

---

## Stochastic Frontier Analysis of the Efficiency of Czech Grammar Schools\*

MICHAL FRANTA\*\*

CERGE-EI, Prague

TOMÁŠ KONEČNÝ

Institute of Sociology AS CR and CERGE-EI, Prague

**Abstract:** The study focuses on the evaluation of the relative efficiency of Czech grammar schools ('gymnázium') at preparing their graduates for admission to university programmes, taking into account the relative demand for the programmes, grammar school endowments, and a number of other relevant external factors. The authors argue that a comparison of secondary schools based exclusively on university acceptance rates or other direct measures of study achievements might be misleading, given that such approaches ignore many aspects related to the educational process, such as differences in the level of non-cognitive skills of students, family background, overall living standards in a region, and the specific focus of the given secondary school. The authors derive novel indirect measures of grammar school efficiency that take into account relative demands for university study programmes. Relying on high-quality panel data that cover the period 2000–2004, and using the stochastic frontier methodology commonly applied in efficiency evaluation in a wide number of sectors, including education, the study also argues that the divergence between selected direct measures of school performance and indirect indicators has been increasing over time.

**Keywords:** education, secondary education, performance evaluation, technical efficiency

*Sociologický časopis/Czech Sociological Review, 2009, Vol. 45, No. 6: 1265–1282*

### Introduction

The Czech Republic has recently witnessed an extensive public discussion on the optimal structure of secondary education and the design of education policies in general. The existence of six-year and eight-year grammar schools, the role of private schools, and the perspective of apprenticeship schooling belonged

---

\* This article was prepared with the support of the Grant Agency of the Czech Republic, grant no. GA403/08/0109.

\*\* Direct all correspondence to: Michal Franta, CERGE-EI, PO Box 882, Politických vězňů 7, 11121 Prague, Czech Republic, e-mail: michal.franta@cerge-ei.cz; Tomáš Konečný, Institute of Sociology AS CR, Jilská 1, 110 00 Prague 1, Czech Republic, e-mail: tomas.konecny@cerge-ei.cz.

among the most contentious issues. The present study intends to point to a related though rather standalone topic: the evaluation of the relative efficiency of schools at producing successful graduates. Our interest here is highly selective, as we focus exclusively on Czech grammar schools – ‘gymnázium’ – and we judge them successful based on how many of their graduates go on to university and how many of them are admitted to the top universities most in demand.

This topic is important because it relates to the planned reform of the Czech educational system. This reform intends to introduce uniform state school-leaving-exams that can serve as a basis of comparison of grammar school efficiency. In Great Britain, for example, SAT-like scores were used not only for admission to the tertiary level of education, but also as the main input into rankings of secondary schools [Griffith 2008].

Below we argue that comparing secondary schools by study achievements or by the success of their graduates in the labour market can be misleading from a public policy perspective because such approaches ignore many other aspects related to the educational process, such as differences in students’ non-cognitive skills, family background, overall living standards in a region, and the specific specialisation of a given secondary school. The aim of this article is to apply a method that takes into account inputs into the schooling process alongside other relevant external factors.

Controlling for several secondary school factors in determining the efficiency of Czech secondary schools is provided by Stupnytskyy [2002]. He applied the Data Envelopment Analysis (DEA) to provide a measure of efficiency adjusted for the basic characteristics of a secondary school (e.g. the number of students per teacher, the presence of a guidance counsellor).

In this article, we use the method of stochastic frontier analysis (SFA) to measure the efficiency of grammar schools in the Czech Republic. With the SFA approach we can relate efficiency estimates even to those factors that do not directly pertain to a particular secondary school. Using panel data on grammar schools, data on the local economic and social environment, and data on applicants to the tertiary level of education over the period 2000–2004, novel measures of technical efficiency are obtained and compared to a simple unconditional measure – the probability of being admitted to the tertiary level of education.

We show that environmental characteristics (e.g. the local unemployment rate and skilled/unskilled wage ratio) significantly affect the efficiency of the educational process. For example, the mean of grammar school inefficiency is positively related to the local wage ratio of university graduates to high school graduates. The flow of FDI has a negative effect on inefficiency means. The effect of direct inputs on the educational process is found to be intuitive. The students per teacher ratio, for example, relates negatively to the performance of grammar school graduates – one more student per teacher decreases the normalised probability of being admitted by more than 0.1 percentage point. We also demonstrate that a direct assessment of grammar schools based on the success of graduates in the admission process to university might in fact be rather misleading from

a public policy perspective, given that the divergence between popular measures (such as simple shares of students admitted to universities) and indicators capturing the broader family of factors within the educational process increases over time.

The rest of this article is organised as follows. First we describe the system of institutions that provide secondary education in the Czech Republic and the role of grammar schools in the system. Next we introduce the basic principles of SFA, followed by a discussion of data and of the estimation results, and then we show how much the measure of efficiency based on SFA differs from the measure based on university admission probability.

### **Grammar schools in the Czech Republic**

In the Czech Republic, the completion of primary nine-year school attendance is compulsory. Afterwards a student can continue at the secondary level of education at various types of schools: grammar schools,<sup>1</sup> specialised secondary schools, and vocational schools. In 2007, approximately one-quarter of all secondary school students were enrolled on a grammar school. In addition to four-year grammar schools, there also exist six- and eight-year grammar schools that provide part of the compulsory nine years of education.

Grammar schools are institutions that provide general study programmes. Their primary task is to prepare students for study at the tertiary level of education. It is the success at gaining admission to tertiary study programmes that is often viewed as an important criterion for the assessment of grammar schools. The table below shows the average shares of grammar school graduates applying for and being admitted to a university between 1999 and 2006. The table also shows the increase in the shares of admitted students over time. The growing demand for tertiary education has translated into increases in the relative numbers of applicants. The growing supply of tertiary education is reflected in increases in the share of students admitted. In the subsequent analysis of grammar school efficiency supply-side factors should not influence the estimates. The change in the share of grammar school graduates caused by the establishment of new tertiary institutions should not affect the efficiency of a grammar school.<sup>2</sup> Finally, the changes in the number of grammar schools in Table 1 reflect, for example, the administrative merging of grammar schools and the introduction of private institutions.

---

<sup>1</sup> Grammar schools in the Czech Republic correspond to category 3A in the ISCED classification.

<sup>2</sup> A tertiary institution can affect grammar schools by channels other than the simple supply of study programmes. For example, grammar school students in the neighbourhood of a university can be influenced by face-to-face contacts with university students, by being able to use the university's facilities, and so on. We will return to this effect in the subsection discussing estimation results.

**Table 1. Gymnasiums and gymnasium graduates entering faculties**

Year	1999	2000	2001	2002	2003	2004	2005	2006
Number of grammar schools	285	323	310	331	336	334	341	346
Share of applicants (%)	71.46	92.32	88.05	89.16	90.65	92.18	92.64	91.79
Share of admitted (%)	51.76	65.73	65.68	69.52	72.58	77.14	79.27	78.95

Source: 'Applicant' database of the Institute for Information in Education.

### Stochastic frontier analysis (SFA)

SFA was introduced in Aigner, Lovell and Schmidt [1977] and Meeusen and van den Broeck [1977].<sup>3</sup> The method is used to measure efficiency in regulated or public sectors such as the education sector.<sup>4</sup> The principle it is based on is the concept of production function, that is, the relationship between the set of inputs and the set of outputs in production. In the classic perspective, producers produce the maximum attainable output for a given technology and level of inputs (or, equivalently, they attain given output with the minimum level of inputs). Such an optimal relationship between inputs and outputs defines the production possibility frontier. The potentially sub-optimal behaviour of a producer is modelled using the concept of technical efficiency.

The technical efficiency of a given producer is defined as the ratio of the producer's actual (observed) production and the optimal level of production given available inputs and the present state of technology. The greater the discrepancy between the two production figures, the greater the inefficiency and the lesser the technical efficiency. In this basic formulation the notion of technical inefficiency is purely non-random.<sup>5</sup> A more realistic approach to describe the production relationship allows for non-randomness that accounts for the possible external influences producers have to face. These influences (such as shocks to unemployment or bureaucratic procedures) do not enter their production function directly and producers take them as given. External shocks may affect the technical efficiency despite optimal utilisation of inputs in the production process.

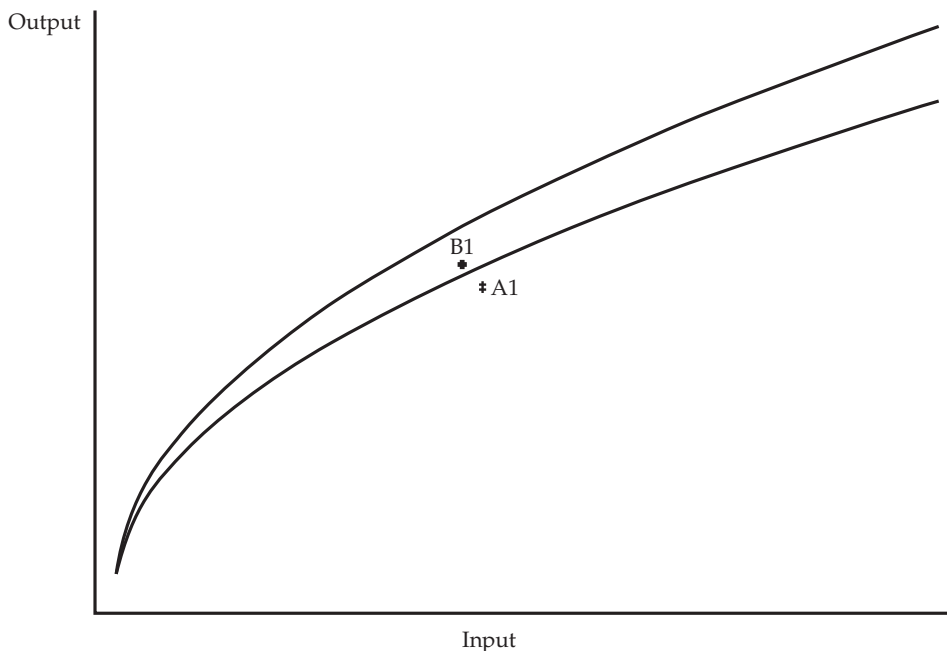
Figure 1 provides some more insight into the deterministic notion of technical efficiency. Figure 1 shows the situation of two producers, A1 and B1, in the case of a single input. Both curves represent production frontiers defined by

<sup>3</sup> Economic efficiency has also been examined using other methods: Data Envelopment Analysis and Free Disposable Hull Analysis. A comparison of the methods can be found in Fried, Lovell and Schmidt [1993] and Kumbhakar [2000].

<sup>4</sup> See Bonnacorsi and Daraio [2003], Guan and Wang [2004], Mensah and Werner [2003], Mizala, Romaguera and Farren [2002], and Warning [2004].

<sup>5</sup> For a more formal description of technical efficiency, see the Appendix and the references in this section.

**Figure 1. Production frontier for type A units and type B units, a single input case**



Source: Pereira [2007].

optimal allocations of inputs given the different technologies available to each producer, reflected in different slopes. The outer frontier stands for the optimal production allocations for producer B1, the inner frontier refers to producer A1.

Taking our definition of technical efficiency, producer A1 is relatively closer to his (inner) production frontier as compared to producer B1. On the other hand, the output of producer B1 is greater than that of A1, so that by simple absolute measures producer B1 outperforms producer A1, while by the efficiency criteria the ranking would be reversed. The illustration assumes each producer uses different technology.

SFA generates so-called *modified* efficiency predictors by comparing the two producers with respect to a common frontier that has been estimated from the sample of both producers, under the assumption that the observed production is always less than or equal to optimal production. In other words, SFA in its simplest set-up assumes identical technology (production function) for all producers. The rankings of output performance in such a set-up would depend on the inputs employed by individual producers and on the external shocks each producer faced.

The general issue of applying SFA in the area of efficiency of secondary education involves an assumption of the existence of one particular perfectly measurable output of educational process at a secondary school.<sup>6</sup> There are many views on the purpose of secondary education and thereby on the output of the educational process. With regard to the often cited primary mission of grammar schools of 'preparing graduates for the tertiary level of education', we consider the share of grammar school graduates admitted to a faculty as the relevant output variable.

There are, however, several drawbacks to the suggested output variable. First, data on the admission process to the tertiary level of education usually cover only a subset of grammar school graduates – those who apply. The efficiency measure based on shares of *applicants* admitted to the tertiary level of education would thus ignore those who do not intend to continue their studies at the tertiary education level. Second, the comparison of the shares (regardless of whether related to all graduates or applicants only) across grammar schools is problematic since graduates from different grammar schools can in general apply for study programmes with different access demands. The shares of students admitted are then influenced by the individual universities (programmes at universities) and do not relate to the efficiency of grammar schools. The next section focuses on the construction of the output variable that deals with the two above-mentioned issues.

## Data

The numbers of students and other characteristics of grammar schools for the period 2000–2004 were obtained from the database Applicant (Uchazeč) involving applicants and those admitted to a university which is collected yearly by the Institute for Information in Education.<sup>7</sup> In the article, we restrict our attention to grammar school graduates who have applied to full-time study in a programme leading to a bachelor or master's degree.

The district data on external variables assumed to influence the educational process at grammar schools and graduates' application decisions come from several sources. The average hourly wages for the secondary- and tertiary-educated populations for a particular year and district were estimated from the database Information System on Average Wage. Data on foreign direct investments were taken from the data published by the Czech National Bank. Data on unemploy-

---

<sup>6</sup> In contrast to SFA, DEA can deal with more outputs when assessing efficiency. However, in case of panel data SFA is preferable.

<sup>7</sup> In most cases a school corresponds to an administrative unit defined by its IZO (Identifikační znak organizace – The Identification Code of an Organisation). In some cases, the IZO consisted of two different schools, which were distinguished using their identification codes of study

**Table 2. Basic characteristics of variables, n=1,572**

Variable	Mean	Standard deviation	Minimum	Maximum
Probability of admission index	0.63	0.17	0.00	1.00
Normalised probability of admission	1.88	0.41	0.00	3.28
Students/teacher	14.12	2.29	2.14	38.36
Students/class	27.33	3.68	6.50	31.93
Students/school management	185.03	69.79	7.27	998.00
≥3 foreign languages	0.02	0.09	0.00	1.00
Number of students	421.93	186.23	8.00	1,089.00
Share of women in teaching staff	0.69	0.10	0.20	1.00
Unemployed/population	0.05	0.02	0.01	0.13
Unemployed white collars/blue collars	0.13	0.03	0.08	0.25
Vacancies/1000 inhabitants	1.53	0.62	0.44	3.24
Hourly wage tertiary educated/secondary educated	1.71	0.36	0.60	3.98
Foreign direct investment/inhabitant	127.79	155.74	0.00	528.39
District>80 000 pop.	0.11	0.31	0.00	1.00
District >170 000 pop.	0.09	0.28	0.00	1.00
District >500 000 pop.	0.16	0.37	0.00	1.00
Faculty in a district	0.40	0.49	0.00	1.00

Source: 'Applicant' database, Information System on Average Wages, Czech National Bank, Ministry of Education and Labour Offices (authors' calculations).

ment are regularly published by the district Labour Offices. The indicators of a presence of a faculty within a particular district were constructed from lists of tertiary education institutions maintained by the Ministry of Education. Table 2 reports the list of variables and their basic characteristics.

The output variable closely relates to the share of grammar school graduates admitted to a study programme. We have constructed this share by merging two datasets – one including the number of students in each class and one including applicants and those admitted from a particular grammar school. These two data sets were merged at the level of grammar schools. Relating the share of admitted to all graduates from a grammar school thus also takes into account those not applying. Furthermore, since the admission processes are generally not

comparable across university programmes, they do not provide comparable information about an average grammar school graduate from a particular grammar school and thus about the particular grammar school. For that reason we consider two output variables – the normalised probability of admission and the admission probability index.

‘Normalised admission probability’ is the share of admitted graduates from a grammar school divided by the average probability of admission for all the university programmes they apply to. That is, the original share of those admitted is normalised to account for different average preferences of graduates from different grammar schools. Assume, for instance, two grammar schools of the same efficiency (i.e. the same level of direct inputs and external factors) along with the same number of students and those applying to a study programme at a university. Graduates from the first grammar school apply solely to the programmes that admit all applicants. On the other hand, graduates from the second grammar school apply to highly in-demand programmes with a low probability of being admitted. Were the different excess demands for different programmes neglected, our analysis could falsely suggest the first grammar school to be more efficient. Normalisation accounts for the different admission probabilities between study programmes and ensures comparability of admittance shares. Normalised admission probability is computed as follows:

$$p_{i,t}^{norm} = \frac{admitted_{i,t} / graduates_{i,t}}{\sum_{j(i,t)} \left( \frac{\sum_{m(j(i,t))} excess\_demand_{m(j(i,t))}}{M(j(i,t))} \right)},$$

where  $j(i,t)$  stands for graduate  $j$  from grammar school  $i$  in year  $t$ ,  $excess\_demand_{m(j(i,t))}$  represents the excess demand for study programme  $m$  applied for by student  $j$ , and  $M(j(i,t))$  denotes the number of study programmes applied for by student  $j$ .

Another approach is pursued in the construction of the ‘admission probability index’. Like the normalised index, for each university study programme  $m$  opened in a given year, we derived the excess of demand above the programme’s capacity.<sup>8</sup> Afterwards, we calculated the average excess demand for the study programmes the graduate applied to and used it as a weight for the derivation of the modified ratio of accepted to applying grammar school graduates. This procedure generated variable that reflects both different chances of admission for individual study programmes and admission probabilities for students of the particular grammar school in question. Given the fact that the admission probability index is an indicator conditioned by the application decision of fresh grammar school graduates, we consequently multiplied it by the likelihood of application:

<sup>8</sup> The excess demands included *all* applicants, not just fresh grammar school graduates. Less in-demand study programmes with a capacity exceeding the number of applicants were assigned a value of 1.

the ratio of university programme applicants to the total number of grammar school graduates for each grammar school and year. Holding the cohort size, the number of graduates and accepted fixed, the resulting index variable rises, once the graduates on average apply for more in-demand study programmes. The admission probability index is formally expressed as

$$p_{i,t}^{index} = \frac{\sum_{j(i,t)} \left( \frac{\sum_{m(j(i,t))} excess\_demand_{m(j(i,t))}}{M(j(i,t))} * Ind\_admitted_{j(i,t)} \right)}{\sum_{j(i,t)} \left( \frac{\sum_{m(j(i,t))} excess\_demand_{m(j(i,t))}}{M(j(i,t))} \right)} * \frac{J_{i,t}}{N_{i,j}},$$

where,  $Ind\_admitted_{j(i,t)}$  corresponds to a binary indicator equal to 1 if individual  $j$  has been accepted to a university programme in a given year  $t$ ,  $J_{i,t}$  is the total number of applicants for university programmes for given grammar school  $i$  and year  $t$ , and finally  $N_{i,t}$  is the total number of graduates from given grammar school  $i$  and year  $t$ .

Both indicators reflect the differences between the entrance exams applied by individual study programmes and the specific preferences for particular places of study. To check for the robustness of our results, the analysis will be performed on both output variables.

The dependent variables include factors directly entering the educational process and external factors that work indirectly through it, such as local labour markets and the corresponding availability of qualified teaching staff, or socio-economic background and the aspirations of secondary school students. Factors directly entering the educational process include the teacher/student ratio,<sup>9</sup> the average class size, the number of directors and deputy directors per student as a proxy for management size, and the total number of students. Variables typically employed within the context of education production, such as the average experience of teaching staff, schools' capital investments over recent years, the number of specialised classes, or the number of teachers under 30 years of age could not be included given the data restrictions.

External factors relate to the school-specific effects that are assumed not to directly enter the educational process, the economic situation in a district, and the university presence in the area. The former are represented by the share of women in teaching staff and the number of students learning more than two foreign languages. The economic situation in a district is approximated by the ratio of the hourly wage of university to secondary-school graduates, the district unemployment rate,<sup>10</sup> the unemployment rates for white- and blue-collar workers,

<sup>9</sup> For the above-mentioned variable the full-time equivalents have been employed.

<sup>10</sup> The yearly district unemployment rates are not available owing to the absence of data on the district labour force. Therefore, we chose a proxy assuming a similar demographic structure in individual districts.

and foreign direct investments. The university presence is expressed by a binary indicator equal to 1 if a given district is home to a university. The three dummy variables specifying four location sizes capture those characteristics that have not been explicitly accounted for in our model, yet might work towards the technical effectiveness of grammar schools (e.g. the population's education structure or the expansion of tertiary education).<sup>11</sup>

The last explanatory variables represent dummies for individual years, which should capture forces simultaneously affecting all grammar schools within a given year; such forces, for example, as the increasing capacity of tertiary education sector Matějů and Straková [2006] or the business cycle.

As mentioned above, the choice of explanatory variables has been to some extent constrained by data restrictions. A key dimension in this respect constituted the absence of information on the innate and retrieved abilities of students and on their socio-economic background. Selection in education [Willis and Rosen 1979] can lead to the seemingly superior results and support of a sub-group of schools, which could in fact be attributable to the better starting position, abilities, and/or socio-economic background of their students, rather than to the schools' actual higher quality.<sup>12</sup> The analysis of technical efficiency within the context of Czech grammar schools is nonetheless less likely to be sensitive to the above-mentioned concerns. There are several reasons for this assertion. First, one might expect that larger discrepancies exist between grammar schools and other types of secondary schools (vocational schools, specialised secondary schools) than between grammar schools themselves. Second, the basic unit of observation is not the individual, but the school, so that part of the differences at the individual level is cancelled out. Despite these facts we still lack information on the potential effect on efficiency of the individual-level heterogeneity within schools.

The resulting specification is thus formulated as follows:

$$\ln(y_{i,t}) = \beta_0 + \beta_1 \ln(students/teacher)_{i,t} + \beta_2 \ln(students/class)_{i,t} \\ + \beta_3 \ln(students/management)_{i,t} + \beta_4 \ln(no\_of\_students)_{i,t} + v_{i,t} - u_{i,t}$$

where the dependent variable represents either the normalised admission probability (Model 1) or the admission probability index (Model 2). We employ the natural-logarithm specification when it is reasonable to expect a diminishing marginal contribution of a variable to output.

The random term  $v_{i,t}$  is assumed to be independently distributed, follow-

<sup>11</sup> We defined four location sizes:  $\geq 500\,000$  inhabitants (Prague); 170 000–500 000 (Brno and Ostrava), 80 000–170 000,  $< 80\,000$ . These dummy indicators capture all the residual effects at the relevant town size above those already modelled by regional economic conditions variables.

<sup>12</sup> This mechanism could also partially explain superior results of grammar schools in comparison to the remaining secondary schools in the Czech Republic [Münich 2004].

ing the normal distribution  $v_{i,t} \sim N(0, \sigma_v^2)$ . The disturbance  $u_{i,t}$ , is assumed to be independent of  $v_{i,t}$  and to follow the left-truncated normal with the mean  $\mu_{i,t}$ , i.e.  $u_{i,t} \sim N^+(\mu_{i,t}, \sigma_u^2)$ . The mean value of the left-truncated normal distribution is modelled as:

$$\begin{aligned} \mu_{i,t} = & \delta_0 + \delta_1(\text{teachers : ratio\_of\_women})_{i,t} + \delta_2(\geq 3\_frgn\_languages)_{i,t} \\ & + \delta_3(\text{un\_rate})_{i,t} + \delta_4(\text{vacancies}/1000inh)_{i,t} + \delta_5(\text{un\_white\_coll/blue\_coll})_{i,t} \\ & + \delta_6(\text{wage\_univ/second})_{i,t} + \delta_7(\text{fdi}/1000inh)_{i,t} + \delta_8(\text{city}<80thsnd)_{i,t} \\ & + \delta_9(\text{city}<170thsnd)_{i,t} + \delta_{10}(\text{city}<500thsnd)_{i,t} + \delta_{11}(\text{univ\_in\_city})_{i,t} + \sum_{k=2004}^{2009} \delta_k(I_{year}) \end{aligned}$$

### Estimation results

Table 3 presents the estimation results for the specifications employing both the normalised admission probability (Model 1) and the admission probability index (Model 2). Both models have been evaluated with the econometric software Frontier 4.1<sup>13</sup> using the maximum likelihood method, which searches for a maximum of the likelihood function derived from the distribution of random variable  $\varepsilon_{i,t} = v_{i,t} - u_{i,t}$  [Pereira 2007]. The relative importance of the individual parts – i.e. the symmetric part  $v_{i,t}$  and the technical inefficiency component  $u_{i,t}$  – can be scrutinised with the parameter estimate  $\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$ . If  $\gamma \rightarrow 0$ , the technical inefficiency variance  $\sigma_u^2$  constitutes a relatively negligible part as compared to the variance of the symmetric (random) part  $v_{i,t}$ , and the model could thus be estimated with the traditional OLS. Regardless of specification, however, the estimate  $\gamma$  is close to one and it is possible to reject the hypothesis  $\gamma=0$  at 1-percent level of significance. The application of the OLS thus seems to be inappropriate.

Our results are robust with respect to the output variable. As Table 3 shows, signs and statistical significance remain isomorphic regardless of specification for the majority of the coefficient estimates. In this respect thus selected output indicators seem to capture basic characteristics of education production on Czech grammar schools. One should nonetheless keep in mind that the coefficient levels should not be directly compared, given the different left-hand-side variable.

Turning to the individual coefficients, the upper part of Table 3 lists the explanatory variables that directly enter the production function. The signs and statistical significance of the coefficients suggest the relative stability of the estimated coefficients. An increasing number of students per teacher lowers the production of high quality graduates. In other words, within the context of our output variables, and all other things held constant, grammar schools with on average a smaller student/teacher ratio tend to be more successful at getting their fresh graduates into Czech universities. On the other hand, relatively close to zero and statistically insignificant are the variables approximating the size of management (*student/mng*).

<sup>13</sup> More details on the actual functioning of the software can be found in Coelli [1996].

**Table 3. Regression results for Model 1 and Model 2, i=344, t=5, N=1,572.**

	<i>Model 1 Normalised admission probability</i>			<i>Model 2 Admission probability index</i>		
	Coefficient	Standard error	t-ratio	Coefficient	Standard error	t-ratio
$\beta_0$	-0.27***	0.11	-2.24	-1.72***	0.13	-13.34
Direct inputs						
students/teacher	-0.18***	0.03	-5.5	-0.14***	0.04	-3.51
students/class	0.22***	0.05	4.68	0.39***	0.06	7.11
students/mng	-0.01	0.02	-0.3	0.00	0.02	-0.2
no of students	0.14***	0.02	8.4	0.1***	0.02	5.21
External factors						
$\delta_0$	-1.88***	0.31	-6.04	-3.93***	0.88	-4.47
teachers' gender	-2.87***	0.33	-8.78	-3.76***	1.03	-3.67
$\geq 3$ languages	-3.04***	0.25	-12.25	1.03	1.12	0.93
unempl. rate	5.19***	1.13	4.6	15.82***	1.29	1.23
vacations/ 1000 inhabitants	0.03	1.00	0.03	0.46	1.00	0.46
unemployed white collars/blue collars	-7.81***	1.35	-5.78	0.5	1.27	0.4
wage univ.edu./ high school edu	0.83***	0.08	10.01	1.02***	0.14	7.53
FDI/ 1000 inh.	-2.36***	0.45	-5.26	-0.55	0.61	-0.9
city>80thd.	-1.01***	0.15	-6.79	-0.55	0.35	-1.55
city>170thd.	0.36***	0.13	2.7	-0.26	0.38	-0.69
city>500thd.	1.99***	0.22	9.19	2.07***	0.2	10.23
univ. in city	2.56***	0.3	8.43	2.29***	0.31	7.31
$\Sigma^2$	0.55***	0.05	11.59	1.96***	0.15	12.86
$\gamma$	0.98***	0.00	362.46	0.99***	0.00	1315.21
Maximum likelihood			356.96			-75.94

Notes: \*, \*\*, \*\*\* correspond to 10%, 5% and 1% significance level.

The coefficient  $\sigma^2$  equals a variance estimate of the random term  $\varepsilon_{it}$ .

The coefficients and other statistics for individual years can be provided by the authors upon request.

Source: 'Applicant' database, Information System on Average Wages, Czech National Bank, Ministry of Education and Labour Offices (authors' calculations).

The coefficient estimate on the average class size (*students/class*) turned out to be positive and statistically significant at a 1-percent level. One explanation for the positive relationship between the average number of students in class and their success rate at the entrance exams could be a positive correlation between the average class size and the total number of students at a given grammar school, which in turn partly approximates a given school's capital endowments. One could make a similar argument in the case of the total number of students, which remains again significant at a 1-percent level.

One of the major motivations of our analysis lies in the quantification of the role of external factors on the technical efficiency of grammar schools in the Czech Republic. The resulting estimates can be found in the lower part of Table 3. Note that the negative signs in front of external factors indicate a negative link between a given external (indirect) characteristic and technical *inefficiency*, that is, a shift in the technical inefficiency mean closer to zero.

The coefficients at variables related to school-specific effects (*teachers' gender* and  $\geq 3$  *languages*) are negative and statistically significant. As already noted, the problems with individual school data availability weaken the interpretation of the estimated school-specific effects. So, estimated coefficients probably capture broader links than the mere effect of feminisation and the effect of foreign languages taught at grammar schools. A high share of students studying more than two foreign languages can, for example, capture a wider spectrum of subjects that a particular grammar school provides, leading to more opportunities in terms of the university programmes to which a student can apply and consequently to the greater efficiency of such schools.

Regarding the influence of local economic and other conditions, the situation is easier because we can control for a variety of such effects. Grammar schools located in districts with a higher unemployment rate report higher expected inefficiency relative to otherwise identical grammar schools that differ exclusively in terms of a lower district unemployment rate. One explanation for the dominating negative impact of the unemployment might be the combination of a difficult economic situation and the migration of a part of the educated population to more prosperous regions, which in turn translates into students' socio-economic background. A related mechanism could be the meagre attractiveness the region holds for well-educated teachers and consequently the shortage of such teachers.

The above-mentioned relationship can be partly supported by the estimated negative relationship between FDI and technical inefficiency (Model 1 specification only). A higher rate of FDIs within a particular district and the related growth of economic activity can boost incomes and the socio-economic background of students, and could even enhance the quality of the teaching staff.

The coefficient estimates for the non-manual/manual occupations unemployment ratio (*unempl. white collars/blue collars*) and for the relative wage of university-educated employees to those with secondary as the highest education

achieved point to other possible mechanisms at work. Our estimates are consistent with the hypothesis that qualified teaching staff in districts with a relatively abundant supply of non-manual job opportunities and/or higher relative wages might be more motivated to take the opportunity and look for employment in other sectors than education, which in turn translates into an increase in technical inefficiency.

Despite the fact that the model specification with the admission probability index fails to reach as unequivocal results as Model 1, it still remains in line given the statistical insignificance of the white collars/blue collars unemployment ratio, and a positive and significant coefficient estimate for relative wages. The coefficient estimates for the unemployment ratio and relative wages thus suggest a notable influence of external factors such as, for example, alternative market opportunities that tend to be met by the well-qualified grammar school staff, while the mechanism working through students' socio-economic background or aspirations seems to be less pronounced.

A rather surprising result seems to be the estimated dummies for grammar school location size. Taking into account other available inputs and varying external environments, grammar schools from relatively smaller towns on average outperform their counterparts from larger localities (Model 1 specification). These discrepancies might be a consequence of, for example, the different aspirations of grammar school students from smaller municipalities, whose source has not been captured by other variables in our model.

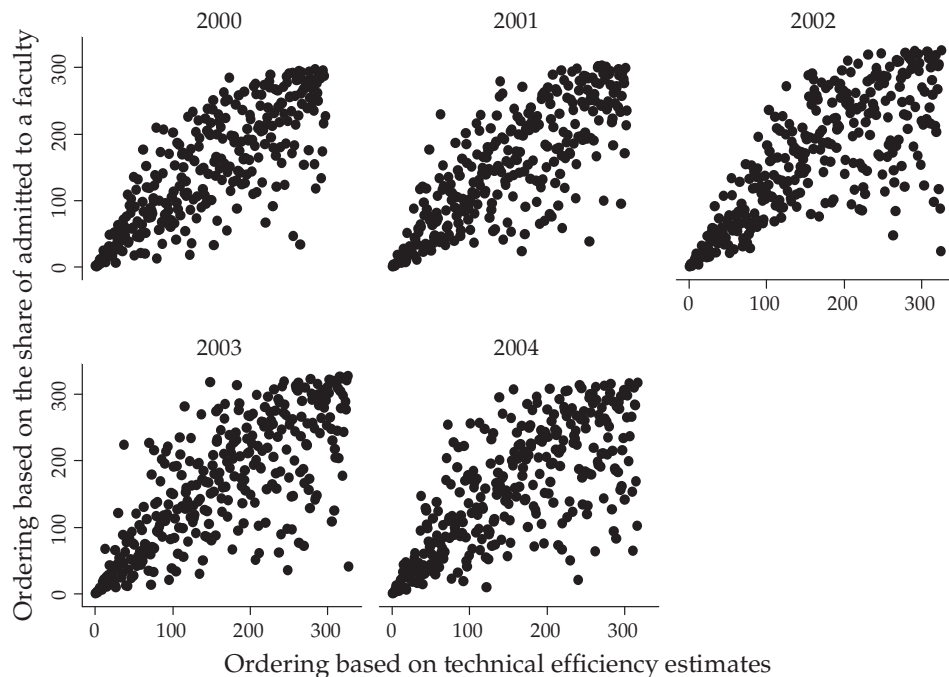
### **Do we need new measures of technical efficiency?**

In this sub-section we compare the measure of technical efficiency of grammar schools based on SFA with a simple measure based on the share of students admitted to a faculty. Figure 1 shows how the ordering of grammar schools according to the success of their graduates in the admission process differs from the ordering based on SFA. Both orderings are averages over the years 2000–2004. Less efficient grammar schools (according to SFA) and grammar schools with low shares of graduates admitted are ordered lower than more efficient grammar schools and grammar schools with high shares of admitted students respectively. The presented estimates of technical efficiency are based on the Model 1 specification.

Points on the 45-degree line starting at origin represent grammar schools that are in the same position according to both orderings. Points below the line represent grammar schools that rank higher by technical efficiency than by the share of admitted students, and vice versa.

The Spearman's rank correlation coefficient of the two orderings is 0.77. So, the two measures are not as close and to build the ordering of grammar schools on the share of admitted students could be misleading. Figure 1 suggests that if

**Figure 2. Comparison of gymnasium ordering over 2000–2004**



Source: 'Applicant' database, Information System on Average Wages, Czech National Bank, Ministry of Education and Labour Offices (authors' calculations).

we take into account the preferences of grammar school graduates for faculties, grammar school characteristics, and the external factors that influence the educational process, some grammar schools may prepare their graduates to enter the tertiary level of education better than others. Grammar schools should be, therefore, assessed using more sophisticated methods such as SFA.

The difference in both types of ordering increases over time, as suggested by Figure 2. The variance in the technical efficiency of grammar schools that exhibit similar ratios of graduates admitted to a university is higher in 2004 than in 2000.<sup>14</sup> The need to assess the efficiency of grammar schools in a more complex way therefore increases.

<sup>14</sup> The Spearman's correlation coefficient of the two rankings of a grammar schools decreases from 0.8 in 2000 to 0.72 in 2004.

## Conclusion

Our study has presented the prospects for applying stochastic frontier modelling within the domain of secondary school evaluation. Since the pending education policy reform in the Czech Republic foresees the introduction of a unified school-leaving-exam system and the relative performance of secondary schools would thus likely become evaluated (like in the UK, see Griffith [2008]) based on average school-leaving grades and/or the university enrolment rate of the school's graduates, we focused on the potential performance indicators that would also reflect different student abilities and economic situation within the school's district such as the local unemployment rate or relative wages, that is, factors that usually lie out of reach of individual schools. Plain comparisons of student results might in similar situations provide a misleading picture of school's performance.

We focused on grammar schools in the Czech Republic over the years between 2000 and 2004. Apart from common school characteristics (student/teacher ratio, average class sizes, etc.), the enrolment success of grammar school graduates becomes likewise indirectly linked to the economic situation within the district. The negative role of the unemployment is generally consistent with the previous findings on the substantial role of students' socio-economic background in the Czech education system [Matějů, Řeháková and Simonová 2003; Matějů and Straková 2006]. The exodus of qualified staff from the education sector represents yet another potentially significant indirect mechanism related to the general economic situation. In this case and consistent with expectations, our results suggest that, other things held equal, grammar schools located in districts with higher relative wages and a comparatively low unemployment rate in non-manual occupations report relatively lower transition rates from secondary to tertiary education.

The main motivation of this study is to contribute to the ongoing discussion on effectiveness and performance measurements in the Czech education system. We showed that if one ignores the varying input quantities entering the education process and the effects of the environment in which the particular secondary school units operate, the ranking of schools based exclusively on the schools' graduates' success rate in the university application process can vary significantly from the rankings constructed using the technical efficiency estimation. This discrepancy has been moreover increasing over the observed period.

MICHAL FRANTA *is completing his PhD in economics at CERGE-EI. He currently works at the Czech National Bank.*

TOMÁŠ KONEČNÝ *obtained his PhD in economics in 2009 from CERGE-EI. His recent publications include an article titled 'Student Finance and Access to Higher Education' (in Czech), co-written with P. Matějů, S. Weidnerová and H. Vossensteyn in Sociologický časopis/Czech Sociological Review.*

## References

- Aigner, D. J., C. A. Lovell and P. Schmidt. 1977. 'Formulation and Estimation of Stochastic Frontier Production Functions.' *Journal of Econometrics* 6: 21–37.
- Bonaccorsi, A. and C. Daraio. 2003. 'A Robust Nonparametric Approach to the Analysis of Scientific Productivity.' *Research Evaluation* 12 (1): 47–69.
- Coelli, T. J. 1996. 'A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation.' *CEPA Working Papers* 7/96. Armidale, Australia: Department of Econometrics, University of New England.
- Fried, H. O., C. A. Knox Lovell and Shelton S. Schmidt. 1993. *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.
- Griffith, A. S. 2008. 'A Proposed Model for Assessing Quality of Education.' *International Review of Education* 54 (1): 99–112.
- Guan, J. C. and J. X. Wang. 2004. 'Evaluation and Interpretation of Knowledge Production Efficiency.' *Scientometrics* 59 (1): 131–155.
- Kumbhakar, C. and C. Lovell. 2000. *Stochastic Frontier Analysis*. Cambridge, UK: Cambridge University Press.
- Matějů, P., B. Řeháková and N. Simonová. 2003. 'Transition to University under Communism and after Its Demise. The Role of Socio-economic Background in the Transition between Secondary and Tertiary Education in the Czech Republic 1948–1998.' *Sociologický časopis/Czech Sociological Review* 3: 301–324.
- Matějů, P. and J. Straková (eds.). 2006. *Nerovné šance na vzdělání. Vzdělanostní nerovnosti v České republice*. (Unequal Educational Prospects: Educational Inequalities in the Czech Republic) Prague: Academia.
- Meeusen, W. and J. van den Broeck. 1977. 'Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error.' *International Economic Review* 18 (2): 435–444.
- Mensah, Y. M. and R. Werner. 2003. 'Cost Efficiency and Financial Flexibility in Institutions of Higher Education.' *Journal of Accounting and Public Policy* 22 (4): 293–323.
- Mizala, A., P. Romaguera and D. Farren. 2002. 'The Technical Efficiency of Schools in Chile.' *Applied Economics* 34 (12): 1533–1552.
- Münich, D. 2004. 'Estimating the Impact of School Quality, Selection, and Supply on Student's Achievements: Evidence from The Czech Nation-wide Testing of Youth.' Retrieved 31 December 2008 ([www.cerge.cuni.cz/pdf/gdn/RRCIII\\_26\\_paper\\_01.pdf](http://www.cerge.cuni.cz/pdf/gdn/RRCIII_26_paper_01.pdf)).
- Pereira, M. C. 2007. 'A Stochastic Frontier Analysis of Secondary Education Output in Portugal.' Working paper series (Federal Reserve Bank of Atlanta) 6.
- Stupnytskyy, O. 2004. 'Secondary School Efficiency in the Czech Republic.' CERGE-EI, mimeo.
- Warning, S. 2004. 'Performance Differences in German Higher Education: Empirical Analysis of Strategic Groups.' *Review of Industrial Organization* 24 (4): 393–408.
- Willis, R. J. and S. Rosen. 1979. 'Education and Self-Selection.' *Journal of Political Economy* 87 (5): 7–36.

## Appendix: A mathematical derivation of the estimated equation

Technical efficiency of the producer  $i$  in the period  $t$ ,<sup>15</sup>  $TE_{i,t}$ , relates the optimal level of production (given technology and inputs),  $f(\beta, x_{i,t})$ , and observed production  $y_{i,t}$ :

$$TE_{i,t}(y_{i,t}, \beta, x_{i,t}) = y_{i,t} / f(\beta, x_{i,t}) \leq 1, \quad (1)$$

where  $\beta$  is a vector of  $k+1$  technology parameters and  $x_{i,t}$  denotes a vector of  $k$  direct inputs.

The production function  $f(\beta, x_{i,t})$  is referred to as a *deterministic* production function since there is no role for frictions. A more realistic approach to describe the production relationship allows for a stochastic component capturing external shocks producers are exposed to. The stochastic frontier of production possibilities is defined as  $f(\beta, x_{i,t}) \cdot \exp(v_{i,t})$ , where  $v_{i,t} \sim \text{i.i.d. } N(0, \sigma_v^2)$ . Technical efficiency is then:

$$TE_{i,t}(y_{i,t}, \beta, x_{i,t}) = y_{i,t} / (f(\beta, x_{i,t}) \cdot \exp(v_{i,t})). \quad (2)$$

Assuming Cobb-Douglas functional specification of the deterministic production function, defining technical efficiency as  $TE_{i,t} = \exp(-u_{i,t})$  and transforming equation (2) into the log form yields estimation equation:

$$\ln y_{i,t} = \beta_0 + \sum_k \beta_k \ln x_{k,i,t} + v_{i,t} - u_{i,t}, \quad (3)$$

where  $u_{i,t} \geq 0$ .<sup>16</sup>

In what follows we assume that  $u_{i,t} \sim \text{i.i.d. } N^+(\mu, \sigma_u^2)$  (i.e. normal distribution truncated from left at 0). The specification (3) is estimated by maximum likelihood. The inputs entering directly into the education process ( $x_{i,t}$ ) are distinguished from the external factors that may influence the output of the education process indirectly. These external effects enter the model by modelling the mean of the inefficiency  $\mu$ . We assume a linear specification, i.e.,

$$\mu_{i,t} = \delta_0 + \sum_l \delta_l z_{l,i,t}, \quad (4)$$

where  $\delta_l$  denotes a parameter capturing the effect of  $l$ -th external variable,  $z_{l,i,t}$ , for grammar school  $i$  and the period  $t$ .

<sup>15</sup> In contrast to cross-section data, panel data make it possible to obtain more precise estimates of the technical efficiency as the estimated parameters are consistent over time and not over the number of production units.

<sup>16</sup> Note that  $u_{i,t} = -\ln TE_{i,t} \sim 1 - TE_{i,t}$ . The term  $u_{i,t}$ , therefore, refers to inefficiency.