

School staff autonomy and educational performance: within school type evidence*

Marijn Verschelde[†] Jean Hindriks[‡] Glenn Rayp[§] Koen Schoors[¶]

September 20, 2013

Abstract

This paper shows the effect of school staff autonomy on educational performance. The distinctive feature with existing literature is that we employ variation in autonomy within the same country and within the same school type to reduce the omitted variable problem. To fully capture the informational advantage of local actors, we define autonomy as the operational empowerment of the school's direction and teachers. The Flemish secondary school system in Belgium is analyzed as it displays unique within school type variation in school staff autonomy. This variation originates from autonomously operating school governing bodies that can group multiple schools and are free to delegate responsibilities to the school staff; This variation is likely exogenous as it is mainly driven by approximately random variation in characteristics and composition of the school boards. Combining detailed school level and pupil level data from the PISA 2006 study with a semiparametric hierarchical model, we find strong positive effect of school staff autonomy on educational performance. The result is shown to be robust to problems of reverse causality and simultaneity.

*We sincerely thank the participants of the Workshop 'Educational Systems and Human Capital' in Trondheim (March 2011), Bart Cockx, Gerdie Everaert, Meryem Fethi, Kristiaan Kerstens, Dirk Van de gaer, Sietse Bracke, Klaas Mulier and Joost Vandenbossche for their helpful suggestions and insightful comments on an earlier draft of this paper. Further, we thank Inge De Meyer of PISA Vlaanderen for her support. Marijn Verschelde acknowledges financial support from the Fund for Scientific Research Flanders (FWO-Vlaanderen). All remaining errors are ours.

[†]SHERPPA, Department of General Economics, Ghent University, Tweekerkenstraat 2, 9000 Gent, Belgium. email: marijn.verschelde@ugent.be.

[‡]Corresponding author: CORE, Department of Economics, Université catholique de Louvain, Voie du Roman Pays, 34, Louvain-la-Neuve, Belgium, e.mail: jean.hindriks@uclouvain.be.

[§]SHERPPA, Department of General Economics, Ghent University, Tweekerkenstraat 2, 9000 Gent, Belgium. email: glenn.rayp@ugent.be.

[¶]CERISE, Department of General Economics, Ghent University, Tweekerkenstraat 2, 9000 Gent, Belgium. email: koen.schoors@ugent.be.

Keywords educational performance, PISA, school autonomy, educational production function, semiparametric

JEL-classification I28, H52

1 Introduction and related literature

A remarkable stylized fact of educational economics is that higher school resources do not necessarily yield higher pupil performance. Evidence for an overall large effect of school resource policies on pupil performance is largely missing (Hanushek, 2003; Wößmann, 2003). It cannot be excluded however that the effect of input-based policies on pupil performance is moderated by the incentive structure of school actors (Hanushek, 2003). Since the incentive structure of school actors is embedded in school institutions, insight in the latter may be crucial to understand the relation between school resources and pupil performance.

There is accumulating cross-country evidence that getting incentives right by a combination of monitoring and autonomy is beneficial for educational efficiency (Wößmann, 2008). Since these studies compare institutions across countries the results may be biased because of an obvious omitted variable problem. Any other source of cross-country variation, like legal or cultural differences, may indeed bias the results. To identify the effect of school autonomy from within-country changes, Hanushek et al. (2011) use a panel estimation with country fixed effects on student level data from 42 countries from the Programme for International Student Assessment (PISA) 2000, 2003, 2006 and 2009 dataset. As school autonomy is considered as a country level average, within-country selection does not affect their estimates. Moreover, the use of country fixed effect estimation ensures that estimates are not affected by time-invariant country-specific heterogeneity. The authors find that the effect of school autonomy depends on the level of economic and educational development. In other words, when institutions are strong (weak), considerable academic gains (losses) are found from decentralized decision-making. In addition, the authors find larger gains from school autonomy if an adequate accountability system is in place such as central

examination.

However, we may question whether institutional differences in a cross-country analysis can be adequately captured by including country fixed effects in the estimation model. As secondary education is decentralized to the regional level in several countries (e.g. Belgium, Germany, Spain, United Kingdom), there is still regional level institutional variation that can affect the results from country fixed effects estimations. Moreover, the assumption that institutional features such as the awareness of the importance of education and academic culture are time-invariant is hard in the specific setting. One example of time-varying institutional features is the change in institutional settings at country-level after the publication of dramatic scores for some countries in the first round of PISA (i.e., PISA 2000). e.g. in Germany, this publication was the start of an “*an intense political debate that spread over almost all areas of the political and economic life, as the human capital acquired in a nation’s schooling system is generally regarded as the most valuable resource of society*” (Ammermüller, 2004). Not without reason, the PISA 2000 publication was called there the ‘*PISA shock*’.

In studies that use within-country variation of monitoring and autonomy, the problem of adequately controlling for (time-varying) institutional variation at country level is of course avoided. However, almost all student level within-country evidence of positive effects of school autonomy comes from studies comparing different types of schools (see e.g. the charter school literature, referenced in e.g. Abdulkadiroglu et al. (2011)). As noted by Hanushek et al. (2011), it is difficult to extract school autonomy effects from school type effects such as parental choice, quality of information or constraints on school location. Clark (2009) uses a different kind of variation, namely the variation in autonomy by the formation of a new type of public school with more autonomy (i.e., the ‘*Grant Maintained*’ schools) in the UK between 1988 and 1997. The authors use a regression discontinuity design to compare public schools in which the vote to become a more autonomous ‘*grant maintained*’ public school barely won with those in which the vote barely lost. Significant and persistent achievement gains are found for schools that opted to become ‘*grant main-*

tained' schools. Still, also in this study, school autonomy effects are not fully separated from other school type specific effects as variation in school autonomy is between school type.

The first contribution of this paper is that we employ within-country, *within school type* variation of autonomy on a dataset of pupil level performance, which gives us the degrees of freedom needed for statistical inference, and allows us to better isolate the school autonomy effects from school-type and country-specific effects.

This within school type variation comes from the particular structure of secondary education in Flanders, one of the three regions to which education is decentralized in Belgium.

The second contribution of this paper is that we restrict ourselves to a narrow definition of autonomy. We only look into the effects of autonomy of principals and teachers, the local agents that, through their local informational advantage, are supposed to boost educational outcomes. This identification strategy brings our work closer to a clean test of the supposed effects of autonomy in a principal-agent framework (see below), where the government is the principal and the local school actors (school's direction and teachers) are the agents.¹

The remainder of this paper is structured as follows. In a second section, we discuss the theoretical background for the expected effect of school autonomy on educational performance. Identification is discussed in section 3. In section 4 and 5, we present the semiparametric multi-level analysis and the PISA data. In section 6, we discuss the results on the effect of school staff autonomy on educational achievement. Section 7 concludes.

¹Hallinger et al. (1996) were the first to measure principals' activities in key dimensions of a school's instructional program and to relate these to student outcomes such as reading achievement. Wößmann (2003) is one of the first to look into the effect of individual teacher influence over teaching on student performance.

2 Theoretical background

The impact of school autonomy is linked to several strands of the literature. The decentralization of education may boost efficiency and productivity by eliminating unnecessary bureaucratic burdens (see Niskanen (1971) and Niskanen (1991), for seminal work on budget maximizing bureaucrats). School autonomy may help schools to overcome bureaucratic rigidity and in this way impact student performance positively (Bottani and Favre, 2001; Chubb and Moe, 1990).

Entrusting the provision of education to local agents may also lead to more efficient provision because local agents will be closer and more responsive to student needs and preferences since students can ‘vote with their feet’ by changing school or even community. Tiebout (1956) shows that decentralized public good provision may, under certain conditions, yield the efficient provision of public goods like education. Hoxby (1999) confirms this Tiebout hypothesis for local school productivity under much less restrictive conditions. This suggests that the combination of decentralization and free school choice may indeed provide greater opportunities for local citizens and students to monitor and discipline the local agents that are responsible for educational policy, thereby creating greater efficiency and productivity.

If the decentralization of education is accompanied by public information on school performance, it may also be conducive to yardstick competition (see Shleifer (1985), and Besley and Case (1995)) among schools, in this way encouraging the adoption of more effective teaching methods and more efficient operational procedures. Card et al. (2010) for example recently find significant effects of enhanced competition on the test score gains of students in Canada in all studied school systems. They however also point at a possible negative effect of this yardstick competition. It cannot be excluded that “in more competitive markets teachers and principals spend more time and effort preparing for standardized tests, and less on other aspects of learning. If “test skills” have limited intellectual value, the effort devoted to competing over test outcomes is socially wasteful, and the higher test

score gains observed in more competitive markets may be counter-productive” (Card et al. (2010), p. 29-30).

In weak institutional environments, decentralization may however have negative implications, like (i) increased levels of uncoordinated rent-seeking and corruption as government structures become more complex and devoluted (Fan et al., 2009), (ii) increased coordination costs and (iii) slower institutional reform.

The most important negative consequence of increased autonomy may lie in a potential principal-agent problem (see Wößmann et al. (2007)). The government (principal) tries to improve cognitive skill creation by delegating responsibilities to schools (agents) that are assumed to have a local information advantage over the principal. A principal-agent problem appears when the interests of the government and the school diverge and information is asymmetric. Interests typically diverge for decisions that influence the financial position of the school or the workload for school actors. Budget formation and curriculum content are therefore policy areas with a high probability of divergence between the interests of the government and the school. In process and personnel decisions on the other hand, little divergence of interest is expected. This principal-agent problem requires to put in place appropriate accounting systems. With effective accountability, autonomy is expected to enhance educational performance. Central examinations are a widely used accountability mechanism to align incentives between schools and the government (Wößmann et al., 2007), but other mechanisms can be used to attain this goal.

In Flanders, instead of central examinations, inspection teams investigate on a regular basis whether the curriculum and teaching process are aimed at reaching the centrally imposed final standards and whether budget formation is in accordance with the posed requirements. Benchmarking by parents is possible as the inspection reports are publicly available. In addition, freedom in budget formation is limited to additional funding, above the centrally imposed funding system. The size of these additional budgets is very small in comparison to the school budget. It mainly consists of revenues from student enterprises (such as a bakery in a bakery school) and donations by parents to finance e.g. school trips

or school material (Poesen-Vandeputte and Bollens, 2008). Consequently, discretionary power of schools are limited on divisive issues like budget formation and curriculum development. We therefore expect that the principal-agent problem will be limited in our case, and that the institution of school autonomy, through improved incentives for schools and teachers, will affect resource-allocation decisions and ultimately the educational performance of students positively.

It is worth noting that the positive effect of autonomy can also be supported by Oates' efficiency theorem (see e.g. Hindriks and Myles (2006, chapter 17) for an overview). Indeed under Oates' approach, autonomy is beneficial to better match local preferences and needs (the preference matching benefit) but could be detrimental in terms of lack of coordination and spillovers (the spillover costs). To determine if autonomy is beneficial it is necessary to compare the magnitude of the costs and benefits. It is easily seen that most items in school policy (such as budget formation, course content, teacher selection, disciplinary policies, student admission, etc.) display both preferences matching benefits and spillover costs. However the relative magnitude of these cost-benefits vary from one item to the other. Student and teacher selection is probably the one with highest spillover costs and the lowest preference matching benefit. It is therefore natural that we observe little school autonomy on such issues. On the other hand, the budget allocation presents low spillover costs and high preference matching benefits so autonomy is expected to produce better outcome than centralized decision making with uniform policy choice. In fact without spillover, decentralization is always superior. With spillover, decentralization can still dominate if there is sufficient difference across school in terms of needs and preferences. This argument relies heavily on the assumption that autonomy leads to better differentiation in educational policy to match local needs, and that the spillover effects are limited.

Last but certainly not least, autonomy is linked with intrinsic motivation. Human behaviour is driven by both intrinsic and extrinsic motivation. Both economic and psychological literature (i.e., Frey (1993), Frey (1994), Frey and Oberholzer-Gee (1997) and Deci and Ryan (1985)) pinpoint the so called "*hidden cost of reward*", and cost of control by

destroying the “*psychological contract*”. Frey (1993) shows in a principal-agent framework that monitoring by the principal can be perceived by the agents as an indication of distrust, with lower work effort by the agents as result. Frey and Oberholzer-Gee (1997) shows that extrinsic motivation (such as price incentives) has a crowding-out effect on intrinsic motivation. Differently put, price incentives or external intervention that is perceived to be controlling can reduce the feeling of living up to the civic duty and diminish altruistic behaviour. This implies that the intrinsic motivation of the school staff and education quality is expected to be higher in schools where the school staff perceives to have considerable responsibilities.

3 Identification

The constitutionally guaranteed freedom of education is the cornerstone of the Flemish landscape.² Freedom of education implies free school choice and freedom to start a school and autonomously organize education for any natural or legal person. In practice, freedom of education has resulted in large autonomy for non-profit school governing bodies that can group multiple schools.

The non-profit governing boards can either determine school policy themselves or delegate operational autonomy to the school principals (principals and vice-principals are usually grouped in a school direction) and teachers. Between school type variation exists because there is considerable variation in the composition of school boards. In addition to between school type variation, there is variation in operational autonomy that takes place within the four main school types: 1) the publicly funded, privately managed education network, 2) the Flemish Community education network, 3) the Flemish Provincial education network and 4) the education network of Flemish cities and municipalities.

In subsidized education there are no rules that exhaustively define the role of school princi-

²See Devos and Tuytens (2006), OECD (2008a) and OECD (2011) for a complete overview of the Flemish education system.

pals. For the Flemish Community Education Network (GO!), the Special Decree of 14 July 1998 stipulates which responsibilities should be delegated to the school direction.³ However, there is large variation in the interpretation of the law and the actual autonomy that is delegated to the school principals. Some governing boards take all financial decisions, while others stand aloof (Van Petegem et al., 2009, p. 20-22).⁴ OECD (2008b) concludes that *“Even in countries that do have legislative frameworks to define the responsibilities of school principals, there are concerns about the relevance and clarity of the legislation...in Flemish Belgium there is no comprehensive statutory description of the post of the school leader.”*

Ample qualitative research documents the extreme differences in the responsibility structure within school governing bodies. Panel discussions between school principals (Van Petegem et al., 2009), case studies of schools (among others Caenepeel (1988), Devos et al. (1989) and Devos et al. (1999) and Van Petegem et al. (2009)) and an OECD review team (Day et al., 2008)) document that in some school governing bodies, the school direction is delegated full operational responsibility over its own school and the school governing body only intervenes when large problems occur. In other governing bodies, the school principal is considered as a mere employee of the school governing body, with virtually no autonomy.

The PISA 2006 data used in this study (discussed in detail in section 5), support the idea of considerable within school type variation in school staff autonomy. Of the 92 sampled publicly funded, privately managed schools, 12 schools have no considerable responsibility

³Examples of school direction responsibilities are the general and pedagogic organization of the school, the organization of extramural activities and the allocation of the by the school group formulated school budget. Besides these by law delegated responsibilities, the school principal can be delegated extra responsibilities by the school group.

⁴In particular, Devos et al. (1999) and year reports of schools in GO! education show that in some school groups, each school is free to formulate and allocate the full budget for its own school and the role of the school board is only to ratify the budgets if there are no large problems. In other schools, financial cells of the governing board formulate the school budgets and leave little room for budget allocation by the school principals (e.g. only small purchases of educational material).

for within-school budget allocation, while 34 schools have full autonomy. Of the 34 publicly managed schools (organized by the Flemish Community, provinces or cities), 7 publicly funded schools have no considerable responsibility for budget allocation, while 13 schools have full autonomy. As there is also considerable variation in school staff autonomy within the three public education networks (which we cannot document because of anonymity reasons), we thus have both qualitative and quantitative support for unique within school type variation in school staff autonomy. The within school type variation in autonomy is not related to educational tracks in the sense that general education schools are given more autonomy than technical-arts or vocational schools. Levels of autonomy are similar across educational tracks and overall slightly higher in technical-arts and vocational education.⁵

The variation in school staff autonomy is likely to be random as it is mainly driven by approximately random variation in characteristics and composition of the school board, resulting from:

1. Historical reasons: some governing boards in publicly funded, privately managed education originate from e.g. dioceses that were very actively involved in education and centrally structured, whereas others stem from orders, dioceses and congregations that left school policy in the hands of educational personnel, i.e. the school principals. (Devos et al., 1999).
2. Discontinuities caused by political elections and political appointments: school case studies in Devos et al. (1999) illustrate that changes in the political composition of local authorities and provincial authorities (the governing bodies of respectively city schools and provincial schools) resulted in new responsibility structures within school governing bodies.
3. Differences in the interpretation of the law and job description: some school governing bodies restrict themselves to daily management of schools, while others intervene in

⁵In general education, 83,41 percent of the pupils are in schools with considerable autonomy. In technical-arts and vocational education, this is respectively 88,02 and 88,29.

all facets of school policy. Other school governing bodies only have a formal existence. In one extreme case example, the governing board did not deliberate on school policy in the last three years.(Devos et al., 1999)

4. Other unobserved heterogeneity such as geographical development and budgetary tensions that can, in the Flemish context, be considered as approximately random.

The Flemish education system is, besides this unique within school type variation, a good testing ground to study the effects of school autonomy as we do not expect significant influences of issues of reverse causality, simultaneity and self selection.

As regards reverse causality, first, pupil examination is not centralized, implying many school governing boards have no information on pupil achievement that can be used in school principal evaluations. The PISA 2006 dataset indicates that in only a few schools (9 out of the 126 sampled schools), performance data (e.g. mean exam scores, average test results, rates of success in higher education) are used to evaluate the school direction.

Second, school principals have many and complex tasks. Therefore, the criteria to evaluate school principals are complex and case specific. In the words of a Flemish policy maker: “*No one knows about the quality of the principal*” (Day et al., 2008). It is thus probable that perceived principal performance is only weakly related with actual principal performance. Third, as discussed in OECD (2008a), in the period studied, there is a lack of school self-evaluation in Flemish secondary education. Only recently, programs for self-evaluation are installed (OECD, 2011). Thus even if there would be knowledge on the school principal’s performance, in many governing bodies, this information will not be used in evaluations, as evaluations are more focused on being in line with regulatory requirements.

As regards simultaneity, we do not find indications of simultaneity issues, caused by dynamic principals and teachers that struggle for both more autonomy and higher educational achievement in the school, without there needs to be a causal relation between school autonomy and educational achievement. As documented in Table 8 of Verschelde et al. (2012), there is little association between observed school staff dynamism and school staff empow-

erment, indicating that our variation in staff autonomy is not simply reflecting the school staff dynamism. Therefore staff autonomy per se matters for educational performance.

As regards the self-selection problem, as discussed and argued above, approximately random school board characteristics drive the variation in school autonomy. We thus expect no endogeneity of assignment of treatment of school principals. Further, as it is unobserved for students and teachers that consider to enter a school whether or not the school governing body or school's direction is in charge of budgeting issues, we do not expect that school choice by pupils and teachers is related to school staff empowerment in our specific setting.

4 Methodology

We hypothesize that a combination of school staff empowerment in budgeting and effective accountability will, after controlling for socio-economic and school-level institutional variation, have a positive effect on educational performance.

Educational settings are complex and heterogeneous. First, the largest part of the empirical data have a multilevel structure (pupils are nested within classes, classes within schools, schools within regions and school types, etc.). It is necessary to include this highly multilevel data structure into the empirical analysis to obtain unbiased estimates (Raudenbush and Bryk, 2002). This can be done by the use of a so called 'hierarchical' or 'mixed' model. This implies that the intercept - and in some models also the slopes - is allowed to randomly vary between groups. To estimate the effects of school-level institutional factors and family background on student achievement, a multilevel regression analysis is carried out where covariates are distributed at two levels: the students and schools. In an educational setting, unobserved school effects are expected from school-level disparities in e.g. the unobserved academic culture of school staff. As students are clustered in different schools, the assumption of independent noise is violated. It is thus necessary to include random

school effects into the empirical analysis to obtain unbiased estimates. However, to include random effects, we need to make the assumption that the unobserved school effects are not correlated with the included school-level variables. We expect that this assumption holds as we include a wide variety of rich school characteristics.

Second, as result of the complex, heterogeneous nature of the data structure, imposing parametric assumptions on the relationship between educational inputs and output can lead to biased estimates if there is misspecification. As it is unclear how all variables affect educational performance, it is advisable to use a more flexible approach. Nonlinearities can be addressed in different ways. First, polynomial expansions can be considered. This would be easy to implement, but the risk of introducing multicollinearity is very high. Second, nonparametric approaches can be considered. Fully non-parametric approaches do not impose parametric assumptions on the functional form, but imply the so called ‘curse of dimensionality’- i.e. including a large amount of regressors dramatically slows down convergence speed - and involves practical difficulties to include random effects. To avoid the ‘curse of dimensionality’, we use a semiparametric additive mixed model approach. We define pupil test scores of pupil i (with $i = 1, \dots, n$) in school j (with $j = 1, \dots, m$) as a function of socio-economic, institutional predictors and unobserved determinants such as innate ability and random noise at the pupil level $\epsilon_{i,j}$. To allow for hierarchically clustered noise, we define θ_j as the random effect of school j . Almost all schools come from different school governing bodies in our sample. There are only two governing bodies that have two schools in the sample. Therefore, we did not include random or fixed effects at the level of the governing body. As expected, there is no variation in school autonomy in budgeting issues within these two governing bodies. Hindriks et al. (2010) show the large within-school, between educational track differences in the social composition of students. To control for this social segregation by school track (general, technical-arts and vocational education), we include an aggregate of the economic and socio-cultural status index, as explained in more detail in section 5 at the ‘*sub-school*’ level. A sub-school is defined as a unit that provides either general, technical-arts, or vocational education. When a school

provides both general and technical or arts education, then the school is treated as two separate (sub-)schools. To capture the within school variation between educational tracks, we include the mean Economic and Socio-Cultural Status (ESCS) within a sub-school and dummies for the educational tracks. Besides the sub-school average ESCS, all included variables are at the school or pupil level. The semiparametric varying-intercept model is defined as:

$$\begin{aligned}
\text{PISA test score}_{i,j} = & \beta_0 + \beta_1 \text{School staff autonomy}_j \\
& + \sum_{p=2}^{p=k} \beta_p \text{Student characteristic (categorical)}_{p,i,j} \\
& + \sum_{q=k+1}^{q=k+l} \beta_q \text{School characteristic (categorical)}_{q,j} \\
& + \sum_{g=1}^{g=m} s_g \text{School characteristic (continuous)}_{g,j} \\
& + s_h \text{ESCS}_{i,j} \\
& + s_r \text{Sub-school ESCS}_{t,j} \\
& + \theta_j + \epsilon_{i,j},
\end{aligned} \tag{1}$$

where β_f , with $f = 1, \dots, k+l$ are the *fixed parameters* of the categorical variables related to school staff autonomy, student characteristics and school characteristics and s_g , with $g = 1, \dots, m$ and s_h are the *smooth functions* for respectively the additive continuous variables and the pupil-level Economic and Socio-Cultural Status (ESCS) index. s_r is the smooth function for the mean of ESCS at sub-school level (with school track t).⁶

Semiparametric regressions can be estimated by the use of kernel weights or by using piecewise polynomial functions - splines. Each approach has its own merits and drawbacks in a particular setting. We opt for a spline based approach as it is less cumbersome to use

⁶Multivariate smooths can also be introduced, but are not used in this analysis as multivariate tensor products of B-splines (see further) imply a dramatic loss of degrees of freedom.

with large datasets and allows the inclusion of random effects. In particular, to smooth the continuous variables, we opt for the penalized splines (P-splines) approach of Eilers and Marx (1996), discussed in detail in Verschelde et al. (2012).

The interest is not in the control variables per se. Therefore, if we find a non-linear effect of a smoothed variable, we only include information on the direction of influence. If the semiparametric model pinpoints towards a linear relationship between educational performance and a specific continuous variable, we drop the smooth term and include the variable parametrically.

5 Data

We discuss in this section the used data, the dependent variable, the unique within school type variation in budgeting autonomy and the controls.

5.1 PISA 2006

We focus on the educational setting in Flanders, as its specific characteristics of education allow for a within school type analysis of school autonomy as we show below.⁷ We use the PISA 2006 dataset, for which we were able to group the schools in the four main school types in Flanders (as discussed, education organized by cities, provinces, the Flemish Community and publicly funded, privately operating entities (mainly catholic education)), whereas PISA only groups the schools in ‘public’ and ‘publicly funded, privately operating schools’. PISA Flanders prohibits comparison between public school types (i.e., education organized by cities, provinces, the Flemish Community). Therefore, we include the four main school types as control variables in the empirical analysis, but only provide summary statistics for public school types at the aggregated level ‘public education’.

⁷Since 1989, Belgian education is organized by the Flemish community, the French-speaking community and the German-speaking community. Hirtt (2007) argues that Flanders has the most effective accountability system of the three communities.

In 2006, the PISA survey was implemented in 57 countries. The main focus of PISA 2006 is on science, however all pupils are also requested to complete a standardized test on math, science and reading and fill out a survey with questions related to their family background, views on issues related to science, the environment, careers, learning time and teaching and learning approaches of science.

Tests are typically constructed to assess between 4,500 and 10,000 students of age 15 in each country. To sample the target population of 15-year old pupils that are at least in grade 7, PISA 2006 has implemented a two-stage stratified sample design. In stage 1, for each stratum⁸, schools are sampled proportionally to size from a list of schools in the region (PPS sampling). The target was 150 schools in each region. In stage 2, 35 pupils are randomly drawn from a list of 15-year old pupils in the school.⁹ Final student weights are constructed to correct for varying selection probabilities of the students.¹⁰ In PISA 2006 the plausible value approach is used to estimate the pupil performance in respectively mathematics, science and reading literacy. Plausible values are multiple random draws from the unobservable latent student achievement and cannot be aggregated at pupil level (OECD, 2005). Therefore, in what follows, we use the first plausible value component to estimate educational outcomes in math, science and reading at pupil level.¹¹ A Balanced Repeated Replication (BRR) procedure with 80 replication estimates - described in OECD (2005)- is used to construct standard errors and to account for sampling variation (OECD, 2009).

Pupils in special education or part-time education are dropped from the sample. Pupils in private-funded schools or with missing values for some variables are also dropped from the sample. In addition, we do not take pupils in schools with less than four pupils per teacher

⁸A group of schools, formed to improve the precision of sample based estimates.

⁹If the school has less than 35 pupils, all pupils are included in the sample.

¹⁰This occurs because certain subgroups that are over- or under-sampled, the information of school size at the time is not completely correct, school non-response, student non-response and the inclusion of trimming weights to ensure stable estimates.(OECD, 2009)

¹¹As plausible values are random draws, the choice to take the first plausible value is arbitrary. Other plausible values could be taken as well.

into account as we expect that these schools have among others a different educational approach.¹² By this, the sample is reduced to 3603 observations. As discussed in section 4, a sub-school is defined as a unit that provides either general, technical-arts, or vocational education. The sample consists of 126 schools and 245 sub-schools.

5.2 Student performance

Table 1 shows descriptive statistics of the educational achievement of pupils in Flanders. Standardized test scores for math are high in Flanders (PISA average is 500). But the high standard deviation of educational outcomes indicates that the inequality in individual test scores is considerable. On average, pupils in privately operating schools have higher PISA scores than the pupils in public education.

<Table 1 about here >

5.3 School staff autonomy

School autonomy is a rather vague concept. In our study we will explicitly focus on autonomy of principals and teachers, using the data made available by PISA (2006). The PISA dataset among other things looks specifically into the roles that principals and teachers might play in educational decision-making and contains measures of centralization and decentralization for these different functions.

Table 2 illustrates how the level of school autonomy varies item-by-item in Flanders. In particular, the principal is asked who has the main responsibility for any specific item. The principal can tick multiple levels if there is joint decision making on a particular item. In line with Eurydice (2007) and Eurydice (2008), the PISA 2006 data show that Flanders is

¹²However, sensitivity tests, available on request, show that our results still hold when we do consider these schools with very few pupils per teacher as comparable and take them into the analysis.

characterized by considerable school (group) autonomy in staffing, budget allocation and formation, assessment and discipline of pupils and textbook choice. Neither schools, nor intermediate government institutions have the autonomy to set the salaries of teaching and non-teaching staff. Although selection of students by schools is restricted to avoid exclusion of minority groups, the school's principals remains largely in charge of approving the admission of pupils to the school. As in any survey, there are clear cases of mis-reporting (e.g., 1 out of 126 schools reports that it has autonomy to set salaries). Still, there is noteworthy meaningful variation in school staff autonomy for the items on firing teachers, budget allocation, budget formation, course content and courses offered.

We only consider in our analysis the variation in budget formation and allocation. It is for these items that the school governing bodies have not equally decentralized decision making to the school staff. Since the autonomy in budget formation is rather limited (as it needs to be in line with the centrally imposed funding system), our main focus is on budget allocation. We defined school autonomy in budget allocation (formation) as a binary indicator, equal to one if the school principal ticks that he/she or the teachers have a considerable responsibility in the task of budget allocation (formation), zero otherwise. The reviewed theoretical studies do not pinpoint perverse effects of joint-decision making, i.e. when the school governing board and the school principal have considerable responsibility. Therefore, we did not focus on a strict definition of school autonomy that excludes joint-decision making. Table 3 shows that over 80 percent of the sampled schools in both public and private education have considerable autonomy on budget allocation. Less than 40 percent of the sampled schools do not share this responsibility with higher levels. However, as shown further, all results hold when we alter the definition of school autonomy.

We did not focus on autonomy in curricular content and courses offered as we expect interpretation issues for these items. As discussed in the Introduction, the Flemish government sets the final standards, but schools have considerable autonomy in how to reach the final standards. In practice, this amounts to a centrally imposed programme of basic courses, and considerable autonomy in the curricular content of optional subjects. As the final stan-

dards are detailed and well defined (see Hirtt (2007)), the principals can either interpret they have considerable autonomy (in reaching the posed final standards) or little autonomy (as there is little room to teach other things than the centrally imposed programme). As a result, we note that principals tick both themselves and/or the educational authorities as decision makers in course content and courses offered.

In addition, we did not include autonomy in firing teachers into the analysis as we don't believe school actors can actually make use of the large authority to fire teachers, as regulations strongly limit the possibility to fire a teacher, unless a serious fault is established.

<Table 2 about here >

<Table 3 about here >

5.4 Control variables

Student characteristics We control for student-level heterogeneity by including family background, gender, education track, grade retention and the weekly followed lessons of the test subject in our model (see Table 5 and 6 in Verschelde et al. (2012)). To relate variation in outcomes to family background, we consider 2 socio-economic variables: socio-economic status and migration status. Family socio-economic status is estimated by PISA 2006 as a composite index of the Economic and Socio-Cultural Status (ESCS) of a pupil, derived from (1) the highest occupational status of each student's parents, (2) their highest educational level, and (3) a summary measure of household possessions. For the sampled students of all participating OECD countries, the mean is 0 and standard deviation is 1. The ESCS score shows substantial variation across pupils in Flanders.

For migration status, three proxies are used. First-generation immigrants and second-generation immigrants are respectively defined as pupils that are not born in Belgium and pupils that are born in Belgium, but are children of immigrants. Pupils that are first- or second-generation immigrant and do not speak an official Belgian language at home are grouped in a third variable. The proportion of non-native pupils is around 5 percent. 2

percent of the pupils in the sample do not speak an official Belgian language at home.

Pupils are tracked in the first year of secondary education in either general, technical-arts or vocational education based on academic records. In our filtered sample, 50 percent of pupils are in general education (high track), 33 percent are in technical-arts education (middle track) and 17 percent in vocational education (low track). If a pupil has not reached the basic skills determined by the posed final standards in a school year, grade repetition and re-orientation to lower tracks are used. In our sample, 21 percent of pupils are lagging behind.

School characteristics To control for school-level heterogeneity, we include controls for variation in educational resources, teacher shortage, class and school size, school type, social segregation, selectivity by schools and urbanization (see Table 5 and 6 in Verschelde et al. (2012)).

1. *Educational resources.* Schools receive funding and ‘teaching hours’ according to the number of pupils. Schools with more disadvantaged pupils receive additional resources (‘GOK policy’). On average, schools have a modest lack of educational resources (e.g. instructional material, labs) (the average is above the PISA 2006 average of 0).
2. *Teacher shortage.* In line with Rivkin et al. (2005) and Kane et al. (2006), a preliminary analysis showed that the relation between formal teacher quality and educational performance is not significant. As almost all teachers in Flanders are certified, we dropped this variable. In contrast, shortage in educational personnel can have severe negative effects on the true teacher quality and teaching process in a school. Therefore, a negative sign is expected. For math and science, there is a systematic shortage in teachers.¹³ Therefore, we focus on shortage of adequate teaching staff in

¹³In a preliminary analysis, we did not find effects of shortages in math, science or language teachers.

areas other than math, science and reading.

3. *Class and school size.* The effect of school and class size on performance is difficult to disentangle from the selection effect (i.e. parents choose better schools that are consequently of larger size). Therefore, the school size and student-teacher ratio are included as control variables to control for heterogeneity in class and school size.
4. *School type.* We control for the 4 main school types in Flanders: Flemish community schools, city schools, provincial schools and subsidized private-operating schools. As discussed, private-granted schools are only a negligible proportion of the school population.
5. *Social segregation.* When schools organize different tracks, each school track is considered as a distinct sub-school. There is considerable variation in sub-school average ESCS, indicating social segregation between (sub-)schools. A quarter of the students are in sub-schools with an average ESCS below or equal to the OECD student average of 0, while the most elite sub-school groups students with an ESCS which is on average 64 percent higher than the OECD student standard deviation. As shown in Hindriks et al. (2010), there is less social segregation between school types than between school tracks.
6. *Selectivity by schools.* Selection by schools is officially not allowed within a track. However, the PISA data show that selection on academic record or recommendation is used in 81 schools of the sampled 126 schools.
7. *Urbanization.* 60 percent of the sampled pupils receive education in a town with 15,000 up to 100,000 inhabitants.

Variables related to school competition could be included. In Hanushek and Luque (2003), a significant positive effect is found of competition of private schools. Hoxby (2000) finds that Tiebout choice leads to better school performance in the US.¹⁴ However, as we did not

¹⁴In Hoxby (2003), an overview of the economics of school choice can be found.

find any effect of a proxy for the number of competing schools in a preliminary analysis, we dropped this variable to reduce the number of missing values.

Overall, we try to obtain insight into the relation between school autonomy -adequately measured - and educational performance, while controlling substantively for heterogeneity in student composition and institutional settings between schools.

6 Empirical results

Table 4 shows six different models in which we measure the effect of autonomy of the school's staff in budget allocation on its educational performance in math. Flanders is in this subject persistently ranked as a top performer (see e.g. De Meyer and Warlop (2010)).

In a first model (i.e., Math I, first column of Table 4), we only control for student-level heterogeneity, school type and urbanization. Model Math I explains 55 percent of the variation of educational performance between pupils. We find a clear positive effect of the autonomy in budget allocation on educational performance. Math performance is 13 points higher in schools with considerable autonomy in budgeting. As in all the following models, we do not find significant interaction between the effect of autonomy in budgeting and respectively school type or the average social position of pupils in a sub-school.

If we control additionally for school-level heterogeneity - by including teacher shortage, selectivity of the school, sub-school average ESCS, school educational resources, student-teacher ratio and school size- the relation of interest is still significantly positive, but is estimated to be only 6.73 points.

<Table 4 about here >

In Model Math III, we include random school effects to control for unobserved school-level heterogeneity. By taking this unobserved random school level heterogeneity into account, the estimates indicate it is optimal to employ no smoothing; we return to a fully parametric model. Results still hold.

Our main results are documented in Model Math IV. As discussed, the PISA dataset groups pupils in ‘Public schools’ and ‘Publicly financed, privately operating schools’ to control for school type. However, as there are 3 public school types in Flanders (i.e., education organized by cities, provinces and the Flemish Community), we use anonymous data to control for the school type heterogeneity within public education. The model controls for a wide variety of family background and rich-school characteristics. In line with Hindriks et al. (2010), we find that the largest part of the variation of rich school characteristics is captured by the educational tracks and school types, resulting in an insignificant effect of the average sub-school ESCS after controlling for school type and educational tracks. Additionally, in line with Rivkin et al. (2005) and Kane et al. (2006), we find a significant negative relation between shortage of adequate teaching staff and educational performance. Model Math IV shows that the finding of a positive relation between educational performance and autonomy in budget allocation is not driven by school type differences. In particular, we find that variation in budgeting autonomy explains 0.1 standard deviation of pupil performance in math. The positive effects of school autonomy (a better match with local preferences, yardstick competition, local information advantage, higher intrinsic motivation) dominate the potential negative effects (principal-agent problem, uncoordinated rent-seeking or corruption, spillover costs). This finding supports the cross-country and between school type evidence for a positive effect of a combination of monitoring and school autonomy on educational achievement (referenced in Section 1) by separating school staff autonomy from school type specific effects.¹⁵

Overall, we extensively controlled for the heterogeneity in student population and institutions between schools. As we use cross-sectional pupil level data, estimates can be sensitive for influences of reverse causality, simultaneity and unobserved selection. In section 3, we argued we do not expect our results are substantially affected by these issues. Additional sensitivity tests supports that autonomy per se matters for educational performance. Model

¹⁵Results available upon request show that the findings are not sensitive for excluding the two smaller public education networks from the analysis (city and provincial schools).

Math V shows that the significant positive relation between educational performance and school autonomy in budgeting also holds for the subsample of schools in which there is no evaluation of the school's principals, based on achievement data.

Further, if we include the proxies for school dynamism (separately or grouped), documented in Table 8 in Verschelde et al. (2012), into our analysis, results for school autonomy do not change.¹⁶ In particular, Model Math VI shows both a significant effect of two proxies for school dynamism (i.e., extracurricular science projects and lectures and/or seminars on environmental topics) and school autonomy in budget allocation. No interactions were found. The effect found of school autonomy in budgeting is only significant at the 10 percent level when we consider the first plausible value of math as dependent variable. However, if we use the four other plausible values, the effect is significant at the 5 percent level, indicating a robust significant positive effect of autonomy in budget allocation on math performance (see Table 6 in Appendix).

As discussed, autonomy in budget formation is limited to additional funding above a centrally imposed funding scheme. Table 5 shows a non-robust positive effect of school autonomy in budget formation. Once we control for the four main school types in Flanders, we do not find a (robust) significant effect of school staff empowerment in budget formation on math performance.

Finally, we test the sensitivity of our findings for test scores in different subjects and different definitions of school autonomy. We ran Model VI for science and reading (see Table 5). For science, we find for both autonomy in budget allocation and budget formation a robust positive effect. As the additional budgets are very small in comparison to the school budget, we expect that the effect of budget formation is not mainly driven by an informational advantage of the agents. It is more plausible that higher perceived budget formation autonomy motivates the school staff (see Frey (1993)), with better education quality as result. For reading, we find the relation to be insignificant for both budget allocation and

¹⁶Results available on request.

budget formation. A possible explanation is that autonomy in budgeting is not used to promote reading literacy. Another interpretation is that for reading, the unobserved heterogeneity is larger. For reading, only 52 percent of variation can be explained by Model Read VI, whereas Math VI and Science VI explain respectively 57 and 58 percent.

As discussed above, we defined school autonomy in budget allocation (formation) as a binary indicator, equal to one if the school principal ticks that school direction or teachers have a considerable responsibility in the task of budget allocation (formation), zero otherwise. If we exclude joint-decision making (with the school governing board or higher levels) from the definition of school autonomy, results still hold.¹⁷

We have tested the robustness of the findings for different model specification. We have regressed a fully parametric model with random school effects, included variation in probability weights and clustering within strata. In addition, we have tested the effect on the results of including the ‘dropped’ variables. The findings remain robust to all such specification changes. Results are available upon request.

To test for a possible unequal effect of autonomy on educational performance, we compared the effects at the bottom and the top of test scores distribution by a quantile regression approach. Results, documented by Figure 2 and 3 in Verschelde et al. (2012), show no indications that school staff autonomy affects top and low performers differently.

<Table 5 about here >

7 Conclusion

In this paper we have tested whether a combination of accountability and autonomy in education provision can improve educational performance as measured by PISA test scores.

¹⁷Results available upon request. We find no indications that joint-decision making is less or more effective than strict autonomy. The coefficient of school autonomy is not significantly different, but results are statistically more significant. This is probably due to the fact that there is more variation in autonomy when using the strict definition.

There is ample cross-country and cross-school type (e.g. charter school literature) evidence for this. But these studies cannot separate the effects of school autonomy from (time-varying) country-specific or school type specific effects.

We test this conjecture on the PISA-dataset for Flanders, where there is substantial variation within school-type in autonomy and other institutional settings. In Flanders, the government delegates a lot of budgeting responsibilities to non-profit school governing bodies that can group multiple schools. There is variation in the extent this budgeting authority is further delegated to the school staff (the school's direction and teachers). The reviewed qualitative research shows that this variation in school autonomy mainly originates from approximately random school governing board characteristics. Higher school staff empowerment should lead to better use of local information and better performance, if the central government can align incentives properly. In Flanders, there are no central examinations, but inspection teams investigate on a regular basis whether the curriculum and teaching process are aimed at reaching the centrally imposed final standards, and whether budget formation is in keeping with the posed requirements. In addition, freedom in budget formation is limited to additional funding, above the centrally imposed funding system. This ensures that information asymmetries are not misused by local staff and so that autonomy should improve performance. Therefore Flanders is a very good testing ground for the theory that the institution of school autonomy, through improved incentives for schools and teachers, will affect resource-allocation decisions and ultimately the educational performance of students positively.

Our findings support this hypothesis. While including a large set of student-level and school-level controls, we find indeed that local staff empowerment clearly and significantly boosts educational outcomes. Results are robust for controlling for reverse causality or variation in school dynamism. Overall, we confirm cross-country evidence and within-country between school type evidence for a positive effect of school (staff) autonomy on educational performance by separating the effect of school staff autonomy from school type specific effects.

From a policy perspective, this study shows first of all that even in countries with large school autonomy, there is room for optimizing the delegation of powers as not all autonomy trickles down to the lowest level, the principals and teachers. Second, the results support policy oriented towards delegation of operational responsibilities to the school staff, as it are these local agents which have the information advantage, can respond to local preferences and are expected to have high returns of intrinsic motivation because of their direct contact with students. However, an effective accountability system is required for the positive effects to dominate the negative effects. The expected effects of delegation of responsibilities are thus country-specific. The results are most of all supportive for large operational autonomy for school staff in countries that have an effective central examination system at place.

References

- Abdulkadiroglu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., 2011. Accountability and flexibility in public schools: evidence from Boston's charters and pilots. *Quarterly Journal of Economics* 126 (4), 2133–2134.
- Ammermüller, A., 2004. PISA: What makes the difference? explaining the gap in PISA test scores between Finland and Germany. ZEW Discussion Papers 04-04, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.
- Besley, T., Case, A., 1995. Incumbent behaviour: vote-seeking, tax-setting, and yardstick competition. *American Economic Review* 85 (2), 25–45.
- Bottani, N., Favre, B., 2001. School autonomy and evaluation. *Prospects - Quarterly Review of Comparative Education* 31 (4), 166, issue 120.
- Caenepeel, L., 1988. Scholen besturen. Samenstelling, structuur en functioneren van inrichtende machten van katholieke scholen. Leuven: Acco.

- Card, D., Dooley, M., Payne, A., 2010. School competition and efficiency with publicly funded catholic schools. Canadian Labour Market and Skills Researcher Network, Working Paper 66, 61.
- Chubb, J., Moe, T., 1990. Politics, markets, and America's schools. Brookings Institution paper, 318.
- Clark, D., 2009. The performance and competitive effects of school autonomy. *Journal of Political Economy* 117 (4), 745–783.
- Day, C., Moller, J., Nusche, D., Pont, B., 2008. Chapter 6. the flemish (belgian) approach to system leadership. In: *Improving School Leadership VOLUME 2: Case studies on system leadership*. OECD Publishing.
- De Meyer, I., Warlop, N., 2010. PISA - Leesvaardigheid van 15-jarigen in Vlaanderen. De eerste resultaten van PISA 2009.
- Deci, E., Ryan, R. M., 1985. *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Devos, G., Tuytens, M., 2006. *Improving school leadership - OECD Review. Background report for Flanders*. Tech. rep., Improving School Leadership Activity Education and Training Policy Division.
URL <http://www.oecd.org/edu/schoolleadership>
- Devos, G., Vandenberghe, P., Verhoeven, J., 1989. *Schoolbeleid; mogelijkheden en grenzen; een empirisch onderzoek*. Leuven: Acco.
- Devos, G., Verhoeven, J., Beuselinck, I., Van den Broeck, H., Vandenberghe, R., 1999. *De rol van schoolbesturen in het schoolmanagement*. Garant, , Leuven/Apeldoorn.
- Eilers, P., Marx, B., 1996. Flexible smoothing with B-splines and penalties. *Statistical Science* 11, 89–121.

- Eurydice, 2007. School autonomy in Europe. Policies and Measures. Eurydice, ISBN 978-92-79-07522-3.
- Eurydice, 2008. Levels of autonomy and responsibilities of teachers in Europe. Eurydice.
- Fan, C., Lin, C., Treisman, D., 2009. Political decentralization and corruption: Evidence from around the world. *Journal of Public Economics* 93 (1-2), 14–34.
- Frey, B., 1993. Does monitoring increase work effort - The rivalry with trust and loyalty. *Economic Inquiry* 31 (4), 663–670.
- Frey, B., 1994. How intrinsic motivation is crowded out and in. *Rationality and Society* 6 (3), 334–352.
- Frey, B. S., Oberholzer-Gee, F., 1997. The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review* 87 (4), 746–755.
- Hallinger, P., Bickman, L., Davis, K., 1996. School context, principal leadership, and student reading achievement. *Elementary School Journal* 96 (5), 527–549.
- Hanushek, E., Link, S., Wößmann, L., 2011. Does school autonomy make sense everywhere? Panel Estimates from PISA. CESifo Working Paper, No. 3648.
- Hanushek, E. A., 2003. The failure of input-based schooling policies. *Economic Journal* 113 (485), F64–F98.
- Hanushek, E. A., Luque, J. A., 2003. Efficiency and equity in schools around the world. *Economics of Education Review* 22 (5), 481–502.
- Hindriks, J., Myles, G., 2006. *Intermediate Public economics*. MIT Press.
- Hindriks, J., Verschelde, M., Rayp, G., Schoors, K., 2010. School tracking, social segregation and educational opportunity: Evidence from Belgium. CORE Discussion Paper 2010/81.

- Hirtt, N., 2007. Pourquoi les performances PISA des élèves francophones et flamands sont-elles si différentes ? Appel pour une école démocratique, 27.
URL http://www.skolo.org/IMG/pdf/PISA_F.pdf
- Hoxby, C., 1999. The productivity of schools and other local public goods producers. *Journal of Public Economics* 74 (1), 1–30.
- Hoxby, C., 2003. *The economics of school choice*. University of Chicago Press.
- Hoxby, C. M., 2000. Does competition among public schools benefit students and taxpayers? *American Economic Review* 90 (5), 1209–1238.
- Kane, T., Rockoff, J., Staiger, D., 2006. What does certification tell us about teacher effectiveness? Evidence from New York City. National Bureau of Economic Research Working Paper (12155).
- Niskanen, W., 1971. *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton.
- Niskanen, W., 1991. The budget-maximizing bureaucrat: Appraisals and evidence. In: Blais, A., Dion, S. (Eds.), *A reflection on Bureaucracy and Representative Government*. pp. 13–31.
- OECD, 2005. *PISA 2003 Data Analysis Manual: SAS users*. OECD Programme for International Student Assessment.
- OECD, 2006. *PISA 2006: Science Competencies for Tomorrow's World*. OECD Programme for International Student Assessment.
- OECD, 2008a. *Improving School Leadership, Volume 2: Case Studies on System Leadership*. OECD Publishing, Paris.
- OECD, 2008b. *Improving School Leadership, Volume 1: Policy and Practice*. OECD Publishing, Paris.

- OECD, 2009. PISA 2006 Technical report. OECD Programme for International Student Assessment.
- OECD, 2011. School Evaluation in the Flemish Community of Belgium 2011, OECD Reviews of Evaluation and Assessment in Education. OECD Publishing.
URL <http://dx.doi.org/10.1787/9789264116726-en>
- Poesen-Vandeputte, M., Bollens, J., 2008. Studiekosten in het secundair onderwijs. Wat het aan ouders kost om schoolgaande kinderen te hebben. Leuven: Katholieke Universiteit Leuven. Hoger instituut voor de arbeid.
- Raudenbush, S. W., Bryk, A. S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage.
- Rivkin, S. G., Hanushek, E. A., Kain, J. F., 2005. Teachers, schools, and academic achievement. *Econometrica* 73 (2), 417–458.
- Shleifer, A., 1985. A theory of yardstick competition. *RAND Journal of Economics* 16 (3), 319–327.
- Tiebout, C., 1956. A pure theory of local public expenditures. *Journal of Political Economy* 64, 416–424.
- Van Petegem, P., Devos, G., Mahieu, P., Dang Kim, T., Warmoes, V., 2009. Hoe sterk is mijn school? Het beleidsvoerend vermogen van Vlaamse scholen. Plantyn, Mechelen.
- Vershelde, M., Hindriks, J., Rayp, G., Schoors, K., 2012. School staff autonomy and educational performance: within school type evidence. *ECORE Discussion Paper* 2012/111.
- Wößmann, L., 2003. Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics* 65 (2), 117–170.
- Wößmann, L., 2008. Efficiency and equity of European education and training policies. *International Tax and Public Finance* 15 (2), 199–230.

Wößmann, L., Lüdemann, E., Schütz, G., West, M., 2007. School accountability, autonomy, choice, and the level of student achievement: International evidence from PISA 2003. Unclassified OECD Education Working Paper 13, 1–85.

Appendix

<Table 6 about here >

Tables and Figures

Table 1: Pupil performance

Variable	Mean	St.Dev.
PISA 2006 Performance in math	560.3	(87.3)
PISA 2006 Performance in math, private education	573.3	(83.3)
PISA 2006 Performance in math, public education	522.5	(87.7)

Note: A SAS procedure for a Balanced Repeated Replication procedure with 80 replication estimates and 5 plausible values for each subject, described in OECD (2005), is used to construct the mean and standard error.

Table 2: School-level variation in perceived school autonomy, privately and publicly operating schools between brackets. 126 schools whereof 92 privately operating and 34 publicly operating (organized by either the Flemish Community (24), provinces (5) or cities (5))

Who has a considerable responsibility for the following tasks? (multiple ticks are allowed)	School's direction or teachers	Non-profit school governing body	Regional or local education authorities	National education authorities
Selecting teachers for hire	120 (90 - 30)	51 (34 - 17)	4 (0 - 4)	3 (0 - 3)
Firing teachers	101 (77 - 24)	87 (64 - 23)	1 (1 - 0)	2 (1 - 1)
Establishing teachers' starting salaries	1 (1 - 0)	1 (0 - 1)	20 (14 - 6)	104 (76 - 28)
Determining teachers' salaries increases	1 (1 - 0)	1 (0 - 1)	19 (13 - 6)	106 (77 - 29)
Formulating the school budget	110 (81 - 29)	88 (64 - 24)	7 (1 - 6)	9 (6 - 3)
Deciding on budget allocations within the school	107 (80 - 27)	72 (53 - 19)	4 (1 - 3)	10 (6 - 4)
Establishing student disciplinary policies	125 (91 - 34)	33 (21 - 12)	6 (3 - 3)	9 (6 - 3)
Establishing student assessment policies	121 (89 - 32)	31 (19 - 12)	10 (3 - 7)	10 (8 - 2)
Approving students for admission to the school	117 (86 - 31)	16 (11 - 5)	11 (6 - 5)	24 (18 - 6)
Choosing which textbooks are used	125 (92 - 33)	6 (3 - 3)	1 (0 - 1)	2 (1 - 1)
Determining course content	88 (67 - 21)	7 (2 - 5)	28 (17 - 11)	64 (47 - 17)
Deciding which courses are offered	108 (80 - 28)	37 (26 - 11)	31 (16 - 15)	58 (46 - 12)

Table 3: Summary statistics on school autonomy

Variables	Students (proportions)	Schools	Private	Public
Autonomy in budget allocation	0.85	107	80	27
- for school direction, not school board	0.40	49	35	14
- for school direction, not any other higher level	0.38	47	34	13
Autonomy in budget formation	0.88	110	81	29
- for school direction, not school board	0.28	36	26	10
- for school direction, not any other higher level	0.27	35	25	10
Number of observations	3603	126	92	34

Table 4: Effect of school autonomy in budget allocation, full model

Model	Math I	Math II	Math III	Math IV	Math V	Math VI
Plausible Value	PV1	PV1	PV1	PV1	PV1	PV1
(Intercept)	362.26*** (11.71)	362.62*** (13.45)	332.30*** (24.69)	327.36*** (23.91)	320.14*** (23.98)	310.20*** (23.29)
Autonomy in budget allocation	13.43*** (2.80)	6.73* (2.92)	10.68* (5.36)	9.78 ^o (5.25)	10.84 ^o (5.59)	10.45 ^o (5.38)
Privately operating school	18.48*** (2.40)	12.44*** (2.69)	14.78** (4.72)			
General education track	117.06*** (3.57)	107.68*** (4.72)	103.07*** (5.54)	103.90*** (5.50)	100.99*** (5.74)	99.34*** (5.72)
Technical-arts education track	65.90*** (3.28)	66.22*** (3.49)	65.74*** (3.64)	65.57*** (3.62)	63.93*** (3.79)	63.48*** (3.78)
Gender (female=1)	-20.81*** (1.97)	-20.34*** (1.97)	-18.47*** (2.07)	-18.19*** (2.07)	-19.06*** (2.15)	-19.19*** (2.15)
Small town	54.89*** (10.12)	74.50*** (11.75)	69.95*** (20.73)	68.55*** (19.80)	72.16*** (19.62)	80.29*** (18.97)
Town	47.30*** (10.05)	67.37*** (11.91)	61.75** (20.66)	63.29** (19.74)	66.03*** (19.57)	73.34*** (18.93)
City	43.63*** (10.50)	61.43*** (12.36)	56.00** (21.32)	54.48** (20.36)	57.71** (20.16)	64.32** (19.62)
Lessons test subject, less than 2 hours	2.85 (5.45)	1.13 (5.42)	-2.12 (5.36)	-2.50 (5.36)	-1.22 (55.11)	-1.49 (5.51)
Lessons test subject, 2 up to 4 hours	3.95 (5.51)	1.56 (5.48)	-2.98 (5.41)	-3.34 (5.41)	-3.05 (55.62)	-3.33 (5.56)
Lessons test subject, 4 up to 6 hours	27.00*** (5.68)	24.82*** (5.64)	20.20*** (5.55)	19.83*** (5.55)	21.23*** (56.98)	21.11*** (5.70)
Lessons test subject, 6 or more hours	33.66*** (8.02)	31.71*** (7.94)	24.78** (7.80)	24.33** (7.80)	23.93** (7.97)	24.56** (8.00)
First-generation immigrant	-21.42** (7.02)	-17.51* (7.00)	-10.76 (6.96)	-10.86 (6.95)	-12.06 ^o (7.09)	-11.56 (7.09)
Second-generation immigrant	-19.36** (7.25)	-21.51** (7.47)	-15.50* (7.34)	-15.24* (7.33)	-25.87** (8.26)	-26.46** (8.25)
Immigrant, no off. Belgian language at home	-15.84 ^o (8.86)	-16.36 ^o (8.89)	-18.46* (8.70)	-18.77* (8.70)	-14.44 (9.02)	-13.96 (9.02)
Not lagging behind	50.14*** (2.61)	49.34*** (2.59)	47.51*** (2.54)	47.41*** (2.54)	47.99*** (2.62)	47.62*** (2.63)
ESCS	6.05*** (1.27)	Pos.***	5.40*** (1.27)	5.40*** (1.27)	5.55*** (1.32)	5.51*** (1.32)
Teacher shortage		-12.09*** (2.68)	-9.37 ^o (4.81)	-9.65* (4.69)	-12.51* (4.95)	-11.92* (4.78)
Selectivity of school		1.64 (2.19)	4.73 (4.01)	6.21 (3.936)	7.04 ^o (4.07)	5.72 (3.94)
Average ESCS (sub-)school		2.33 (3.83)	2.14 (5.06)	2.21 (4.97)	3.21 (5.05)	4.47 (5.00)
School educational resources		Neg.***	-2.73 (2.21)	Neg.	-4.57* (2.27)	-3.70 ^o (2.20)
Student-teacher ratio		Pos.***	3.28*** (0.96)	3.07*** (0.92)	3.44*** (0.94)	3.09*** (0.91)
School size		Wiggly***	0.01 (0.01)	0.00 (0.01)	0.000 (0.01)	0.00 (0.01)
Extracurricular science projects (including research)						9.08* (3.79)
Lectures and/or seminars on environmental topics						8.86* (4.24)
Student-level control variables	Yes	Yes	Yes	Yes	Yes	Yes
School-level control variables	No	Yes	Yes	Yes	Yes	Yes
Random school effects	No	No	Yes	Yes	Yes	Yes
Detailed school type FE	No	No	No	Yes	Yes	Yes
Control for reverse causality	No	No	No	No	Yes	Yes
Control for school dynamism	No	No	No	No	No	Yes
R ² (adj.)	0.554	0.566	0.56	0.565	0.567	0.569
Obs.	3603	3603	3603	3603	3346	3318

Significance levels : o : 10% : * : 5% ** : 1% *** : 0.1%

Table 5: Effect of school autonomy in budgeting

Model	Math I	Math II	Math III	Math IV	Math V	Math VI	Science VI	Read VI
Plausible Value	PV1	PV1	PV1	PV1	PV1	PV1	PV1	PV1
Autonomy in budget allocation	13.43*** (2.80)	6.73* (2.92)	10.68* (5.36)	9.78° (5.25)	10.84° (5.59)	10.45° (5.38)	10.91* (4.78)	11.12 (6.97)
R^2 (adj.)	0.55	0.57	0.56	0.57	0.57	0.57	0.58	0.52
Autonomy in budget formation	8.09** (3.00)	7.26* (3.11)	10.21° (5.73)	7.32 (5.60)	7.63 (5.82)	7.02 (5.61)	14.62** (4.86)	10.77 (7.28)
R^2 (adj.)	0.55	0.57	0.56	0.56	0.57	0.57	0.58	0.52
Student-level control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School-level control variables	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Random school effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Detailed school type FE	No	No	No	Yes	Yes	Yes	Yes	Yes
Control for reverse causality	No	No	No	No	Yes	Yes	Yes	Yes
Control for school dynamism	No	No	No	No	No	Yes	Yes	Yes
Obs.	3603	3603	3603	3603	3346	3318	3318	3318

Significance levels : ◦ : 10% * : 5% ** : 1% *** : 0.1%

Table 6: Effect of school autonomy in budgeting, different plausible values

Model	Math VI	Math VI	Math VI	Math VI
Plausible Value	PV2	PV3	PV4	PV5
Autonomy in budget allocation	13.99*	11.72*	11.04*	11.04*
	(5.54)	(5.54)	(5.54)	(5.56)
R^2 (adj.)	0.57	0.57	0.57	0.58
Autonomy in budget formation	11.63*	8.59	9.34	8.07
	(5.80)	(5.79)	(5.77)	(5.81)
R^2 (adj.)	0.57	0.57	0.57	0.58
Model	Science VI	Science VI	Science VI	Science VI
Plausible Value	PV2	PV3	PV4	PV5
Autonomy in budget allocation	12.82**	10.72*	12.20*	11.00*
	(4.87)	(4.82)	(4.94)	(4.88)
R^2 (adj.)	0.58	0.58	0.57	0.59
Autonomy in budget formation	17.01***	13.90**	15.79**	15.60**
	(4.93)	(4.92)	(5.02)	(4.94)
R^2 (adj.)	0.58	0.58	0.58	0.59
Model	Read VI	Read VI	Read VI	Read VI
Plausible Value	PV2	PV3	PV4	PV5
Autonomy in budget allocation	8.74	8.67	10.77	9.79
	(6.81)	(6.49)	(6.87)	(6.92)
R^2 (adj.)	0.52	0.53	0.51	0.52
Autonomy in budget formation	10.69	6.73	9.80	10.98
	(7.05)	(6.77)	(7.17)	(7.18)
R^2 (adj.)	0.52	0.53	0.51	0.53
Student-level control variables	Yes	Yes	Yes	Yes
School-level control variables	Yes	Yes	Yes	Yes
Random school effects	Yes	Yes	Yes	Yes
Detailed school type FE	Yes	Yes	Yes	Yes
Control for reverse causality	Yes	Yes	Yes	Yes
Control for school dynamism	Yes	Yes	Yes	Yes
Obs.	3318	3318	3318	3318

Significance levels : ◦ : 10% : * : 5% ** : 1% *** : 0.1%