

IZA DP No. 2477

## Decentralization and the Productive Efficiency of Government: Evidence from Swiss Cantons

Iwan Barankay  
Ben Lockwood

December 2006

# Decentralization and the Productive Efficiency of Government: Evidence from Swiss Cantons

**Iwan Barankay**

*University of Warwick, CEPR  
and IZA Bonn*

**Ben Lockwood**

*University of Warwick and CEPR*

Discussion Paper No. 2477  
December 2006

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Decentralization and the Productive Efficiency of Government: Evidence from Swiss Cantons<sup>\*</sup>**

Advocates of fiscal decentralization argue that amongst other benefits, it can increase the efficiency of delivery of government services. This paper is one of the first to evaluate this claim empirically by looking at the association between expenditure decentralization and the productive efficiency of government using a data-set of Swiss cantons. We first provide careful evidence that expenditure decentralization is a powerful proxy for legal local autonomy. Further panel regressions of Swiss cantons provide robust evidence that more decentralization is associated with higher educational attainment. We also show that these gains lead to no adverse effects across education types but that male students benefited more from educational decentralization closing, for the Swiss case, the gender education gap.

JEL Classification: H40, H52, H70, I20

Keywords: decentralization, productive efficiency, local public goods

Corresponding author:

Iwan Barankay  
Department of Economics  
University of Warwick  
Coventry CV4 7AL  
United Kingdom  
E-mail: [iwan.barankay@warwick.ac.uk](mailto:iwan.barankay@warwick.ac.uk)

---

<sup>\*</sup> We thank Imran Rasul, Michael Smart, Kolemian Strumpf, Daniel Sturm, Alois Stutzer, the editor Sören Blomquist and two anonymous referees for comments. Thanks also go to seminar participants at the Workshop on Federalism and Decentralization at ZEI Bonn 2003, the Public Economics Weekend Leicester 2003, the Canadian Public Economics Group Annual Conference Toronto 2004, CEPR/SNS Public Policy Symposium Stockholm 2004, CESifo Munich, CORE, Essex, and Warwick for useful comments. The first author thanks the ESRC for financial support. We are indebted to the Swiss Federal Statistical Office for invaluable data support. All errors remain our own.

## 1. Introduction

Fiscal decentralization, the allocation of tax and spending powers to lower levels of government, is now an established policy objective, in many developed and developing countries. Moreover, it is actively promoted as a development strategy by organizations such as the World Bank (Azfar et al., 2001, World Bank, 2000). The usual advantages that are claimed for decentralization that one can find in the literature include the following (Azfar et al., 2001, Lockwood, 2005, Oates, 1999). First, decentralization is claimed to improve *allocative efficiency*, in the sense that the goods provided by governments in localities will be better matched to the preferences of the residents of those localities. This is sometimes known as the *preference-matching* argument. Second, decentralization is argued to increase the *efficiency* of delivery of government services. In this literature, production efficiency is interpreted in a wide sense, to accommodate inefficiencies like corruption, waste, and poor governance. There is now quite a large literature on decentralization and allocative efficiency<sup>1</sup>. By contrast, the literature on decentralization and productive inefficiency is small.

The theoretical literature identifies three mechanisms by which decentralization may lead to increased efficiency<sup>2</sup>. The first is that decentralization may give voters increased electoral control over incumbents. For example, in Seabright (1996) and Persson and Tabellini (2000, Chapter 9), decentralization is shown, under some conditions, to reduce the incentives for incumbents to divert rents from tax revenue, because under decentralization, the (negative) link between such rent diversion and the probability of re-election is stronger. Hindriks and Lockwood (2005) extend this argument to show how decentralization may increase the equilibrium probability that corrupt incumbents are voted out of office (a stronger *selection effect* in the terminology of Besley and Smart, 2004).

The second mechanism is that decentralization enables yardstick competition between sub-national governments: voters can use the performance of neighbouring governments (or governments of regions that are otherwise similar to their own) to make inferences about the competence or benevolence of their own local politicians (Besley and

---

<sup>1</sup>See for example, Alesina and Spolare(1997), Besley and Coate (2003), Bolton and Roland (1997), Ellingsen (1998), Gilbert and Picard (1996), Lockwood (2002), Oates (1972), Oberholzer-Gee and Strumpf (2002), Wallis and Oates(1988).

<sup>2</sup>Note that along with all of the related literature, we estimate a reduced form relationship between fiscal decentralization and productive efficiency. In our data set, we have no way of distinguishing between the lobbying channel, and the political agency channel. In particular, data on any dimension of lobbying activity in any country, including Switzerland, is very hard to find, and does not exist at the cantonal level.

Smart(2004), Bordignon, Cerniglia, and Revelli (2004))

The third mechanism is that decentralization may lead to a decrease in lobbying by interest groups, which both distorts policy choice and increases waste of public funds. Here, a small number of contributions by Bardhan and Mookherjee (2003), Bordignon, Colombo, and Galmarini (2003), and Redoano (2003) emphasize that the link between decentralization and lobbying is ambiguous. Indeed, under some conditions, there can be more lobbying and distortion of policy choice under decentralization, confirming the belief, going back to the US Federalist Papers in the 18th century, that local government is more susceptible to “capture” by lobbies.

The existing empirical literature does not try to precisely identify these mechanisms. Rather, the approach is to look at a reduced-form relationship between fiscal decentralization and some indicator of the efficiency of government. This literature is, to our knowledge, principally based on cross-country data. The level of fiscal decentralization is usually<sup>3</sup> measured by the percentage of government expenditures made, or taxes collected, at sub-national level, as recorded by the IMF’s Government Financial Statistics. These papers then run regressions where the dependent variable is some easily measured and internationally comparable outcome of government activity against the preferred decentralization measure together with a set of controls. For example, in Khaleghian, 2003, the outcome is immunization coverage rate in the population Treisman, 2002, used immunization coverage also, along with basic drug availability, youth illiteracy rates, and the number of kilometers of paved road per resident of the country. Alternatively, some papers use as the dependent variable some more general indicator of government effectiveness (Huther and Shah, 1998), or corruption (Mello and Barenstein, 2001, Fisman and Gatti, 2000).

In our view, there are two main problems with this literature. First, many of these papers rely on the IMF measure of fiscal decentralization which - it is widely recognized - does not measure very accurately the true autonomy of sub-central government to choose expenditures and set taxes<sup>4</sup>. Second, these regressions do not estimate government “pro-

---

<sup>3</sup>Triesman (2002) is an exception here: he uses several different constitutional indicators of decentralization, such as a dummy variable for a federal country, the number of different tiers of government, etc.

<sup>4</sup>For a critique of the IMF’s Government Finance Statistics. the use of these statistics as measures of the true autonomy of local governments, see Ebel and Yimaz(2002). These statistics tend to overestimate the share of government expenditure and tax revenues that are under the control of sub-national government, and they do so in a way that varies widely across countries. For example, consider a country (e.g. Germany) where some taxes are set nationally but where the revenues are shared with sub-national governments via a fixed formula. The share of tax revenue going to sub-national government is measured

duction functions”, because they do not control for the *inputs* to the output that is the dependent variable. For example, several papers that study health outputs do not control for health expenditures, number of doctors, etc. In the absence of controls for these inputs, these regressions can not tell us much about the efficiency of government as any observed correlation between decentralization and government output can be attributed to omitted variable bias.<sup>5</sup>

To confront these criticisms, our paper uses data on fiscal decentralization and educational inputs and outputs from Swiss cantons over the period 1982-2000. There are several advantages of these data. First, there are data on the level of fiscal decentralization of spending on education within each Canton, collected on a consistent basis over Cantons and years. So, this indicator is likely to be a much better indicator of true fiscal autonomy than in the cross-country case. Moreover, the spending relates to education only, and so there is no danger of the kind of aggregation bias that arises when using the decentralization of *total* expenditure as an indicator, as do the studies cited above. Our paper also relies on specific features of the Swiss context that allow us to treat the variation in decentralization to be exogenous with respect to educational attainment, discussed further in Section 2.3 below<sup>6</sup>. Second, in contrast to the existing literature, as well as measuring “output,” we can control for the influence of inputs on educational outputs, such as educational expenditure per pupil and class size. In our sample the education output in a given year is the fraction of the 19 year old population that obtains the Maturité certificate that allows continuation to university.<sup>7</sup>

---

in the IMF’s statistics as sub-national revenue, even though the lower level of government may have little or no control over choice of the rate or the base. Similar problems appear on the expenditure size from spending that is mandated by central government but implemented by lower-level governments.

<sup>5</sup>For example, if it is found that decentralization is positively related to immunization rates (as does Khaleghian, 2003), this could simply reflect the fact that decentralized countries spend more on immunization, not that they can deliver this service with greater efficiency.

<sup>6</sup>As an additional check on this, we study the relationship across Cantons between the spending measure of decentralization and various direct measures of sub-cantonal autonomy in educational policy, specifically, which level of government has the authority to appoint teachers and/or determine their pay level, whether the local government has the power to set incentive pay, and whether they have some powers over the organization of the school or curriculum. There is a strong positive correlation between the expenditure measure and an index of these direct autonomy measures.

<sup>7</sup>This measure is a natural choice for Switzerland as standardized test scores are not available. Other papers (Jacob and Lefgren, 2005), employs measures of educational attainment combined with subjective teacher performance evaluations. Using US data they show that higher poverty schools strongly value student achievement and are essentially indifferent to the principal’s report of a teacher’s ability to promote student satisfaction.

Our main finding is that there is a robust positive relationship between fiscal decentralization and productive efficiency of public good provision in the case of education. This is present even when canton and year effects and further control variables are allowed for. Finally we find no empirical support that the gains were accompanied by losses in terms of other measures of educational attainment. However, we should note that because the education “production function” is being estimated at the level of the canton, rather than the individual school, our finding might be interpreted as partly due to better preference-matching with fiscal decentralization. For example, more decentralized cantons may be better able to direct resources towards schools where they know there is a stronger demand for education up to maturité level.

We also take our investigation further by asking whether other – possibly time-invariant – features of cantonal and local government interact with decentralization of education to affect the Maturité pass rate. We find that a five-year moving average of local budgetary surplus (which we take to measure “good governance”, following Galiani. and Schargrodsky, 2002) tends to have a positive interaction effect on the pass rate. So, a given amount of decentralization will lead to greater efficiency gains if the local government is more competent.

The rest of the paper is laid out as follows. Section 2 describes the context of decentralization in Switzerland and why the variation in decentralization can arguably taken to be exogenous. Section 3 assesses the extent of local autonomy across cantons and its relationship to expenditure decentralization. Section 4 then turns to the empirical evidence of expenditure decentralization and educational attainment. Section 5 concludes and discusses the results.

## 2. Education Decentralization in Switzerland

### 2.1. Some Background

Switzerland is a Confederation of 26 cantons. These cantons are independent from the federal government in terms of school-level education and most aspects of the day-to-day life in which the state is involved. These 26 Cantons are further divided into 2896 local counties. Table 1 gives for each canton the number of local counties, the average population size per county and the average surface per local county.<sup>8</sup> As it is apparent, the number of local counties per canton, also referred to as fragmentation, is associated

---

<sup>8</sup>The average surface is based on the total polygonal surface for each canton minus the non-productive surfaces, i.e. lakes, glaciers, rocks, etc.

with population size. A simple correlation between population size in a canton and the number of local counties is 0.69, so more populous cantons have more local counties. Despite this high correlation between population size and the number of counties there is still a large variation in the degree of *administrative decentralization*, as measured by the average population per local county in a canton. As the number of counties per region does not vary over time we control for this aspect of administrative decentralization by fixed effects.

Our focus is on school education. In Switzerland, there is a primary school level, a lower secondary school level, and, finally, an upper secondary school level. The first two school levels comprise the nine years of compulsory education<sup>9</sup>. The third level, with a duration of four to five years, offers the university entry qualification, called the Maturité, to students who successfully completed this school and past the final exams.<sup>10</sup> This last level of education will be used as a proxy for educational attainment.

Many responsibilities are common across all cantons in the way they are shared between the local and the central government. Common features among all cantons are that the primary school is the exclusive responsibility of the local communities and many aspects of the upper secondary school is under the jurisdiction of the cantonal government. School material and the definition of aims, scope, and structure of school at all three levels are regulated by cantonal, i.e. central, law. Organizational issues concerning the day-to-day running of the schools like the allocation of pupils to classes, the enforcement of discipline at the school, and ensuring that pupils attend class are the responsibility of the local counties. As in most OECD countries pupils are guaranteed a place in a school within the 'catchment area' of residence.<sup>11</sup> Furthermore pupils can only attend schools of another area under exceptional circumstances. Education is free for residents of a canton but school fees can be levied for pupils residing outside the canton; 95% of pupils in Switzerland attend public schools.

So as such pupils and parents only have a choice of school via the choice of residence.

---

<sup>9</sup>Many cantons however give pupils the option to stay on for one more year after the ninth year without giving an additional qualification.

<sup>10</sup>There also exist professional schools that qualify students for field-specific tertiary education. We return to those type of schools in the section assessing adverse effects of decentralization.

<sup>11</sup>See for instance in Appenzell-Ausserrhoden art. 20 Schulgesetz, par 1. (Also in art 20.3 states that when pupils from another local county attend school than financial support can be requested from that county to help in the financing of those pupils; in Bern art. 7, Volksschulverordnung (VSV) 432.211; in Freiburg art. 8, Schulgesetz and also "Gesetz über den Mittelschulunterricht"; in Nidwald art. 11 Volksschulgesetz (312.1); in St. Gallen art.s 52 and 53 Volksschulgesetz; in Solothurn art. 45, Volksschulgesetz, or in Schwytz art. 32 Verordnung über die Volksschulen.

However unlike the US and the UK where the relative performance of schools across and within areas is well documented and made public, no generalized information of this kind exists in Switzerland. Even if parents get informed through casual observations and experience, no objective evaluation is available which suggests that school quality is not of first order importance in residential choice. We return to this point in more detail below.

Inspections and auditing of schools is present in all cantonal legislation and is made operational through external inspectors appointed by the central cantonal government.<sup>12</sup> They are in charge of training and evaluation of teachers, the supervision of school management, the observation of school curricula, and the use of teaching material, and the inspection of school locations.

## 2.2. Decentralization in Education Expenditures

As noted above, in Switzerland, responsibility for school-level education divided between the cantons and the local counties, with the federal government playing no role. Thus, a natural measure of decentralization is the degree at which local counties are in charge of public expenditure. We use data from the Swiss Federal Statistical Office. Our measure of expenditure decentralization in year  $t$  and in canton  $c$ ,  $D_{ct}$  is

$$D_{ct} = \frac{LE_{ct}}{LE_{ct} + CE_{ct}} \quad (2.1)$$

where  $LE_{ct}$  is the sum of education expenditures in all counties in canton  $c$  in year  $t$ , and  $CE_{ct}$  is the sum of education expenditures at the cantonal level in year  $t$ . When all expenditures are carried out on the local level  $D_{ct} = 1$ , and when the cantonal government is solely in charge of expenditure then  $D_{ct} = 0$ .<sup>13</sup>

For the education sector we see in Table 1 that the average level of  $D_{ct}$  in Switzerland is 0.55, that is about half of all expenditure on primary and secondary education in a canton is spent by the local counties. Whereas Basel-City is a very centralized canton along this measure, Obwald is almost entirely decentralized. The last column in Table 1 gives the standard deviation of education decentralization for each canton which reflects the variation in decentralization within a canton over time. Two facts can thus be highlighted.

---

<sup>12</sup>For instance see the legislations of the cantons of Thurgau (art 5, 410.1 Unterrichtsgesetz, www.tg.ch), Valais (Titre 2, Chapitre 1, 400.1 Loi sur l'instruction publique, www.vs.ch), or Zug (art 67, 411.11 Schulgesetz, www.zg.ch)

<sup>13</sup>An issue that is raised in the literature on federalism is the transfers across states via the federal government - or in our context between the cantonal governments and the local counties. In Switzerland, the local counties principally raise their own taxes to cover expenses and the transfers play, in general, a minor role in adjusting living standards across regions.

First there is sufficient time variation of  $D_{ct}$  within each canton for meaningful fixed effects regressions. Second, there is no general discernible time trend in decentralization.

A key question, of course, is what  $D_{ct}$  is really measuring. First, in the empirical literature on fiscal decentralization, there is always a concern that fiscal indicators may not measure the true autonomy of sub-central government over educational policy decisions. (Ebel and Yimaz(2002)). In the next section, we address this issue in some detail.

### **2.3. What is our Decentralization Indicator Measuring?**

The expenditure decentralization measure,  $D_{ct}$ , exhibits both time-series and cross-section variation. In this section, we argue the following. First, in the cross-section,  $D_{ct}$  is measuring cantonal variations in legislation on local autonomy in education policy. The latter does not change over time in the sample period, and thus cannot be used directly.<sup>14</sup> Yet, we show that in the cross section, expenditure decentralization is correlated with local autonomy. Second, regarding the time-series variation in  $D_{ct}$  within a canton, we argue that this is driven by changes in expenditure caused by arguably exogenous variation in student cohort size.

First, we explain how we can measure legislation on local autonomy in education policy. As the primary school is always under local and the upper secondary always under central jurisdiction we focus on the delegation of decisions at the lower secondary school. Specifically we identified who is in charge of:

1. appointing teachers,
2. determining the pay level of teachers,
3. teachers' incentives, and
4. structural school organization.

In each of these four aspects of education policy, we made a judgement, based on our reading of legal sources, as to whether the policy was wholly under the control of central (cantonal) government or local government.<sup>15</sup>

It is probably helpful to go into more detail on the last three dimensions of autonomy. The second dimension asks whether pay can be set at the local level. In most cantons,

---

<sup>14</sup>As autonomy measures are not time variant therefore their effects are indistinguishable from canton fixed effects, which we want to include in our regressions to control for important unobserved heterogeneity at the canton level.

<sup>15</sup>See the working paper version (Barankay and Lockwood, 2006) for legal sources.

teachers' pay is regulated by the law on civil servants. Each teacher is allocated a pay grade (or spinal point) depending on qualification, the type of occupation, and work experience. However in a few cantons the local counties can make additional payments to attract teachers or can independently design their own pay systems.

The third dimension of local autonomy concerns the presence of incentive pay set at the local level. Usually teachers, after being hired, are automatically promoted at the beginning of each academic year to the next pay grade (or spinal point). However in some cantons<sup>16</sup>, this progression can be put to a halt should the performance of teachers be insufficient. Then teachers can either be kept on the same pay level or even relegated to a step further down the salary scale.

The fourth dimension concerns local autonomy in terms of school organization. This does not refer to routine management of schools, but whether local counties can in fact make important *structural decisions* of some kind. Here we find that four cantons can indeed make such choices. In two cantons the local counties can choose between different school models to provide a specific level of education (Appenzell-Ausserrhoden and Zürich). In a further two cantons - Solothurn and Valais - local counties can decide whether they want to introduce an additional 10th year of education.<sup>17</sup> Finally, the canton of Valais entitles the local counties to decide if they want to regroup pupils from different school years for reasons of effectiveness.

Table 2 indicates the correlation between  $D_{ct}$  and legal autonomy. First, cantons are ranked in descending order in terms of their time-averaged value of  $D_{ct}$ , the level of education decentralization defined as the sum of local expenditures divided by all education expenditure, local and central, in a canton. In columns (3) to (6) we show if a canton allows for local autonomy in any of the four types of dimension discussed above. This table reveals an interesting pattern: cantons with high levels of decentralization are

---

<sup>16</sup>This form of autonomy is present in the cantons of Appenzell-Ausserrhoden, Glarus, Nidwalden, Zug and Zurich and to a certain extent (punishment requested by local authority and granted by cantonal government) in Basel-Landschaft. In the canton of St. Gallen teachers can be dismissed by the local counties. Apart from such "sticks," some cantons allow the local counties to award its teachers "carrots" in the form of performance related pay. This can take on different forms. In Appenzell-Ausserrhoden, local governments can spend up to 0.2% of their wage bill on performance related pay. In Schwytz local councils can budget a specific credit - referred to as *Spontanhonorierung* (spontaneous reward) - of up to 0.3% of gross total wage pay. If this credit is granted by the local legislative body, school councils can then reward teachers for their exceptional performance. In Zug, local councils are generally also allowed to make such bonus payments.

<sup>17</sup>Except for these two cantons the presence of the 10th year is regulated by cantonal law and local counties can not choose to introduce when it is not present or opt out of it when it is.

more likely to have broader local autonomy. In particular one can see that cantons with more decentralization have a higher probability to give their local counties autonomy over teachers' incentive pay. A simple cross section regression analysis - not reported - makes this point more formally where we regress the average level of decentralization per canton over time on a set of dummy variables equal to one if in a local autonomy is present. Two categories have statistically significant explanatory power: autonomy over teacher appointment and on teacher incentive pay is positively correlated with higher decentralization and over 70% of the variation in decentralization across cantons can be explained by a variation in autonomy.<sup>18</sup>

We now turn to the time-series variation in  $D_{ct}$ . We argue that an important part of the time-series variation within a Canton can be explained by variations in school student numbers, and also that variation in student numbers is likely to be independent of the maturité rate. Note first that the local and cantonal components of school education spending can be written by definition as:

$$\sum_l LE_{lct} \equiv \sigma LTS_{ct} + LOC_{ct}, \quad CE_{ct} \equiv (1 - \sigma_c) LTS_{ct} + HTS_{ct} + COC_{ct} \quad (2.2)$$

where  $LTS_{ct}$  denotes the total teachers' salary bill in primary and lower secondary education (lower-school) in canton  $c$  at time  $t$ ,  $HTS$  denotes the total teachers' salary bill in higher secondary education (higher-school) in canton  $c$  at time  $t$ , and finally  $LOC_{ct}, COC_{ct}$  denote other education costs incurred at local and cantonal level which may be a function of the number of students.

These formulae describe several important stylized facts of the Swiss context. First, in all cantons, the cost of teachers' salaries at the lower level are shared between local counties and the canton. In particular, the share  $\sigma_c$  varies across cantons but is fixed over time during our sample period<sup>19</sup>: it is given for all cantons in column (2a) of Table 2. Note that there is quite a lot of cross-section variation in  $\sigma_c$ : for example, in the canton Obwald the full cost of teacher salaries is financed at the local level, whereas in Zürich 67% of the cost is borne by the local counties and finally in Geneva all costs are covered by the central government. Second, the cost of cost of teachers' salaries at the higher secondary level is paid entirely by cantonal government. Third, all other costs are shared between local and cantonal government. We do not have details on the sharing of

---

<sup>18</sup>This result is also robust to the introduction of other control variables: cantons that are more fragmented in the administrative divisions - as measured by the population per local county - do have higher decentralization.

<sup>19</sup>These local shares do not vary over time for the sample period. Legal sources for this data are available from the authors.

other costs, but we do know that other costs in the aggregate are relatively unimportant compared to teachers' salaries - the latter comprise 50-70% of spending in the aggregate (Plotke, 1979, 2003)

Note that (2.1) and (2.2) must explain fully the variation over time in the decentralization indicator, in particular,

$$D_{ct} \equiv \frac{\sigma_c LTS_{ct} + LOC_{ct}}{LTS_{ct} + HTS_{ct} + LOC_{ct} + COC_{ct}} \quad (2.3)$$

This shows that mathematically, variation over time in the indicator must come from either (i) fluctuation in lower-school teacher salaries  $LTS_{ct}$ ; (ii) fluctuation in local and central other costs, and (iii) cantonal expenditure on higher-school teacher salaries and other costs which may be related to the number of students.

Moreover, as the cost of lower-school teacher salaries is shared between the central and the local governments in a fixed ratio, a change in the number of teachers in a canton generically induces a change in the degree of decentralization. Indeed when the local cost share  $\sigma_c$  in a canton is higher (lower) than the degree of decentralization,  $\sigma > D_{ct}$ , then more teachers leads to a subsequent increase (decrease) in the level of decentralization<sup>20</sup>. Mathematically, from (2.3):

$$\frac{\partial D_{ct}}{\partial LTS_{ct}} = \frac{\sigma - D_{ct}}{LTS_{ct} + HTS_{ct} + LOC_{ct} + COC_{ct}}$$

Thus the foremost source of year-to-year variation of decentralization in a canton is induced by changes in the number of teaching personnel and their compensations.

Second, all cantons regulate the class sizes for all school levels by imposing minimum and maximum class sizes.<sup>21</sup> When a class in a school goes above or below these thresholds readjustments are made either by relocating students to other classes within the same school or across nearby schools. So, ultimately changes in student numbers in a canton will lead to straightforwardly changes in teacher salary costs (though cantons have the discretion to adjust to changes in student numbers<sup>22</sup> by hiring more part-time teachers or not renewing their contracts, or hiring new full-time permanent teachers). Thus, in

---

<sup>20</sup>For example in the canton of Fribourg where on average the degree of decentralization is 0.50 and the local cost share is 0.65 an increase in the number of teachers leads on average to an increase in decentralization whereas in Zug, where decentralization is at 0.74 and the local cost share is 0.50, more teachers leads to lower decentralization.

<sup>21</sup>For reasons of space the legal sources have been omitted but are available upon request.

<sup>22</sup>It can be argued that schools have differential incentives to exhaust the possibilities of adjusting classes through relocating students rather than by hiring more teachers if they have to face more of the financial burden themselves when the cost shares are high. We return to this point in section 4.6.

principle, an important source of variation in the variable  $LTS_{ct}$  over time is variation in the size of the student population in canton  $c$  over time.

This shows up in our data. We can also establish empirically that there is a positive relationship between student numbers and the level of decentralization in those cantons where the local cost shares in teacher salaries is higher than the level of decentralization. As the level of expenditures on education other than salaries varies across cantons the effect of changes in student numbers on the level of decentralization will vary, too. Furthermore changes in student numbers may entail other expenditures that are financed by the local or the cantonal governments depending on each canton. As we do not disaggregated data decompose expenditure types within the education category we follow a more general approach by estimating;

$$D_{ct} = \sum_c \theta_c F_c + \sum_c \kappa_c (F_c \times S_{ct}) + \xi_{ct} \quad (2.4)$$

where  $F_c$  are canton dummies which model fixed effects in the level of decentralization and  $S_{ct}$  are the student numbers in a canton and year which is interacted with the canton fixed effects  $F_c$ . As no intercept is present in this model, the  $\kappa_c$  measure the effect of changes in the number of students separately for each canton. The prediction is that the coefficient  $\kappa_c$  is positive (negative) when the local cost shares are higher (lower) than the level of decentralization. We estimated this model and report the  $\kappa_c$  coefficients in column (2b) of Table 2. As can be seen the coefficient has the predicted sign in 13 cantons, in 11 cantons the coefficient is insignificant and but in 2 cantons is the coefficient of the opposite sign than predicted. Thus we find broad empirical support that changes in student numbers induce changes in the level of decentralization.

Note that – in contrast to the U.K. or the U.S.A. – there is no school level or even canton level ranking of schools that are made available to students, parents, or teachers which would allow them to update over time their information about school quality. Thus parents, for instance, do not have yearly updates on school performance on which they can base their decision where to move to enrol their children<sup>23</sup>.

---

<sup>23</sup>There may well be time-invariant differences across schools due to historic differences in the level of development across cantons which parents are aware of. Yet, these permanent differences are controlled for by fixed-effects.

### 3. Decentralization and Educational Attainment

#### 3.1. Empirical Strategy and Data

We now turn to an econometric estimation of the relationship between decentralization and efficiency of public good provision in the educational sector. We will approach the estimation with a panel data set of 26 Swiss Cantons over the period 1982-2000. We have for each canton yearly observations on fiscal decentralization  $D_{ct}$ , a measure of educational output, the maturité rate  $E_{ct}$ , and various input measures. All these variables are defined in Table A1 and are discussed in more detail below.

We estimate the effect of decentralization in canton  $c$  in year  $t$ , on the educational attainment  $E_{ct}$  with

$$E_{ct} = \alpha_c + \beta_t + \gamma D_{ct}^{t-1,t-k} + \eta X_{ct}^{t-1,t-k} + u_{ct}. \quad (3.1)$$

All variables are converted into natural logarithms. The variable  $D_{ct}^{t-1,t-k}$  contains a measure of the degree of decentralization that affected a cohort in period  $t$  in canton  $c$  over the past  $k$  periods. This allows to capture the effect a cohort of students experienced over their whole schooling career. Specifically in the regression we will model  $D_{ct}^{t-1,t-k}$  by

$$D_{ct}^{t-1,t-k} = \frac{D_{ct-1} + \dots + D_{ct-k}}{k}, \quad (3.2)$$

i.e. as a moving average of the past  $k$  periods.<sup>24</sup> The next section will vary  $k$  to filter out the appropriate specification. In particular we will identify separately the effect during the whole schooling career, i.e. during the past 12 years, and separately from the effect during the post-compulsory upper-secondary education, i.e. during the past 5 years, and during the primary and lower secondary education period, i.e. during the first seven years of education.

Similarly the vector  $X_{ct}^{t-1,t-k}$  contains moving averages of the past  $k$  periods of further control variables capturing the quality of the human resources, schooling infrastructure, and per student expenditure. Finally  $\alpha_c$  are canton and  $\beta_t$  are year fixed effects and  $u_{ct}$  are unobservable disturbance terms clustered at the cantonal level. The precise specification of  $u_{ct}$  will be explained and discussed below.

So, (3.1) can be interpreted as a canton-level production function relating educational "output"  $E_{ct}$  to "inputs"  $X_{ct}^{t-1,t-k}$  and the decentralization indicator (3.2). So a positive sign for coefficient  $\gamma$  indicates that increased fiscal decentralization leads to greater "output" for given "inputs" i.e. implies greater productive efficiency.

---

<sup>24</sup>We also allowed discounting values further in the past which yield similar results.

The main<sup>25</sup> output variable  $E_{ct}$  is the *Maturité rate* which is the number of students who obtain the university entrance level qualification deflated by the size of the 19 year old population. Overall in Switzerland in 2000 17% of the 19 year old population obtained the *Maturité* which entitles them to attend university. This level of education is four to five years beyond the compulsory level of education in Switzerland. Numerous studies have shown how students who obtain this level of education have higher future income, better choice of jobs and subjective well-being.

Turning to choice of input variables, there is a vast literature on the economics of education that concentrates on the question as to which input measures affect educational attainment (Hanushek, 1997 and 2003). Our choice of variables is constrained by data availability but includes such standard variables such as class size and expenditure per student at primary and secondary school level excluding tertiary education. To capture the effect of the social composition of the student population we have the proportion of students whose first language is different to the language of instruction: on average 16% of students are not instructed in their first language. By this we want to control for the fact that more foreign students can reduce the educational attainment of a cohort as they may be more difficult to teach due to the language barrier; in a way we thereby capture the quality of the input factor.

Of course, many other covariates contribute to the level of educational attainment of a specific individual student, such as neighborhood characteristics. However the focus of this paper is to identify the effect of the level of decentralization, which is measured at the cantonal level, and as such the identification and the bias of that coefficient is not sensitive to the omission of individual level data (Hoxby (2000)).

### 3.2. Regression Results

Table 3 turns to a set of panel regressions. In the first column we report a first cut at the data regressing the educational attainment among the 19 year old population in a year as a function of decentralization during the past 12 years, i.e. the average effect of exposure to variations in decentralization during the time this cohort spent in school. This regression is thus in the spirit of cross section regressions that ignore the problem of omitted variable bias induced by unobserved heterogeneity (Bardhan 2002). There is

---

<sup>25</sup>However, not all students who continue their education at the upper secondary school level, i.e. beyond compulsory education, attend schools that provide the *Maturité*, but rather attend *professional schools* that also lasts four to five years past compulsory schooling. In the next section we will also address the effect of decentralization on *professional school degrees*.

a significant *negative* correlation between decentralization and educational attainment. However this can be due to omitted factors that are specific to a region. The importance of this omitted variable bias is revealed when we add canton and year fixed effects in column (2). Note that now the coefficient is significantly *positive* at the 10% level. This illustrates that it is not innocuous to ignore the potential for unobserved heterogeneity stemming from historical or cultural differences that are correlated with the degree of decentralization. This can explain why in some studies using cross-section regressions a negative correlation has been found.

In column (3) we introduce our set of control variables. Per pupil expenditure, class size,<sup>26</sup> and the share of non-native speakers are not related to educational attainment. The coefficient on educational expenditure has the expected positive sign just as larger classes are related to lower educational attainment. When we omit the least significant variables, class size and foreign language speakers, we find that expenditure per student is significant at the 10% level. Most notably however is that even after controlling for this last set of variables we find that decentralization is positively related to educational attainment.

So far the right hand side variables are the moving averages of the past 12 years, i.e.  $D_{ct}^{t-1,t-12}$  and  $X_{ct}^{t-1,t-12}$ . In column (5) we identify separately the effect of changes in decentralization during the last 5 years of schooling, i.e. during the post-compulsory education at the upper secondary school, from the effect of changes during the first seven years of schooling;

$$E_{ct} = \alpha_c + \beta_t + \gamma_1 D_{ct}^{t-1,t-5} + \gamma_2 D_{ct}^{t-6,t-12} + \eta X_{ct}^{t-1,t-12} + u_{ct}. \quad (3.3)$$

We find that it is in fact the changes in decentralization experienced during the last five years of education that matter. The coefficient  $\gamma_1$  is significant at the 5% level but  $\gamma_2$  is not significant. The coefficient on expenditure per student is not significant at the 5% level.

A further concern are local business cycles which could drive educational attainment directly thus making our estimates of  $\gamma_1$  biased. We thus constructed a measure of the local business cycle by taking the cyclical component of per capita GDP per capita using the Hodrick and Prescott (1997) filter. As can be seen in column (6) of Table 3 the coefficient of decentralization is robust and the coefficient on our business cycle measure is not significant.

---

<sup>26</sup>This refers to total expenditure, local and central, per pupil. Equally, class size refers to cantonal yearly averages. Note that the effect is estimated only by the variation within a canton over time.

Finally it has been shown in other contexts that there are important cohort effects (Jacobsen, 2004, Card and Lemieux, 2000). First, an increase in the number of students entering maturite schools may affect the amount of resources available to each student and thus the probability of obtaining the final Maturité certificate. On the one hand, more students in Maturité schools may “crowd out” those students whose success is most sensitive to obtaining resources. This channel may be attenuated by the fact that class-size is adjusted to student number by legislation as discussed above. On the other hand, scale economies may not have been exploited by schools with few students. Indeed, many Swiss communities that offer the Maturité have small schools. Second, the presence of more students entering the Maturité school may generate social externalities in communities. This may be through an increased awareness among parents in a community as to what type of support and information these students need and that this support is found locally. Also it can lead to the acceptance that aiming for a Maturité is the new educational norm rather than just finishing compulsory education. Third, controlling for the number of students in maturiete schools we get closer to a measure completion probabilities as we explain the share of 19 year olds in a population who obtain a maturite means conditional of the number of students who attend it. Yet, as discussed in the previous section there is a specific relationship between student numbers and decentralization. Thence the question arises whether decentralization just captures changes in cohort effects. We investigate this potential source of omitted variables bias in Column (7) of Table 3 where we add the average of the past five years of the log of the number of students in Maturité schools but we continue to find that the coefficient on decnetralization is significant at the 5% level. We also investigated non-linear effects of student numbers by adding polynomials of the log of student numbers. In column (8) of Table 3 we added up to the fifth power of the log of student numbers. These non-linear student numbers variables were jointly significant at the 1% level. The level of the coefficient on decentralization is 0.34 which is not significantly different from the coefficient reported in Column (7) of Table 3 but the p-value on of the coefficient on decentralization increases to 0.074. Given the importance of these non-linear cohort effects, however, we control non-linearly for student numbers in the models of the following sections.

Looking at column (8) of Table 3, we see that if decentralization increases by ten percentage points it leads to a 3.5% higher share of 19 year olds obtaining the Maturité. This effect is quite large given that a one standard deviation increase in expenditure is only associated with 1.5% increase in educational attainment. Also, for comparison, to generate a 3.5% increase in educational attainment would require an increase by about 900 in the number of students attending Maturité schools at the mean of student numbers.

Given that the average number of students attending Maturité schools is 2206 and that the within canton standard deviation for Maturité students is only 174 the effect of changes in decentralization on educational attainment can be judged to be very important.

### 3.3. Economies of Scale

It has often been argued that one crucial advantage of centralized provision of public goods is that it can benefit from economies of scale in the production process of public goods: it may be more efficient to focus the design, implementation, and maintenance of public goods in one place rather than have several jurisdictions simultaneously engage in the same production process. To assess if this claim holds in our context we proxy for the scope for economies of scale in two ways.

First we by look at the number of jurisdictions in a canton.<sup>27</sup> We thus estimate the following model:

$$E_{ct} = \alpha_c + \beta_t + \gamma D_{ct}^{t-1,t-5} + \rho(D_{ct}^{t-1,t-5} \times J_c) + \eta X_{ct}^{t-1,t-12} + u_{ct}. \quad (3.4)$$

The variable  $J_c$  measures the number of jurisdictions in a canton. Note that even though the number of jurisdictions is time-invariant, the interaction term is identified by the time-series variation in  $D_{ct}$ . The idea of estimating the interaction term is to capture scale economies that are not a function of population size but only driven by the number of administrations in a canton. This is arguably a plausible measure of economies of scale as an important part of administration takes place at the school level whose size is determined by local factors and is difficult to vary. The empirical prediction is that the interaction term is negative, i.e.  $\rho < 0$ : decentralizing is more efficient when there are fewer jurisdictions involved in the process. Among the Swiss cantons the number of jurisdictions varies considerable between 3 in Basel-Stadt and 400 in Bern. In Table 4 column (1) we report results using the same sample as in Table 3. We restrict the specification to those variables that were significant in Table 3, column (7). As can be seen the coefficient on the interaction term is negative but not significant. The coefficient on decentralization  $\gamma_1$  now increases to 0.52. This is due to the presence of the interaction term: the mean number of jurisdictions is 111 in the data set and therefore the average effect of decentralization is 0.38 which is not significantly different from the coefficient found in column (7) of Table 3.

---

<sup>27</sup>An alternative measure is to use population per jurisdiction which yields qualitatively similar results to the one presented in Table 6. Similarly the estimated effect is robust to the inclusion of population size as an additional control variable.

A second measure of economies of scale is to look at the average population size per local jurisdiction. This measure captures those aspects of scale economies that are a function of population size. We find no evidence, as reported in column (2) of Table 4, for a differential effect of decentralization when the population per local county changes.

### 3.4. Decentralization and Budgetary Competence

A further criticism of decentralization is that the competence of local politicians standing for election may be lower than those standing for positions in the central government. This may be due to the fact that holding an office at the local government is less prestigious than at the central government. These issues can be particularly acute in the context of developing countries, as discussed in Bardhan (2002), where the competence of local public officials is often low. To assess this second argument we follow the methodology in Galiani and Schargrotsky (2001). In that paper, competence of a government is proxied by the size of the budgetary surplus. Low or negative surplus, i.e. deficit, is interpreted to be associated with less competent governments. In contrast to Galiani and Schargrotsky (2001) we have data not only on the level of budgetary surplus at the central but also at the local level. To assess the effect of decentralization jointly with the level of competence we estimate the following model:

$$E_{ct} = \alpha_c + \beta_t + \gamma D_{ct}^{t-1, t-5} + v_L (D_{ct}^{t-1, t-5} \times SL_{ct}) + v_C (D_{ct}^{t-1, t-5} \times SC_{ct}) + \eta X_{ct}^{t-1, t-12} + u_{ct} \quad (3.5)$$

Here  $SL_{ct}$  measures the budgetary surplus of all *local* governments in a canton as a percent of cantonal GDP;  $SC_{ct}$  measures the budgetary surplus of the *central* government as a percent of cantonal GDP. Thus  $v_L$  measures the effect of decentralizing *towards* local governments with a relatively high level of competence and  $v_C$  measures the effect of decentralizing *away from* a central government with a relatively high level of competence. Note that we continue to control for canton fixed effects which holds fixed levels of deficits that are determined by the financing structure in a canton. As there is a strong cyclical component in the budget data we first smoothed the underlying budget deficit data series using the Hodrick and Prescott (1997) filter.

In column (3) of Table 4 we report regression results on examining budgetary competence. We see that  $v_L$  is significantly different from zero and positive and  $v_C$  is insignificant. This is evidence that decentralizing *towards* a competent local government, increases the gains from expenditure decentralization. In column (4) we add a measure of canton level business cycle to address a potential further omitted variable bias. We measure the business cycle as the cyclical part of the per capita GDP in a canton using

the Hodrick and Prescott (1997) filter. The results also remain the same if do not control for business cycles directly or in the underlying budget data. Thus there is evidence that decentralization is more effective when the local government is incompetent. This result gives support to the notion that competence – even in a developed country like Switzerland – needs to be taken into account when deciding on the degree of decentralization.

### **3.5. Adverse Effects of Decentralization**

So far we have found evidence that decentralization is associated with better educational attainment as measured by Maturité rates. It is important, however, to test if these gains were accompanied by losses along other dimensions. We focus on two types of adverse effects.

#### **3.5.1. Gender**

First, we ask if these gains in educational attainment are gender specific, i.e. if the gains to one gender has been accompanied by losses for the other. There are various reasons why decentralization leads to more targeted outcomes. If decentralization is increasing responsiveness to the median voter in each region then the preferences of this voter will change the policy choice. Central governments on the other hand are argued to be more able to redistribute gains to minority interests that are less well represented at the local level. In the case of Switzerland, there are more women than men obtaining the Maturité as a share of 19 year old female and male population respectively. Table 5 presents panel regressions where now the dependent variable is the Maturité rate among women in column (1) and the Maturité rate among men in column (2).<sup>28</sup> For the case of female educational attainment, decentralization has a positive effect but it is not statistically significant at conventional levels.<sup>29</sup> In column (2) we repeat the same exercise but now for male students only. Here decentralization has a strong and significant effect on men; equally the budgetary competence effect only comes into play for male students but not for female students. Given that the gains have accrued more to men than to women we also estimated a gender education gap model - not reported but available on request - where we do indeed find some evidence that decentralization reduced the gender gap in education.

---

<sup>28</sup>The sample is somewhat smaller as educational attainment by canton and gender is not available for the year 2000.

<sup>29</sup>The coefficient is only significant at the 11% level.

### 3.5.2. Professional degrees

Second, we assess if the gains in terms of Maturité rates has been accompanied by a decline in other degrees. Among post-compulsory education the modal group of students attends professional schools rather than Maturité schools. For instance in 2000 17% of 19 year old students obtained the Maturité but 57% obtained a degree from a professional school. In column (3) of Table 5 we estimate a model where the dependent variable is the share of 19 year olds that obtain a degree from a professional school. We find that decentralization had no effect on that level of education.

In summary, we find no evidence for adverse effects of decentralization. More local expenditure is associated with better education among men but that has not been to the detriment of education among women and we find evidence that is reduced the gender gap in education for the Maturité. Similarly other degrees like those obtained from professional schools are not related to decentralization. Therefore it seems that decentralization is associated with a net gain in terms of educational attainment.

## 4. Conclusions

We investigated the empirical evidence on the relationship between decentralization and the efficiency of public good provision. As a first step we looked at the current legislation in the Swiss cantons to provide careful evidence on the positive relation between expenditure decentralization and local autonomy in decision, and also describing the sources of variation in expenditure decentralization over time. We then proceeded to panel regressions of data from Swiss cantons for the last 20 years. We saw that even after controlling for other input variables the degree of decentralization is positively related to educational attainment. We take this to be novel and consistent evidence that contradicts earlier cross-section findings that decentralization worsens governance (Treisman, 2002). Indeed it is the quality of data and the correct specification of the model that helps to uncover the precise effect of decentralization on the efficiency of public good provision.

We also find evidence that expenditure decentralization is more beneficial when central governments are less competent. These results shed new light on the empirical relevance of decentralization and the conditions under which it can attain its often claimed aim of improving the efficiency of public good provision. Future research should aim to further refine the institutional details under which decentralization is beneficial and to also take the data to a micro level to see which people benefit or lose most and under what conditions from decentralization.

## References

- [1] Alesina, A. and E. Spolare (1997), "On the Number and Size of Nations," *Quarterly Journal of Economics*, 112, 1027-56.
- [2] Azfar, O., S. Kahkonen, and P. Meagher, (2001) "Conditions for Effective Decentralised Governance: A Synthesis of Research Findings," IRIS Center, University of Maryland.
- [3] Barankay, I. and B. Lockwood (2006): "Decentralization and the Productive Efficiency of Government: Evidence from Swiss Cantons," CEPR Discussion Paper 5639, London.
- [4] Bardhan, P. (2002). "Decentralization of Governance and Development," *Journal of Economic Perspectives*, 16(4), Fall.
- [5] Bardhan, P. and D. Mookerjee (2000), "Capture and Governance at Local and National Levels," *American Economic Review*, 90(2), 135-139.
- [6] Bordignon, M, Colombo, and Galmarini (2003), "Fiscal Federalism and Endogenous Lobbies' Formation", CESifo Working Paper 1017
- [7] Bordignon, M F. Cerniglia, and R. Revelli (2004), "Yardstick Competition in Intergovernmental Relationships: Theory and Empirical Predictions", *Economics Letters* 83, 325-333
- [8] Besley, T. and S. Coate (2003), "Centralised vs. Decentralised Provision of Public Goods: a Political Economy Analysis," *Journal of Public Economics*, forthcoming.
- [9] Besley, T. and M. Smart (2003): "Fiscal Restraints and Voter Welfare," typescript, University of Toronto.
- [10] Bolton, P. and G. Roland (1997), "The Break-Up of Nations: A Political Economy Analysis," *Quarterly Journal of Economics*, **112**, 1057-90.
- [11] Card, D. and Th. Lemieux, "Dropout and Enrollment Trends in the Post-War Period: What Went Wrong in the 1970s?" National Bureau of Economic Research Working Paper 7658.
- [12] Ebel, R.D., and S. Yimaz (2002), "On the Measurement and Impact of Fiscal Decentralization," unpublished paper, World Bank.

- [13] Ellingsen, T.(1998), “Externalities vs. Internalities: A Model of Political Integration,” *Journal of Public Economics*, **68**, 251-268.
- [14] Fisman, R., and R.Gatti (2001) “Decentralization and Corruption: Evidence Across Countries,” *Journal of Public Economics*, 83(3), 325-45.
- [15] Galiani, S. and E. Schargrodsky (2002), “Evaluating the Impact of School Decentralization on Education Quality,” *Economia, Journal of the Latin American and Caribbean Economic Association*, 2 (2), pp. 275-314.
- [16] Gilbert, G. and P.Picard (1996), “Incentives and the Optimal Size of Local Jurisdictions,” *European Economic Review* **40**, 19-43.
- [17] Hanushek, E. (1997), “Assessing the effects of school resources on student performance: an update,” *Educational Evaluation and Policy Analysis*, 19, 141-64
- [18] Hanushek, E. (2003), “The failure of input-based schooling policies,” *The Economic Journal*, 103, F64-F98.
- [19] Hindriks, J. and B.Lockwood (2005), “Decentralization and Electoral Accountability: Incentives, Separation and Voter Welfare,” CEPR Discussion Paper 5125.
- [20] Hodrick, R.J. and E.C. Prescott (1997), “Postwar U.S. Business Cycles: An Empirical Investigation,” *Journal of Money, Credit, and Banking*, Vol. 29(1), 1-16.
- [21] Hoxby, Caroline M. (2000): “Does Competition among Public Schools Benefit Students and Taxpayers?” *American Economic Review*, 90(5), 1209-1238.
- [22] Huther, J. and A.Shah (1998): “Applying a Simple Measure of Good Governance to the Debate on Fiscal Decentralisation,” World Bank Policy Research Working Paper 1894.
- [23] Jacob, Brian A. and Lars Lefgren (2005): “What Do Parents Value in Education? An Empirical Investigation of Parents’ Revealed Preferences for Teachers,” NBER Working Paper 11494.
- [24] Khaleghain, P. (2003), “Decentralisation and Public Services: the Case of Immunisation”, World Bank Policy Research Working Paper 2989.
- [25] Lockwood, B., (2002), “Distributive Politics and the Costs of Centralisation,” *Review of Economic Studies*, 69(2), 313-337.

- [26] Lockwood, B., (2005), "Fiscal Decentralization: A Political Economy Perspective", in *The Handbook of Fiscal Federalism* (ed. E.Ahmad and G.Brosio), Edward Elgar, forthcoming.
- [27] Mello, L. and M. Barenstein (2001), "Fiscal Decentralization and Governance: A Cross-Country Approach," IMF Working Paper 01/71.
- [28] Jacobsen, Mireille (2004), "Baby Booms and Drug Busts: Trends in Youth Drug Use in the United States, 1975-2000," *Quarterly Journal of Economics*, 119(4), 1481-1512.
- [29] Oates, W. (1972), *Fiscal Federalism*, Harcourt-Brace, New York.
- [30] Oates, W (1999), "An Essay on Fiscal Federalism," *Journal of Economic Literature*, 37(3), 1120-49.
- [31] Oberholzer-Gee, Felix and Koleman S. Strumpf (2002): "Endogeneous Policy Decentralization: Testing the Central Tenet of Economic Federalism," *Journal of Political Economy*, 110, 1-36.
- [32] Persson, T and G. Tabellini (2000), *Political Economics: Explaining Economic Policy*, MIT Press
- [33] Plotke, Herbert (1979): *Schweizerisches Schulrecht*, Paul Haupt, Bern.
- [34] — (2003): *Schweizerisches Schulrecht (2nd edition)*, Paul Haupt, Bern.
- [35] Redoano, M., (2003), "Does Centralization Affect the Number and Size of Lobbies?," Warwick Economic Research Paper 674
- [36] Seabright, P. (1996), "Accountability and Decentralization in Government: an Incomplete Contracts Model", *European Economic Review* **40**, 61-91.
- [37] Treisman, Daniel (2002), "Decentralization and the Quality of Government," unpublished paper, Department of Political Science, UCLA.
- [38] Wallis, J. and W.E.Oates (1988), "Decentralisation in the Public Sector: an Empirical Study of State and Local Government," in Rosen, H.J. (ed.), in *Fiscal Federalism: Quantitative Studies*, University of Chicago Press.
- [39] World Bank (2000), *Entering the 21st Century*, World Development Report.

**Table 1: Descriptive Statistics by Canton**

Canton	Number of local counties	Population in 1000s	Average population per local county	Average surface (km <sup>2</sup> per county) <sup>1</sup>	Average education expenditure decentralization over time <sup>2</sup>	
					Mean	Standard Deviation
Aargau	232	539	2325	5.91	0.37	0.024
Appenzell-Auserrhoden	20	54	2688	11.98	0.78	0.014
Appenzell-Innerrhoden	6	15	2488	26.35	0.73	0.040
Basel-Stadt	3	190	63338	11.79	0.01	0.010
Basel-Landschaft	86	258	2996	5.98	0.63	0.030
Bern	400	942	2356	12.02	0.48	0.025
Fribourg	242	233	965	6.30	0.50	0.030
Genève	45	402	8938	5.37	0.15	0.014
Glarus	29	39	1334	14.95	0.68	0.031
Graubünden	212	186	878	19.55	0.76	0.032
Jura	83	69	831	10.02	0.45	0.022
Luzern	107	345	3224	13.01	0.71	0.035
Neuchâtel	62	166	2672	11.46	0.66	0.012
Nidwald	11	38	3416	18.97	0.89	0.031
Obwald	7	32	4586	56.99	0.93	0.012
Schaffhausen	34	74	2165	8.67	0.66	0.020
Schwyz	30	128	4250	24.22	0.79	0.016
Solothurn	126	243	1931	6.22	0.75	0.017
St. Gallen	90	447	4963	19.63	0.81	0.024
Ticino	245	308	1257	7.84	0.44	0.038
Thurgau	80	227	2832	10.64	0.77	0.030
Uri	20	36	1776	23.88	0.67	0.023
Vaud	384	615	1601	6.99	0.45	0.013
Valais	160	275	1719	15.07	0.40	0.018
Zug	11	97	8834	18.38	0.74	0.021
Zürich	171	1196	6996	9.53	0.64	0.045
<b>Switzerland</b>	<b>2896</b>	<b>7185</b>	<b>2481</b>	<b>14.26</b>	<b>0.55 <sup>3</sup></b>	<b>0.215</b>

Notes: Decentralization is defined as the ratio between the sum of all local expenditure in a canton and the sum of all local plus cantonal expenditures. <sup>1</sup>Based on total surface per canton minus non-productive surface (lakes, rock, glaciers). The population data is based on an average between 1997 and 2001. <sup>2</sup>Calculated for average education expenditure between 1981 and 1999. <sup>3</sup>Average decentralization rate for 1982 to 1999 defined as the sum of all local expenditures across cantons on public education as a share of all expenditure by local and cantonal governments across cantons. The unweighted average across cantons is 0.61.

**Table 2: Sources of Variation in Education Decentralization**

	(1)	(2a)	(2b)	(2c)	(3)	(4)	(5)	(6)
Canton	Education Decentralization	Local Cost Share of Teacher Salaries	Marginal effect of student numbers	In line with predictions?	Teacher Appointment	Teacher Salary	Teacher Incentives	School Organization
Obwald	0.93	1.00	0.0313***	Y	x		x	
Nidwald	0.89	1.00	0.0325*	Y	x	(x)	x	
St. Gallen	0.81	0.75	0.00118*	N	x		(x)	
Schwyz	0.79	0.80	0.00684**	Y	x		x	
Appenzell-Ausserrhoden	0.78	1.00	-0.0296		x		x	x
Thurgau	0.77	0.75	0.00048		(x)		x	
Graubünden	0.76	0.63	0.00048		x		x	
Solothurn	0.75	0.54	-0.000952		x			x
Zug	0.74	0.50	-0.0141***	Y	x	x	x	
Appenzell-Innerrhoden	0.73	0.80	0.3087***	Y	x		x	
Luzern	0.71	1.00	0.0005		x	(x)	x	
Glarus	0.68	0.50	0.0317		x		x	
Uri	0.67	0.65	-0.0832		x	(x)	x	
Neuchâtel	0.66	0.55	-0.0192***	Y	x		(x)	
Schaffhausen	0.66	0.57	0.0230		x			
Zürich	0.64	0.67	0.00168***	Y	x		x	x
Basel-Landschaft	0.63	0.50	-0.00197*	Y	x		(x)	
Fribourg	0.50	0.65	0.00377***	Y	x			
Bern	0.48	0.70	0.0000122***	Y	x			
Jura	0.45	0.00	0.0451		x			
Vaud	0.45	0.00	-0.00283***	Y	x			
Ticino	0.44	0.57	0.0184***	Y	x			
Valais	0.40	0.50	0.00382**	Y	x			x
Aargau	0.37	0.25	-0.00115		x			
Genève	0.15	0.00	-0.00168					
Basel-Stadt	0.01	0.00	0.00957***	N	(x)			

*Note: Education decentralization is defined as the sum of all local expenditure divided by cantonal expenditure plus the sum of all local expenditure in a canton. Local share of teacher costs refers to the share of teacher salaries and compensations that are financed at the local level. The correlation between the degree of decentralization and the cost share is 0.81. Legal sources on cost shares are available upon request. In column 2b we report the marginal effect of changes in student numbers in a canton on the level of decentralization. In 15 out of 26 cantons the effect significant and in 13 cantons the coefficient is of the expected sign. An 'x' denotes that a decision is taken at the local rather than at the central level; (x) decision is shared between local and central government. Teacher Incentives refers to decisions about performance related pay (pay level increment, bonus etc.); School Organization refers to decisions about significant organizational matters (class structure, schemes). See Appendix 1 and 2 for definitions and legal sources on autonomy.*

**Table 3 – Decentralization and Educational Attainment in Swiss Cantons**

Dependent variable = Share of 19 year old population obtaining university entry qualification (Maturité rate)

Robust standard errors reported in parentheses, allowing for clustering at canton level

	Base regression (1)	Canton & year effects (2)	Further controls (3)	Further controls (4)	Effect over time (5)	Business Cycle (6)	Student Numbers (linear) (7)	Student Numbers (non-linear) (8)
Decentralization <i>average of past 12 years</i>	-0.2147*** (.0607)	0.3767* (.2185)	0.4085* (.2323)	0.4123* (.2133)				
Decentralization <i>average of past 5 years</i>					0.3749** (.1740)	0.3766** (.1714)	.3443** (.1656)	.3354* (.1801)
Decentralization <i>average of 6 to 12 years lagged</i>					0.0811 (.1003)			
Expenditure per pupil <i>average of past 12 years</i>			0.0842 (.0537)	0.0937* (.0483)	0.0994** (.0450)	0.1051** (.0434)	.0826** (.0404)	.0675 (.0477)
Class size in upper secondary school <i>average of past 12 years</i>			-0.0288 (.0580)					
Non-native speakers <i>average of past 12 years</i>			0.0105 (.0236)					
Business Cycle						0.0004 (.0006)		
Student Numbers in Maturité schools <i>average of past 5 years</i>							.0510*** (.0116)	
Canton fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-linear student numbers	No	No	No	No	No	No	No	Yes
Errors clustered at canton level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.4480	0.9516	0.9539	0.9536	0.9544	0.9544	0.9571	0.9579
Number of observations	208	208	208	208	208	208	208	208

Notes: \* significant at 10% level, \*\* at 5% level, \*\*\* at 1% level. See Table A1 for definition of variables. Right hand side variables in column (1)-(4) are average values of past 12 years. In column 5 decentralization is split into average of past 5 years - proxying for the period during upper secondary school - and 6 to 12 years in the past - proxying for the period during primary and lower secondary school. In column (6) we measure business cycles as the *cyclical* part of canton-level GDP per capita using the Hodrick-Prescott (1997) filter with  $\lambda=1600$ . Results in column (4) and (5) are robust to using the raw budget data instead of the smoothed data. In column 8 we entered the log of students in Maturite schools up to the fifth power. The coefficients on these non-linear student number variables are jointly significant at the 1% level with an F-test of 4.85.

**Table 4 – Decentralization, Economies of Scale, Local Autonomy, and Budgetary Competence**

Dependent variable = Share of 19 year old population obtaining university entry qualification (Maturité rate)  
 Robust standard errors reported in parentheses, allowing for clustering at canton level

	Number of local jurisdictions (1)	Population size per jurisdiction (2)	Budgetary Competence (3)	Business Cycle (4)
Decentralization <i>average of past 5 years</i>	0.4818 (.3593)	0.3370* (.1753)	0.3821** (.1685)	0.3862** (.1677)
Decentralization x Number of local jurisdictions <i>average of past 5 years</i>	-0.00092 (.00133)			
Decentralization x Average population per local jurisdiction <i>average of past 5 years</i>		-0.0144 (.0285)		
Decentralization x Dummy =1 when teacher incentives decided at local level <i>average of past 5 years</i>				
Decentralization x Smoothed budget surplus at local level <i>average of past 5 years</i>			0.0174** (.0081)	.0152** (.0077)
Decentralization x Smoothed budget surplus at cantonal level <i>average of past 5 years</i>			0.0026 (.0055)	0.0027 (.0053)
Expenditure per pupil <i>average of past 12 years</i>	0.0771 (.0492)	0.0818* (.0441)	0.0635 (.0458)	0.0710 (.0436)
Business Cycle				0.0005 (.0004)
Canton fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Errors clustered at canton level	Yes	Yes	Yes	Yes
Non-linear student numbers	Yes	Yes	Yes	Yes
Joint F-test on non-linear student numbers	3.08**	4.61***	4.11***	3.31**
Adjusted R-squared	0.9581	0.9580	0.9589	0.9592
Number of observations	208	208	208	208

Notes: \* significant at 10% level, \*\* at 5% level, \*\*\* at 1% level. In column 3 and 4 we first eliminated the cyclical component of the underlying budget surplus data using the Hodrick-Prescott (1997) filter with  $\lambda=1600$ , where  $\lambda$  is a parameter penalizing variability in the underlying growth component series. Results are robust to alternative values of  $\lambda$ . In column (4) we measure business cycles as the *cyclical* part of canton-level GDP per capita using the Hodrick-Prescott (1997) filter with  $\lambda=1600$ . Results in column (3) and (4) are robust to using the raw budget data instead of the smoothed data. In all specifications student number are controlled for linearly and up to the fifth power.

**Table 5 – Any Adverse Effects of Decentralization?**

Robust standard errors reported in parentheses, allowing for clustering at canton level

	Dependent Variable: Maturite rate among women (1)	Dependent Variable: Maturite rate among men (2)	Dependent Variable: Professional school degrees (3)
<b>Decentralization</b> <i>average of past 5 years</i>	<b>0.4212</b> (.2540)	<b>0.4711**</b> (.2122)	<b>0.7645</b> (.6328)
<b>Decentralization x Budget Surplus at the local level</b> <i>average of past 5 years</i>	<b>14.42</b> (8.953)	<b>28.98***</b> (8.583)	<b>-41.55</b> (71.24)
<b>Decentralization x Budget Surplus at the cantonal level</b> <i>average of past 5 years</i>	<b>-7.993</b> (6.878)	<b>-4.812</b> (4.683)	<b>33.10</b> (25.53)
<b>Expenditure per pupil</b> <i>average of past 12 years</i>	<b>0.0485</b> (.0359)	<b>0.0818**</b> (.0356)	<b>-0.0965</b> (.2418)
<b>Canton fixed effects</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<b>Year fixed effects</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<b>Errors clustered at canton level</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<b>Non-linear student numbers</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<b>Joint F-test on non-linear student numbers</b>	<b>22.55***</b>	<b>28.16***</b>	<b>6.65***</b>
<b>Adjusted R-squared</b>	<b>0.9748</b>	<b>0.9414</b>	<b>0.7756</b>
<b>Number of observations</b>	<b>182</b>	<b>182</b>	<b>205</b>

Notes: \* significant at 10% level, \*\* at 5% level, \*\*\* at 1% level. See Table A1 for definition of variables. The dependent variables are: in column (1) the number of women obtaining the maturite degree as a share of 19 year old women; in column (2) the number of men obtaining the maturite degree as a share of 19 year old men; in column (3) the number of students obtaining the professional school degree as a share of 19 year old population. Student numbers in (1) are male students in maturite schools, in (2) female students in maturite schools, and in (3) students in professional schools – in all three columns controlled for linearly and up to the fifth power.

## Table A1 – Summary Statistics

The Data covers 26 Swiss Cantons during 1982-2000

<b>Performance Measure for Public Education</b>	<b>Mean</b>	<b>St. Dev</b>
<b>Maturité rate</b>		
Share of 19 year population with University entry level qualification	.164	.057
Female Maturité rate as a share of 19 year old women	.165	.074
Male Maturité rate as a share of 19 year old men	.161	.046
Professional School Degrees as a share of 19 year old population	.593	.134
<b>Decentralization</b>		
Share of local expenditure to all expenditure (local and central) <i>5 year moving average</i>	.610	.215
Number of local jurisdictions in a canton	111.38	111.74
<b>Control Variables</b>		
Education expenditure per pupil* (12 year moving average) at all levels of government in 1990 1000s of Swiss Francs	14.933	2.926
Class size in schools (12 year moving cantonal average)	32.60	3.29
Non-native speakers (12 year moving average) percent of students whose first language is not the language of instruction	16.30	8.13
<b>Budget surplus as a percent of cantonal GDP (5 year moving average)</b>		
-at the central (cantonal) government level	-.7825	.9249
-at the local government level	-.3655	.3825
<b>Business Cycle</b>		
Cyclical component of the canton-level GDP per capita using the Hodrick-Prescott (1997) filter	-789.44	2104.86

Notes: All data is from the Swiss Federal Statistical Office (various departments); <http://www.statistik.admin.ch>. \*primary and secondary education excluding tertiary education.