# School accountability: (how) can we reward schools and avoid pupil selection?\*

Erwin Ooghe and Erik Schokkaert<sup>†</sup>

Preliminary version (September 2012)

#### Abstract

To be completed.

Keywords:

JEL-codes:

## 1 Introduction

Public education used to have some common features around the world. Schools roughly received funding per pupil and had limited autonomy, an inspectorate controlled the quality of education, and school choice by parents was often restricted. Critics argued that these features explained the poor performance of (some) public schools.

<sup>\*</sup>We would like to thank Dolors Berga, Geert Dhaene, Carmen Herrero, Iñigo Iturbe-Ormaetxe, Dirk Van de gaer, Frank Vandenbroucke, and Carine Van de Voorde for their useful comments, Ides Nicaise and Jan Van Damme for their permission to use the SiBOdata, and Frederik Maes and Peter Helsen for their valuable help with the data. The usual disclaimer applies.

<sup>&</sup>lt;sup>†</sup>Erwin Ooghe: Department of Economics, KU.Leuven and research fellow at IZA, Bonn, erwin.ooghe@kuleuven.be. Erik Schokkaert: Department of Economics, KU.Leuven, and CORE, Université Catholique de Louvain, erik.schokkaert@kuleuven.be.

School accountability increased in several countries to improve student learning. In the U.S., for example, the 'No Child Left Behind Act of 2001' forced all states to set up an accountability system for public schools. In some states schools had to publish report cards, information about their performance based on pupil test scores, to inform parental school choice. Other states used bonuses (sanctions) for well (poorly) performing schools. But could these and similar incentive-based reforms turn the tide?

School accountability improves pupil test scores, but it is unclear whether explicit financial bonuses and sanctions are necessary (Wössmann, 2003; Hanushek and Raymond, 2004, 2005; Jacob, 2005; Figlio and Rouse, 2006; West and Peterson, 2006; Burgess et al., 2007; Chiang, 2009). Accountability can also result in potentially undesirable strategic reactions such as teaching to the rating, student retainment, removal of low-achieving students, and even adapting the caloric content of the school lunches at the testing date (Jacob, 2005; Figlio and Winicki, 2005; Burgess et al., 2005; Reback, 2008). In a nutshell, the overall success of incentive-based reforms crucially depends on the design.

We focus here on another strategic reaction of schools, pupil selection. The average test score in a school strongly depends on the characteristics of the pupil population. Insufficiently correcting for pupil characteristics may lead to a biased evaluation of school performance (Meyer, 1997; Ladd and Walsh, 2002; Hanushek and Raymond, 2003; Taylor and Nguyen, 2006; Neal, 2008). Moreover, it can seduce schools to appear more attractive for specific student groups. Pupil selection may improve the measured performance of a school without adding real skills.

A key question follows: is it possible to reward schools for good administration and avoid pupil selection?<sup>1</sup> The answer is negative if the funding scheme must satisfy both principles for *all* educational production functions. For specific educational production functions both principles can be reconciled, but the functional form restriction is rejected by the data for Flanders (the northern

<sup>&</sup>lt;sup>1</sup>Although we focus on school funding, the question is also relevant for the design of report cards and differentiated vouchers (Epple and Romano, 2008).

part of Belgium). It turns out that a trade-off is inevitable.

We propose some compromise solutions. One family of solutions rewards schools for good administration, but does not necessarily eliminate all pupil selection; the other family avoids pupil selection, but does not necessarily reward schools. The performance of the proposed solutions is an empirical question. We therefore illustrate the relevancy of the trade-off with the data.

## 2 Accountability and incentives

To bring the key question into focus we start from the most favourable assumptions, at the cost of neglecting other important issues. The selection of relevant pupil test scores and its aggregation (over different dimensions and pupils) into a cardinal and comparable indicator of school output is assumed to be settled before (Cawley et al., 1999; Neal, 2008). We also neglect that school output measures are typically less reliable for small schools (Kane and Staiger, 2002). Sufficient data are available at the pupil level, as informationally less demanding accountability schemes cannot sufficiently correct for differences in pupil characteristics (Meyer, 1997; Hanushek and Raymond, 2003). We do not explicitly model school or teacher behaviour (as, e.g., in Barlevy and Neal, 2011), but rely on reduced form equations for educational production. We also do not model an overall social objective, but focus on partial objectives.

#### 2.1 Preliminaries

The agreed measure of school output  $y \in \mathbb{R}$  is a function of several school variables denoted by  $x \in X$ ; we write y = f(x). School variables consist of administration variables  $a \in A$  and background variables  $b \in B$ ; we write x = (a, b), and define the set X as the product  $A \times B$ .

The classification of a school variable as an administration or background variable is simple in theory. Endogenous variables that can be chosen by a school are attributed to administration; for example, the number of instruction hours, the level of remediation per pupil, and teacher motivation. Exogenous variables that cannot be set by a school—but its distribution at school can possibly be influenced—belong to background; think of initial test scores, innate intelligence, and socio-economic status of the pupils at school. Because background only consists of pupil variables in the empirical part, we call it pupil background from now on.

The classification is less evident in practice. The function f will be estimated in the empirical part via a standard explanatory model of test scores; see, e.g., Hanushek (2006) for an overview. A typical estimation includes observable characteristics, unobserved pupil and school effects, and idiosyncratic error terms. Each right-hand side variable, observed and unobserved, must be classified as a administration or a background variable. The empirical part contains a specific proposal.

As stated before, we do not explicitly model school behaviour. The output function f is a reduced form equation that reflects educational production. We implicitly assume that f does not change under the incentive scheme. Changes in subsidies can of course motivate schools to be more effective, otherwise the whole exercise would be meaningless. But that effect is fully captured in our framework by a change in the way the school is administered, i.e., a change in a.

We use subscripts j = 1, 2, ..., J to denote schools. A school subsidy scheme  $s : X^J \to \mathbb{R}^J$  maps all information about the different schools  $\mathbf{x} = (x_1, x_2, ..., x_J)$  into a vector of school subsidies  $s(\mathbf{x}) = (s_1(\mathbf{x}), s_2(\mathbf{x}), ..., s_J(\mathbf{x}))$ . Simple outputrelated subsidy schemes are one specific example. We look for a funding scheme that rewards schools for good policy without providing incentives to attract (or discourage) pupils with specific characteristics. What form should  $s(\mathbf{x})$  take?

To tackle this question we formulate two principles. These principles, and some of the results later on, are inspired by the theory of fair allocation (see, e.g., Fleurbaey, 2008, for an overview) and its application to health insurance (Schokkaert et al., 1998; Schokkaert and Van de Voorde, 2004).

#### 2.2 Getting the incentives right

We start with the reward principle. If an increase in the output of a school is *only* caused by a change in administration, then the subsidy of the school must increase as well. Let  $(\mathbf{a}, \mathbf{b})$  be the decomposition of  $\mathbf{x}$  (with obvious notation).

INCENTIVES FOR GOOD ADMINISTRATION: For all  $\mathbf{x}, \mathbf{x}'$  in  $X^J$ , for all j = 1, 2, ..., J, if  $a_k = a'_k$  for each school k except j, and  $\mathbf{b} = \mathbf{b}'$ , then there exists a strictly increasing function  $\phi$ , with  $\phi(0) = 0$ , such that  $s_j(\mathbf{x}') - s_j(\mathbf{x}) = \phi(y'_j - y_j)$ .

The axiom does not say that the subsidy increase should be sufficiently large to make the cost (if any) of the change in administration worthwhile. It simply says that good administration should be financially encouraged. It can be interpreted as a minimalist necessary condition for efficiency. For later use, if the subsidy functions  $s_j$  and the output function f are differentiable with respect to some administration variable  $a_{j,k}$  (an element of  $a_j$ ), then the axiom would require

$$\partial s_j(\mathbf{x})/\partial a_{j,k} = \phi(\partial f(x_j)/\partial a_{j,k}),$$
 (1)

for all profiles and schools.

We now turn to pupil selection. Changes in the background of pupils without changes in administration must not be rewarded in the funding scheme. Otherwise schools would have an incentive to attract pupils with a specific background and discourage others.

NO INCENTIVES FOR PUPIL SELECTION: For all  $\mathbf{x}, \mathbf{x}'$  in  $X^J$ , for all j = 1, 2, ..., J, if  $\mathbf{a} = \mathbf{a}'$ , and  $b_k = b'_k$  for each school k except j, then  $s_j(\mathbf{x}') = s_j(\mathbf{x})$ .

The principle clearly wipes out all financial incentives for pupil selection. But a normative trade-off can arise if the segregation or integration of pupils over schools would increase average school output. Incentives for pupil selection could then be desirable from an efficiency point of view. We will discuss the issue when we discuss compromise solutions that allow for pupil selection. Given differentiability with respect to a pupil background variable, say  $b_{j,k}$ , the axiom implies

$$\partial s_j(\mathbf{x}) / \partial b_{j,k} = 0, \tag{2}$$

for all profiles and schools.

#### 2.3 Performance incentives create selection incentives

It makes sense to impose both principles if the aim is to create incentives for good administration and to avoid pupil selection at the same time. It is therefore striking that it is not possible to design a funding scheme that satisfies both in general, i.e., for all possible output functions f.

The impossibility result is well known (in many variants) in the social choice literature (Fleurbaey, 2008). But it remained largely unnoticed in the literature on school accountability. Meyer (1997) raises a related impossibility result. If the effect of background variables on output differs between schools, then schools cannot be ranked according to performance without ambiguity. He claims that the empirical relevance is limited, because "the assumption that slopes do not vary across schools is often a very reasonable assumption." In the empirical part we falsify the claim for Flemish data.

We provide a simple proof of the incompatibility between the two incentive axioms. We focus on an arbitrary school, keeping information in all other schools constant. We suppress subscripts, and the output and the subsidy of the school under consideration are denoted by f(a, b) and, with slight abuse of notation, s(a, b). Let  $b \in B = \mathbb{R}$  be an index of pupil background at the school. Figure 1 presents school output as a function of pupil background for two types of administration a and a'. Using Meyer's (1997) terminology, the slopes differ between the different administration styles.

#### Figure 1

Start at situation 1 with administration a and pupil background b. An increase in the background index from b to b' leads us to situation 2. The axiom

NO INCENTIVES FOR PUPIL SELECTION requires the same subsidy in both situations, thus s(a, b) = s(a, b'). If the school would now change administration from a to a', then we go from situation 2 to 3 with a lower output. The axiom INCENTIVES FOR GOOD ADMINISTRATION requires a lower subsidy leading to s(a, b') > s(a', b'). If the school sticks to administration a', but the pupil background index changes back to b, then we arrive in situation 4. Again the same subsidy should apply, so s(a', b') = s(a', b). Finally, a change in administration back to a lowers output again, and the subsidy must follow, or s(a', b) > s(a, b). All things together we get a cycle; we summarize:

PROPOSITION 1. There is no subsidy scheme that satisfies INCENTIVES FOR GOOD ADMINISTRATION and NO INCENTIVES FOR PUPIL SELECTION in general, i.e., for each possible output function f.

Proposition 1 has to be interpreted carefully: the general impossibility result only holds if we look for a subsidy scheme satisfying both axioms for *all* possible output functions f. It is obvious that the incompatibility disappears in Figure 1 if slopes do not intersect. We can generalize the observation (a proof can be found in the appendix).

PROPOSITION 2. A subsidy scheme can satisfy INCENTIVES FOR GOOD ADMIN-ISTRATION and NO INCENTIVES FOR PUPIL SELECTION if and only if there exist functions  $g : \mathbb{R} \times B \to \mathbb{R}$  and  $h : A \to \mathbb{R}$ , with g strictly increasing in its first argument, such that f(a, b) = g(h(a), b), for all x = (a, b) in X.

The intuition is again easy. The separability condition allows to classify schools according to the performance index h(a): a higher index corresponds with a higher output irrespective of the pupils' background. If we define each subsidy  $s_j(\mathbf{x})$  to be a strictly increasing function of the performance index  $h(a_j)$  only, then both requirements will be satisfied by the resulting subsidy scheme.

The separability condition of proposition 2 is satisfied by the simple linear models that are typically used to estimate educational production functions. In our empirical work, however, we falsify the separability condition. Proposition 1 suggests that relying on a linear form may have undesirable consequences when the true model is non-separable. We show in the empirical part that the problem is also empirically relevant.

#### 2.4 Compromise solutions

We can keep the incentives for good administration intact, but then we may introduce incentives for selecting pupils with a certain background. Or we can make sure that we avoid selection, but then the incentives to improve pupil learning can be very different for different pupils and may even become negative.

For ease of exposition, we suppress the dependency on the profile  $\mathbf{x}$ . We focus on linear subsidy schemes and write the per pupil subsidy for school j as

$$s_j = \text{constant} + \underbrace{\text{slope}}_{>0} \times \widetilde{y}_j,$$
 (3)

with  $\tilde{y}_j$  the (possibly corrected) output of school j (as defined later on). In the empirical part, the **constant** is chosen to satisfy the budget constraint of the regulator, and the **slope** is set to guarantee a minimal subsidy to all schools. At this stage, we simply neglect the role of both parameters.

Before we propose two families of compromise solutions, we discuss some benchmark subsidy schemes, being per capita (PC), and an uncorrected output (UO) funding. Later on, we also introduce a typical value-added (VA) model, and show that it is a special case of a reference administration model.

In many countries school funding is simply per capita, i.e.,

$$s_j^{PC} = \text{constant.}$$
 (4)

A per capita scheme does not provide any incentives, neither for good administration, nor for pupil selection.

An uncorrected output scheme fully rewards schools for output increases, without any correction for pupil background. The subsidy is equal to

$$s_j^{UO} = \text{ constant} + \text{ slope } \times \underbrace{f(a_j, b_j)}_{y_j}.$$
 (5)

The scheme gives incentives for good administration, because changes in administration that lead to higher output clearly will be rewarded. With differentiability we obtain

$$\partial s_j^{UO} / \partial a_{j,k} = \text{slope} \times \partial f(a_j, b_j) / \partial a_{j,k},$$
 (6)

and condition (1) is satisfied. For the same reason, also changes in background that lead to higher output will be rewarded. Schools have an incentive to attract pupils with a background that is 'favourable' to output. Given differentiability the subsidy change is equal to

$$\partial s_j^{UO}/\partial b_{j,k} = \text{slope} \times \partial f(a_j, b_j)/\partial b_{j,k},$$
(7)

violating condition (2) if  $\partial f(a_j, b_j) / \partial b_{j,k}$  differs from zero.

A first family of compromise solutions is based on a reference administration (*RA*), denoted  $\tilde{a}$ , to correct output. Define corrected output as  $\tilde{y}_j = y_j - f(\tilde{a}, b_j)$ , and the subsidy is equal to

$$s_j^{RA} = \text{constant} + \text{slope} \times (\underbrace{f(a_j, b_j)}_{y_j} - f(\widetilde{a}, b_j)).$$
 (8)

Schools are rewarded if their output is higher than the hypothetical output that would occur if the school had chosen the reference administration level, ceteris paribus. It creates incentives for good administration, because changes in administration that are favourable to output translate into higher subsidies. Assuming differentiability of the reference scheme, the incentive for good administration  $\partial s_j^{RA} / \partial a_{j,k}$  is exactly equal to the one for uncorrected output. The selection incentive is equal to

$$\partial s_j^{RA} / \partial b_{j,k} = \text{slope} \times (\partial f(a_j, b_j) / \partial b_{j,k} - \partial f(\widetilde{a}, b_j) / \partial b_{j,k}),$$
 (9)

and will typically be different from zero. But because one can expect the derivatives  $\partial f(a_j, b_j) / \partial b_{j,k}$  and  $\partial f(\tilde{a}, b_j) / \partial b_{j,k}$  to be of a similar magnitude, also  $|\partial s_j^{UO} / \partial b_{j,k}|$  will typically be larger than  $|\partial s_j^{RA} / \partial b_{j,k}|$ . Summing up, reference administration schemes provide similar incentives for good administration compared to uncorrected output schemes, but lower incentives for pupil selection. The empirical part confirms the analysis here. The mirror image of the previous scheme is to choose a reference pupil background (RB), say  $\tilde{b}$ . If we correct output as  $\tilde{y}_j = f(a_j, \tilde{b})$ , then the school will be rewarded on the basis of the hypothetical output that would arise if its actual administration were applied to the reference pupil. This yields

$$s_j^{RB} = \text{constant} + \text{slope} \times f(a_j, \widetilde{b}).$$
 (10)

 $s_j^{RB}$  does not depend on the school background  $b_j$  anymore, removing selection incentives. With differentiability, we indeed obtain  $\partial s_j^{RB} / \partial b_{j,k} = 0$  for the different schools. But actual output does not appear in equation (10). We immediately derive that

$$\partial s_{j}^{RB} / \partial a_{j,k} = \text{slope } \times \partial f(a_{j}, b) / \partial a_{j,k},$$
 (11)

and condition (1) is not guaranteed anymore if a change in the true output  $\partial f(a_j, b_j)/\partial a_{j,k}$  has a different sign compared to a change in the hypothetical output  $\partial f(a_j, \tilde{b})/\partial a_{j,k}$ . Because we can expect that the signs of  $\partial f(a_j, \tilde{b})/\partial a_{j,k}$  and  $\partial f(a_j, b_j)/\partial a_{j,k}$  often coincide, the reference background scheme will provide more incentives for good administration compared to a per capita scheme. Summing up, reference background schemes provide no incentives for pupil selection like per capita schemes, but can be expected to provide some incentives for good administration.

Table 1 summarizes the different schemes and their properties, i.e., is the axiom satisfied, how large do we expect the incentives to be, and how many schools will satisfy the axioms.

#### Table 1

Per capita financing does not give incentives for good administration nor incentives for pupil selection to *any* school. The uncorrected output scheme gives both incentives to *all* schools. We expect the reference schemes to do better. More precisely, a reference administration scheme outperforms the uncorrected output schemes, because it gives the same incentives for good administration to all schools, but with a lower incentive for pupil selection. The reference background scheme outperforms the per capita scheme, because it provides no incentives for pupil selection, but some incentives for good administration to some schools.

Finally, compare the reference subsidy schemes in (8) and (10). To implement the schemes, the regulator must have (an estimate) of the educational production function f. A reference administration (RA) scheme requires output  $y_j$  and background variables  $b_j$  in addition, while a reference background (RB) scheme also needs administration variables  $a_j$ . The different informational requirements have practical consequences. A reference background (RB) scheme offers scope for strategic behaviour, e.g., increasing instruction time without any real results. Even worse, it may create incentives for misreporting variables, like instruction time, that are difficult to verify. Strategic behaviour is less problematic in a reference administration scheme. Test scores are collected in a standardized way, and the background variables typically consist of pupil characteristics that can more easily be controlled by the regulator.

### 3 Empirical illustration

The aim of the 'SiBO'-project is to describe and explain differences in the primary school curriculum of Flemish pupils. Pupils were tested in mathematics at the start of the first grade (in September-October 2003 when (most) pupils were 6 years old) and at the end of grade 1 and 2 (in May-June of 2004 and 2005). The test scores have been standardized and calibrated over time to measure progress.

Other pupil data include the gender of the pupil, the language they speak with each of the parents, and the education level of the parents. Classroom data consist of the total experience of the teacher, the class size, the instruction time for mathematics, and the number of teachers in a class. We also include the average initial test score of the peers, defined as the fellow pupils in the same class. Table 2 provides an abbreviation and a description of each variable.

#### Table 2

We focus on schools with at least 10 pupils tested in each grade. We have 5817 pupil-time observations—2239 pupils appearing in both grades, 628 in grade 1 only and 711 in grade 2 only—distributed over 111 schools. The main reasons for attrition and replenishment is student retainment. We come back to this potential source of selection bias. Tables 3a and 3b contain summary statistics for the pupil and classroom data.

#### Table 3a and 3b

#### 3.1 Explaining test scores

Let  $y_{ijt}$  be the (standardized) math test score of pupil *i* at school *j* at time *t* and let  $z_{ijt}$  be the vector of observable regressors. To explore the data, we start with a standard linear panel model, i.e.,

$$y_{ijt} = \beta'_{a} z_{a,ijt} + \beta'_{b} z_{b,ijt} + u_{i} + v_{j} + w_{ijt}, \qquad (12)$$

with  $u_i$  a 'random' pupil-level effect,  $v_j$  a 'fixed' school level effect, and  $w_{ijt}$ an idiosyncratic error term. The specification (12) satisfies the separability condition in (2), irrespective of how the right-hand variables are assigned to administration or background.

Because of attrition and replenishment in the data, we must check and, if needed, correct estimates for selection bias. To check for selection bias we use a variable addition test; see, Verbeek and Nijman (1992) and Wooldridge (1995). The results indicate that missingness might be informative. To check whether a selection correction influences the estimation results, we add a selection equation to each period in the spirit of Hausman and Wise (1979); we allow for correlation between the individual level effects in the selection and the output equation. The corrected estimates do not statistically differ from the uncorrected estimates, allowing us to ignore selection issues in the sequel.<sup>2</sup>

<sup>&</sup>lt;sup>2</sup> The selection correction model assumed random (rather than fixed) school effects, leading to a so-called multilevel model. An attempt with fixed school effects did not converge, probably due to the high number of dummies in the selection equation.

Table 4 reports estimates for (12).<sup>3</sup>

#### Table 4

The initial test score plays an important role in all models. Its coefficient is rather robust and smaller than 1, indicating that the added value, i.e., the gain in test scores, is larger for pupils with a lower initial test score. The background variables play a more modest role and their effects depend on whether or not the initial test score is taken up as a covariate. In model (b) without initial test score, boys do better than girls, being ahead of age is not significant while lagging behind is correlated with a lower math performance, having Dutchspeaking and better educated parents improve test scores and these effects are stronger and more significant for mothers compared to fathers. In model (c) with initial test scores as an additional regressor, some of the estimated coefficients for the background variables change in magnitude and even in sign. We provide two striking examples.

First, once we correct for initial test scores, having Dutch-speaking parents gets a negative coefficient. Indeed, pupils with non-Dutch speaking parents have (on average) a worse preparation before starting primary education. Therefore their initial test score underestimates their potential, leading to a catching-up effect in the first grades. Second, the effect of father education is now stronger than that of mother education. One hypothesis could be that mothers have a larger effect on initial test scores (during the pre-primary education period), while fathers have a larger effect on the primary education growth of their children.

Comparing model (c) and (d), adding class data does not change the coefficient estimates for the individual-specific variables much. Among the class variables, instruction time and class size have a significant and positive effect, while having two teachers reduces test scores on average. (The positive effect for class size is not exceptional in correlation studies like ours that do not control

<sup>&</sup>lt;sup>3</sup>We add a dummy 'missing' to each covariate group (to limit the reduction in total sample size). We do not report the corresponding estimates which are, as expected, never significant.

for endogeneity.)

#### 3.2 Separability

We split observables  $z_{a,ijt}$  into administration and background variables  $z_{a,ijt}$ and  $z_{b,ijt}$ . It is natural to assign variables at the class and school level to administration, except the peer variable; all other variables—the pupil-level variables, the time dummy and the peer variable—are classified as background. To test separability we generalize (12), allowing the pupil background coefficients to vary over schools, i.e.,

$$y_{ijt} = \beta'_{a} z_{a,ijt} + \beta'_{b,j} z_{b,ijt} + u_i + v_j + w_{ijt},$$
(13)

For the purpose of illustration, we define school output as the expected average pupil output. If a bar  $\overline{\cdot}$  denotes averages and a hat  $\hat{\cdot}$  refers to the OLS estimates obtained from estimating (13), then school output is equal to

$$\overline{y}_j = \widehat{\beta}'_a \overline{z}_{a,j} + \widehat{\beta}'_{b,j} \overline{z}_{b,j} + \widehat{v}_j.$$
(14)

To apply the theory, we must classify all right-hand variables that vary at the school level. The school averages  $\overline{z}_{a,j}$  and  $\overline{z}_{b,j}$  have been classified before. The school-specific constant  $\hat{v}_j$  and the slope coefficients  $\hat{\beta}_{b,j}$  tell us how pupils with a certain background perform at each school. It seems natural to assign these coefficients to administration. We then get

$$\overline{y}_{j} = \underbrace{\widehat{\beta}'_{a}\overline{z}_{a,j} + \widehat{v}_{j}}_{\text{pure administration}} + \underbrace{\widehat{\beta}'_{b,j}\overline{z}_{b,j}}_{\text{mixture}}$$

The non-linear terms in  $\hat{\beta}'_{b,j}\bar{z}_{b,j}$  mix administration and background. They are crucial to test the separability condition of proposition 2. More precisely, separability is satisfied if the slope coefficients in  $\hat{\beta}_{b,j}$  would be the same for all schools. Table 5 summarizes the separability tests based on model (13).

#### Table 5

The 'equal slope'-hypothesis is statistically rejected for each background variable separately, as well as for all background variables jointly.

The theoretical consequences of non-separability have been described before. Incentives for good administration may create incentives for pupil selection and vice-versa, no incentives for pupil selection may create incentives for bad administration. We now turn to the empirical relevance.

#### 3.3 The trade-off in practice

Recall the linear subsidy scheme defined in equation (3). We specify the constant and the slope. Suppose the regulator faces a budget constraint, i.e., the average subsidy per pupil has to be equal to the available budget per pupil. If we normalize the available budget to be 1 unit per pupil, then the per-pupil subsidy at school j becomes

$$s_j = 1 + \text{slope} \times (\widetilde{y}_j - \widetilde{y}),$$

with  $\overline{\widetilde{y}}$  the average (corrected) output.

A natural additional constraint is to guarantee each school a minimal subsidy per pupil, say <u>s</u>, with  $0 < \underline{s} < 1$ . A minimal subsidy requirement imposes an upper bound on the **slope**. If we arbitrarily fix the minimal subsidy to be half the average subsidy ( $\underline{s} = 0.5$ ) and choose the maximal **slope** possible, then we get

slope 
$$= 0.5/(\overline{\tilde{y}} - \min \tilde{y}_j).$$
 (15)

The **slope** can be calculated for each scheme separately or for all schemes together. We choose the latter route (leading to **slope=0.4**), and comment on it later on.

We provide a formula for each subsidy scheme here; the derivation can be

found in the appendix. If a tilde  $\widetilde{\cdot}$  denotes a reference level, then we get

$$\begin{array}{lll} s_{j}^{PC} & = & 1, \\ s_{j}^{UO} & = & 1 + \texttt{slope} \times (\overline{y}_{j} - \overline{y}), \\ s_{j}^{RA} & = & 1 + \texttt{slope} \times \{(\overline{y}_{j} - \overline{y}) - \widetilde{\beta}_{b}'(\overline{z}_{b,j} - \overline{z}_{b})\}, \\ s_{j}^{RB} & = & 1 + \texttt{slope} \times \{(\overline{y}_{j} - \overline{y}) - \widetilde{\beta}_{b,j}'(\overline{z}_{b,j} - \widetilde{z}_{b}) + \overline{\beta}_{b,j}'(\overline{z}_{b,j} - \widetilde{z}_{b})\}. \end{array}$$

We also consider a value added (VA) model. Suppose one would stick to the (rejected) separable model

$$y_{ijt} = \beta_a^{+\prime} z_{a,ijt} + \beta_b^{+\prime} z_{b,ijt} + u_i^+ + v_j^+ + w_{ijt}^+, \tag{16}$$

and school output is then equal to

$$\overline{y}_j = \widehat{\beta}_a^{+\prime} \overline{z}_{a,j} + \widehat{\beta}_b^{+\prime} \overline{z}_{b,j} + \widehat{v}_j^+.$$

The part  $\hat{\beta}_a^{+'} \bar{z}_{a,j} + \hat{v}_j^+$  is usually considered to be the value-added of the school; see, e.g., Meyer (1997). If we equate corrected output  $\tilde{y}_j$  with value added, we show in the appendix that the per pupil subsidy can be written as

$$s_j^{VA} = 1 + \texttt{slope} \times \{ (\overline{y}_j - \overline{y}) - \widehat{\beta}_b^{+\prime} (\overline{z}_{b,j} - \overline{z}_b) \}.$$

Compare it with a reference administration subsidy and note indeed that it is a special case, with reference coefficients  $\tilde{\beta}_b$  equal to the estimated coefficients  $\hat{\beta}_b^+$  obtained via (16).

A final step is to choose the reference levels. The reference levels for  $\beta_b$ and  $\tilde{z}_b$  are based on the distribution of the estimated coefficients  $\hat{\beta}_{b,j}$  and the averages  $\bar{z}_{b,j}$ . We choose the 5th percentile (low), the median (mid) and the 95th percentile (high).

We consider two simulations. A first simulation focusses on the subsidy change as a consequence of a change in administration  $(\Delta \hat{\beta}_{b,j}, \Delta \hat{v}_j)$  that does not change school output  $(\Delta \bar{y}_j = 0)$ . For example, consider a school that spends more effort to help initially stronger students at the cost of the weaker ones without changing output. If subscript 0 refers to initial test scores, then  $\Delta \hat{\beta}_{0,j} > 0$  and  $\Delta \hat{v}_j < 0$ , with  $\Delta \overline{y}_j = \Delta \hat{\beta}_{0,j} \overline{z}_{0,j} + \Delta \hat{v}_j = 0$ , given (14). A second simulation looks at the subsidy change as a result of a change in one of the background characteristics in the vector, denoted  $\Delta \overline{z}_{b,j}$ . For example, if a school attracts pupils with higher initial test scores, then  $\Delta \overline{z}_{0,j} > 0$ .

Ideally, the school subsidies should not change in both simulations. But the ideal subsidy scheme does not exist, neither in theory, nor in practice. Table 6a focusses on initial test scores and shows how the subsidies would change at the different schools for the different schemes.

#### Table 6

We do not report the per-capita scheme, because per-capita subsidies do not respond to the simulations.

The reference background schemes do not provide good incentives. Schools with a pupil population stronger than the reference population will loose money, while schools with a weaker pupil population will gain. If the reference background is a weak school, then most schools will loose, and vice-versa if the reference is a strong school. If the reference is the median school, then roughly half of the schools gain and the other half looses. Note that the percentage of schools that gain and loose roughly reflect the way the reference level is constructed (the 5th percentile, median, and the 95th percentile).

It is interesting to wonder what happens if school behaviour would be introduced. The reference background level will drive behaviour. A sufficiently low reference implies that all schools loose if they choose a more elitarian school policy (that benefits the stronger at the cost of the weaker pupils). And vice-versa, all schools gain when choosing the opposite egalitarian policy. In general, the reference level will imply that schools below the reference benefit if they choose a more elitarian policy, while those above the reference benefit if they choose a more egalitarian one.

The reference administration and uncorrected output schemes do provide incentives for pupil selection. Incentives are especially strong for uncorrected output schemes. For all other schemes the incentives are moderate, especially for the median reference and the value added model. The reference slope plays a crucial role again if we turn to behaviour. A low reference implies that all schools benefit from attracting stronger pupils, while a high reference implies that schools gain from attracting weaker pupils. If school segregation is an issue, the latter solution provides incentives for more integration.

## 4 Conclusion

To be completed.

## References

- Barlevy, G., and D., Neal, 2012, Pay for percentile, American Economic Review 102(5), 1805-1831.
- [2] Burgess, S., C., Propper, H., Slater, and D., Wilson, 2005, Who wins and who loses from school accountability? The distribution of educational gain in English secondary schools. CMPO, Bristol: Working Paper 05/128.
- [3] Burgess, S., C., Propper, and D., Wilson, 2007, The impact of school choice in England. *Policy Studies* 28(2), 129-143.
- [4] Cameron, C., and P., Trivedi, 2005, *Microeconometrics*, Cambridge University Press.
- [5] Cawley, J., J., Heckman, and E., Vytlacil, 1999, On policies to reward the value added by educators, *Review of Economics and Statistics* 81(4), 720-727.
- [6] Chiang, H., 2009, How accountability pressure on failing schools affects student achievement, *Journal of Public Economics* 93, 1045-1057.
- [7] Del Rey, E., 2004, Funding schools for greater equity, Regional Science and Urban Economics 34, 202-224.
- [8] Epple, D., and R., Romano, 2008, Educational vouchers and cream skimming, *International Economic Review* 49(4), 1395-1435.
- [9] Figlio, D., and C., Rouse, 2006, Do accountability and voucher threats improve low-performing schools? *Journal of Public Economics* 90, 239-255.
- [10] Figlio, D., and J., Winicki, 2005, Food for thought: the effects of school accountability plans on school nutrition, *Journal of Public Economics* 89, 381-394.

- [11] Fleurbaey, M., 2008, Fairness, Responsibility and Welfare, Oxford University Press.
- [12] Fleurbaey, M., and F., Maniquet, 2011, A Theory of Fairness and Social Welfare, Cambridge University Press.
- [13] Hanushek, E., 2006, School resources, chapter 14 in, E., Hanushek, and F., Welch, eds., *Handbook of the Economics of Education* vol 2, Elsevier.
- [14] Hanushek, E., and M., Raymond, 2003, Lessons about the design of state accountability systems, in, P., Peterson, and M., West, eds., No Child Left Behind? The Politics and Practice of Accountability, Brookings.
- [15] Hanushek, E. and M., Raymond, 2004, The effect of school accountability systems on the level and distribution of student achievement, *Journal of* the European Economic Association 2(2/3), 406-415.
- [16] Hanushek, E., and M., Raymond, 2005, Does school accountability lead to improved student performance? Journal of Policy Analysis and Management 24(2), 297-327.
- [17] Hausman, J, and D., Wise, 1979, Attrition bias in experimental and panel data: the Gary income maintenance experiment. *Econometrica* 47(2), 455-473.
- [18] Holmstrom, B., and P., Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica* 55(2), 303-328.
- [19] Jacob, B., 2005, Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago public schools, *Journal of Public Economics* 89, 761-796.
- [20] Kane, T., and D., Staiger, 2002, The promise and pitfalls of using imprecise school accountability measures, *Journal of Economic Perspectives* 16(4), 91-114.

- [21] Little, R., and D., Rubin, 1987, Statistical Analysis with Missing Data, Wiley.
- [22] Ladd, H., and R., Walsh, 2002, Implementing value-added measures of school effectiveness: getting the incentives right, *Economics of Education Review* 21, 1-17.
- [23] Laffont, J.-J., and D., Martimort, 2002, The Theory of Incentives, Princeton University Press.
- [24] Longford, N., 1994, Random Coefficient Models, Oxford University Press.
- [25] Milgrom, P., and J., Roberts, 1992, Economics, Organization and Management, Prentice Hall.
- [26] Meyer, R., 1997, Value-added indicators of school performance: a primer, Economics of Education Review 16(3), 283-301.
- [27] Neal, D., 2008, Designing incentive systems for schools, in, M., Springer, ed., Performance Incentives: Their Growing Impact on American K-12 Education, Brookings.
- [28] Reback, R., 2008, Teaching to the rating: school accountability and the distribution of student achievement, *Journal of Public Economics* 92, 1394-1415.
- [29] Schokkaert, E., G., Dhaene, and C., Van de Voorde, 1998, Risk adjustment and the trade-off between efficiency and risk selection: an application of the theory of fair compensation. *Health Economics* 7, 465-480.
- [30] Schokkaert, E., and C., Van de Voorde, 2004, Risk selection and the specification of the conventional risk adjustment formula, *Journal of Health Economics* 23, 1237-1259.
- [31] Taylor, J., and A., Ngoc Nguyen, 2006, An analysis of the value added by secondary schools in England: is the value added indicator of any value? Oxford Bulletin of Economics and Statistics 68(2), 203-224.

- [32] Verbeek, M., and T., Nijman, 1992, Testing for selectivity bias in panel data models, *International Economic Review* 33, 681-703.
- [33] West, M., and P., Peterson, 2006, The efficacy of choice threats within school accountability systems: results from legislatively induced experiments, *Economic Journal* 116, 46-62.
- [34] Wooldridge, J., 1995, Selection corrections for panel data models under conditional mean independence assumptions, *Journal of Econometrics* 68, 115-132.
- [35] Wössmann, L., 2003, Schooling resources, educational institutions and student performance: the international evidence, Oxford Bulletin of Economics and Statistics 65(2), 117-170.

## Proof of proposition 2

A subsidy scheme can satisfy INCENTIVES FOR GOOD ADMINISTRATION and NO INCENTIVES FOR PUPIL SELECTION if and only if there exist functions g:  $\mathbb{R} \times B \to \mathbb{R}$  and  $h: A \to \mathbb{R}$ , with g strictly increasing in its first argument, such that f(a, b) = g(h(a), b), for all x = (a, b) in X.

If the separability condition holds, it is possible to define a subsidy scheme s such that each school subsidy  $s_j$  is a strictly increasing function of  $h(a_j)$  only. Such a scheme satisfies both axioms. We show the opposite.

Consider a subsidy scheme that satisfies INCENTIVES FOR GOOD ADMIN-ISTRATION and NO INCENTIVES FOR PUPIL SELECTION. We show that, for arbitrary administrations  $a, a' \in A$  and backgrounds  $b, b' \in B$ , we have

$$f(a,b) \ge f(a',b) \Leftrightarrow f(a,b') \ge f(a',b').$$
(17)

This would indeed allow to properly define functions

1.  $h: A \to \mathbb{R}$  with  $h(a) \ge h(a')$  if  $f(a, b) \ge f(a', b)$  for some  $b \in B$ , and

2. 
$$g : \mathbb{R} \times B \to \mathbb{R}$$
 with  $g(h(a), b) = f(a, b)$  for all  $x = (a, b)$ ,

and g will be strictly increasing in its first argument.

We proceed by contradiction. Suppose equation (17) does not hold, e.g., both  $f(a,b) \ge f(a',b)$  and f(a,b') < f(a',b') are true for some  $a, a' \in A$  and  $b,b' \in B$ . (It is easy to verify the other direction using the same logic.) We can use these  $a, a' \in A$  and  $b, b' \in B$  to construct four states—(a,b), (a',b), (a,b'), and (a',b')—for some school (tacitly assuming that school information remains constant for all other schools). We suppress subscripts and use f(a,b) and (with slight abuse of notation) s(a,b) to refer to the output and the subsidy of the school under consideration. Applying INCENTIVES FOR GOOD ADMINISTRATION twice, we must have

$$s(a,b) - s(a',b) \ge 0$$
 and  $s(a,b') - s(a',b') < 0.$  (18)

Applying NO INCENTIVES FOR PUPIL SELECTION twice, we obtain

$$s(a, b) = s(a, b')$$
 and  $s(a', b) = s(a', b')$ ,

and, subtracting both equations, we get:

$$s(a,b) - s(a',b) = s(a,b') - s(a',b').$$
(19)

Equation (18) and (19) are incompatible, a contradiction.

## A derivation of the empirical subsidy schemes

The per-capita and uncorrected output schemes are straightforward. We discuss the reference administration, reference background and value added scheme. A subsidy scheme is defined as

$$s_j = 1 + \text{slope} \times (\widetilde{y}_j - \widetilde{y}),$$

with the slope defined by (15) for each scheme. We focus here on the difference  $\tilde{y}_j - \overline{\tilde{y}}$ .

We start from the empirical model

$$\overline{y}_j = \widehat{\beta}'_a \overline{z}_{a,j} + \widehat{v}_j + \widehat{\beta}'_{b,j} \overline{z}_{b,j} = f(\underbrace{\overline{z}_{a,j}, \widehat{v}_j, \widehat{\beta}'_{b,j}}_{a_j}, \underbrace{\overline{z}_{b,j}}_{b_j}),$$

$$= f(\underbrace{a_j}_{a_j}, \underbrace{b_j}_{b_j}).$$

The RA models use a reference administration, say  $\tilde{a} = (\tilde{z}_a, \tilde{v}, \tilde{\beta}_b)$ , to define the hypothetical output as

$$\widetilde{y}_j = \overline{y}_j - f\left(\widetilde{a}, b_j\right) = \overline{y}_j - \left(\beta'_a \widetilde{z}_a + \widetilde{v} + \widetilde{\beta}'_b \overline{z}_{b,j}\right).$$

The average hypothetical output is equal to

$$\overline{\widetilde{y}} = \overline{y} - (\beta'_a \widetilde{z}_a + \widetilde{v} + \widetilde{\beta}'_b \overline{z}_b),$$

and the difference  $\widetilde{y}_j - \overline{\widetilde{y}}$  is indeed equal to

$$(\overline{y}_j - \overline{y}) - \widetilde{\beta}'_b(\overline{z}_{b,j} - \overline{z}_b).$$

Starting from the same empirical model, the RB models replace  $\overline{z}_{b,j}$  by a reference background  $\tilde{b} = \tilde{z}_b$  to get

$$\widetilde{y}_j = f(a_j, \widetilde{b}) = \widehat{\beta}'_a \overline{z}_{a,j} + \widehat{v}_j + \widehat{\beta}'_{b,j} \widetilde{z}_b.$$

The OLS estimate for  $\hat{v}_j$  is

$$\widehat{v}_j = \overline{y}_j - \widehat{\beta}'_a \overline{z}_{a,j} - \widehat{\beta}'_{b,j} \overline{z}_{b,j},$$

and we can rewrite the hypothetical output as

$$\widetilde{y}_j = \overline{y}_j - \widehat{\beta}'_{b,j}(\overline{z}_{b,j} - \widetilde{z}_b).$$

The average is given by

$$\overline{\widetilde{y}}_j = \overline{y} - \overline{\widehat{\beta}'_{b,j}(\overline{z}_{b,j} - \widetilde{z}_b)},$$

and the difference  $\widetilde{y}_j - \overline{\widetilde{y}}_j$  indeed becomes

$$(\overline{y}_j - \overline{y}) - \widehat{\beta}'_{b,j}(\overline{z}_{b,j} - \widetilde{z}_b) + \overline{\widehat{\beta}'_{b,j}(\overline{z}_{b,j} - \widetilde{z}_b)}.$$

Finally, for the value-added (VA) model we have

$$\widetilde{y}_j = \widehat{\beta}_a^{+\prime} \overline{z}_{a,j} + \widehat{v}_j^+,$$

with the OLS estimate of  $v_j^+$  in (16) to be

$$\widehat{v}_j^+ = \overline{y}_j - \widehat{\beta}_a^{+\prime} \overline{z}_{a,j} - \widehat{\beta}_b^{+\prime} \overline{z}_{b,j}.$$

Plugging in the OLS estimate, corrected output becomes

$$\widetilde{y}_j = \overline{y}_j - \widehat{\beta}_b^{+\prime} \overline{z}_{b,j}.$$

Averaging the corrected output, we get

$$\overline{\widetilde{y}} = \overline{y} - \widehat{\beta}_b^{+\prime} \overline{z}_b,$$

and the difference  $\widetilde{y}_j - \overline{\widetilde{y}}$  indeed reduces to

$$(\overline{y}_j - \overline{y}) - \widehat{\beta}_b^{+\prime} (\overline{z}_{b,j} - \overline{z}_b).$$

## Figures and tables

Figure 1. Aligning performance and selection incentives: mission impossible.



incentive	for	good adminis	tration		for pupil selec	etion
	always	magnitude	# schools	never	magnitude	# schools
PC		0	0	$\checkmark$	0	0
RB		low	some	$\checkmark$	0	0
RA	$\checkmark$	high	all		low	some
UO	$\checkmark$	high	all		high	all

 Table 1. Different schemes provide different incentives (in theory).

5		TRADEATERING WITH COORTINATION OF ANTRONO.
math		test score result in mathematics at the end of grade 1 and 2
time2	= 1,	if pupil is in second grade, 0 otherwise
$\mathrm{math}_0$		initial test score result in mathematics at the start of grade 1
girl	= 1,	if girl, 0 otherwise
$m\_dutch/f\_dutch$	= 1,	if mother/father speaks dutch with pupil, 0 otherwise
m_edu_sec/f_edu_sec	= 1,	if mother/father has a secondary education degree, $0$ otherwise
m_edu_high/f_edu_high	= 1,	if mother/father has a tertiary (short type) education degree, $0$ otherwise
m_edu_uni/f_edu_uni	= 1,	if mother/father has a tertiary (long type) education degree, 0 otherwise
duo	= 1,	if there are two teachers, 0 otherwise
peer		average initial test score of peers, i.e., fellow pupils in the class
$time\_math$		mathematics instruction in the classroom in hours per week
experience		(average) teaching experience of teacher(s) in years
class_size		number of pupils in the classroom

**Table 2.** Abbreviation and description of the variables.

math score	mean	std.dev.	p10	median	p90
grade 1	8.75	1.00	7.44	8.77	10.06
grade 2	9.71	1.00	8.43	9.71	11.04
initial math score	mean	std.dev.	p10	median	p90
grade 1	8.05	1.02	6.71	8.13	9.31
grade 2	8.16	0.97	6.85	8.22	9.37
sex	= boy	= girl			
grade 1	50.54%	49.46%			
grade 2	50.58%	49.15%			
language mother	= dutch	$\neq \mathrm{dutch}$	miss.		
grade 1	86.82%	8.61%	4.57%		
grade 2	85.69%	9.12%	5.19%		
language father	= dutch	$\neq \mathrm{dutch}$	miss.		
grade 1	84.80%	10.29%	4.91%		
grade 2	84.54%	9.90%	5.56%		
mother's highest	$< 2^{ary}$	$2^{ary}$	$3^{\mathrm{ary}}$	$3^{\mathrm{ary}}$	miss.
degree			$(\neq \text{univ.})$	(=univ. $)$	
grade 1	19.25%	33.94%	29.06%	8.72%	9.03%
grade 2	16.54%	33.56%	30.41%	10.00%	9.49%
father's highest	<2 <sup>ary</sup>	$2^{ary}$	$3^{\mathrm{ary}}$	$3^{\mathrm{ary}}$	miss.
degree			$(\neq univ.)$	(=univ $.)$	
grade 1	19.29%	34.74%	21.00%	12.07%	12.90%
grade 2	17.49%	34.78%	22.10%	13.39%	12.24%

 Table 3a. Summary statistics for pupil variables.

# of teachers	= 1	= 2			
grade 1	89.33%	10.67%			
grade 2	86.27%	13.73%			
instruction time	mean	std.dev.	p10	median	p90
grade 1	6.17	0.86	5	6	7
grade 2	6.30	0.87	5.5	6	7
total experience	mean	std.dev.	p10	median	p90
grade 1	15.15	8.95	4	15	28
grade 2	17.67	9.37	4	18	30
class size	mean	std.dev.	p10	median	p90
grade 1	20.12	3.80	15	20	26
grade 2	20.24	4.08	15	20	26
peer effect	mean	std.dev.	p10	median	p90
grade 1	8.05	0.47	7.48	8.13	8.55
grade 2	8.16	0.48	7.62	8.27	8.64

Table 3b. Summary statistics for class variables.

math	mod	lel a	mod	lel b	mod	lel c	mod	lel d
	coeff.	p >  t						
constant	3.35	0.00	8.34	0.00	3.75	0.00	2.15	0.00
time2	0.82	0.00	0.82	0.00	0.82	0.00	0.80	0.00
$\mathrm{math}_0$	0.67	0.00			0.65	0.00	0.65	0.00
girl			-0.26	0.00	-0.24	0.00	-0.24	0.00
m_dutch			0.16	0.00	-0.13	0.00	-0.14	0.00
f_dutch			0.10	0.03	-0.07	0.07	-0.07	0.08
m_edu_sec			0.17	0.00	0.00	0.86	0.01	0.85
m_edu_high			0.46	0.00	0.10	0.00	0.10	0.00
m_edu_uni			0.55	0.00	0.19	0.00	0.19	0.00
f_edu_sec			0.08	0.02	0.06	0.01	0.07	0.01
f_edu_high			0.25	0.00	0.14	0.00	0.14	0.00
f_edu_uni			0.37	0.00	0.24	0.00	0.24	0.00
duo							-0.10	0.01
peer							0.10	0.27
$time_math$							0.08	0.00
experience							0.00	0.25
class_size							0.01	0.03

 Table 4. Explaining math test scores.

	<i>F</i> -value	$\mathrm{Prob} > F$
initial test score	4.46	0.00
girl	3.89	0.00
mother dutch	8.46	0.00
father dutch	12.07	0.00
education mother	7.37	0.00
education father	6.16	0.00
all variables	$9 \times 10^{7}$	0.00

Table 5. Educational production is not likely to be separable.

practice
incentives
different
provide
schemes
Different
6.
Table

	Table	<b>6.</b> Diff	erent sc	hemes	provide .	different	incenti	ves (in	practic	æ).	
				6(a).	Initial to	est score	10				
incentive	for		good	adminis	stration			nd	pil selec	ction	
change	II	$\Delta\widehat{eta}_{j,j}$	$_{k}=\sigma(\beta_{j,.})$	$_k$ ) & $\Delta i$	$\widehat{v}_j$ s.t. $\Delta_{\overline{i}}$	$\overline{y}_j = 0$		$\Delta \overline{z}_l$	$\sigma_{k}=\sigma_{k}$	$(\overline{z}_{b,k})$	
measure	II			$\Delta s_j$					$\Delta s_j$		
ideally	II		zer	o everyv	vhere			zerc	every	where	
statistic	11	p10	p50	p90	%<0	%>0	p10	p50	$^{ m p90}$	%<0	%>0
	$\widetilde{eta}$ low		zero	o everyv	vhere		0.00	0.03	0.07	5.3%	94.7%
RA, with	$\widetilde{eta}$ mid		zero	o everyv	vhere		-0.03	0.00	0.03	50.7%	49.3%
	$\widetilde{\beta}$ high		zero	o everyv	vhere		-0.07	-0.04	0.00	95.4%	4.6%
VA			zero	o everyv	vhere		-0.03	0.00	0.03	47.6%	52.4%
	$\widetilde{z}  \log$	-0.08	-0.05	-0.02	95.9%	4.1%		zerc	everyv	$_{\rm where}$	
RB, with	$\widetilde{z}$ mid	-0.02	0.00	0.03	51.3%	48.7%		zerc	every	$_{\rm where}$	
	$\widetilde{z}$ high	0.00	0.03	0.06	5.6%	94.4%		zerc	everyv	where	
ΟŪ			zer	o everyv	vhere		0.11	0.14	0.17	0%	100%