ISSN 1503-299X

WORKING PAPER SERIES

No. 7/2003

ESTIMATING THE ELASTICITY OF LABOUR SUPPLY TO AN ENTERPRISE UTILIZING A QUASI–NATURAL EXPERIMENT

Torberg Falch

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
www.svt.ntnu.no/iso/wp/wp.htm

Estimating the elasticity of labour supply to an enterprise utilizing a quasi–natural experiment

TORBERG FALCH

Department of Economics, The Norwegian University of Science and Technology,

N–7491 Trondheim, Norway

E-Mail: Torberg.Falch@svt.ntnu.no

May 2003

Abstract

This paper utilizes institutional features to identify the supply of labour directed towards individual enterprises. The labour market for Norwegian teachers is characterized by a high degree of central regulations. In the empirical period, the only variation in the wage level was determined centrally, and together with information on whether there is excess demand, this identifies the elasticity of labour supply. Using a sample selection model with fixed school effects, the estimated supply elasticity faced by the individual schools is close to unity and seems to be robust with respect to the model specification.

Keywords: Labour supply elasticity; teacher supply; sample selection

JEL-codes: C23; C24; I29; J22

* This paper is part of a project funded by the Norwegian Research Council. The Ministry of Education and Research and the government educational offices in Nordland, Troms and Finnmark are acknowledged for kindly making school level data available for this analysis. The rest of the data are obtained from the Norwegian Social Science Data Service. Comments from Kåre Johansen, Alan Manning, Bjarne Strøm, Jean–Pierre Urbain, and participants at the congress of the European Economic Association in Venice, the conference of the European Association of Labour Economists in Jyväskilä, and the Nordic Econometric Meeting in Sandbjerg are greatly acknowledged. The author bears the responsibility for the analysis and the conclusions that are drawn.

1. Introduction

The shape of the supply curve directed towards individual firms is the core of the recent debate on the effects of minimum wages. If minimum wages have a positive effect on employment, the firms must face an upward sloping supply curve. But it is inherent difficulties in estimating the structural parameters of the labour supply function. In the minimum wage literature, reduced form equations are estimated, which may give some indication of the form of the supply curve. For example, Card and Krueger (1994) find that a 19 percent increase in the minimum wage in New Jersey increased the employment in the fast-food restaurants by about 13 percent. If one assumes that the employment in all stores were on the supply curve both prior and after the minimum wage rise, the implied elasticity of labour supply towards this labour market area is about 0.7. The evidence in more recent studies of the minimum wage increase in New Jersey, however, indicates that the effect was close to zero, see Card and Krueger (2000) and Neumark and Wascher (2000). Boal and Ransom (1997) and Dickens et al. (1999) show that a minimum wage increase will have a nonlinear effect on employment in a monopsony market because more firms will be "demand constrained" as the minimum wage increases. Dickens et al. (1999) find an average elasticity of the minimum wage with respect to UK employment of about 0.3, but the elasticity seems to be above 2 in the industries with the lowest minimum wage.

The first attempt to estimate the structural parameters of labour supply directed towards individual firms seems to be Nelson (1973). Using cross-section data, Nelson identifies the labour supply by utilization of a population density measure. Later studies argue that it is important to have data with a panel structure in order to control for unobserved factors as remote location and other potential sources of compensating differentials, see for example Sullivan (1989) and Boal (1995). Sullivan (1989) estimates the supply elasticity of nurses directed towards individual hospitals, and identifies shifts in demand by assuming an exogenous level of hospital caseload.[1] Staiger et al. (1999) argue that the caseload was endogenous in the period studied by Sullivan (1989). The main problem with studies of this

[1] The estimated elasticity of labor supply varies considerably between the different studies. Nelson (1973) finds large elasticities for most US states in 1929, with a national average of 19. In a study of coal mining in the first decades of the 20th century, Boal (1995) finds the labor supply elasticity to be in the range 1.9–6.8 in the short run and infinite in the long run. The supply elasticity of hospital nurses found by Sullivan (1989) is in the range 1.3–3.8.

kind is that one typically does not have good instruments for wages. To identify the supply curve, it thus seems attractive to study specific institutional settings and experiments with exogenous wage changes. Staiger et al. (1999) claim that they identify the supply curve for nurses by a legislated increase in the wage level at Veteran Administration hospitals. However, their approach is not much different from the minimum wage studies and other studies of exogenous wage changes as Manning (1996). Staiger et al. (1999) argue that the employment of nurses is likely to be "supply constrained" because nearly continuous reports of shortages of nurses since World War II. Because this is aggregated information, one cannot, however, infer whether the actual adjustment is on the supply curve at all hospitals or whether this is the case only for a fraction of the hospitals. Their estimate of the supply elasticity of hospital nurses is around 0.1 and insignificant, in line with the main results in the minimum wage literature.[2]

This paper utilizes a quasi–natural experiment with centralized wage determination and a distinct measure of labour supply in the Norwegian public-sector primary and lower secondary schools to estimate the school level elasticity of teacher supply. A clear–cut rule regarding the appointment of teachers makes it possible to identify teacher supply. According to the rule, the schools have to appoint the best-qualified applicant in a vacant teacher post. When it is impossible to get a certified teacher to a vacant teacher post, the schools must employ an individual without a teacher credential on a short-term contract. Thus, utilizing data on the number of certified and non-certified teachers, I observe whether there is excess demand for teachers at each individual school, and in the case of excess demand, the size of the teacher supply is given by the number of certified teachers.

During a period of more than 40 years, the Norwegian teacher wages was solely decided by central wage bargaining. With one exception, the only sources of variation in the wage across

[2] It should be noted that the exogenous wage increase of the individual Veteran Administration (VA) hospitals studied by Staiger et al. (1999) was related to the local relative wages of nurses. In 1991, the VA went from paying the nurses based on a national scale to a system where the nurses' wages are based on local surveys. As a result, the wages rose in about two-thirds of the VA hospitals, while at the remaining hospitals, paying at the same level or above the surrounding hospitals, the wages did not change. By this rule, one would guess that the hospitals most likely to be demand constrained in their employment of nurses are not used to identify the effect of the wage change. However, variations in working conditions and workload may imply that the employment was demand constrained also in hospitals initially paying a lower wage than surrounding hospitals.

teachers have been teaching experience and the amount of formal education. The exception from this rule is the experiment utilized to identify the effect of wages on the supply. In three out of the 19 counties of the country, covering several local labour markets, teachers in individual schools with a high degree of teacher vacancies in the past get about 10 percent higher wage. Because the criteria to get the wage premium is teacher shortage of a certain extent (20–30 percent), excess teacher demand is also observed in a range of other schools both in the same geographical areas and in the rest of the country.

The estimated supply elasticity in the present paper can be interpreted as a partial short run elasticity. It is a partial effect, or a first order effect, because schools and school districts cannot influence the wage level. With decentralized wage setting, the final effect is smaller in the case of wage spillovers. The estimated elasticity is a short run elasticity because at the large majority of the schools, the teachers receive a wage premium only in a limited time period. With expectations about a permanent wage premium, the response to a wage rise would be greater, and due to frictions in the labour market, more teachers may react on wage differentials in the medium run than in the short run. In addition, an increased average teacher wage would stimulate more students to undertake teacher education.

The next section discusses the institutional setting more closely. Section 3 develops a supply function directed towards particular enterprises using a model that aggregates individual supply decisions. The empirical strategy and the data are discussed in the following sections, while the main empirical findings are reported in Section 6. I find a teacher supply elasticity close to unity. Several sensitivity tests are undertaken, and I also test whether the supply elasticity is heterogeneous. Section 7 concludes.

2. The quasi–natural experiment

The wage determination of teachers is almost completely centralized with basically a common wage schedule at each school in several European countries as for example France, Germany, Italy and the UK. In Norway, the wage of an individual teacher was solely determined by central wage bargaining up to the school year 2000–01.[3] The wage varied across teachers only with respect to education level and teaching experience, but with one exception. This exception was teachers in schools located in one of the three northernmost counties in the country (out of a total of 19 counties) with particular recruitment problems. The eligible schools within these counties were selected by a criterion based on previous teacher shortage set by the central government, and the central government paid a wage premium of about 10 percent. The school districts had no influence on which schools that were eligible of a higher wage and the selection had no financial implications for them. The details are described below.

To evaluate the effect on teacher supply, the supply must be identified. When the wage is equal at each school (except the experimental schools), it is likely to be excess demand at some schools and excess supply at other schools. With excess demand, the employment of teachers with a teacher certificate, hereafter denoted qualified teachers, is given by the teacher supply. The appointment rule of teachers is crucial in this regard as discussed in Bonesrønning et al. (2003). First, the teachers are linked to the schools and not to the school district. Second, the school districts have to appoint the best-qualified applicant in a vacant teacher post. If at least one qualified teacher wants the vacant position, a qualified teacher will fill the post. A person without the required qualifications to be certified as a teacher, hereafter denoted a nonqualified teacher, can only be employed if no qualified teachers apply for a vacant teacher post. Nonqualified teachers can only be hired for one school year. The subsequent year, the school must make the vacant post public again in order to get qualified teachers to apply. According to the national contract, representatives of the teacher union must be informed prior to every hiring decision. In this way the union is able to closely monitor that the schools act in accordance with the rule, which have been one of the cornerstones in the teacher trade union policy. Thus, observed shortage of qualified teachers a particular year accordingly reflects the state of the teacher labour market that particular year, and since one observes whether there is excess demand, one can also observe for which schools the actual adjustment is on the supply curve. In that case, the supply is identified by the number of qualified teachers.

[3] Some very limited local wage flexibility was introduced in 2001, and the wage setting will be further decentralized from 2004.

The situation is illustrated in Figure 1. The figure compares two schools with equal teacher demand, but different teacher supply. Given the centrally determined wage W^*, school 1 (described by the teacher supply function TS^1) has excess supply. The supply is greater than observed employment, and cannot be identified. The employment of qualified teachers at school 2 is determined by the size of the teacher supply, S^{2*}, and the supply is identified. If teachers at school 2 are eligible to receive a higher wage, illustrated by W^{**}, the teacher shortage is reduced if the supply curve is upward sloping as in the figure. Since the central government pay the wage rise, the teacher shortage is reduced from $S^*–S^{2*}$ to $S^*–S^{2**}$. The excess demand can be eliminated if the wage rise is high or if the teacher supply is highly elastic.

Figure 1 about here

The institutional setting described identifies the supply, and the quasi–natural experiment with exogenous determined wage variation identifies the effect of wages on the supply. The experiment includes several schools during the 1990s. Schools at which teachers have received a wage premium at least once will be denoted experimental schools. Table 1 presents the development for two stylized experimental schools. Three different systems to reduce teacher shortage have been in place during the empirical period. In 1993–94 to 1995–96, teachers in schools with more than 30 percent teacher shortage the past school year received a wage premium in nominal terms, which implies that the percentage wage premium varied with the initial wage. At mean, the wage premium was about 10 percent. For schools with 20–30 percent teacher shortage in the past school year, the rules differed in different parts of the counties. In some parts the teacher wage premium was the same as for schools with more than 30 percent teacher shortage, while in other schools the teacher wage premium was only half as large.[4] Because it is previous teacher shortage that is the criterion for a higher wage, it is known well in advance of the school year which schools that will be eligible to pay a higher wage. This also implies that there is no direct causal effect of current supply on whether there is a wage premium.

Table 1 about here

[4] The schools with the lowest wage premium are located in areas with a lower national income tax rate.

A new system took place in the school year 1996–97. Now only teachers in schools with more than 30 percent teacher shortage in the past school year were eligible for a higher wage. Thus, fewer schools had a wage premium in this period. This is illustrated by School A in Table 1. School A is included (IN) in first system in 1993–94 to 1995–96, but not included (NIN) in the second system in 1996–97 to 1997–98. Within the last system, continuing to the school year 2002–03, teachers in schools with more than 20 percent teacher shortage on average during the four last school years get a wage premium, where the wage premium is independent of region.[5]

Since increased wage is expected to raise the supply, schools with teacher shortage marginally above the criterion for paying higher wage are expected to increase employment of qualified teachers such that the school is not eligible for a higher wage the next school year. This is illustrated by School B in Table 1. School B was an IN school in 1993–94, but not the following years. In this case, the teachers at the school in 1993–94 kept their wage premium as long as the system was in place, while new teachers at the school did not get the wage premium. In such semi–included (SIN) schools, only quits and not hires are expected to differ from other schools. When a new system starts up, and the school is still not eligible to be included in the system, none of the teachers receive the wage premium. Thus, if there is no serious teacher shortage at School B in 1994–95 to 1997–98, the school will be a SIN school in 1994–95 to 1995–96 and a NIN school in 1996–97 to 1997–98.

To summarize, the experimental schools can be in three different states in a particular year; all teachers receive a wage premium, the incumbent teachers receive a wage premium, or none of the teachers receive a wage premium. Due to at least four different reasons, several schools shifted status during the empirical period. First, the requirement to be an IN school varies over time. Second, a higher wage is expected to increase the supply, and hence, may change the status for a school from IN to SIN. Third, other factors than the wage level influencing the supply may change, and finally, purely stochastic factors are likely to play a role at least in schools close to the criteria to be eligible to pay the teachers a wage premium.

[5] Up to 1997–98, the teachers could choose between the wage premium described and one year of studies at a college with full pay. In the latter case, they had to commit themselves to work at the school the following five years. Those choosing College considered the benefit from a year of study to be larger than the wage premium. Thus, using the wage premium as the measure of the benefit of being employed at an IN school may underestimate the true benefit.

The systematic factors explaining why some schools have a low teacher supply can be hard to observe. The attractiveness with this experiment is that several schools shifted status. Then it is possible to investigate the within school variation in teacher supply and wage level in an analysis with fixed school effects capturing the major systematic factors important for low supply.

3. An applied theory of labour supply

This section develops a log–linear expression of the labour supply directed towards a school utilizing McFaddens (1974) random utility function approach and an assumption on the distribution of the individual utility functions that makes aggregation simple. Because the model does not include a particular institutional setting, the model can be regarded as describing the supply towards any types of enterprises.

Assume that the well-behaved utility function of teacher b in the set of teachers B working at establishment e in the choice set E is $U_e^b = U^b(C^b, H^b; X_e)$, where C is consumption, H is hours worked, and X_e is a vector of characteristics of the enterprise (not necessarily a school because the teacher can work outside the school–sector). The budget constraint is simply written $W_e H^b = C^b$, where W_e is the wage level. Maximizing the utility with respect to C^b and H^b, subjected to the budget constraint, yields the traditional indirect utility function $U_e^b = v^b(W_e, X_e)$. Although deterministic for the individuals, the indirect utility function is assumed to be random from the observers' point of view.

$$u_e^b = v^b(W_e, X_e) + \varepsilon_e^b, \tag{1}$$

where $\varepsilon^b{}_e$ is an independent error term with mean zero.

Each teacher wishes to work at the establishment supplying the highest utility. The probability that a teacher prefers to work at school i is

$$P_i^b = P\left(u_i^b = \max_{e \in E} u_e^b\right). \tag{2}$$

If the stochastic part of the random utility function has an extreme value distribution,

$P\left(\varepsilon_i^b \le z \middle| v^b(\cdot)\right) = \exp\left(-e^{-z}\right)$, and that the property of independence of irrelevant alternatives holds, McFadden (1974, 1984) shows that (1) and (2) yields

$$P_i^b = \frac{\exp\left(v_i^b(\cdot)\right)}{\sum_{e \in E}\left(\exp\left(v_e^b(\cdot)\right)\right)}. \tag{3}$$

This is a model in which it is possible to identify the parameters of $v^b(\cdot)$.

The empirical counterpart to the supply directed towards a particular school, S_i, is the sum of the probabilities that the individual teachers wish to work at school i times their preferred working time.

$$S_i = \sum_{b \in B}\left(P_i^b H_i^{b*}\right) = \sum_{b \in B}\left(H_i^{b*}\frac{\exp\left(v_i^b(\cdot)\right)}{\sum_{e \in E}\left(\exp\left(v_e^b(\cdot)\right)\right)}\right), \tag{4}$$

where H^{b*} is the optimal hours worked. At this general form, the supply to the individual schools depends on characteristics of the individual teachers. A meaningful aggregation requires some assumptions on the distribution of the teacher utility functions. Denote the set of distinct different teacher types M, where M ≤ E. Assume that only one teacher type works in school i with more than an indefinite probability. Then (4) can be approximated by

$$S_i \approx H_i^{m*}\sum_{b \in m} b\,\frac{\exp\left(v_i^m(\cdot)\right)}{\sum_{e \in E}\left(\exp\left(v_e^m(\cdot)\right)\right)}. \tag{5}$$

The supply directed to school i depends both on the wage level and other characteristics at this particular school and at other schools with type m teachers. Normalizing the supply to school i with the supply in another school with type m teachers, say school j, yields

$$\frac{S_i}{S_j} = \frac{H_i^{m*}}{H_j^{m*}}\frac{\exp\left(v_i^m(\cdot)\right)}{\exp\left(v_j^m(\cdot)\right)}, \tag{6}$$

By a similar procedure for all schools, the model can be written

$$\ln S_i = v^m\left(W_i, X_i\right) + \ln H_i^{m*} - \left[v^m\left(W_j, X_j\right) + \ln H_j^{m*} - \ln S_j\right] = v^m\left(W_i, X_i\right) + \ln H_i^{m*} + \delta_m, \tag{7}$$

where $\delta_m = -\left[v^m\left(W_j, X_j\right) + \ln H_j^{m*} - \ln S_j\right]$ is equal for all schools employing teachers of similar type.

Even though the observed utility level of being employed at school j is higher than at school i for teachers of type m, the model implies that some of those teachers will prefer to work at the latter school. This is for some random reason ε^b that is not observable. The aggregation procedure requires that the deterministic part the utility function of type m teachers is equal, but they may differ in tastes that are unobservable. Another interpretation is that frictions in the labour market hinder the employees to fully respond to utility differences. In models with costs of searching in the market, the employers have some monopsony power, see for example Mortensen (1986) and Mortensen and Pissarides (1998). In the present model, the error term ε^b may be interpreted as following from presence of searching costs or other transaction costs.

Assuming that the relative risk aversion is equal to unity, $U_i^m = \alpha \ln C^m - e^{-\beta(\ln X_i)} (1-\gamma)^{-1} (H^m)^{1-\gamma}$, where α, β and γ are parameters, (7) may be written

$$s_i = \delta_c + \delta_m + \alpha w_i + \beta_x x_i, \tag{8}$$

where small letters denote logarithmic values, $\beta_x = \frac{1+\alpha}{1-\gamma}\beta$, and $\delta_c = \frac{1}{1-\gamma}\left((1-\alpha)\ln\alpha - \alpha\right)$ is a constant term. Equation (8) is a simple log-linear estimable model with labour supply elasticity $\alpha > 0$. It follows from the model that increased wage in school i has a positive effect on the supply of teachers because the school gets more attractive for teachers initially working at another school or outside the school-sector. With the chosen utility function, the optimal number of hours of work is independent of the wage level.

4. Empirical strategy

It is likely that whether or not a school is an experimental school is related to a large set of partly unobserved factors as location characteristics. In order to control for such time invariant variables, a model with fixed school effects is necessary. By partitioning the vector of school characteristics X into time invariant, school invariant, and time-school variant variables, (8) can be written

$$s_{it}^* = \delta_i + \delta_t + \alpha w_{it} + \beta_x x_{it} + \eta_{it}^{s*}, \tag{9}$$

where s^*_{it} is the true supply towards school i at time t, δ_t is time specific effects, δ_i is school specific effects (including the parameter δ_m), the vector x_{it} only includes variables with within school variation, and η is assumed to be a normally distributed i.i.d. error term.

Teacher supply is only observed when there is excess demand for teachers,

$$s_{it} = \begin{cases} s_{it}^{*} \text{ if } s_{it}^{*} < d_{it} \\ d_{it} \text{ if } s_{it}^{*} \geq d_{it}, \end{cases} \tag{10}$$

where d_{it} is the log of teacher demand. The model that applies to the observed sample is

$$\begin{aligned} s_{it} \left| \left(s_{it} - d_{it} < 0 \right) \right. &= \delta_i + \delta_t + \alpha w_{it} + \beta_x x_{it} + E\left[\eta_{it}^{s*} \left| \left(s_{it} - d_{it} < 0 \right) \right. \right] + \eta_{it}^{s} \\ &= \delta_i + \delta_t + \alpha w_{it} + \beta_x x_{it} + \beta_\lambda \lambda_{it} + \eta_{it}^{s}, \end{aligned} \tag{11}$$

where λ is denoted the inverse Mills ratio, see Heckman (1979). Because a large literature clearly indicates that local government income and population influences the demand for school expenditures, the selection process will not simply be captured by the constant terms of the model.

The selection process can be identified utilizing that the demand behaviour differs from the determination of supply. Generally, the selection equation can be written (notice that all schools pay the same "out of pocket" wage for equal teachers)

$$z_{it}^{*} = d_{it} - s_{it} = \mu_i + \mu_t - \alpha w_{it} + \left(\varphi_x - \beta_x \right) x_{it} + \varphi_y y_{it} + \eta_{it}^{z} = \mu_i + \theta \Upsilon_{it} + \eta_{it}^{z}, \tag{12}$$

where the φ's are parameters in the demand equation. The sample selection is identified through the variables y_{it} specific to the demand equation.

In modelling demand for teachers, the institutional setting must be taken into account. The compulsory education in Norway is the responsibility of the local governments providing several other local public services as care of the elderly, pre–schools and different local utilities. The local governments have limited possibility to influence their income level. Due to central regulation, all local governments have the same income tax rate in a system of income revenue sharing. The second most important revenue is lump-sum grants from the central government. I will use the sum of these components as the exogenous local government income. The per capita income varies between local governments due to different size (and assumed economics of scale) and an active regional policy.

Different studies have found that the age composition of the population is important for the allocation of resources in multi–purpose authorities, see for example Borge and Rattsø (1995) and Poterba (1997). In addition, there may be congestion in the consumption of the local public services. Thus, the instruments used to identify the selection of schools into the state of excess demand are local government income per capita, the shares of pre–school children and elderly in the population, and population size. Because the model includes fixed school effects, the instruments are valid only to the extent that they can explain the different developments over time in the local governments.

The selection rule is $z^*_{it} > 0$. Following Heckman (1979), the sample selection model can be estimated by estimating (13) in a first stage using a probit technique and calculate $\lambda_{it} = \left[\phi\left(\mu_i + \theta\Upsilon_{it}\right)\right] / \left[\Phi\left(\mu_i + \theta\Upsilon_{it}\right)\right]$, where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal cumulative and marginal distribution functions, respectively. However, the incidental parameter problem, derived by Neyman and Scott (1948) and elaborated in for example Chamberlain (1984) and Lancaster (2000), implies that the maximum likelihood estimator of (13) is inconsistent. A fixed effects probit model is only asymptotically consistent in the time dimension because the fixed effects are not asymptotically independent of the other coefficients of the model. As a result, the fixed effects cannot be differenced out of the equation.

Some recent papers have proposed different estimators of panel data selection models. Vella and Verbeek (1999) discuss a random effect approach. The selection equation can be estimated consistently by a probit model if the individual effects δ_i have a known distribution. This requires that the regressors are uncorrelated with the individual effects, probably not a valid assumption in the present case because the experimental schools are likely to have less favourable location than other schools. Kyriazidou (1997) discusses the case when the fixed effect selection equation is estimated in a consistent way, for example by a conditional logit model. In this case, non–censored observations for which $\Upsilon_{it} \simeq \Upsilon_{is}$ can be used to estimate the supply equation consistently. In the present paper, the supply elasticity is identified by schools in which the teachers have a wage premium some years but not all years, implying that $\Upsilon_{it} \neq \Upsilon_{is}$. By the method proposed by Kyriazidou (1997), it will be hard to identify the supply elasticity.

The problem of estimating the selection equation can also be seen as a high dimensionality problem. When the fixed effects cannot be differenced out of the equation, the only way to estimate the model, although inconsistent, is to include dummy variables for each school. Because there is a large number of schools in the present data, this is merely intractable, a problem similar to the analysis of Abowd et al. (1999). Abowd et al. (1999) seek to estimate a wage equation including both fixed individual and fixed firm effects. While one set of fixed effects can be differenced out of the model, it is not possible to difference out two sets of fixed effects. They solve the econometric problem due to high dimensionality by imposing some orthogonality conditions, which make it possible to estimate some of the coefficients of the model consistently. The idea is to find some variables $\bar{Z}$ such that $E\left[\Upsilon'\left(I-\bar{Z}\left(\bar{Z}'\bar{Z}\right)\bar{Z}'\right)\Upsilon\right]=0$. Utilizing the linearity of their model, the last fixed effects can be revealed in a second step estimation, which, however, is not possible in a nonlinear model. Instead I will here model the fixed effects. The fixed effects are approximated by

$$\mu_i = \mu\bar{Z}_i + \eta_i^{\mu}. \tag{13}$$

Sufficient conditions for unbiased estimators are $\mathrm{Cov}(\eta_{it}^{z}, \eta_i^{\mu}) = 0$ and that $\eta_{it} = \eta_{it}^{z} + \eta_i^{\mu}$ is an i.i.d. error term. In the vector $\bar{Z}$ I will include local government specific effects for local governments with more than 20 schools at average, dummy variables for the average number of schools, a full set of county dummies, three dummy variables for school type, the local government mean over the empirical period in the share of students at experimental schools, and a dummy variable for whether the school is an experimental school. With this specification, the number of observations to estimate the effect of each dummy variable is never below 100.

With this approach, the variance–covariance matrix cannot be derived exactly as suggested by Heckman (1979) and Greene (1981) because the explanatory variables enter at level in the selection equation and can be considered as entering as deviation from school mean in the supply equation. The approximation to their matrix used is described in the Appendix.

As time varying characteristics influencing teacher supply (x), I include school size measured as the logarithm of the number of students. In a sensitivity analysis, some other variables are included.

In addition to the occurrence of a wage premium, the mean wage level varies across schools because the mean teacher experience and amount of formal education vary. In the regressions below, three different specifications are used to identify the effect of the wage premium. First I simple use a dummy variable for experimental schools in which there is a wage premium. Second, I include the actual wage premium at the schools, which is independent of the basic wage and equal for all. Lastly, to estimate the supply elasticity directly, I use the percentage wage premium. The percentage wage premium is calculated based on information of the mean wage level of qualified teachers at each school.

5. The data

All Norwegian primary and lower secondary public–sector schools (first through tenth grade) are included in the sample,[6] approximately 3300 schools in 435 local governments followed over the five school years 1995–96 to 1999–2000.[7] The experimental schools are located in three counties consisting of 89 local governments. Experimental schools exist in 65 of these local governments, indicating that they are spread in many different local labour markets. It is, however, not straightforward to define the borders of labour markets. In the relevant counties, relatively few workers reside in another local government than their working place, indicating that the local governments are not too narrow measures of the local labour market.[8] Based on commuting statistics, Statistics Norway has classified 17 different labour markets areas in the relevant counties. There are experimental schools in all these areas, ranging from three to 48 percent of the schools.

Table 2 shows the number of schools with a wage premium in the different school years.[9] Few schools were eligible of a higher wage during the relatively restrictive system in 1996–97 to 1997–98, while about three times as many schools were included in the following years. The change in the criteria to be included in the system over time implies that most schools

[6] Private schools exist to a very small degree. In 1995, only 0.5 percent of the students in the counties with experimental schools were enrolled at private schools.

[7] Employment data are not available before the school year 1995–96.

[8] In 1990, 85 percent of the workers in the relevant counties worked in the local government they resided.

[9] Schools with only one observation in the empirical period (new schools or closed schools) are excluded from the analysis because they yield no information in a fixed effect model.

only had a wage premium in a short period. Table 3 shows that out of the 139 experimental schools, teachers in 60 schools received a wage premium only in one year during the empirical period. As much as 35 schools have never been IN schools in the empirical period, they were SIN schools in 1995–96.

Table 2 and 3 about here

Table 4 presents some descriptive statistics.[10] The Norwegian school system is characterized by a large number of small schools due to scattered population, and the experimental schools are smaller than other schools. While the mean size of the experimental schools is 38 students, the mean size of the other schools is 170 students. This is also reflected in smaller teacher supply measured in full–time equivalent employment of qualified teachers in the cases the supply is identified. I will perform a sensitivity analysis to investigate whether the empirical results are stable across school size. The income per capita of the local governments wherein experimental schools are located are higher than in other local governments, reflecting high grants to small and northern local governments. The average wage premium in the experimental schools is 920–NOK (about 100–USD) per month, or 4.2 percent. But since only 43 percent of the observations of experimental schools are IN– or SIN schools, the average wage premium for those schools is 9.6 percent. The wage premium ranges from 5.0 percent to 12.2 percent, with the lowest and highest mean in 1997–98 (8.4 percent) and 1998–99 (11.1 percent).

Table 4 about here

6. Results

6.1. Baseline results

[10] Since I will use a logarithmic specification of the supply equation, schools with no qualified teachers (supply equal to zero) are excluded from the analysis. That is 8 and 9 observations of non–experimental schools and experimental schools, respectively. The model has also been estimated with the dependent variable at level. The estimated supply elasticity at mean values was similar to those reported below.

Table 5 presents the main results. Using a dummy variable approach (IN + SIN), I find that a wage premium increases teacher supply by 9 percent, and the effect is significant at 5 percent level. At mean percentage wage premium, the elasticity of teacher supply is equal to 0.94. Since the nominal wage premium is independent of the initial wage, the percentage wage premium differs across teachers. For newly educated teachers with the lowest qualifications, the percentage wage premium is largest. Using the centrally decided wage frame, the model implies that the supply elasticity is in the range 0.60–1.46.

Table 5 about here

Column (3) in Table 5 reports the result of the model using the actual wage premium as an explanatory variable. With this specification, the supply elasticity at mean is equal to 0.91. The model in column (4) estimates the supply elasticity directly by using the percentage wage premium at each school, which vary with the average experience and education level of the teachers. The estimated elasticity of 0.99 is close to the mean elasticities in the previous models. This is the baseline result undertaken several sensitivity tests below.

The selection equation is reported in Appendix Table A1. The results are reasonable because variables that are expected to increase demand mainly have a positive effect on the probability of being IN or SIN school. To calculate the p–value of joint significance of the instruments, the selection equation includes school mean values of the instruments because identification is related only to within school variation.

The elasticity of school size is 0.61, and is highly significant. Given school specific factors, the teachers seem to prefer large schools. The effect of the inverse Mills ratio is also highly significant, implying that the selection bias could be severe.

I next ask whether the response to the wage premium differs between IN– and SIN schools. Since only the incumbent teachers receive a wage premium in SIN schools, one would expect the effect of the wage premium to be largest in IN–schools. However, if schools that change from IN to SIN are hit by a positive shock to a larger degree than other schools, I may find a larger elasticity for SIN schools due to a kind of selection bias. The model in column (5) in Table 5 indicates that the elasticity for SIN schools is larger than for IN schools, but the difference is clearly insignificant.

6.2. Sensitivity analysis

Table 6 considers several different alternatives to the baseline model repeated in row (1). First, the result for the non–censored sample estimated by ordinary least square is reported. The estimated elasticity of 0.61 indicates that it is important to take sample selection into account. One would expect, however, that since the experimental schools have a serious teacher shortage problem at least one year, there is never real excess supply at these schools. The model in row (3) assumes that cases without observed excess demand at the experimental schools (22 percent of the observations), the supply equates the demand. Using OLS, the estimated supply elasticity is almost identical to the baseline model, indicating that the supply never markedly exceeds demand in experimental schools. Notice that the standard error of the supply elasticity is about three times smaller in the OLS regressions than in the two–step sample selection model, indicating that the upward bias of the standard errors in the two–step model compared to an efficient model may be quite large.

Table 6 about here

Is the results flavoured by the fact that school specific effects are not included in the selection equation? Row (4) in Table 6 reports the result when school specific effects for the experimental schools are included in the selection equation. The estimated elasticity increases somewhat, but only by one fourth of a (conservative) standard error. The result differs not much in two models with different shortcomings. The selection equation in the baseline model may have an omitted variable bias, while the estimators of the extended model are biased because the estimates of some variables are based on few observations. The model in row (5) includes a full set of local government fixed effects in the selection equation with no effect on the estimated elasticity. Excluding the number of students from the model neither has any marked impact.

Rows (7)–(9) extend the model by including some new variables. Both the share of language and ethnic minority students and the share of students with special needs are shown to influence the teacher supply by Bonesrønning et al. (2003). They also show that the number of teacher education hours at the school, which may be important for teacher effort given the number of students, has a strong effect on teacher supply. However, including measures of the student body composition or teacher effort in the model does not change the estimated supply elasticity, neither does a more flexible functional form of the number of students and teacher education hours.[11]

Nonqualified teachers are utilized to a small degree in several schools. I have assumed that teacher supply is observable in such schools, but different kinds of noise may cause schools to employ nonqualified teachers even in cases without excess demand. To test whether such misclassification is a problem, row (10) in Table 6 assumes that the supply is only observed when nonqualified teachers account for at least one full–time equivalent teacher.[12] Again, the elasticity changes very little.

Lastly, I estimate the model for different sub–samples. When the sample is restricted to only the experimental schools, the supply elasticity increases, but by less than half of a standard error. Again the estimated elasticity is relatively independent of whether the selection equation includes school fixed effects or not. Are the results sensitive with respect to school size? This is an important question given that most of the experimental schools are small schools. The model in the last row of Table 6 restricts the sample to schools with more than 60 students at average during the empirical period,[13] but again the elasticity changes very little. A more thorough evaluation of the stability of the model is evident from the recursive estimates presented in Figure 2. The schools are sorted by the average number of students in a descending way, and initially, only schools with more than 130 students are included in the sample.[14] Figure 2 shows how the estimated parameter changes when smaller schools are included in the sample successively. The estimated supply elasticity varies to some degree, but there is no evidence of different supply elasticity for the largest school, and the estimates are always within the 80 percent confidence interval of the baseline model.

Figure 2 about here

---

[11] Since the number of teacher education hours is the main determinant of teacher demand, this variable is not included in the selection equation.

[12] In this case, the share of censored observations increases from 0.22 to 0.56 for experimental schools and from 0.48 to 0.76 for other schools.

[13] The number of observations falls with 28 percent, and the number of observations of experimental schools falls with 79 percent.

[14] In the initial sample, seven experimental schools are included.

6.3. Variation in the supply elasticity across school locations

A classic example of labour monopsony is an isolated firm in a remote, sparsely populated area. In this case, the incumbent workers have high mobility costs and can accept a relatively low wage without quitting. If the firm wants to recruit new workers, the firm must pay the mobility costs of the new hires. Thus, remote firms are faced with a less elastic labour supply than other firms. In this section I perform crude tests on whether the teacher supply elasticity depends on the size of the local labour market or on whether teachers at neighbouring schools have a wage premium. The local governments are used as the local labour market in this section.

First I consider to which extent new hires due to increased wage are recruited from other schools in the same local labour market. A simple examination is to estimate the effect of the weighted mean wage premium at the local government level, weighted by the number of students. A positive effect of the weighted mean percentage wage premium implies that some teachers are recruited from other local governments or from a job outside the school–sector. The result is presented in column (2) in Table 7. The elasticity of the mean wage premium is close to 0.5, and must be interpreted as the elasticity of teacher supply for local governments where all teachers receive the wage premium. Thus, given that the elasticity of the individual school is equal to unity at mean, half of the new hires seem to be teachers not previously working in a school in the same local government. This result must, however, be interpreted carefully because it is far from statistical significant.

Table 7 about here

The effect of the mean wage premium captures both the increased supply at schools with wage premium and the decreased supply at neighbouring schools without a wage premium. The model in column (3) in Table 7 seeks to disentangle these effects by including the wage premium at the school level in the model. To interpret the results, consider a local government where half of the students are enrolled at schools in which the teachers receive a wage premium of 10 percent. Then the model implies that the supply increases by 8.5 percent in schools with a wage premium and decreases by 4.5 percent in schools without a wage premium. The supply at the local government level increases by 2 percent, which implies a teacher supply elasticity of 0.4. The latter elasticity is independent of the share of the teachers with a wage premium, but is again highly insignificant.

Lastly I investigate whether the supply elasticity depends on the size of the local labour market, measured by the number of students in the local government. I expect as Boal (1995) that the labour supply is less elastic in thin local labour markets. Column (4) in Table 7 reports the results when the baseline model is extended with an interaction term between the percentage wage premium and the log of the number of students in the local government. The interaction term is positive as expected, although insignificant at conventional levels with the conservative standard errors estimated. Taking at face value, and using the sample variation in the number of students, the result implies that the supply elasticity varies from zero to 2.4, with a mean of 1.0. Small local governments seem to have greater problems with recruitment of new teachers than large local governments, which, however, also imply that they seem to have larger potential monopsony power.

7. Conclusion

Given the problem of identifying good instruments of labour demand, I have argued that utilization of experiments and institutional features is a fruitful way to establish evidence on the elasticity of labour supply directed towards individual enterprises. This paper develops a log–linear supply equation from a random utility function approach aggregating the individual choices to the enterprise level by assuming that all workers at a particular enterprise are observational identical. According to the model, the elasticity of labour supply is equal to the parameter determining the marginal utility of consumption. To estimate the elasticity of labour supply, I utilize a quasi–natural experiment in the Norwegian public–sector schools ensuring that the variation in wages across schools for identical teachers is solely determined by the central government. In addition, and in contrast to the existing literature, I am able to identify the enterprises with excess demand. In such enterprises, the size of the supply is equal to the actual employment. Using a sample selection model with fixed effects, the elasticity of labour supply is estimated to be close to unity. The results also indicate that the elasticity is largest in schools located in densely populated areas in which few other teachers have a wage premium.

While the estimated elasticity of labour supply is a structural parameter, one should be careful in interpreting the estimate as a behavioural parameter. The stylized theoretical model includes several critical assumption, where the most severe assumption in order to interpret the supply elasticity as a behavioural parameter is the aggregation rule. In addition, it is in general extremely difficult to disentangle the income and substitution effects, which in the present model are assumed to be of equal magnitude.

The elasticity of labour supply estimated must be interpreted as a partial short run effect due to the particular experiment considered. In a situation of decentralized wage setting, wage spillovers would make the effect smaller. On the other hand, if frictions in the labour market are present, the supply is likely to be more elastic in the medium– and long run when workers can fully react to wage differentials. Regardless of the long run elasticity, however, upward sloping supply curves in the short run implies that some monopsony power exists in the hand of the enterprises. But the possible exploration of monopsony power in the short run must be balanced against long run considerations as discussed by Boal and Ransom (1997).

References

Abowd, John M., Francis Kramarz and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 67 (1999), 252–333.

Boal, William M., "Testing for Employer Monopsony in Turn-of-the-Century Coal Mining," *RAND Journal of Economics*, 26 (1995), 519–536.

Boal, William M., and Michael R. Ransom, "Monopsony in the Labor Market," *Journal of Economic Literature*, 35 (1997), 86–112.

Bonesrønning, Hans, Torberg Falch and Bjarne Strøm, "Teacher Sorting, Teacher Quality, and Student Composition,“ *European Economic Review*, forthcoming 2003.

Borge, Lars–Erik, and Jørn Rattsø, "Demographic Shift, Relative Costs and the Allocation of Local Public Consumption in Norway,“ *Regional Science and Urban Economics*, 25 (1995), 705-726.

Card, David, and Alan B. Krueger, "Minimum Wages and Employment: A Case Study of the Fast–Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84 (Sep. 1994), 772–793.

Card, David, and Alan B. Krueger, "Minimum Wages and Employment: A Case Study of the Fast–Food Industry in New Jersey and Pennsylvania: Reply," *American Economic Review*, 90 (Dec. 2000), 1397–1420.

Chamberlain, Gary, "Panel Data", In Z. Griliches and M.D. Intriligator (Eds.) *Handbook of Econometrics*, Volume 2, Elsevier Science B.V., 1984.

Dickens, Richard, Stephen Machin, and Alan Manning, "The Effects of Minimum Wages on Employment: Theory and Evidence from Britain," *Journal of Labor Economics*, 17 (1999), 1–22.

Greene, William H., "Sample Selection Bias as a Specification Error: Comment," *Econometrica*, 49 (May 1981), 795–798.

Heckman, James J., “Sample Selection Bias as a Specification Error,” *Econometrica*, 47 (Jan. 1979), 153–161.

Kyriazidou, Ekaterini, “Estimation of a Panel Data Sample Selection Model,” *Econometrica*, 65 (1997), 1335–1364.

Lancaster, Tony, “The Incidental Parameter Problem Since 1948,” *Journal of Econometrics* 95 (2000), 391–413.

Manning, Alan, “The Equal Pay Act as an Experiment to Test Theories of the Labour Market,” *Economica*, 63 (1996), 191–212.

McFadden, Daniel, “Conditional Logit Analysis of Qualitative Choice Behavior,” In Paul Zarembka (ed.) *Frontiers in Econometrics*, Academic Press, New York 1974.

McFadden, Daniel, “Econometric Analysis of Qualitative Response Models,” In Z. Griliches and M.D. Intriligator (Eds.) *Handbook of Econometrics*, Volume 2, Elsevier Science B.V., 1984.

Mortensen, Dale T., ”Job Search and Labor Market Analysis,” In Orley Ashenfelter and Richard Layard (Eds.) *Handbook of Labor Economics*, Volume 2, Elsevier Science B.V., 1986.

Mortensen, Dale T., and Christopher A. Pissarides, “New Developments in Models of Search in the Labor Market,” Working Paper, Northwestern University, 1998.

Nelson, Philip, “The Elasticity of Labor Supply to the Individual Firm,” *Econometrica* 41 (Sept. 1973), 853–866.

Neumark, David, and William Wascher, “Minimum Wages and Employment: A Case study of the Fast–Food Industry in New Jersey and Pennsylvania: Comment,” *American Economic Review*, 90 (Dec. 2000), 1362–1396.

Neyman, Jerzey, and Elizabeth L. Scott, “Consistent Estimation from Partially Consistent Observations,” *Econometrica*, 16 (1948), 1–32.

Poterba, James, “Demographic Structure and the Political Economy of Public Education,” *Journal of Policy Analysis and Management*, 16 (1997), 48-66.

Staiger, Douglas, Joanne Spetz and Ciaran Phibbs, “Is There Monopsony In the Labor Market? Evidence from a Natural Experiment,” NBER Working Paper 7258, 1999.

Sullivan, Daniel, “Monopsony Power in the Market for Nurses,” *Journal of Law & Economics*, 32 (Oct. 1989), S135–S178.

Vella, Francis, and Marno Verbeek, “Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias,” *Journal of Econometrics*, 90 (1999), 239–263.

Appendix: The computation of the variance–covariance matrix.

Rewrite the regression model (11) and the selection model (12) and (13) on compact forms,

$$s_{it} = \theta_s \Xi + \eta^s_{it}, \tag{A.1}$$

$$z^*_{it} = \theta_z \Psi + \left(\eta^z_{it} + \eta^\mu_i\right), \tag{A.2}$$

where Ξ and Ψ are the explanatory variables (including λ and the fixed effects) and θ_s and θ_z are the parameters. Heckman (1979) and Greene (1981) derive the following approximately correct variance–covariance matrix.

$$\mathrm{Var}\left[\theta_s\right] = \hat{\sigma}^2_{\eta^s}\left[\Xi'\Xi\right]^{-1}\left[\Xi'\left(I - \hat{\rho}^2\hat{\Delta}\right)\Xi + Q\right]\left[\Xi'\Xi\right]^{-1}. \tag{A.3}$$

$\hat{\sigma}^2_{\eta^s}$ is the variance of η^s_{it}, $\hat{\rho} = \hat{\beta}_\lambda / \hat{\sigma}_{\eta^s}$, $\left(I - \hat{\rho}^2\hat{\Delta}\right)$ is a diagonal matrix with $\left(1 - \hat{\rho}^2\hat{\delta}_i\right)$ on the diagonal, $\hat{\delta}_{it} = \hat{\lambda}^2_{it} + \hat{\lambda}_{it}\tilde{\Upsilon}\hat{\theta}_z$, $Q = \hat{\rho}^2\left[\Xi'\hat{\Delta}\Psi\right]\Sigma\left[\Psi'\hat{\Delta}\Xi\right]$, and Σ is the asymptotic covariance matrix of the probit coefficients. With a large number of units in a fixed effects model, the matrix $\Xi'\Xi$ is intractable to estimate. Since differencing solves the problem for the supply equation, I difference both Ξ and Ψ in the Q-matrix to make it meaningful. The covariance matrix used in this paper is given by

$$\mathrm{Var}[\theta_s] = \hat{\sigma}^2_{\left(\eta^s - \bar{\eta}^s\right)}\left[\left(\Xi - \bar{\Xi}\right)'\left(\Xi - \bar{\Xi}\right)\right]^{-1}\left[\left(\Xi - \bar{\Xi}\right)'\left(I - \hat{\rho}^2\hat{\Delta}\right)\left(\Xi - \bar{\Xi}\right) + Q\right]\left[\left(\Xi - \bar{\Xi}\right)'\left(\Xi - \bar{\Xi}\right)\right]^{-1} \tag{A.4}$$

and

$$Q = \hat{\rho}^2\left[\left(\Xi - \bar{\Xi}\right)'\hat{\Delta}\left(\Psi - \bar{\Psi}\right)\right]\Sigma\left[\left(\Psi - \bar{\Psi}\right)'\hat{\Delta}\left(\Xi - \bar{\Xi}\right)\right]. \tag{A.5}$$

For small samples, it is possible to compare the proposed standard errors calculated from (A.4) with those calculated by the traditional approximation (A.3). As an illustration, consider the sample including only experimental schools, for which the supply elasticity is given in row (12) in Table 6. In this case, the standard errors of the traditional and proposed approximations are equal at the precision reported in the table.

Table 1. The possible states of the experimental schools

School year	School A	School B
1993–1994	IN	IN
1994–1995	IN	SIN
1995–1996	IN	SIN
1996–1997	NIN	NIN
1997–1998	NIN	NIN
1998–1999	IN	IN
1999–2000	IN	SIN

Note: IN denotes that all teachers at the school receive a wage premium, SIN denotes that only the incumbent teachers at the school receive a wage premium, while NIN denotes that none of the teachers receive a wage premium.

Table 2. The number of included experimental schools

	1995–96	1996–97	1997–98	1998–99	1999–2000
IN	34	21	15	61	74
SIN	61	0	16	0	15

Note: IN denotes that all teachers at the school receive a wage premium and SIN denotes that only the incumbent teachers at the school receive a wage premium.

Table 3. The number of experimental schools in years of different states

	0 years	1 year	2 years	3 years	4 years	5 years
IN	38	45	32	16	7	4
SIN	57	79	5	1	0	0
NIN	14	11	30	33	54	0
IN + SIN	0	63	35	25	6	13

Note: IN denotes that all teachers at the school receive a wage premium, SIN denotes that only the incumbent teachers at the school receive a wage premium, while NIN denotes that none of the teachers receive a wage premium.

Table 4. Descriptive statistics, mean values (standard deviation)

Sample	Non–experimental schools	Experimental schools	IN and SIN schools
Number of students	170.2 (130.6)	38.7 (40.4)	31.1 (37.8)
Teacher supply in the cases of excess demand	15.6 (10.2)	5.2 (4.7)	4.5 (4.6)
Local government income per capita[1]	16471 (4401)	24898 (6862)	25292 (6494)
IN	0	0.297	0.692
SIN	0	0.132	0.308
Monthly wage premium[1]	0	900 (1085)	2095 (481)
Percentage wage premium	0	0.041 (0.049)	0.096 (0.019)
Observations	15230	673	289
Proportion censored observations	0.475	0.217	0.242

[1] Measured in 1998–NOK. The exchange rate NOK/USD ≈ 9.

Table 5. Estimated labour supply elasticity. Dependent variable is the log of teacher supply

(1)	(2)	(3)	(4)	(5)
ln(Number of students)	0.608 (0.037)*	0.607 (0.038)*	0.607 (0.038)*	0.605 (0.039)*
(IN + SIN)	0.090 (0.041)*	-	-	-
Monthly wage premium	-	0.042 (0.020)*	-	-
(Percentage wage premium) * (IN + SIN)	-	-	0.993 (0.439)*	0.919 (0.472)
(Percentage wage premium) * SIN	-	-	-	0.372 (0.764)
Inverse Mills ratio	-0.425 (0.132)*	-0.440 (0.138)*	-0.436 (0.136)*	-0.446 (0.138)*
Time fixed effects	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
Minimum elasticity	0.598	0.620	-	-
Mean elasticity	0.944	0.912	-	-
Maximum elasticity	1.463	1.053	-	-
Observations	15 903	15 903	15 903	15 903
Non–censored observations	8 518	8 518	8 518	8 518
Standard error of equation	0.14815	0.14821	0.14817	0.14813
Test for joint significance of the instruments in the selection equation, p-value	0.010	0.012	0.012	0.009

Note: Standard errors in parentheses, * denotes significance at 5 percent level. The selection equation for the model in column (4) is reported in Appendix Table A1.

Table 6. Sensitivity analysis

Model	Specification changes compared to the model in column (4) in Table 5		Schools included	Elasticity (St. error)
	Supply equation	Selection equation		
(1)	None	None	All	0.993 (0.437)
(2)	None	Irrelevant	Censored observations excluded	0.574 (0.161)
(3)	None	Excluded	Censored observations of non-experimental schools excluded	1.005 (0.144)
(4)	None	School specific effects for experimental schools included	All	1.101 (0.306)
(5)	None	Full set of local government specific effects included	All	0.949 (0.369)
(6)	ln(Number of students) excluded	ln(Number of students) excluded	All	1.100 (0.577)
(7)	The share of students from ethnic minorities and the share of students with special needs included	The share of students from ethnic minorities and the share of students with special needs included	All	0.970 (0.402)
(8)	ln(Teacher education hours) included	None	All	0.979 (0.317)
(9)	ln(Number of students) and ln(Teacher education hours) included at squared, cubic and differenced form	ln(Number of students) included at squared, cubic and differenced form	All	1.001 (0.266)
(10)	Only observations with at least one full–time equivalent nonqualified teacher are regarded as non–censored	Only observations with at least one full–time equivalent nonqualified teacher are regarded as non–censored	All	1.036 (0.300)
(11)	None	None	Only experimental schools	1.165 (0.321)
(12)	None	School fixed effects included	Only experimental schools	1.097 (0.349)
(13)	None	None	Average no. of students > 60	1.000 (0.480)

Note: * denotes significance at 5 percent level.

Table 7. The supply elasticity and concentration of schools

(1)	(2)	(3)	(4)
ln(Number of students)	0.604 (0.038)*	0.605 (0.039)*	0.604 (0.039)*
Percentage wage premium	-	1.304 (0.509)*	-1.761 (2.303)
Weighted mean percentage wage premium in the local government	0.529 (0.982)	-0.906 (1.133)	-
(Percentage wage premium) * ln(number of students in the local government)	-	-	0.461 (0.371)
Inverse Mills ratio	-0.435 (0.134)*	-0.450 (0.139)*	-0.453 (0.142)*
Time fixed effects	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes
Observations	15 903	15 903	15 903
Uncensored observations	8 518	8 518	8 518
Standard error of equation	0.14827	0.14787	0.14786

Note: * denotes significance at 5 percent level.

Appendix Table A1. The selection equation, specification as column (4) in Table 5.

Variable	Mean probability derivatives (standard error)
ln(Local government income per capita)	0.216 (0.124)
The local government mean during the empirical period in the variable above	0.191 (0.135)
The share of the population between 0 and 6 years of age	-0.948 (1.796)
The local government mean during the empirical period in the variable above	1.056 (1.921)
The share of the population above 80 years of age	5.475 (2.865)
The local government mean during the empirical period in the variable above	-7.170 (2.934)*
ln(population size)	1.016 (0.335)*
The local government mean during the empirical period in the variable above	-0.990 (0.335)*
ln(Number of students)	0.095 (0.008)*
Percentage wage premium	-0.786 (0.465)
Experimental school	0.250 (0.027)*
The local government mean during the empirical period in the share of the students at experimental schools	0.567 (0.161)*
Students only at grades 1 to 7, and there are separate classes for each grade	-
Students only grades 8 to 10, and there are separate classes for each grade	-0.000 (0.013)
Students at all grade levels and there are separate classes for each grade	0.045 (0.015)*
The school has at least one class that include students at different grade levels	0.021 (0.016)
Mean number of schools in local government is $\left[1,2\right\rangle$	-
Mean number of schools in local government is $\left[2,3\right\rangle$	0.042 (0.047)
Mean number of schools in local government is $\left[3,4\right\rangle$	0.054 (0.045)
Mean number of schools in local government is $\left[4,5\right\rangle$	0.025 (0.047)
Mean number of schools in local government is $\left[5,6\right\rangle$	0.065 (0.047)
Mean number of schools in local government is $\left[6,7\right\rangle$	0.031 (0.049)
Mean number of schools in local government is $\left[7,8\right\rangle$	0.068 (0.048)
Mean number of schools in local government is $\left[8,9\right\rangle$	0.109 (0.049)*
Mean number of schools in local government is $\left[9,10\right\rangle$	0.068 (0.055)
Mean number of schools in local government is $\left[10,11\right\rangle$	0.008 (0.055)
Mean number of schools in local government is $\left[11,12\right\rangle$	0.014 (0.056)
Mean number of schools in local government is $\left[12,13\right\rangle$	-0.061 (0.058)
Mean number of schools in local government is $\left[13,14\right\rangle$	0.041 (0.059)
Mean number of schools in local government is $\left[14,15\right\rangle$	0.104 (0.063)
Mean number of schools in local government is $\left[15,16\right\rangle$	-0.162 (0.068)*
Mean number of schools in local government is $\left[16,17\right\rangle$	-0.166 (0.066)
Mean number of schools in local government is $\left[17,18\right\rangle$	-0.038 (0.063)
Mean number of schools in local government is $\left[18,19\right\rangle$	-0.119 (0.067)
Mean number of schools in local government is $\left[19,20\right\rangle$	-0.023 (0.070)
Local government specific effects for local governments with at least 20 schools at mean	Yes
Year specific effects	Yes
County specific effects	Yes
Observations	15 903

Note: * denotes significance at 5 percent level.

Figure 1. Identification of teacher supply

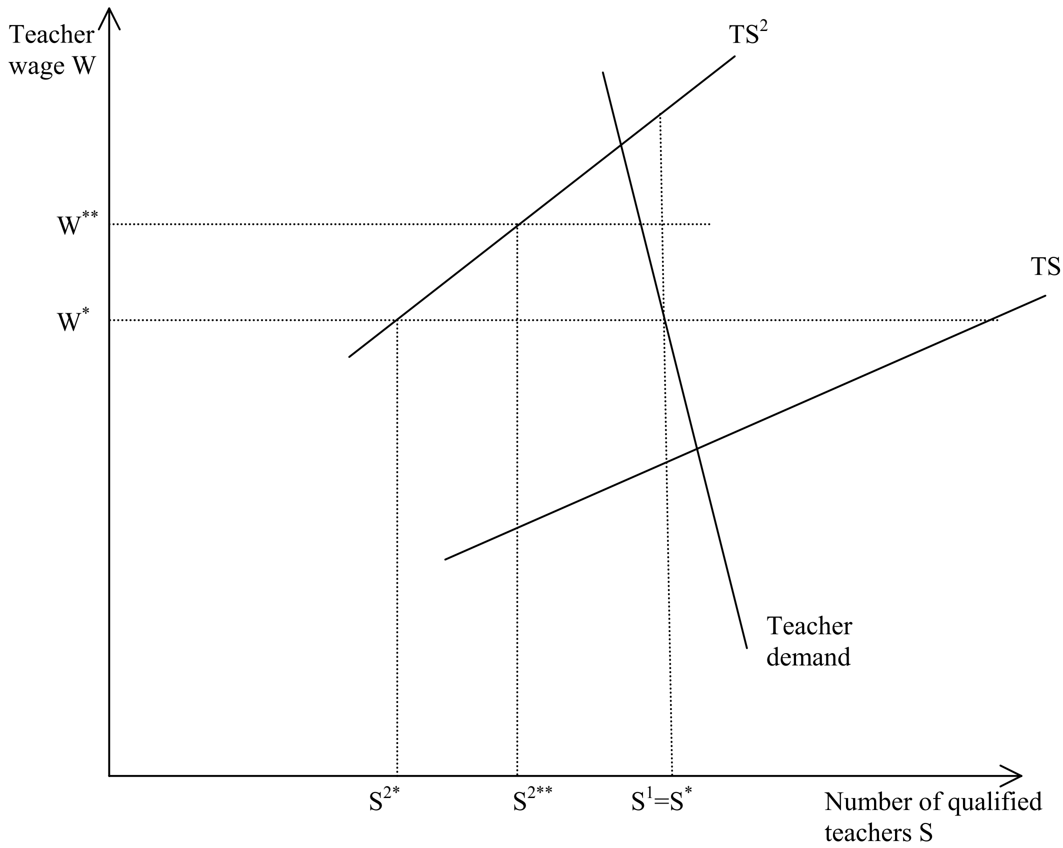

Figure 2. Recursive estimates ± 2 standard errors of the model in column (4) in Table 5. The initial sample is schools with more than 130 students on average during the empirical period.