	4.288*** (4.19)	5.379*** (5.04)	7.403*** (6.17)
CEE[1)]							42.570*** (2.95)	49.798*** (3.50)	45.556*** (3.33)
DISP[1)]								1.410*** (2.75)	2.290*** (4.08)
FEE2[1)]									-6.739*** (-3.22)
Adj. R^2	0.488	0.547	0.601	0.637	0.653	0.690	0.711	0.728	0.751
Log L	642.679	650.082	657.726	663.501	666.600	673.372	677.886	681.854	687.279

Annotation: t-values are given in brackets.
Legend: * p < 0.1; ** p < 0.05; *** p < 0.01
1) Values multiplied by 10^5
2) Values multiplied by 10^7

Table A.3: Institutional, political and socio-economic determinants of educational output (TE-Model)

	1	2	3	4	5	6
HOUR[1)]	1.824** (2.15)	1.958** (2.49)	1.914*** (2.81)	1.821*** (2.70)	1.313** (2.05)	1.343** (2.13)
FEM[2)]					1.159*** (3.69)	2.100*** (4.66)
MIG[2)]						1.348*** (2.75)
SEG[2)]		2.964*** (2.62)	2.024*** (2.87)	1.874*** (2.68)	2.362*** (3.56)	2.337*** (3.48)
CONS[2)]			-1.110*** (-5.36)	-1.370*** (-6.05)	-1.322*** (-6.24)	-1.452*** (-6.79)
NTP[2)]				-0.893*** (-2.82)	-0.810** (-2.01)	-0.950*** (-3.20)
∂_0 [2)]	8.916 (0.13)	-53.543 (-1.10)	5.477 (0.24)	23.186 (0.98)	7.776 (0.35)	49.310 (2.14)
γ	-0.755 (-0.26)	-0.353 (-0.19)	-12.333 (-0.51)	-11.161 (-0.62)	-11.699 (-0.47)	26.200 (0.06)
σ^2	-7.304*** (-12.79)	-7.524*** (-39.46)	-7.790*** (-78.91)	-7.809*** (-99.12)	-7.959*** (-99.09)	-7.996 (-755.47)
Log L	259.548	265.157	277.559	281.610	288.369	292.357

Annotation: t-values are given in brackets.
Legend: * p < 0.1; ** p < 0.05; *** p < 0.01
1) Values multiplied by 10^6
2) Values multiplied by 10^3

Tabelle A.4: Final Model-Specifications

	OLS	ED (normal / exponential)	TE
HOUR[1)]	0.794* (1.82)		0.134** (2.13)
DEBT[2)]	-0.385*** (-10.19)	-2.137*** (-10.82)	
FEE[1)]	23.300*** (7.09)	68.600*** (3.28)	
DUR[3)]	-0.195*** (-3.38)	-0.966*** (-4.45)	
NTP[4)]	3.586 (1.64)	0.402*** (4.31)	-0.095*** (-3.20)
FOR[4)]	-0.325*** (-4.79)	-2.902*** (-6.53)	
DENS[1)]	8.978*** (3.69)	0.740*** (6.17)	
CEE[4)]	1.118*** (3.37)	4.556*** (3.33)	
DISP[5)]		2.290*** (4.08)	
FEE2[5)]		-6.739*** (-3.22)	
FEM[3)]			2.100*** (4.66)
MIG[3)]			1.384*** (2.75)
SEG[3)]			2.337*** (3.48)
CONS[3)]			-1.452*** (-6.79)
∂_0			49.310** (2.14)
σ^2			-7.996***
γ			26.200
Log L	326.026	687.279	292.357

Annotation: t-values are given in brackets.
Legend: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
[1)] Values multiplied by 10^6
[2)] Values multiplied by 10^5
[3)] Values multiplied by 10^3
[4)] Values multiplied by 10^4

Table A.5: Statistics

	Variable	Mean	Std.dev.	Min	Max
Endogenous	*ln(RES)*	6.361	0.025	6.304	6.418
Production input	*HOUR*	2646.913	2843.177	1014	10153.2
Path-dependent and Institutional Determinants	*SEG*	28.818	2.754	24.8	34.5
	MIG	1.908	5.600	-6.4	14.51
	CEE	0.438	0.498	0	1
Political Economy Variables	*CONS*	39.884	9.631	22	56.9
	NTP	5.584	6.756	0	24.4
	DEBT	95.4	47.606	29	252
Economic Determinants	*FEM*	69.694	5.968	57.7	81.4
	GDP	244.616	73.423	158	445
	EMP	12.721	4.950	6	21.4
	DISP	161.607	22.152	128	222
Population Variables	*DENS*	666.063	1007.335	75	3800
	FOR	7.757	4.452	1.7	15.4

Table A.6: Abbreviations of variables

	Variable
HOUR	Number of lessons from grade 1 in school year 1991/1992 till grade nine in school year 1999/2000
SEG	Share of 15-year-olds attending Gymnasium
SIZE 1	Teacher-student-relation in secondary school
SIZE 2	Teacher-student-relation in primary school
CEE	Central exit examination
G9	Abitur after 13 years (as opposed to after 12 years)
FEE	Year of abolishment of secondary schooling fees
MIG	Migration balance of the PISA-cohort 1984 between 1990 and 1999. in percent of the age-group-size in1990
CONS	Election result of CDU/CSU. last federal elections before 2000 or 2003, respectively
NTP	Election result of smaller party in government coalition, last federal elections before 2000 or 2003, respectively
DEBT	Public indebtedness per inhabitant, in 100 DM (2000, 2003)
FEM	Empolyment rate of women, in percent (2000, 2003)
GDP	Gross domestic product per inhabitant, in 100 Euro (2000, 2003)
DISP	Disposable income (2000, 2003)
EMP	Unempolyment rate, in percent (2000, 2003)
DENS	Population per square kilometer (2000, 2003)
FOR	Foreign population, in percent (2000, 2003)