Estimating scale economies and the optimal size of school districts: A flexible form approach

BERJ British Educational Research Journal

BERA BRITISH EDUCATIONAL RESEARCH ASSOCIATION

British Educational Research Journal
Vol. 43, No. 6, December 2017, pp. 1048–1067

DOI: 10.1002/berj.3312

Estimating scale economies and the optimal size of school districts: A flexible form approach

Fritz Schiltz[a,⋆] and Kristof De Witte[a,b]
[a]*KU Leuven, Leuven, Belgium;* [b]*Maastricht University, Maastricht, the Netherlands*

This paper investigates estimation methods to model the relationship between school district size, costs per student and the organisation of school districts. We show that the assumptions on the functional form strongly affect the estimated scale economies and offer two possible solutions to allow for more flexibility in the estimation method. First, we introduce a model by adding higher-degree district size polynomials, allowing for multiple optima. Second, we develop a Fourier cost function, innovative in the literature on scale economies in education. We then compare both models to classical approaches in the literature. We illustrate how a minor change in the estimation method can alter policy conclusions significantly using Flemish school district data. In doing so, we find sizeable potential cost savings from the consolidation of school districts, especially at the lower tail of the district–size distribution. The organisational transition from small to large school districts is characterised by an interval between two optima. Beyond an apparent slowdown in cost savings in medium-sized school districts, cost savings from school district consolidation increase again, up to the optimal size of around 6,500 students. Beyond this optimum, school districts incur diseconomies of scale. The commonly used quadratic form ('U'-shaped cost function) overestimates scale economies, and fails to identify the interval between both optima.

Keywords: economies of scale; school district consolidation; costs per student; education economics; Fourier function

Introduction

In line with the growing importance of 'New Public Management' theories, the reorganisation and professionalisation in the public sector is receiving increased attention. This observed trend is spreading towards education (Jarl *et al.*, 2012). With respect to professionalisation, Bloom *et al.* (2015) argue that management practices are a crucial determinant of inter- and intra-country differences in educational outcomes. In their study, one of the main drivers of these management practices is the degree of accountability to an external governing body. This governing body corresponds to the school board in control of a set of schools: a 'school district'.[1] The school district has been shown to have a direct effect on student and financial outcomes (Bidwell & Kasarda, 1975; Heinesen, 2005; Saatcioglu *et al.*, 2011), or an indirect effect on student outcomes through teacher absenteeism (Theobald, 1990) and management practices (De Witte & Schiltz, 2017).

⋆Corresponding author. Katholieke Universiteit Leuven (KU Leuven), Leuven Economics of Education Research, Naamsestraat 69, 3000 Leuven, Belgium. E-mail: fritz.schiltz@kuleuven.be; Twitter: @f_schiltz

With respect to reorganisation, there is a tendency to enlarge the scale of operations in public sector entities (Alonso *et al.*, 2015). The available literature on scale economies in education focuses on three separate levels. First, class size effects are thoroughly studied, leading to a general agreement on significant advantages of smaller classes, especially with respect to students characterised by a lower socio-economic background (e.g. Angrist & Lavy, 1999). Next, the literature focusing on school size effects finds possible economies of scale when increasing the number of students per school. Again, not all students benefit equally from increases in school size. Students with low socio-economic status (SES) are argued to benefit most from schooling in relatively small schools (Luyten *et al.*, 2014; Leithwood & Jantzi, 2016). Bickel and Howley (2000) and Humlum and Smith (2015) stress that school size is context-dependent (e.g. student and neighbourhood characteristics), which makes them sceptical with respect to the 'optimal' school size. Note that the context of schools includes the characteristics of the school board governing its schools. A third branch of the literature studies scale effects at this level: the school district. Despite the direct and indirect link between the school district and student outcomes, this level of analysis has received significantly less attention compared with class and school size effects (see the literature review in the next section).

Moreover, the literature is almost completely US-oriented, which creates issues with external validity (Luyten *et al.*, 2014). Humlum and Smith (2015) confirm this finding and state that 'the issue of school reform is extremely relevant in the European Union due to both the demographic development and the recent economic crisis' and 'the reviewed empirical evidence does not provide a clear roadmap for school reform in the EU countries'.[2] Schütz (2007) concludes that the shape and strength of the size relationship depends on the context and educational setting. Owing to the substantial variation in school and district size both within and between European countries, it is impossible to provide one 'magical number' in the form of the ideal organisation of education (Humlum & Smith, 2015).

For example, in the *school size* literature, a paper by Bradley and Taylor (1998) identified an inverted 'U'-shaped relationship between school size and educational performance in England, attaining an optimum between 1,200 and 1,500 students. In contrast, the 'optimal' size reported by Foreman-Peck and Foreman-Peck (2006) appeared to be 560 students in Wales. Even more contradictory, a study by Sawkins (2002) applied a similar methodology and revealed a 'U'-shaped relationship in Scotland. Luyten *et al.* (2014) offers a possible explanation for these inconsistent findings by pointing to the different average school sizes in these countries. In this paper, we argue that this anomaly might be due to multiple optima in the size–performance relationship. There is some agreement in the literature that economies of scale exist for the smallest districts (and schools), and that these benefits of size turn into diseconomies once a (country-specific) upper threshold is reached. However, as we will develop further below, imposing a smooth U (or inverted-U) curve in between these boundaries, although better than imposing linearity, remains a very restrictive approach.

This paper contributes to the literature by illustrating how policy implications with respect to the optimal scale of school districts are significantly altered when minor changes are made to the empirical strategy. First, we show that more flexibility in specifying the functional form of cost functions is required to obtain more robust estimates of the optimal district scale and its decisive determinants. Overly strong parametric assumptions result in specification bias, and consequent false conclusions. Second, we introduce two estimation methods, building on standard approaches in the literature, which can easily be applied by practitioners in education. However, these minor adaptions to the empirical strategy appear to be of crucial importance. This paper is the first to use a Fourier extension to identify the optimal scale of school districts. Fixed effects estimation of this functional form allows us to identify the major determinants of scale economies and, more importantly, the optimal school district size.

We illustrate our approach empirically by applying these state-of-the-art techniques to (panel) data from schools and districts in a densely populated area in Western Europe (i.e. Flanders, the northern region of Belgium). In doing so, we are able to test the external validity of findings in the international (mainly US) literature, and assess the importance of context dependencies in estimating the optimal district size, as stressed by Luyten *et al.* (2014). Correspondences between the Flemish educational system and other countries are significant as, in line with many other OECD countries like the Netherlands, Germany, the UK, Spain and Italy, it is marked by a recent trend towards decentralisation of (education) governance (e.g. Burns & Köster, 2016).

The remainder of this paper proceeds as follows. The following section presents an overview of the literature. We motivate the need for a more flexible approach to identify the optimal scale of education at the district level. The next two sections present the available data and introduce different specifications used to estimate this optimum. The penultimate section summarises the obtained results, both numerically and graphically. A final section concludes.

Literature review

First, consider the production function literature, which relates school district size to education outcomes. In most production functions, school district size has been included as a secondary control variable. Only a small number of studies considered size as a determinant. Most papers specified a linear functional form, which stands in stark contrast to more flexible forms in cost function studies (Andrews *et al.*, 2002).

Second, consider the cost function literature, which relates the scale of school districts to costs per student. Only a limited number of studies in this literature have simultaneously included measures of school and district size to disentangle both effects. Bickel and Howley (2000) pointed to a joint influence of both levels. Studies that fail to account for both sources of scale economy are likely to report biased effects of the mechanisms under evaluation. Lewis and Chakraborty (1996) concluded that 'when both [school size and district size variables] are included in the regression equation, only school size is significant. Consolidation of schools, not districts, may be the key to achieving lower per unit costs' (p. 23). Another study by Duncombe *et al.* (1995a) also included both school and school district size.[3] Their results indicated sizeable potential cost savings by consolidating (especially small) school districts.[4]

These conflicting findings might be due to either the context dependency of the conclusions, as indicated by Luyten *et al.* (2014) or Leithwood and Jantzi (2016), or it might be due to differences in the chosen cost function specification.[5] Already, in a review by Fox (1981), the significant variation in results has been emphasised. Fox argued that methodological advances could improve the consistency in findings with respect to returns to scale in education. A more recent review of the literature by Andrews *et al.* (2002) concluded that more uniformity in results was observed in post-1980 studies. In general, the available research suggests sizeable cost savings in the smallest (around 500 students) school districts. Cost savings are the largest up to a size of 2,000–4,000 students per district. Once a threshold of approximately 6,000 students is reached, the costs per student tend to stop declining and diseconomies of scale can be observed in the largest districts. Despite improvements in methodology and increasing uniformity in the results, Andrews *et al.* (2002) argue that there is still room for progress in the available literature. More flexible functional forms are needed to minimise the assumptions made on the data and the resulting specification bias.

Methodologically, most cost function studies assume a linear or, since Wales (1973), a quadratic function. Others have estimated a translog cost function which adds flexibility by including a number of interaction terms. However, studies estimating a translog function have not indicated a statistically significant role for interaction terms between SES, district size and costs per student (e.g. Callan & Santerre, 1990). This contrasts with the finding that the specific 'costs and benefits' of consolidation are strongly related to the socio-economic background of students (Friedkin & Necochea, 1988). Heinesen (2005) allows more flexibility in the production function by including size dummies, resulting in scale economies similar to those observed in the cost function literature. However, interval estimations (instead of a continuous approach) cannot reveal anomalies in the cost function. We argue later that this is of major importance in order to explain (dis)economies of scale in education.

To account for district-specific effects, panel data can be used (Downes & Pogue, 1994). Alternative approaches consist of adding an efficiency term as a control variable (e.g. Duncombe *et al.*, 1995b) or using stochastic frontier analysis methods (SFA) (e.g. Johnes & Johnes, 2009). Applying these extensions resolves, at least partially, endogeneity issues due to the simultaneity between district size and quality, as some districts might experience higher growth rates due to higher (perceived) quality (Driscoll *et al.*, 2003).

In addition to the methodological remarks, two major data issues remain problematic in the existing literature. In their review of studies estimating cost and production functions at the school district level, Andrews *et al.* (2002) suggest analysing panel datasets to account for time-invariant district-specific effects. The majority of studies on economies of scale through school district consolidation have analysed cross-sectional data. Some recent papers partially applied these suggestions (e.g. Driscoll *et al.*, 2003; Jones *et al.*, 2008), but these studies focused on production functions instead of cost functions. The second data issue is related to the limited use of non-US data, leaving the external validity question largely unanswered (Luyten *et al.*, 2014; Humlum & Smith, 2015). In sum, a sound flexible analysis of the optimal district size and its determinants in a non-US setting is still lacking in the literature. This paper contributes to the above literature by introducing two flexible approaches towards estimating scale economies in school districts. We illustrate the added value in terms of more nuanced policy implications by estimating the optimal size of school districts in Flanders. In doing so, this paper complements the limited and mainly US-oriented literature focusing on school districts.

Data and setting

Flanders, alongside the Netherlands, has a lengthy practice of freedom of school choice and private (but publicly subsidised) schools have always played a major role in the provision of education. Other countries, such as the UK and Finland, have a long tradition of decentralisation in, among other areas, education (Burns & Köster, 2016). Although the 'point of departure' differs, the majority of OECD countries have experienced a trend towards decentralisation in the past three decades. Despite differences between these countries, policymakers are increasingly looking for tools to evaluate input-based policies related to school and district size. Therefore, the methodology we develop in the next section will be more widely applicable.[6]

We use administrative data at the school district level covering 1,060 school districts during the 2009–2012 period.[7] There is a large heterogeneity in school district size, as school districts are not organised on a geographical basis (i.e. there is no catchment area) and differ by the provider of education. This provider can be the local community government, a private provider (e.g. Catholic education) or the regional government. All providers are free to organise schooling and are funded either in a direct (if organised by the regional government) or indirect way through subsidies. The level of funding differs slightly between providers, and hence dummies are included in all regressions to capture these differences.[8] As there is full freedom in school choice, students can freely choose between schools and districts, resulting in rivalry among schools both within and between school districts. This dual freedom of education creates a heterogeneous educational landscape within a relatively homogeneous and small area. However, in the setting of Flanders, student mobility is rather low and it is unusual for parents to relocate in order to get their children into the schools of a different district—especially when comparing student mobility to the case of the USA.[9] Nonetheless, as high-quality school districts might be endogenously growing, we take this issue into account in the following sections.

Figure 1 presents a histogram to illustrate the dispersion in school district size in Flanders. More than half of the districts are smaller than 500 students, while outliers emerge at the right tail of the distribution (17,383 students in the largest district). More detailed descriptive statistics related to school district size and other variables are presented in Table 1.

As a dependent variable, we consider the operational costs per student. This number is obtained by dividing the available funding at the district level by the number of students in the corresponding district. The natural logarithm of this variable is used to reduce the influence of outliers. Funding is allocated as a lump sum to each district, reflecting the number of students in a district, weighted by their chosen study field and socio-economic characteristics. As some fields of study are more expensive than others (e.g. vocational education requires advanced equipment), the budget per

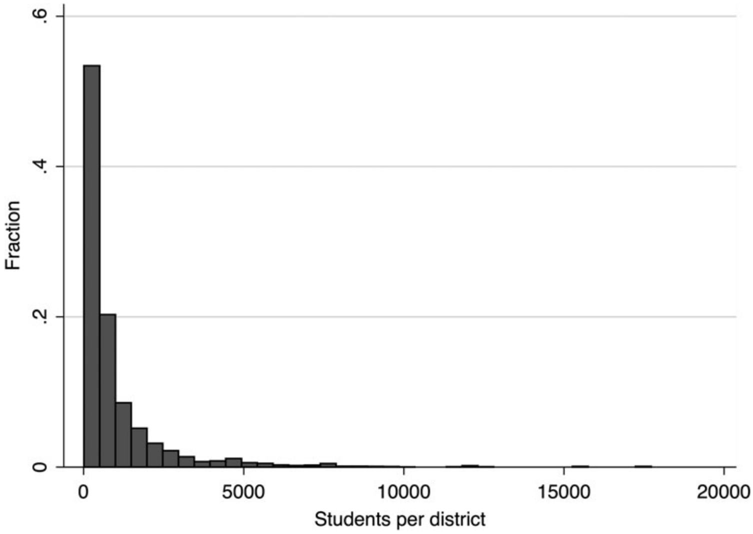

Figure 1. Students per district as a percentage of the population of school districts.

Table 1. Summary statistics of cost-related school district characteristics

Variables	N[a]	Mean	SD	Min	Max	Median
Total resources[b]	1,060	759	238	321	3,315	668
Operational resources[b]	1,060	739	189	353	2,318	661
Students per district	1,060	1,020	1,602	16	17,383	469
School size (mean)	1,060	307	151	16	1,374	284
Class size (mean)	1,060	23	14	3	138	20
Population density	1,060	1,127	2,179	52	23,754	558
Maternal education	1,060	0.069	0.082	0	0.833	0.046
Low-SES students[c]	1,060	0.010	0.010	0	0.092	0.009
% Young (<35) teachers	1,060	0.339	0.113	0	0.864	0.333
% Master's degree	1,060	0.098	0.145	0	0.693	0.019
% Absenteeism[d]	1,060	0.125	0.060	0	0.385	0.117
Gini	1,060	0.348	0.360	0	0.983	0.5
Herfindahl	1,060	0.672	0.345	0.023	1	0.664
Dummy variables	*N*	*Yes*		*No*		
Urban	1,060	22%		78%		
Only high schools	1,060	12%		88%		
Also VET schools	1,060	29%		71%		

[a]All variables are summarised for the year 2012, the most recent year covering all variables. Data for other time periods are not presented here to save space but are included in our analyses.
[b]Resources are expressed at the student level.
[c]Low SES students are receiving additional support.
[d]Teacher absenteeism is measured as the percentage of the total number of teachers, absent due to personal motivation, declining performance, or illness.

student that school districts have (as a lump sum) at their disposal differs significantly. We realistically assume that school districts spend all the funding allocated to them. This assumption is supported by previous research in Flemish schools (Groenez *et al.*, 2015). Also, Flanders is a relatively small region with limited input price variation (Smet & Nonneman, 1998). Consequently, deflating the operational costs by input price differentials would not affect our results. Therefore, by not including input prices, the function we estimate is in fact a 'pseudo-cost function'. By omitting input prices from the cost equation, we implicitly assume that schools try to minimise their total costs. In addition, we assume school districts to be operating as budget maximisers, as they endogenously influence the budget per student by strategic decisions on the number of available places per study field. For example, increasing the available capacity for vocational education and training (VET) students increases the lump sum financing, such that school districts can decide to distribute the resources towards non-VET students. Also, the funding mechanism is linear. This contrasts with the real costs, which are marginally decreasing. In other words, the cost of offering education to an additional VET student does not correspond to the cost of providing a VET study field to the first student in a district. School districts acting as budget maximisers *and* cost minimisers allow us to estimate the optimal scale of education, as our measure of operational costs reflects the true costs incurred by school districts.

The independent variables include both standard control variables and possible determinants of costs per student in a district. First, we include size variables measuring the number of students per district, per school and per class,[10] all expressed in logarithms.

Second, we control for school district structural characteristics that are hypothesised to influence the cost of education. Dummies indicate whether a school district offers only education at the high-school level (=1) or also at other levels (=0), like primary schooling. If schools within a district also offer some sort of VET schooling, then the VET dummy equals 1 and 0 otherwise. This is important to control for cross-subsidies between education levels. A third measure of district structure is captured by Herfindahl and Gini indices applied to school districts, with schools being the 'market players' and the number of students reflecting the 'market share'.[11] This measure of internal organisation might capture difficulties in coordination and communication. Also, the within distribution of school districts reflects the distribution of bargaining power between schools belonging to the same district. We expect more diverse organisations to be less cost-efficient.

As a third group of control variables, we account for differences in student composition. We obtained data on population densities within districts by merging our administrative dataset with statistical geographical data using matches on postal codes. This measure is included to take the sizeable differences in the cost of schooling between urban and rural areas into account (Kenny, 1982). A second variable indicates the percentage of students whose mothers' education is at most a high-school degree. A third variable indicates the percentage of students receiving additional support ('low-SES students'). The share of low-educated mothers and low-SES students captures differences in socio-economic composition between school districts—and their possible impact on school district choice (Urquiola, 2005).

A final set of variables consists of school district personnel characteristics. We include variables measuring the percentage of teachers younger than 35 years old, the percentage of teachers holding a master's degree and the percentage of teachers being absent during a given year. Teacher absenteeism can serve as a proxy for school district management (De Witte & Schiltz, 2017), which in turn affects student outcomes (Bidwell & Kasarda, 1975).[12] By including this variable, we are able to account for quality differences between districts and hence avoid the endogeneity issues mentioned by Driscoll *et al.* (2003).[13] Other (time-invariant) unobservable district-specific characteristics are captured in the fixed-effects term. Year dummies are included to control for general funding changes over time.

Table 1 summarises the data for all variables and districts.[14] Resources per student do not include teacher wages (which are funded in a uniform and centralised way for all school districts), and are therefore only around 750 euros per student per year. Total and operational costs do not differ significantly at the mean and median, but large differences are observed at the maximum. Additional resources, not included in operational costs, are targeted at some school districts to cover expenses linked to student background and type of education, irrespective of scale. The maximum in our dataset equals 3,315 euros per student per year. This high rate is due to a specific type of education, involving blind children, provided by the sole school in this district. Therefore, in the remainder of our analysis, we will mainly focus on operational costs.

District size is measured as the number of students in a district.[15] On average, school districts consist of 1,020 students, while less than half of districts are larger than 500 students, indicating a skewed distribution. Schools and classes have, on average, around 300 and 20 students, respectively. Despite some outliers at the maximum, these distributions are less skewed, with mean and median levels close to each other. The smallest school consists of 16 students. A closer look at the data reveals that this school is the only school in a small community-based district. The maximum observed 'class size' equals 138 and is due to the unit of measurement in our data. Class size is calculated as the number of students in a study field. Some schools provide a specific type of study in multiple classes, which explains the maximum number of 138 students in the study field.

Population density is measured as the number of inhabitants per square kilometre. Approximately 20% of districts are located in urban areas, indicated by a population density above the average value. Owing to large discrepancies in density between rural and urban areas, this value ranges from 52 to 23,754. The majority of school districts offer other levels of schooling (e.g. primary schooling), apart from high schools, and around one in three also provides some type of VET schooling. Gini and Herfindahl indices reflect a significant variation in school district structure, ranging from 0 to 1. In subsequent analyses only the Gini index is included, since both measures are strongly correlated ($\rho = -0.994$).

The variables 'maternal education' (i.e. mother's highest education level is primary education) and 'low-SES students' reveal a low average, displaying a highly educated population in Flanders. However, variation is high in these measures, indicating the need to account for inter-district variation in socio-economic background, which affects the cost of education. For example, 'low-SES students' amounts to 0.092 at the maximum, resulting in additional funding directed towards the corresponding district for almost 10% of its students. School district personnel amounts to one-third of young teachers on average, with some districts consisting of a share of young teachers equal to almost 85%. The percentage of teachers holding a master's degree is generally low within districts, since only teachers instructing in the final grade (17- to 18-year-old students) are required to hold this type of degree. Finally, teacher absenteeism also varies widely across districts, indicating broad quality differences (Bidwell & Kasarda, 1975; Ehrenberg *et al.*, 1991; De Witte & Schiltz, 2017). The average equals 12.5% and is measured as the percentage of teachers absent due to personal motivation (2%), declining performance (5%) or illness (5%). Including this measure as a control variable allows us to 'yield estimates of scale that more appropriately indicate the effects of variation in size' (Duncombe *et al.*, 1995a).

A flexible methodology

Consider a quadratic functional form that includes the natural logarithm of the number of students in district i at time t ($SPD_{i,t}$) and the log of district size squared, in addition to a set of j covariates introduced in the previous section ($X_{j,i,t}$) and an intercept capturing the district-level fixed effects (α_{0i}). The dependent variable ($C_{i,t}$) measures the natural logarithm of operational costs. $u_{i,t}$ represents an independent and identically distributed error term of district i at time t.

$$C_{i,t} = \alpha_{i0} + \theta_1 SPD_{i,t} + \theta_2 SPD_{i,t}{}^2 + \sum\nolimits_j \beta_j X_{j,i,t} + u_{i,t} \qquad (1)$$

Equation (1) can be interpreted as an extension of a standard Cobb–Douglas function, or equivalently, as a translog cost function without interaction terms. A graphical representation of this estimation is presented by the 'quadratic form' line in Figure 2. It can be observed from this illustration that by using a quadratic specification we do not observe a minimal cost, since costs are declining at an increasing rate. This result might be due to two, possibly simultaneously occurring, reasons. On the one hand, there might be no optimal district size and true possibilities of cost savings lie in school or class consolidation (as argued by Lewis & Chakraborty, 1996; Bickel

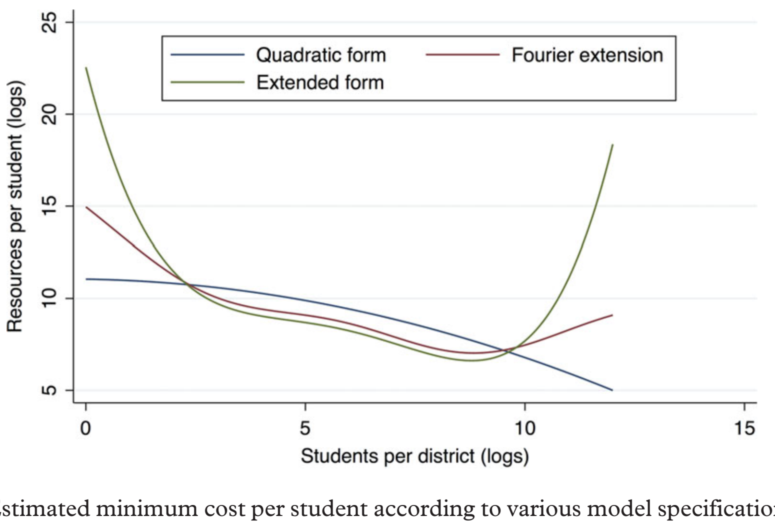

Figure 2. Estimated minimum cost per student according to various model specifications. [Colour figure can be viewed at wileyonlinelibrary.com]

& Howley, 2000), or there might simply be no optimal scale at all (as argued by Wales, 1973). On the other hand, there might be a misspecification of the functional form used to model the relationship between district size and costs per student. We can verify or rule out the latter line of reasoning by introducing more flexible cost-function specifications to estimate school district scale economies.

We extend equation (1) by including a number of interaction terms (as in Callan & Santerre, 1990; Smet & Nonneman, 1998). The resulting model is known as a translog cost function. This specification is still restrictive, but allows more flexibility compared with equation (1).

$$C_{i,t} = (1) + \sum_j \delta_{SPDj} SPD_{i,t} X_{j,i,t} + \sum_j \sum_k \gamma_{jk} X_{j,i,t} X_{k,i,t} \qquad (2)$$

Comparing the F-test of equations (1) and (2) suggests that including the set of interaction terms offers no 'added value'. The obtained *p*-value is too large to reject the null hypothesis that the translog model does not provide a significantly better fit than our baseline model.

Therefore, we offer two additional and complementary extensions to allow for more flexibility in the estimation method: an 'extended form' and estimation by Fourier specification. Both extensions have been overlooked in the previous studies linking district size to the (operational) cost of education. First of all, we extend the assumption of one global optimum towards the possibility of multiple optima by including higher-degree polynomials. This less stringent assumption on the functional form is still consistent with the possibility of a parabolic function, while also permitting local optima. As a result, the presence of mechanisms in between different optima can be identified. Using both AIC (Akaike) and BIC (Schwarz) criteria, we gradually add a polynomial, starting from the base model (without interactions). This stepwise approach results in the base model, extended by a third- and fourth-degree district size polynomial. For simplicity, in the remainder of this paper, we label this specification the 'extended form'.

$$C_{i,t} = (1) + \theta_3 {SPD_{i,t}}^3 + \theta_4 {SPD_{i,t}}^4 \qquad (3)$$

This specification provides a significantly better fit compared with our base model, as indicated by the F-test, in contrast to the translog cost model. Figure 2 provides a graphical representation of the extended form (see the 'extended form' line).

Our second estimation method is known as the Fourier cost function. Estimation using a Fourier cost function is the most flexible parametric approach available. The Fourier cost function extends our baseline model by adding *sine* (sin) and *cosine* (cos) terms. The final term in equation (4) is gradually extended over *N*, based on an F-test to check the improved model fit.

$$C_{i,t} = (1) + \sum_{k=1}^{N} (\delta_{1k} \sin(kSPD_{i,t}) + \delta_{2k} \cos(kSPD_{i,t})) \qquad (4)$$

Despite its parametric form, the Fourier specification provides a framework to estimate cost functions with a flexibility comparable to a nonparametric approach (De

Witte & Dijkgraaf, 2010). Nonparametric approaches let the 'data speak for itself', and do not impose any type of functional form. However, the advantage of parametric analysis lies in the possibility to control for a larger set of covariates and fixed effects, compared with nonparametric models. By estimating the coefficients of these covariates, we can discuss possible mechanisms causing the peculiar form of both the extended form and the Fourier specification, displayed in Figure 2.

School districts might be endogenously growing if they have (or are perceived to have) a higher quality. As a robustness check, and to tackle the issue of simultaneity in district size and quality of schooling,[16] we perform an additional estimation by means of a semi-parametric stochastic frontier analysis (SFA). It can be argued that coefficients obtained by mean regression techniques are affected by inefficiencies associated with the set of covariates. Stochastic frontier methods do not assume school districts to operate at full efficiency, in contrast to mean regression techniques (Duncombe *et al.*, 1995a). To rule out unobserved heterogeneity from the inefficiency component, we re-estimate the extended form model using a fixed-effects stochastic frontier model (i.e. the 'true fixed effects' model; Greene, 2005).[17] In addition, we estimate the model developed by Wang and Ho (2010), which performs a within transformation of the data, avoiding the incidental parameter problem.

Results

The results are presented in Table 2. In both flexible model specifications (i.e. extended and Fourier), school size variables do not appear to be significantly related to costs per student if school district size is included in the regression analysis.[18] In all models, class size (i.e. the size of the study field) is negatively related to the cost per student in a district. The estimation of model specifications that exclude district size returns significant coefficients on class and school size. This finding is contradictory to Lewis and Chakraborty (1996), who are implying a major role of school (and not school district) consolidation as a means to cut costs in education. Considering the Flemish funding mechanism, our opposite finding is intuitive: we focus on the analysis of the relationship between *size* and *operational* costs—excluding personnel costs—which are centrally funded in Flanders. Although cost savings might occur by teaching in larger classes and larger schools, our findings point in the direction of major cost saving possibilities related to school district (re)organisation. We included and excluded the class and school size variables in the different model specifications to emphasise the robustness of district size effects, once school and class size effects are accounted for.[19]

When estimating the quadratic model specification, coefficients of school district size are alternately significant, depending on the inclusion of school and class size variables, implying a misspecification of the underlying cost function. In our two other specifications, all district size variables (higher-degree polynomials, sine and cosine) are significantly related to per student costs.[20] Comparison of the coefficients results in the conclusion that inclusion of school and class size variables mitigates the estimated scale economies, especially at the lower tail of the district size distribution. This finding is supported when we calculate the cost minima for models (1)–(6), or district size optima, indicating lower optima when school and class size variables are

Table 2. Factors affecting school district per student operational costs: fixed-effects regression results

Variable	Quadratic		Extended		Fourier	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	9.725***	11.07***	22.58***	21.53***	13.24***	14.46***
	(0.980)	(1.055)	(3.941)	(4.107)	(1.419)	(1.463)
Size						
Students per district	−0.580*	−0.0450	−10.06***	−8.715***	−1.837***	−1.615***
	(0.339)	(0.468)	(2.803)	(3.192)	(0.501)	(0.624)
Students per district^2	0.0101	−0.0382	2.494***	2.284***	0.111***	0.0930**
	(0.0294)	(0.0324)	(0.723)	(0.788)	(0.0419)	(0.0460)
Students per district^3			−0.276***	−0.269***		
			(0.0800)	(0.0857)		
Students per district^4			0.0111***	0.0113***		
			(0.00321)	(0.00343)		
sin(Students per district)					−0.0413	−0.227*
					(0.0894)	(0.117)
cos(Students per district)					0.446***	0.424***
					(0.143)	(0.144)
School size		−1.111*		−0.656		−0.848
		(0.592)		(0.627)		(0.648)
School size^2		0.118**		0.0815		0.101*
		(0.0481)		(0.0537)		(0.0565)
Class size		−0.116*		−0.125*		−0.132**
		(0.0702)		(0.0671)		(0.0656)

Table 2. (Continued)

Variable	Quadratic		Extended		Fourier	
	(1)	(2)	(3)	(4)	(5)	(6)
Structure						
Also VET schools	0.0391	0.0356	0.0401	0.0362	0.0370	0.0265
	(0.105)	(0.0572)	(0.0905)	(0.0594)	(0.0907)	(0.0570)
Only high schools	−0.0220	−0.0538	−0.0203	−0.0574	−0.0195	−0.0553
	(0.0295)	(0.0411)	(0.0289)	(0.0419)	(0.0288)	(0.0415)
Gini[a]	−0.0332	0.171	−0.0307	0.231	−0.0294	0.263
	(0.123)	(0.192)	(0.121)	(0.185)	(0.120)	(0.191)
Environment						
Population density	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
SES	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Personnel						
Teacher absenteeism	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Composition	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
N	4,212	4,212	4,212	4,212	4,212	4,212
R^2	0.249	0.276	0.282	0.297	0.280	0.298

Note: The dependent variable is the natural logarithm of per student operational costs. For brevity, coefficients of personnel and SES characteristics are not reported in this table, but they are included as control variables. Also, *t*-statistics are not included in this table. Robust standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. All size variables are measured in logarithms. aRobustness checks are performed using Herfindahl indices. Coefficients and optima are consistent when outlying observations (right tail of Figure 1) are excluded from the analysis.

included. The significance of the selected higher-degree polynomials, sine and cosine terms supports our hypothesis of multiple optima, as opposed to the stringent assumption of a ‘U’-shaped cost function.[21] This phenomenon is displayed in Figure 2.

Structural variables related to type (VET or non-VET) and level (primary versus high school) of schooling do not appear to be significantly related to per student costs. The internal organisation (or equivalently, within-district competition) of a school district, captured by the Gini index (or the Herfindahl index), is positively linked to the cost of education in all three specifications. Interestingly, this finding only holds when class and school size variables are included in the equation. The negative coefficient observed when school and class size variables are excluded from the regressions might be due to higher average school and class sizes in more diverse districts. Inspection of the data reveals that half of the school districts consist of only one school and are hence characterised by a Gini index equal to 1. Districts with an index above 0.5 entail higher costs per student, reflecting possible mismanagement or unequal representation of schools’ aspirations in more diverse (with respect to size) school districts. Higher percentages of low-SES students result in higher per student costs, which is in line with the funding mechanism in Flanders. A similar relationship can be observed with respect to more densely populated areas. When we estimate our models separately for both rural and urban areas, we observe different intercepts in addition to a disparity in coefficients, indicating higher overall costs in urban areas. Scale economies of the smallest school districts are more pronounced within urban areas, possibly linked to fewer barriers to coordination and organisation.[22] Finally, data on the age and education of teachers are only observed in 2012 and are dropped accordingly when fixed-effects regressions are carried out. Teacher absenteeism is observed for all time periods, and is found to be higher in districts with a higher per student cost. After controlling for student and school characteristics, this result remains statistically significant. Studies linking teacher absenteeism to school district management could serve as a possible explanation for this higher per student cost (Theobald, 1990; Hartog & Verburg, 2004; De Witte & Schiltz, 2017). Again, sound management performance could explain discrepancies in school district performance with respect to costs, and possibly also with respect to student outcomes, through its impact on human resource management (teacher hiring and retention).

Using the school district size estimates, scale economies are estimated while holding all other variables at the district level to their average values. The results are displayed in Figure 2. The findings are in line with consent in the literature, as the largest potential cost savings can be found within the smallest school districts (<500 students) and diseconomies of scale appear within the largest districts (at the threshold of around 6,000 students). A closer inspection of Figure 2 reveals an apparent slowdown in cost savings in medium-sized school districts (300–1,100 students). Beyond this ‘transition period’, cost savings from school district consolidation increase again, up to the optimal size of around 6,500 students. As indicated before, our presumed mechanism links school district management and organisational structures to school district (financial) performance. This relationship is supported by proxies such as the Gini index and teacher absenteeism trying to capture these school district characteristics. In addition, earlier research identified a connection between school district size and school board organisational characteristics (Sleegers *et al.*, 1994). Transition from small to large school districts is found to be characterised by changes in organisational structure. The observed transition period could hence be due to adaptive difficulties encountered by school boards when transforming small, community-based school districts into relatively large, overarching organisations, while higher costs observed in the smallest and largest school districts could be due to scale (dis)economies.

We observe that the 'transition' is more pronounced in urban areas. In rural areas, the cost function is more flattened out. Once a threshold of around 700 students is reached, both types (urban and rural) entail larger economies of scale until the optimal scale is attained, which is the same for both types of school districts (6,000–7,000 students).

Discussion and conclusion

This paper studies how scale economies in education can be estimated and how policy implications are altered when only slight changes are made to the empirical strategy. We propose two models that offer more flexibility and are comparable to semi- or nonparametric approaches. We apply our methodology to school district data in Flanders to illustrate the added value of this methodology for applied researchers and to complement the mainly US-oriented literature.

Estimating a flexible functional form reveals anomalies in the cost function which cannot be observed when estimating a 'U'-shaped curve. Extended form and Fourier cost functions are able to model the complex cost function and identify multiple optima, whereas classical approaches overestimate scale economies—that is, they fail to identify diseconomies of scale. Alternative flexible estimation methods can be used (see Table A1 in the appendix, using SFA models).

We find that our analysis using Flemish data confirms earlier evidence from the USA in that potential cost savings are the largest within the smallest districts, up to the optimal scale of 6,000–7,000 students per district. Consolidations beyond this optimum result in diseconomies of scale. Our results also indicate a 'transition period' between small and large school districts, characterised by less-pronounced scale economies. Andrews *et al.* (2002) noted that 'principal cost savings of increased enrolments are exhausted by the time a district reaches an enrolment of 500 to 1000 pupils', but the authors did not provide an explanation for this apparent slowdown. Including variables capturing school district structure and management practices (Gini coefficient and teacher absenteeism), we offered a suitable explanation for this transition period. The relationship between organisational effectiveness and district (financial) performance can be seen as complementary to available studies suggesting large cost savings due to economies of scale in small districts (e.g. centralising administrative tasks) and diseconomies of scale in large school districts due to the creation of burdensome bureaucracies (Robertson, 2007).

Further research is needed to identify the drivers of organisational effectiveness (e.g. why do less diverse districts perform better?) and to offer alternative mechanisms driving the curvature of the estimated cost function. Competition between school districts might offer an explanation for the observed trend in scale economies. Leach *et al.* (2010) state the following: 'Although the impact of reduced competition across school boards [due to consolidation] is at least partially offset by greater competition among schools, Urquiola (2005) argues that competition occurs mainly among the boards' (p. 1035). Inclusion of a distance-based competition measure would enable disentangling within- and between-district competition and their effect on school district performance.

Focus should be shifted towards *long-run effects* of district consolidations on costs, and most importantly student outcomes. Brasington (1998) argues that net savings are negative, caused by lower house prices, as a result of weaker student outcomes after district consolidation. If this relationship holds, then the gains of increasing the scale of education are absent and consolidation should not be pursued. Also, analysis based on long-run (panel data) studies enables controlling for selection issues related to consolidation. School districts that have already consolidated are those experiencing the greatest economies, so a short-run analysis tends to overestimate the benefits of scale increases (Leach *et al.*, 2010).

As stressed before, a lack of flexible *production function* specifications might have resulted in conflicting evidence on the impact of school and school district size (Andrews *et al.*, 2002). For example, in the *school* size literature, a paper by Bradley and Taylor (1998) identified an inverted 'U'-shaped relationship between school size and educational performance in England, attaining an optimum between 1,200 and 1,500 students. In contrast, the 'optimal' size reported by Foreman-Peck and Foreman-Peck (2006) appeared to be 560 students in Wales. Both studies impose a quadratic functional form to model the production function, essentially assuming the existence of only one optimum. The methods proposed here could contribute to such papers estimating the optimal scale of education using production functions. In our opinion, releasing the assumption of quadratic (or linear) functional forms will alter policy implications in the same manner as illustrated in this paper. In a recent study using Hungarian data, we find that the existence of multiple optima also holds when student test scores are used as outcome variables (Schiltz *et al.*, 2017). Further research might integrate our approach and resolve the seemingly conflicting evidence in the literature.

It is important to ask to what degree the results in this paper can be generalised to other settings, and there are certainly a number of caveats worth noting. First, the assumptions underlying the data used here are crucial to interpret the extent to which cost savings are possible. School districts acting as budget maximisers and cost minimisers allow us to estimate the optimal scale of education, as our measure of operational costs reflects the true costs incurred by school districts. Second, we do not claim to present causal evidence, but we offer a contribution in terms of a more flexible estimation method which resulted in more nuanced policy implications. Future research might provide causal inference by exploiting exogenous shocks. Third, despite similarities between the education system in Flanders and other OECD countries, results cannot simply be extrapolated to other countries.[23] More studies will be needed to further complement the literature outside the USA. The flexible models proposed in this paper can serve as a robust starting point for applied researchers estimating scale economies in education.

Acknowledgements

We would like to thank Dieter Dohmen, Steven Groenez, Jill Johnes, Iris Kesternich, Mika Kortelainen, Henry Levin, Amrit Thapa and participants of the 2nd Workshop on Economics of Education at Maastricht University, the XXV Meeting of the Economics of Education Association at Universidad de Extremadura for valuable comments and suggestions on earlier drafts of this paper. The authors acknowledge financial support from the Horizon 2020 programme through the EDEN project (Education Economics Network).

NOTES

[1] While the terms 'school board' and 'school district' are used interchangeably in earlier literature, we will use the term 'school district' in this paper.

[2] Although the focus of this report is on school size, comparing the methodologies applied in both school and district size studies, we conclude that our approach is applicable in both.

[3] The authors included the median school size within the evaluated district to capture the extent of district school centralisation.

[4] Apart from the cost and production function literature, others tackled the issue of district size from a different angle using political economy models. Public choice theories point at diseconomies of scale through consolidation due to increasing bureaucracy (Robertson, 2007; Gordon & Knight, 2008) or indirectly through an increase in bargaining power of teachers' unions (Rose & Sonstelie, 2010). This approach is not followed, considering the scope of this paper, but can be most useful to identify alternative mechanisms driving (dis) economies of scale in education.

[5] For example, Duncombe *et al.* (1995a) assume a parabolic function by including students per district squared as an additional variable, while Lewis and Chakraborty (1996) estimate a log–log relationship.

[6] To facilitate applications to other settings, our empirical code is available upon request.

[7] In total, there were 1,099 school districts in Flanders in 2014. However, some districts were established after 2012 and others were eliminated due to missing data. We also dropped some districts due to their different organisational entity structure, which is an executive part of the municipal government ('AGB Antwerp'). All subsequent analyses were replicated including these observations and remain valid.

[8] When estimating a fixed-effects regression, these dummies are omitted and district-specific characteristics (including the type of provider) are captured by the constant term.

[9] In Flanders, official statistics indicate only 4.1% of all people are moving per year. Although no official statistics are available, a report commissioned by the Flemish government indicates that movers motivated by schools appear to be very rare (De Bruyne & Iserbyt, 2011).

[10] School and class size variables were obtained by taking the averages across all schools and classes within these schools belonging to the same school district.

[11] These indices have also been applied in other studies to capture the degree of competition (e.g. Hoxby, 2000). We only calculated these measures within every district, so it cannot be interpreted as a measure of overall competition.

[12] A study by Ehrenberg *et al.* (1991) directly links student achievement to teacher absenteeism.

[13] Note that we also included a set of personnel characteristics (age and education) to reflect the self-selection of teachers into high-quality schools (Falch & Strom, 2005).

[14] For the sake of brevity, data is only presented for the school year 2012. Also, some variables are not observed for all years and could therefore not be aggregated for all time periods.

[15] An alternative is to use schools per district as an indicator of size. However, variation in the number of schools within a district is close to zero and hence less informative when applying a fixed-effects regression.

[16] Note that we have included a measure of teacher absenteeism to control for quality differences between districts. However, frontier regression methods (e.g. SFA) take into account efficiency differences, which are more likely to reflect variation in school and district performance. As we will argue later, consistent results from SFA estimates indicate that issues related to endogeneity might not be of major importance here, once our quality controls are included.

[17] In addition, we apply the random-effects time-invariant frontier model of Battese and Coelli (1988), which returns different results. However, a Hausman test confirms the existence of fixed effects.

[18] Note that we included school size squared to allow for the possibility of a nonlinear relationship between school size and per student costs.

[19] Estimates using SFA do find significant school size effects, although the signs do not differ. The focus of this paper is on scale economies at the district level. Nonetheless, more flexibility could be allowed in modelling the school size and class size effect using the same methods proposed for the district level.

[20] This finding is an immediate result of the selection methods applied to obtain the estimated specifications (AIC, BIC and F-tests).

[21] The existence of multiple optima might explain the seemingly conflicting findings with respect to the size–performance relationship in England, Wales and Scotland (Bradley & Taylor, 1998; Sawkins, 2002; Foreman-Peck & Foreman-Peck, 2006).

[22] Another explanation could be related to increased competition between districts in urban areas, resulting in better financial management by the school boards. Since the Gini index only captures within-district competition, this between component is not accounted for.

[23] To facilitate the use of the model specification in new applications, we have made the STATA code available upon request.

References

Alonso, J. M., Clifton, J. & Díaz-Fuentes, D. (2015) Did new public management matter? An empirical analysis of the outsourcing and decentralization effects on public sector size, *Public Management Review*, 17(5), 643–660.

Andrews, M., Duncombe, W. & Yinger, J. (2002) Revisiting economies of size in American education: Are we any closer to a consensus?, *Economics of Education Review*, 21(3), 245–262.

Angrist, J. D. & Lavy, V. (1999) Using Maimonides' rule to estimate the effect of class size on scholastic achievement, *Quarterly Journal of Economics*, 114(2), 533–575.

Battese, G. E. & Coelli, T. J. (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data, *Journal of Econometrics*, 38(3), 387–399.

Bickel, R. & Howley, C. (2000) The influence of scale on school performance: A multi-level extension of the Matthew principle, *Education Policy Analysis Archives*, 8(22), 1–32.

Bidwell, C. E. & Kasarda, J. D. (1975) School district organization and student achievement, *American Sociological Review*, 40(1), 55–70.

Bloom, N., Lemos, R., Sadun, R. & Van Reenen, J. (2015) Does management matter in schools?, *The Economic Journal*, 125(584), 647–674.

Bradley, S. & Taylor, J. (1998) The effect of school size on exam performance in secondary schools, *Oxford Bulletin of Economics and Statistics*, 60(3), 291–324.

Brasington, D. M. (1998) School district consolidation, student performance, and housing values, *SSRN Electronic Journal*, 6658, 1–25.

Burns, T. & Köster, F. (2016) *Governing education in a complex world* (Paris, OECD Publishing).

Callan, S. J. & Santerre, R. E. (1990) The production characteristics of local public education: A multiple product and input analysis, *Southern Economic Journal*, 57(2), 468.

De Bruyne, L. & Iserbyt, S. (2011) Verhuis- en blijfmotieven van jonge gezinnen en jongvolwassenen, *Vormings- En Onderzoekstraject I.o.v Team Stedenbeleid Vlaamse Gemeenschap*, 1–249.

De Witte, K. & Dijkgraaf, E. (2010) Mean and bold? On separating merger economies from structural efficiency gains in the drinking water sector, *Journal of the Operational Research Society*, 61, 222–234.

De Witte, K. & Schiltz, F. (2017) Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach, Working Paper.

Downes, T. A. & Pogue, T. F. (1994) Adjusting school aid formulas for the higher cost of educating disadvantaged students, *National Tax Journal*, 47(1), 89–110.

Driscoll, D., Halcoussis, D. & Svorny, S. (2003) School district size and student performance, *Economics of Education Review*, 22(2), 193–201.

Duncombe, W., Miner, J. & Ruggiero, J. (1995a) Potential cost savings from school district consolidation: A case study of New York, *Economics of Education Review*, 14(3), 265–284.

Duncombe, W. D., Ruggiero, J. & Yinger, J. (1995b) Alternative approaches to measuring the cost of education, in: *Holding schools accountable: Performance-based reform in education* (Washington, D.C., Brookings Institution Press), 327–356.

Ehrenberg, R. G., Ehrenberg, R. A., Rees, D. I. & Ehrenberg, E. L. (1991) School district leave policies, teacher absenteeism, and student achievement, *Journal of Human Resources*, 26, 72–105.

Falch, T. & Strom, B. (2005) Teacher turnover and non-pecuniary factors, *Economics of Education Review*, 24(6), 611–631.

Foreman-Peck, J. & Foreman-Peck, L. (2006) Should schools be smaller? The size–performance relationship for Welsh schools, *Economics of Education Review*, 25(2), 157–171.

Fox, W. F. (1981) *Size economies in local government services: A review* (vol. 22) (Washington, D.C., US ESCS).

Friedkin, N. E. & Necochea, J. (1988) School system size and performance: A contingency perspective, *Educational Evaluation and Policy Analysis*, 10(3), 237–249.

Gordon, N. & Knight, B. (2008) The effects of school district consolidation on educational cost and quality, *Public Finance Review*, 36(4), 408–430.

Greene, W. (2005) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model, *Journal of Econometrics*, 126(2), 269–303.

Groenez, S., Juchtmans, G., Smet, M. & Stevens, C. (2015) *Analyse van het nieuwe financieringsmechanisme voor de werkingsmiddelen van scholen*. Evaluatie van het financieringsdecreet van 2008: Eindrapport.

Hartog, D. & Verburg, R. (2004) High performance work systems, organisational culture and firm effectiveness, *Human Resource Management*, 14(1), 55–78.

Heinesen, E. (2005) School district size and student educational attainment: Evidence from Denmark, *Economics of Education Review*, 24(6), 677–689.

Hoxby, C. M. (2000) Does competition among public schools benefit students and taxpayers?, *American Economic Review*, 90(5), 1209–1238.

Humlum, M. K. & Smith, N. (2015) *The impact of school size and school consolidations on quality and equity in education*. EENEE Analytical Report No. 24.

Jarl, M., Fredriksson, A. & Persson, S. (2012) New public management in public education: A catalyst for the professionalization of Swedish school principals, *Public Administration*, 90(2), 429–444.

Johnes, G. & Johnes, J. (2009) Higher education institutions' costs and efficiency: Taking the decomposition a further step, *Economics of Education Review*, 28(1), 107–113.

Jones, J. T., Toma, E. F. & Zimmer, R. W. (2008) School attendance and district and school size, *Economics of Education Review*, 27(2), 140–148.

Kenny, L. W. (1982) Economies of scale in schooling, *Economics of Education Review*, 2(1), 1–24.

Leach, J., Payne, A. A. & Chan, S. (2010) The effects of school board consolidation and financing on student performance, *Economics of Education Review*, 29(6), 1034–1046.

Leithwood, K. & Jantzi, D. (2016) A review of empirical evidence about school size effects: A policy perspective, *Review of Educational Research*, 79(1), 464–490.

Lewis, W. C. & Chakraborty, K. (1996) Scale economies in public education, *Regional Analysis & Policy*, 26(1), 23–35.

Luyten, H., Hendriks, M. & Scheerens, J. (2014) *School size effects revisited: A qualitative and quantitative review of the research evidence in primary and secondary education* (Berlin, Springer).

Robertson, F. (2007) Economies of scale for large school districts: A national study with local implications, *Social Science Journal*, 44(4), 620–629.

Rose, H. & Sonstelie, J. (2010) School board politics, school district size, and the bargaining power of teachers' unions, *Journal of Urban Economics*, 67(3), 438–450.

Saatcioglu, A., Moore, S., Sargut, G. & Bajaj, A. (2011) The role of school board social capital in district governance: Effects on financial and academic outcomes, *Leadership and Policy in Schools*, 10(1), 1–42.

Sawkins, J. W. (2002) Examination performance in Scottish secondary schools: An ordered logit approach, *Applied Economics*, 34(16), 2031–2041.

Schiltz, F., Masci, C., Agasisti, T. & Horn, D. (2017) Using Machine Learning to model interaction effects: A graphical approach with an application to education data. Working Paper.

Schütz, G. (2007) School size and academic achievement in TIMMS 2003, in: T. Loveless (Ed.) *Lessons learned. What international assessments tell us about math achievement* (Washington, D.C., Brookings Institution Press).

Sleegers, P., Bergen, T. & Giesbers, J. (1994) The policy-making capacity of schools: Results of a Dutch study, *Educational Management and Administration*, 22(3), 147–159.
Smet, M. & Nonneman, W. (1998) Economies of scale and scope in Flemish secondary schools, *Applied Economics*, 30(9), 1251–1258.
Theobald, N. D. (1990) An examination of the influence of personal, professional, and school district characteristics on public school teacher retention, *Economics of Education Review*, 9(3), 241–250.
Urquiola, M. (2005) Does school choice lead to sorting? Evidence from Tiebout variation, *American Economic Review*, 95(4), 1310–1326.
Wales, T. J. (1973) The effect of school and district size on education costs in British Columbia, *International Economic Review*, 14(3), 710–720.
Wang, H. J. & Ho, C. W. (2010) Estimating fixed-effect panel stochastic frontier models by model transformation, *Journal of Econometrics*, 157(2), 286–296.

Appendix

Table A1. Robustness check using SFA

Variable	SFA (default)	True fixed effects	Wang and Ho (2010)
Intercept	19.34***	28.184***	12.406***
	(2.197)	(0.059)	(0.025)
Size			
SPD	−6.883***	−12.70	−5.813***
	(1.895)	n.a.	(1.582)
SPD^2	2.009***	3.109***	1.979***
	(0.493)	(1.19e−08)	(0.394)
SPD^3	−0.251***	−0.340***	−0.264***
	(0.0584)	(2.19e−09)	(0.046)
SPD^4	0.0110***	0.0136***	0.012***
	(0.00254)	(1.10e−10)	(0.002)
School size	−1.246**	0.0565***	−2.668**
	(0.599)	(1.66e−08)	(0.635)
School size^2	0.123**	0.0157***	0.283**
	(0.0510)	(8.48e−10)	(0.055)
Class size	−0.167***	−0.150***	0.000
	(0.0464)	(8.93e−09)	(0.000)
Observations	4,212	4,212	4,212

Note: SPD stands for 'students per district'. All size variables are measured in logarithms. Robust standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$. Both models estimate coefficients corresponding to the extended and Fourier model. Owing to the incidental parameter problem, all coefficients appear to be significant. However, when computing the optimal district size by plugging in the obtain coefficients, policy recommendations with respect to the optimal district size remain relatively unaltered. This confirms the robustness of our previous findings while it also implies the sensitivity of the SFA methodology when estimating scale economies in large datasets. Therefore, we apply the extended and Fourier model in our main analysis. With respect to the *form* of the cost–size relationship, all flexible models estimated in this paper indicate the same pattern: scale economies for small districts (SPD negative), an apparent slowdown in scale economies (SPD^2 positive), the optimal scale can be identified past 6,000 students (SPD^3 negative) and school districts incurring diseconomies of scale once this optimal scale is exceeded (SPD^4 positive).