

## Efficiency versus Effectiveness: Interpreting Education Production Studies

CHRISTOPHER C. KLEIN\*

*Middle Tennessee State University, Murfreesboro, TN*

### *Abstract*

To gain analytical insight into whether input resources matter in public education, a Becker/Peltzman/Stigler model of the determination of local educational budgets and outputs by political authorities is constructed. The model results are consistent with empirical findings that resources don't matter, even when all schools are efficient, if errors in measurement and specification occur. When all outputs are not observed, one cannot distinguish an inefficient school district from one that chooses an idiosyncratic output mix. Blind application of efficiency measurement techniques in this context yields perverse or counterintuitive findings. Interpretation of feasible approaches to education production studies are discussed.

*Key words:* Education, Efficiency, Productivity

*JEL category:* I12

\* Christopher C. Klein. Associate Professor, Economics and Finance Department, Middle Tennessee State University, Murfreesboro, TN 37132, USA, phone: 615-904-8570, fax:615-898-5045, e-mail:[cklein@mtsu.edu](mailto:cklein@mtsu.edu)

## **1. Introduction**

Failure to measure all outputs of the educational process and to take account of the endogeneity of school outputs and expenditures seriously undermines the integrity of efficiency measures based on estimated cost or production frontiers, or related reduced forms. These problems have been noted in the literature (Hanushek 1979, 1986), yet the flow of research that interprets results of such studies without appreciation of the fundamental difficulties involved continues (Ruggiero & Vitaliano, 1999; Hoxby, 2000; Chakraborty, Biswas, & Lewis, 2001; Grosskopf, Hayes, Taylor, & Weber, 2001; Abbott & Doucouliagos, 2003; Bonesronning, 2003; Dolton, Marcenaro & Navarro, 2003).<sup>1/</sup> My aim here is to build a model in which school budgets and outputs are endogenous and all schools (or districts) are efficient, yet in the presence of measurement and specification error, yields results consistent with the empirical literature – that resources “don’t matter” to achievement as much as demographic characteristics of students do and that variations in measured efficiency exist. This suggests, at worst, that the results of empirical educational production studies are mere artifacts of measurement and specification error or, at best, that considerable caution is called for in interpreting these results.

The same issues could be raised in criticism of most production studies, but in public education the outputs by general consensus are amorphously qualitative and greatly more numerous - compare education to electricity generation, for example – magnifying the potential effect of such errors. Moreover, educational outputs are influenced by a political process that can respond to local differences in demand for public education in both budgetary (input) and output dimensions. Does a rural school offering the minimum academic classes, lacking arts and sports, perform more poorly on standardized tests than a suburban school offering a full complement of academic and nonacademic programs, even though state funding allows the rural school to spend more per student, because of diseconomies of scale? Or is it because few

students from low income households aspire to a college education, preferring local manufacturing jobs or farming?<sup>2</sup> How do social goals, such as racial integration and the mainstreaming of special education, physically handicapped, and English as a second language students factor in? Can the researcher evaluate efficiency in this context?

These concerns differ from those raised in other recent articles. Ruggiero (2003) and Bifulco & Bretschneider (2003) investigate the econometric effect of measuring one observed output with error; I consider two outputs, one of which is observed without error while the other is not observed at all. Pritchett & Filmer (1999) consider production distortions that may arise through teachers' influence on input usage, ignoring output choices, whereas I focus on output choice and measurement assuming efficient input use. Wenger (2000) argues for examining multiple educational outputs, but does not derive the full implications of failure to do so. Further, unlike Bishop and Wößmann (2004), the model is agnostic as to the cause of output choice variation across schools.

The "educational production function" concept was suggested as a viable approach to educational research as early as the late nineteen-sixties.<sup>3</sup> Subsequent studies, briefly reviewed here in the next section, typically viewed test scores as the single output of an educational process characterized as a function of educational inputs and student demographics. These have culminated in the current controversy over whether "resources matter" in education (Hanushek 2003; Krueger 2003). This controversy arises from the frequent, yet seemingly perverse, empirical finding that students' demographic characteristics and family background better explain their performance on standardized tests than do measures of the resources devoted to their education. These demographic and background characteristics include income, wealth, parents' educational attainment and socio-economic status indicators that one normally

associates with determinants of demand. Thus, the results of past studies may have confounded demand-side and supply-side effects.

The success of public policies aimed at raising the educational attainment of the general populace is crucially dependent on a coherent resolution of this controversy. If the resources-don't-matter school is correct, then the current experimentation with improving school performance, as in the U.S. No Child Left Behind initiative, will go for naught. A more effective policy could aim at raising the health, income, and wealth of households. A key component of household income and wealth, however, is educational attainment. Thus, our current state of knowledge pushes policy reasoning in an unproductive circle.

The purpose of this paper is to build a model of public educational funding and production that can generate outcomes consistent with the empirical findings, while providing some insights to guide future research toward an escape from the current cycle. Section 3 presents a Becker/Peltzman/Stigler model (Becker 1983; Peltzman 1976; Stigler 1971) of the political choice of educational outputs subject to a budget constraint, also politically determined. The model allows the educational decision-maker to choose quantities of two outputs simultaneously, recognizing that schools by default or design produce more than mere test scores. A demographic characteristic or index is introduced that shifts the political demand for education in favor of the observed output and increases willingness to pay in the form of the schools budget. This aspect is similar to fixed-effects empirical models such as Ram (2004).

First, to examine the limiting case, efficient production is assumed such that each educational institution operates on the relevant efficient cost frontier. In deriving the implications of the model for measuring the economic efficiency of public schools, only one output is observed, although the results generalize to the case in which only a subset of many

multiple outputs are observed. This simple model generates observations consistent with the empirical literature in which variations in demographics appear to cause variations in educational productivity across institutions even though all production takes place on the efficient frontier. Relaxing the assumption of efficiency allows inefficient schools to appear more efficient than efficient schools when only one output is observed.

The implications of this result for future research are then explored. Techniques such as stochastic cost frontiers and data envelopment analysis can account for multiple outputs; instrumental variables methods can account for the simultaneous determination of outputs and expenditures by school districts. Relative efficiency measures of educational institutions, however, require complete measurement of *all* outputs and inputs (expenditures). This is unlikely due to the broad range of outputs involved. An analysis of the available output measures can identify institutions that are relatively effective, given their student characteristics and expenditure levels, at producing the subset of outputs measured. The relative efficiency of educational institutions in general cannot be inferred from such studies. A conclusion follows.

## **2. Literature on School Productivity**

The “Coleman Report” (Coleman , 1966) suggested that differences in schools had little to do with differences in students’ performance, whereas family background and the characteristics of students’ peers were more important. Hanushek’s (1986) review of studies completed through the mid-1980s came to much the same conclusion: that evidence linking the level of per-student expenditures, or other inputs, to student achievement is extremely weak and disappears when differences in family background are taken into account. A decade later, Card & Krueger (1996) undertook a “meta-analysis” of multiple studies of education, concluding that

“school resources tend to be positively associated with earnings and educational attainment, but the relationship is not always robust to specific features of the data set or empirical specification” (p. 33). Nevertheless, the debate over the role of resources in education continues (Krueger 2003; Hanushek 2003).<sup>4</sup>

Indeed the difficulties in measurement presented by both the dependent and independent variables have long been recognized. Hanushek observes that “adequate measures of innate abilities have never been available” and that whereas “education is cumulative, frequently only contemporaneous measures of inputs are available,” (Hanushek 1986, p. 1156). Not only are test scores unreliable indicators of performance, showing substantial variation in response to small variations in the student sample, but they are notoriously incomplete, assessing math and reading skills, for example, while ignoring science and civics (Kane & Staiger 2002). Further, test scores are an imperfect measure of the *value* of education. Test scores add little or nothing to a standard wage equation, although “a number of studies” find a positive and statistically significant relationship between educational resources and students’ educational attainment and earnings (Card & Krueger 1996, p. 32). These measurement errors, specification errors, and omitted variables can cause biased and inconsistent estimates of the relationships among school resources and student outcomes.

Potential endogeneity raises additional problems. Hanushek (1986) suggests that the observed correlation of teacher experience with student outcomes may reflect a seniority system that allows more senior teachers to choose assignments at schools with better resources or serving higher ability student populations. Card & Krueger (1996) suggest that if wealthier students stay in school longer and earn more later in life due to family connections, regardless of their education level, while also demanding smaller class sizes, even though this has no effect on

school “quality,” then a spurious correlation could be observed between school resources and both educational attainment and earnings. The ability of schools possessing greater resources to attract stronger students, whether through tuition “subsidies” in the case of private schools or Tiebout (1956) effects for public schools, could generate similar spurious associations (Rothschild & White, 1995; Epple & Romano, 1998; Ferris & West, 2002; Hoxby, 1996; Lazear, 2001; Winston, 1999). As Becker (1997, p. 1367) observes with respect to teaching methods, it should not be surprising that “single equation methods, with potentially endogenous regressors, simply may not be able to capture the differences that we are trying to produce.”

Nevertheless, the literature on educational “productivity” continues to grow. Two strands of this literature are relevant here. In the first, a single performance measure (test scores) is related to school district level inputs and student demographics. Parametric or non-parametric (DEA) frontier techniques are used to construct relative efficiency measures from the output and input data and these (in)efficiency measures are then regressed on demographic variables. Variations in the demographic variables are found to “explain” the variations in relative (in)efficiency. Some studies invert this process, relating expenditures per pupil to test scores and demographic variables, with essentially comparable results (Chakraborty et al. 2001; Ruggiero & Vitaliano 1999; Grosskopf et al. 2001; Wenger 2000).

The second strand relates similar performance measures to political variables (local versus state control, school or district choice, unionization) as well as demographic characteristics. Variables reflecting the degree of local control and/or ability of individuals to choose among various school districts are positively related to test scores and negatively related to expenditures, leading some to conclude that competition makes schools more efficient (Grosskopf et al. 2001; Hoxby, 2000; Peltzman, 1993, 1996).<sup>5</sup>

Consequently, what we may know about educational productivity is very limited and subject to many qualifications. Higher ability or better prepared students appear to score higher on tests. Variations in educational inputs do not appear to influence test scores or expenditures per pupil as much as do variations in the mean demographic backgrounds of student populations. Institutional settings in which households may choose among public educational providers, if only in a Tiebout (1956) manner, are associated with higher test scores and lower per pupil expenditures.

Yet all of these conclusions are suspect. For example, Hoxby (2000) pays close attention to endogeneity and takes pains to include multiple output measures. Nevertheless, four of her six achievement measures are math and reading test scores, while one is the highest grade attained. None of these measure achievement in science, social science, or the arts, much less the value of sports or the inculcation of values appropriate for good citizenship. Only her income measure, the log of income at age 32, is more general. Failure to measure all the outputs does not damage her finding that greater school choice (or competition) is associated with higher achievement *along the measured dimensions*. The finding that more choice/competition is associated with lower per student spending, however, is not sufficient to imply that school district choice promotes *economic* efficiency. Reduced spending may be accomplished by reducing the inputs used to produce unobserved outputs, as in times of tight budgets school districts often cut sports and arts programs first, rather than by true efficiency gains that result in producing the same or more output with fewer inputs.<sup>6</sup> If one fails to measure all outputs, these two cases may be indistinguishable.

Future research needs more specific modeling to guide further empirical inquiries. Todd & Wolpin (2003), for example, construct such a model for individual students that provides

substantial insight to guide empirical research using student-level data. In the next section, a similar exercise is undertaken for observations at the public school or public school district level, as a simple political choice model is constructed that allows for multiple school outputs, endogeneity of output and expenditure choices, and failure to observe all outputs in the context of efficient production.

### **3. Choice of Educational Outputs**

Suppose that the relevant governmental architect of education policy, whether school boards and other local officials, or state and national entities, seeks to maximize a political value function,  $V$ , embodying the probability of election/re-election or reappointment, by choosing the output mix of local schools, subject to a budget constraint. One may view  $V$  as a majority generating function as in Peltzman (1976) or as the utility of the median voter (Downes & Pogue; Peltzman, 1993).<sup>7</sup> In any case,  $V$  reflects the underlying demand for educational services by voter/households as viewed by the political system governing public schools.

The following discussion is framed as if the decision-makers operate at the school district level, as this is the level at which public school budgets are set and at which school boards make decisions. Nevertheless, school boards indirectly may assign resources differentially among individual schools within a district by, for example, assigning better educated or more experienced teachers to schools serving areas of high demand for education.<sup>8</sup> In this sense, the model also may apply at the level of individual schools. On the other hand, in some states school district budgets are set at the state level, leaving little budgetary discretion for individual school boards.

Suppose there are two possible outputs – say “academic achievement” vs. art and music instruction, or sports programs, or even graduation rates<sup>9</sup> - then the problem for the decision maker can be written as

$$(1) \quad \text{Max}_{Q_1, Q_2} V(Q_1, Q_2; X) \quad \text{s.t.} \quad C(Q_1, Q_2) = B(X)$$

where  $Q_1$  and  $Q_2$  are the outputs and  $X$  represents a demographic background characteristic of students, such as household income or wealth, parents’ educational attainment, or a composite index of such characteristics.  $B(X)$  allows the budget constraint to shift with demographic factors reflecting the demand for education. The budget is not determined by the same political process as are the output quantities, because the school boards that choose outputs, inputs, and policies often lack budgetary control. School district budgets are often set by city or county executives or governing bodies with some funding provided by the state.  $B(X)$  represents the willingness of the populace to pay for public schools as a function of demographic characteristics.

$V$  and  $C$  are assumed to conform to the usual characteristics of utility and cost functions, respectively, with  $C$  representing cost-minimizing behavior under efficient production given fixed input prices (suppressed):

$$(2) \quad V_i > 0, V_{ii} < 0, V_{ij} \geq 0 \quad \text{for } i, j = 1, 2$$

$$(3) \quad C_i > 0, C_{ii} > 0 \quad \text{for } i, j = 1, 2$$

where subscripts denote partial derivatives. Note that changes in  $C_{ij}$  are unrestricted, allowing for both economies and diseconomies of scope.

Forming the Lagrangean

$$L = V(Q_1, Q_2, X) + u[B(X) - C(Q_1, Q_2)]$$

and differentiating with respect to  $Q_1$ ,  $Q_2$ , and  $u$  yields the first order conditions for a maximum:

$$V_1 - uC_1 = 0$$

$$(4) \quad V_2 - uC_2 = 0$$

$$B(X) - C(Q_1, Q_2) = 0.$$

These imply

$$(5) \quad V_1 / V_2 = C_1 / C_2$$

or that the relative marginal political values are equal to the respective relative marginal costs, or “prices,” of the two outputs. Alternatively, the marginal rate of substitution in political “consumption” is equal to the ratio of the shadow prices of the two outputs. In any case, the outcome is efficient in the sense that the relative marginal benefits equal the relative marginal costs of the two outputs.

Further, we can solve for the reaction of the output choices to a change in demographics by differentiating the first order conditions with respect to  $Q_1$ ,  $Q_2$ ,  $u$ , and  $X$  to get:

$$(6) \quad \begin{pmatrix} V_{11} - uC_{11} & V_{12} - uC_{12} & -C_1 \\ V_{21} - uC_{21} & V_{22} - uC_{22} & -C_2 \\ -C_1 & -C_2 & 0 \end{pmatrix} \begin{pmatrix} dQ_1/dX \\ dQ_2/dX \\ du/dX \end{pmatrix} = \begin{pmatrix} -V_{1X} \\ -V_{2X} \\ -B_X \end{pmatrix}$$

Solving by Cramer’s Rule yields

$$(7) \quad dQ_1/dX = \{C_2[V_{1X}C_2 - V_{2X}C_1] - B_X[-C_2(V_{12} - uC_{12}) + C_1(V_{22} - uC_{22})]\}/[A]$$

$$(8) \quad dQ_2/dX = \{-C_1[V_{1X}C_2 - V_{2X}C_1] + B_X[-C_2(V_{11} - uC_{11}) + C_1(V_{21} - uC_{21})]\}/[A]$$

Where  $[A]$  is the determinant of the matrix of second order partial derivatives of  $L$  in equation (6), which must be positive by the second order condition for a maximum. The signs and/or magnitudes of both equations depend on the signs of  $C_{12}$  and  $C_{21}$ , which in turn determine the signs/magnitudes of the bracketed terms involving  $B_X$ . To simplify the initial analysis, consider two cases:  $B_X = 0$  and  $B_X > 0$ .

### 3.1 CASE I: $B_X = 0$

In this case, the budget is given exogenously and is unaffected by demographic or “demand” factors. This gives

$$dQ_1/dX = \{C_2[V_{1X}C_2 - V_{2X}C_1]\}/[A]$$

$$dQ_2/dX = \{-C_1[V_{1X}C_2 - V_{2X}C_1]\}/[A]$$

Where the signs depend on the signs of the bracketed term in the numerator. These terms can be rewritten by solving the first order condition in equation (5) for  $C_1$  and substituting this expression into the equations above:

$$(9) \quad dQ_1/dX = \{C_2[V_{1X}C_2 - V_{2X}C_2(V_1/V_2)]\}/[A] = \{(C_2)^2 (V_1/X)[\epsilon_{1X} - \epsilon_{2X}]\}/[A]$$

$$(10) \quad dQ_2/dX = \{-C_1[V_{1X}C_2 - V_{2X}C_2(V_1/V_2)]\}/[A] = \{-C_1C_2(V_1/X)[\epsilon_{1X} - \epsilon_{2X}]\}/[A]$$

where  $\epsilon_{iX}$  is the elasticity of the marginal value of output  $i$  with respect to demographics  $X$ .

Equations (9) and (10) have opposite signs, except when the elasticities are equal ( $\epsilon_{1X} = \epsilon_{2X}$ ) and each equation is equal to zero.<sup>10/</sup> Both  $dQ_i/dX = 0$  corresponds to the case in which household demographic characteristics do not affect the demand for schooling. As our interest is in situations in which demographic changes alter education demand, hereafter it is assumed that increases in  $X$  favor output  $Q_1$ , such that

$$(11) \quad \epsilon_{1X} > \epsilon_{2X}, \quad dQ_1/dX > 0, \quad dQ_2/dX < 0.$$

That is, as a district’s demographics change to favor output  $Q_1$ , output of  $Q_1$  will increase and  $Q_2$  will decline, holding the educational budget ( $B$ ) constant.

This result is illustrated in Figure 1. For example, suppose there are two such educational policy “districts” with different demographic characteristics over the same two possible school outputs,

$$V^1 = V(Q_1, Q_2; X^1) \neq V(Q_1, Q_2, X^2) = V^2, \quad X^1 < X^2,$$

but the same cost function and budget constraint. Districts one and two choose different quantities of the two outputs according to their differing valuations of them, but produce these outputs efficiently using the identical production technology as represented by the cost functions  $C^1(\cdot) = C^2(\cdot)$ . District 1 chooses output point (F,D) and District 2 chooses point (G,E),  $G > F$  and  $E < D$ , at equal levels of cost  $C(F,D) = C(G,E) = B$ .<sup>11</sup>

FIGURE 1 ABOUT HERE

Now suppose we wish to compare the technical efficiency of the two districts, but we have data only on output  $Q_1$  and do not observe  $Q_2$ . To do this we construct the efficiency index  $I = Q_1 / (\max Q_1)$ , where  $Q_1$  is the observed output for any district and  $\max Q_1$  is the maximum feasible output of  $Q_1$  at cost level  $B$ .<sup>12</sup>

$$(12) \quad \max Q_1 = Q_1^*: C(Q_1^*, 0) = B$$

This captures the basic concept used by all such efficiency investigations, whether the actual measure is calculated by stochastic or non-stochastic means, using a production function, cost function, or other technique, such as distance functions (Chakraborty et al. 2001; Grosskopf et al. 2001; Ruggiero & Vitaliano 1999). This yields

$$(13) \quad I_2 = G/Q_1^* > I_1 = F/Q_1^*$$

in which district 2 appears to be more efficient than district 1 in producing output  $Q_1$ , even though BOTH districts are producing on the efficient cost frontier at points which are also allocatively efficient given their political preferences. No distortions due to teacher influence

(Pritchett & Filmer 1999) or competition among interest groups (Peltzman, 1993) are necessary to arrive at this result.<sup>13</sup>

Moreover, as  $dQ_1/dX > 0$  for any given budget, the district with the highest value of  $X$  will produce the most “achievement” in terms of  $Q_1$ , and will appear as the most efficient. In other words, “efficiency” ( $E$ ) and “performance” ( $Q_1$ ) appear to be “caused” by demographic factors ( $X$ ), while resources ( $C$ ) “don’t matter.” Resources don’t matter in part, because they are fixed in advance. That assumption is now relaxed.

### 3.2 CASE II: $B_X > 0$

Suppose increases in  $X$  increase the demand for education generally, such that  $B_X > 0$ , and the school district budget increases with  $X$ . If diseconomies of scope are ruled out,<sup>14</sup> such that  $C_{ij} \leq 0$ , then  $B_X > 0$  reinforces the positive magnitude of  $dQ_1/dX$  and offsets the negative magnitude of  $dQ_2/dX$ .<sup>15</sup> In fact, the positive effect of  $B_X$  could dominate the effect of changing demographics on output  $Q_2$ , such that the sign of  $dQ_2/dX$  reverses,  $dQ_2/dX > 0$  as below.

$$(14) \quad dQ_1/dX = \{C_2[V_{1X}C_2 - V_{2X}C_1] - B_X[-C_2(V_{12} - uC_{12}) + C_1(V_{22} - uC_{22})]\}/[A] > 0$$

$$(15) \quad dQ_2/dX = \{-C_1[V_{1X}C_2 - V_{2X}C_1] + B_X[-C_2(V_{11} - uC_{11}) + C_1(V_{21} - uC_{21})]\}/[A] \leq 0$$

If diseconomies of scope are allowed, then these effects of  $B_X$  are partially reversed.<sup>16</sup> Thus, while it is reasonable to expect the sign of  $dQ_1/dX$  to remain positive regardless of the budget effect ( $B_X$ ), the sign of  $dQ_2/dX$  is potentially ambiguous when  $B_X > 0$ .

FIGURE 2 ABOUT HERE

This is illustrated in Figure 2. Suppose we observe two school districts that are identical except for the demographic factor  $X$ :  $X_1 < X_2$ . District 1 chooses  $(Q_1^0, Q_2^0)$  where  $C^1(Q_1, Q_2) = B(X_1)$ , while District 2 chooses more of both outputs  $(Q_1^*, Q_2^*)$  at  $C^2(Q_1, Q_2) = B(X_2)$ .

Now consider the efficiency implications of comparing the two districts. The relationship between the efficiency indices for each district is ambiguous, depending on the proportional change in  $Q_1$  relative to the proportional change in “max  $Q_1$ ” as  $X$  changes. It is a straight forward, if tedious, exercise to show<sup>17</sup> that

$$(16) \quad dI/dX = (e_{QX} - e_{MX})(I/X)$$

where  $e_{QX}$  is the elasticity of  $Q_1$  with respect to  $X$  and  $e_{MX}$  is the elasticity of max  $Q_1$  with respect to  $X$ . Under the assumptions of the model, both elasticities will be positive, as both the quantity of  $Q_1$  produced and the budget will increase as  $X$  increases. The direction of change in the efficiency index as  $X$  increases depends on the relative magnitudes of these two elasticities.

The term  $e_{QX}$  is just equation 14 expressed as an elasticity. The term  $e_{MX}$  shows the relative proportional change in the maximum  $Q_1$  as the budget (cost) changes with  $X$  and is determined by the scale properties of the single-output cost function,  $C(Q_1, 0)$ . For example, if  $C(Q_1, 0)$  displays constant returns to scale,<sup>18</sup> then a proportional change in cost is associated with an equal proportional change in max  $Q_1$ , such that  $e_{MX} = C_1/C = B_X/B$ . Similarly, with economies of scale  $e_{MX} > B_X/B$ ; with diseconomies of scale,  $e_{MX} < B_X/B$ .

TABLE 1 ABOUT HERE

As shown in Table 1, movements in the efficiency index in response to a change in demographics do not necessarily mirror intuitive expectations about what constitutes a more

efficient use of resources. In this context, a finding that the efficiency index for producing output  $Q_1$  decreases as the school budget increases may arise in a number of situations, many of which are counter-intuitive. For example, if the demand response for  $Q_1$  is relatively elastic,  $e_{QX} > 1$ , but the school district is operating in an uneconomic region displaying diseconomies of scale,  $e_{MX} < 1$ , then increases in the schools budget make the district *appear* to become *more* efficient. In contrast, if the demand response is inelastic,  $e_{QX} < 1$ , and there are economies of scale,  $e_{MX} > 1$ , then increases in the schools budget may *appear* to *reduce* efficiency. In both cases, the schools are equally efficient in that they both produce on the efficient frontier at minimum cost.

The possibilities become even more ludicrous if inefficiency is allowed. See figure 3. The point on the frontier produces more output than the point inside the frontier, for the same cost, when both outputs are observed.<sup>19</sup> Thus, the point on the frontier is more efficient, as the true efficiency measures indicate:  $C/C = 1 > B/(A+B)$ . If only output  $Q_1$  is observed, however, then the efficient point on the frontier appears as less efficient:  $D/(D+E+F) < (D+E)/(D+E+F)$ . Clearly, efficiency measures are not reliable if all the relevant outputs are not observed.

#### **4. Discussion and Indications for Future Research**

The results above are driven by special characteristics: the failure to measure all outputs of the educational process; the influence of demographic characteristics on the choice of relative educational outputs holding spending constant; and the influence of demographic characteristics on the overall amount of educational spending. To begin, ignore the output measurement error and consider the effects of the simultaneous determination of a single output and expenditures.

Consider the estimation of a typical so-called “reduced form” educational production function

$$(17) \quad Q = \alpha + \beta'X + \delta E + \varepsilon$$

where  $Q$  is the single educational output,  $X$  is a vector of student and/or household demographic characteristics, and  $E$  is expenditures on schools at the appropriate level (school, school district, etc.) of aggregation. Estimates of  $\delta$  are often statistically insignificant and are partially responsible for the resource controversy. The theoretical model above, however, suggests that output and expenditures are simultaneously determined by the household demographic characteristics,  $X$ , such that

$$(18) \quad E = a + b'X + d'P + u$$

where  $P$  is a vector of input prices, for example.

It is tempting, but incorrect, to merely regress  $Q$  on  $X$  and  $P$  in this case, especially if the estimation of equation (18) finds that  $\delta$  is insignificant, as might be expected in this theoretical context. The proper empirical model should recognize the simultaneity of output and expenditures and employ an Instrumental Variables approach or other suitable technique (Greene 2003). While this point is often ignored in the educational production function literature, it has been recognized in the school choice literature (Houston & Toma 2003). There is no reason to persist in such a specification error and to risk perpetuation of the resource controversy on an issue so easily remedied.

Now consider the multiple educational outputs measurement issue. Although production function estimation as in Equation (18) cannot proceed with multiple outputs, both stochastic cost frontier techniques and Data Envelopment Analysis (DEA) can model and measure the effects of multiple outputs. These capture, directly or indirectly, the technology sets from which

observed multiple output and multiple input combinations are drawn. There is no technical impediment to correcting the measurement error induced by failure to account for multiple outputs. In fact, recent studies have begun to allow for multiple outputs and endogeneity of demand and cost factors (Dodson III & Garrett, 2004), although missing output and input measures remain problematic.

An alternative to a complete neoclassical production function approach to education is to search for the institutions that are effective at producing an observable subset of outputs (say, test scores and graduation rates), while accounting for student characteristics and input use (expenditures), including any simultaneity among them. Such studies are feasible, as multiple output data exist, multiple-output techniques are available, and methods for simultaneous systems are well known. The results can identify the effective institutions and their characteristics, but no relative efficiency or productivity claims can be inferred. Since all outputs are not measured, one cannot distinguish the truly inefficient institutions from those that are efficient, but choose to expend resources on unmeasured outputs. Indeed, this is the safest interpretation of any education production study, as all possible outputs are unlikely to be measured accurately.

## **5. Conclusion**

In the U.S., policies to monitor the effectiveness of schools and promote improvement have become nearly a national obsession. President Bush has vowed to make the data generated as a result of the No Child Left Behind legislation available on the Internet. The inevitable flood of research this will enable should proceed in as productive a fashion as possible. The methods for that research have been identified here, as have the limits on what we can expect to learn.

It should be obvious that the efficiency and economic performance of schools cannot be accurately assessed unless all the relevant outputs and inputs are captured. Nevertheless, past attempts to approach education as a production process have ignored key relationships among the relevant variables, as well as technical aspects of cost and production theory. The resulting specification and measurement errors lead to biased and inconsistent parameter estimates, calling into question the accuracy and reliability of conclusions based on these results. These errors derive from inadequate data and/or from a lack of a guiding theory of the way schools operate.

A simple political-economic theory for public schools was formulated here. This theory suggests that demographic, or “education demand” characteristics, simultaneously determine multiple public school outputs and expenditures. Consequently, accurate identification of the production set available to educational institutions requires measurement of *all* the outputs and inputs, as well as accounting for the simultaneity among the inputs, outputs, and demographic characteristics. Efficiency measures based on partial observation of multiple outputs may possess counterintuitive properties that lead to incorrect policy conclusions.

Fortunately, techniques exist to capture these effects. Stochastic cost frontier and DEA methods can handle multiple outputs that simple production function models cannot. Instrumental variables techniques can be utilized to account for simultaneity.

Unfortunately, even state-of-the-art techniques can falter before inadequate data. Measurement problems in education are legion. Complete and accurate measurement of all outputs is unlikely, especially on a large scale, and data on fixed inputs, such as capital, or their costs, remain scarce. Reliable measures of ability may never be identified. Some measurement techniques, such as the value-added approach to outputs or the variable cost approach to inputs, may mitigate some of these problems, but will not eliminate them.

A feasible procedure may involve applying the appropriate techniques to imperfect data, while recognizing those imperfections when interpreting the results. This would allow identification of those schools/districts that were effective at producing some subset of outputs, given certain student characteristics and expenditures. Relative efficiency or productivity in the economic sense could not be observed due to missing output data and measurement errors.

---

## Notes

<sup>1</sup> Years ago Hanushek (1979, pp. 361-2) noted that, "...with information about only one output, estimation of the reduced form might be quite misleading. The estimated effects of the various inputs will reflect both the production technology (the effect of each input on the single output) and the choice between outputs, not simply the production technology."

<sup>2</sup> Klausnitzer (2004) contrasts two such school systems and their high schools.

<sup>3</sup> Hanushek (1986) traces the suggestion to the Coleman Report (Coleman 1966), but also see Seigfried & Fels (1979).

<sup>4</sup> Although Dustmann, Rajah, & van Soest (2003) in the same volume use a multiple equations approach to find that reductions in class size encourage students to continue their education at age 16 and also increase the wages later in life of students who stay on in school after 16.

<sup>5</sup> A third strand examines individual student level data for evidence that school quality affects earnings. A positive relationship between school resources and future earnings generally obtains, although anomalies – such as positive effects on earnings for college attendees, but not for high school graduates – are found in some studies. See Card & Krueger (1996) and Dustmann et al. (2003).

<sup>6</sup> Dodson III & Garrett (2004) widen the set of outputs to include test scores and graduation rates in the set of outputs and account for endogeneity. As few outputs are measured, however, when some institutions appear to produce higher test scores with fewer resources, one cannot tell if the resource savings are due to true efficiency or to a shifting of resources from unobserved outputs to observed outputs. Moreover, if school districts respond to variations in household demand for educational outputs and spending by choosing different output and spending

---

combinations, then even less can be said about social welfare performance when all outputs are not measured. In this context, as is shown below, the restricted range of output measures precludes economic efficiency conclusions.

<sup>7</sup> An alternative, as in Peltzman (1993, 1996), is to consider a political utility function that is a weighted sum of the utilities of different interest groups. As the influence of competing interest groups is not the focus here, this specification is not pursued.

<sup>8</sup> Iatarola & Stiefel (2003) provide evidence of differential allocation of resources to individual schools within a school district.

<sup>9</sup> Wenger (2000) finds empirical support for the proposition that test scores and graduation rates are substitutes in educational production.

<sup>10</sup> The elasticities are equal if  $V$  is separable in the  $Q$ s and  $X$ , such that  $V(Q_1, Q_2, X) \equiv U(Q_1, Q_2)Z(X)$ . In this case a change in  $X$  causes no change in the first order conditions [ $\partial(V_1/V_2)/\partial X = 0$ ]. That is, changes in school district demographics cause no change in the district's choice of outputs.

<sup>11</sup> The reader may confirm that the result in Figure 1 is exactly the same when changes in  $X$  favor  $Q_2$  instead of  $Q_1$ :  $\epsilon_{1X} < \epsilon_{2X}$ ,  $dQ_1/dX < 0$ ,  $dQ_2/dX > 0$ , for  $X_1 > X_2$ .

<sup>12</sup> One may consider  $\max Q_1$  as the maximum observed or best practice performance with no change to the analysis or the results.

<sup>13</sup> In fact, Pritchett & Filmer (1999) derive superficially similar results, but, since they focus on the empirical finding that "inputs do not matter" in producing educational outcomes, they miss the implications for output measurements as opposed to input measurement.

<sup>14</sup> Although the elimination of diseconomies of scope may seem reasonable on its face, this need not be so. Akerlof & Kranton (2002), for example, suggest that racial integration was detrimental to student performance by, in part, increasing the number of disenchanted students – those who felt as if they were excluded from the school culture. Over time, schools were able to partially reverse the slide in student performance by expending resources to make more students feel included in the school. Thus, increasing “diversity” in the form of racial integration may have increased the cost of student “performance” as measured by standardized tests.

<sup>15</sup> This occurs because the two terms inside the brackets that are multiplied by  $B_X$  in (8) and (9) will have the same sign.

<sup>16</sup> In this case, the two terms inside the brackets that are multiplied by  $B_X$  in (8) and (9) will have opposite signs. Thus, the effect of  $B_X > 0$  is reduced from the “no diseconomies of scale” case, but it seems unlikely in practice that this effect could be large enough to reverse the positive sign of  $dQ_1/dX$  in (10).

<sup>17</sup>

$$\begin{aligned}
 dI/dX &= d(Q_1/\max Q_1)/dX = (d(Q_1/\max Q_1)/dB)B_X \\
 &= [ \{1/(\max Q_1)\}(dQ_1/dB) - \{Q_1/(\max Q_1)^2\}d(\max Q_1)/dB ] B_X \\
 &= [(dQ_1/dB)B_X - \{Q_1/(\max Q_1)\} \{d(\max Q_1)/dB\} B_X] \{1/(\max Q_1)\} \\
 &= [ \{(dQ_1/dX)/Q_1\} - \{d(\max Q_1)/dX\}/(\max Q_1) ] \{Q_1/(\max Q_1)\} \\
 &= [ \{(dQ_1/dX)(X/Q_1)\} - \{d(\max Q_1)/dX\}(X/(\max Q_1)) ] [ \{Q_1/(\max Q_1)\}/X ] = (e_{QX} - e_{MX})(I/X)
 \end{aligned}$$

<sup>18</sup> The scale properties of a cost function are related to the degree of homogeneity of the cost function at any point. If costs are linearly homogeneous,  $C(kQ_1,0) = kC(Q_1,0)$  where  $k$  is a positive constant, there are constant returns to scale. For scale economies,  $C(kQ_1,0) < kC(Q_1,0)$ , and for diseconomies,  $C(kQ_1,0) > kC(Q_1,0)$ .

---

<sup>19</sup> See Coelli, Rao, & Battese (1998) for a summary of efficiency or inefficiency measures in different cost and production contexts.

## **Acknowledgements**

For helpful discussions and comments on earlier versions of the paper, the author thanks Reuben Kyle, Tony Eff, Richard Hannah, and Charles Baum; seminar participants at Middle Tennessee State University; and session participants at the 2002 Southern Economic Association meeting. The author retains responsibility for any remaining errors or omissions.

## **References**

- Abbott. M. & Doucouliagos, C..(2003). The efficiency of australian universities: a data envelopment analysis. *Economics of Education Review* 22, 89-97.
- Akerlof, G. A., & Kranton, R. E. (2002). Identity and schooling: some lessons for the economics of education. *Journal of Economic Literature* 40 (4), 1176-1201.
- Becker, G. (1983). A theory of competition among pressure groups for political influence. *Quarterly Journal of Economics* 98, 371-400.
- Becker, W.E. (1997). Teaching economics to undergraduates. *Journal of Economic Literature* 35 (3), 1347-73.
- Bifulco, R., & Bretschneider, S. (2003). Response to comment on estimating school efficiency. *Economics of Education Review* 22, 635-38.
- Bishop, John H. and Ludger Wößmann (2004). Institutional effects in a simple model of education production, *Education Economics* 12 (1), pp. 17-38.
- Bonesronning, H. (2003). Class size effects on student achievement in Norway: patterns and explanations. *Southern Economic Journal* 69 (4), 952-965.

- Card, D.A., & Krueger, A. B. (1996). School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina. *Journal of Economic Perspectives* 10 (4), 31-50.
- Coelli, T., Rao, D.S.P., & Battese, G.E. (1998). *An introduction to efficiency and productivity analysis*. Boston: Kluwer Academic Publishers.
- Chakraborty, K., Biswas, B., & Lewis, W.C. (2001). Measurement of technical efficiency in public education: a stochastic and nonstochastic production function approach. *Southern Economic Journal* 67 (4), 889-905.
- Chambers, R.E. (1988). *Applied production analysis*. Cambridge: Cambridge University Press.
- Coleman, J.S. (1966). *Equality of educational opportunity*. Washington, DC: U. S. GPO.
- Dee, T.S., Evans, W.N. & Murray, S.E. (1999). Data watch: research data in the economics of education. *Journal of Economic Perspectives* 13 (3), 205-216.
- Dodson III, M.E., & Garrett, T.A. (2004). Inefficient education spending in public school districts: a case for consolidation. *Contemporary Economic Policy* 22 (2), 270-280.
- Dolton, P., Marcenaro, O. D., & Navarro, L. (2003). The effective use of student time: a stochastic frontier production function case study. *Economics of Education Review* 22, 547-560.
- Downes, T., & Pogue, T. (1994). Adjusting school aid formulas for the higher cost of educating disadvantaged students. *National Tax Journal* 47 (1), 89-110.
- Dustmann, C., Rajah, N., & van Soest, A. (2003). Class size, education, and wages. *The Economic Journal* 113, F99-F120.
- Epple, D., & Romano, R.E. (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review* 88 (1), 33-62.

- Ferris, J.S., & West, E.G. (2002). Education vouchers, the peer group problem, and the question of dropouts. *Southern Economic Journal* 68 (4), 774-93.
- Greene, W.H. (2003). *Econometric analysis*. 5<sup>th</sup> ed. Upper Saddle River: Prentice-Hall.
- Grosskopf, S., Hayes, J.K., Taylor, L.L. & Weber, W. L. (2001). On the determinants of school district efficiency: competition and monitoring. *Journal of Urban Economics* 49, 453-478.
- Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources* 14 (3), 351-88.
- Hanushek, E.A. (1986). The economics of schooling. *Journal of Economic Literature* 24 (3), 1141-77.
- Hanushek, E.A. (1996). Measuring investment in education. *Journal of Economic Perspectives* 10 (4), 9-30.
- Hanushek, E.A. (2003). The failure of input-based schooling policies. *The Economic Journal* 113, F64-F98.
- Houston, Jr., R.G., & Toma, E.F. (2003). Home schooling: an alternative school choice. *Southern Economic Journal* 69 (4), 920-935.
- Hoxby, C.M. (1996). Are efficiency and equity in school finance substitutes or complements? *Journal of Economic Perspectives* 10 (4), 51-72.
- Hoxby, C.M. (2000). Does competition among public schools benefit students and taxpayers? *American Economic Review* 90 (5), 1209-38.
- Iatarola, P., & Stiefel, L. (2003). Intradistrict equity of public education resources and performance. *Economics of Education Review* 22, 69-78.

- Kane, T.J., & Staiger, D.O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives* 16 (4), 91-114.
- Klausnitzer, Dorren (2004). Rich school, poor school, *The Tennessean*, 100 (130), A1.
- Koretz, D. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *The Journal of Human Resources* 37 (4), 752-777.
- Krueger, A.B. (2003). Economic considerations and class size. *The Economic Journal* 113, F34-F63.
- Lazear, E.P. (2001). Educational production. *The Quarterly Journal of Economics* 116 (3), 777-803.
- Peltzman, S. (1976). Toward a more general theory of regulation. *Journal of Law & Economics* 19 (2), 211-40.
- Peltzman, S. (1993). The political economy of the decline of american public education. *Journal of Law & Economics* 36 (1 Part 2), 331-370.
- Peltzman, S. (1996). Political economy of public education: non-college-bound students. *Journal of Law & Economics* 39 (1), 73-120.
- Pritchett, L., & Filmer, D. (1999). What education production functions *really* show: a positive theory of education expenditures. *Economics of Education Review* 18, 223-39.
- Ram, Rati (2004). School expenditures and student achievement: evidence for the United States, *Education Economics* 12 (2), pp. 169-76.
- Rothschild, M., & White, L.J. (1995). The analytics of pricing in higher education and other services in which customers are inputs. *Journal of Political Economy* 103, 573-86.
- Ruggiero, J. (2003). Comment on estimating school efficiency. *Economics of Education Review* 22, 631-34.

- Ruggiero, J., & Vitaliano, D.F. (1999). Assessing the efficiency of public schools using data envelopment analysis and frontier regression. *Contemporary Economic Policy* 17 (3), 321-31.
- Siegfried, J.J., & Fels, R. (1979). Research on teaching college economics: a survey. *Journal of Economic Literature* 17 (3), 923-69.
- Stigler, G. (1971). The theory of economic regulation. *Bell Journal of Economics and Management Science* 2, 3-21.
- Tiebout, C.M. (1956). "A pure theory of local government expenditures. *Journal of Political Economy* 64 (5), 416-24.
- Todd, P.E., & Wolpin, K.I. (2003). On the specification and estimation of the production function for cognitive achievement," *The Economic Journal* 113, F3-F33.
- Wenger, J.W. (2000). What do schools produce? implications of multiple outputs in education. *Contemporary Economic Policy* 18 (1), 27-36.
- Winston, G.C. (1999). Subsidies, hierarchy and peers: the awkward economics of higher education. *Journal of Economic Perspectives* 13 (1), 13-36.

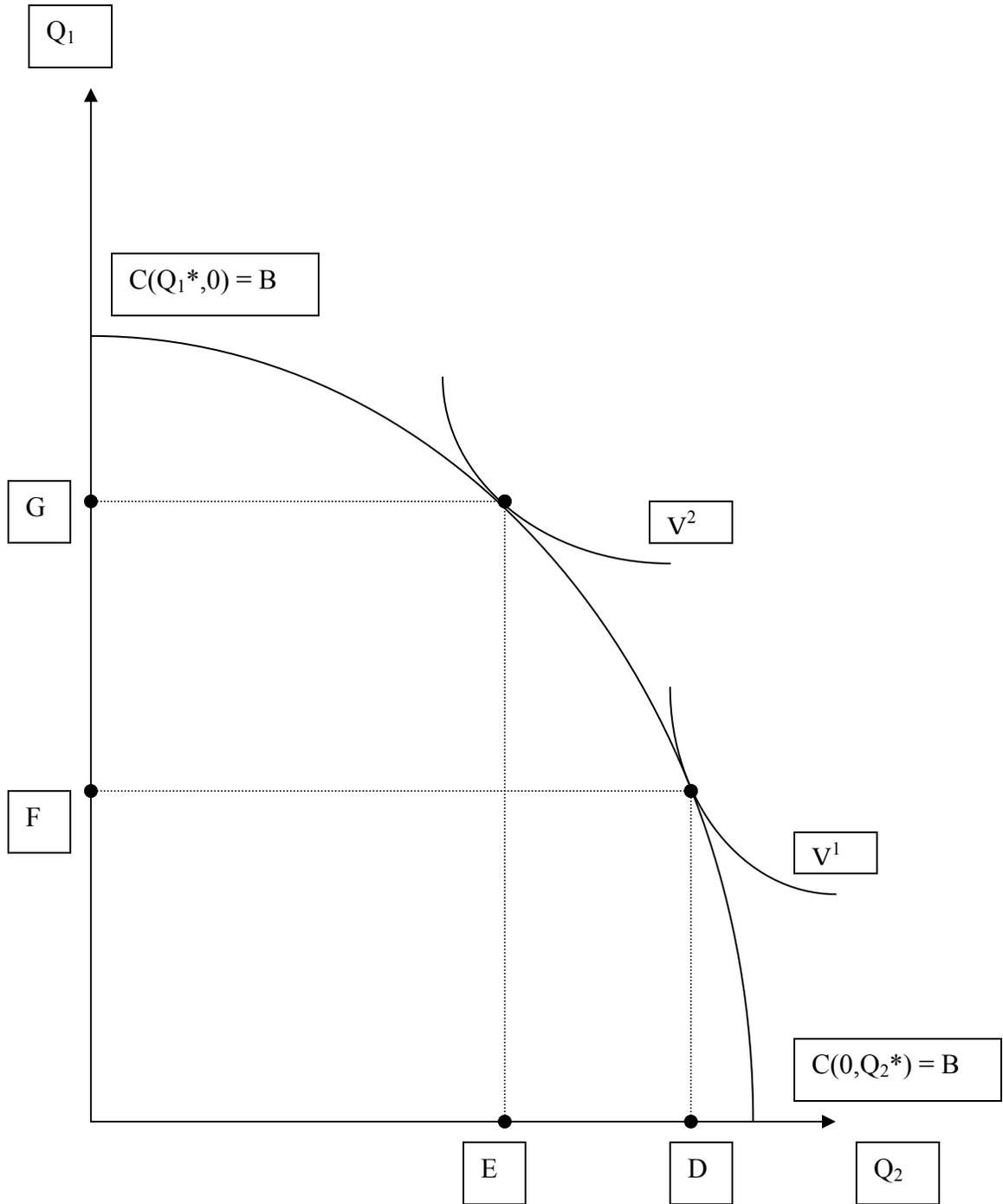
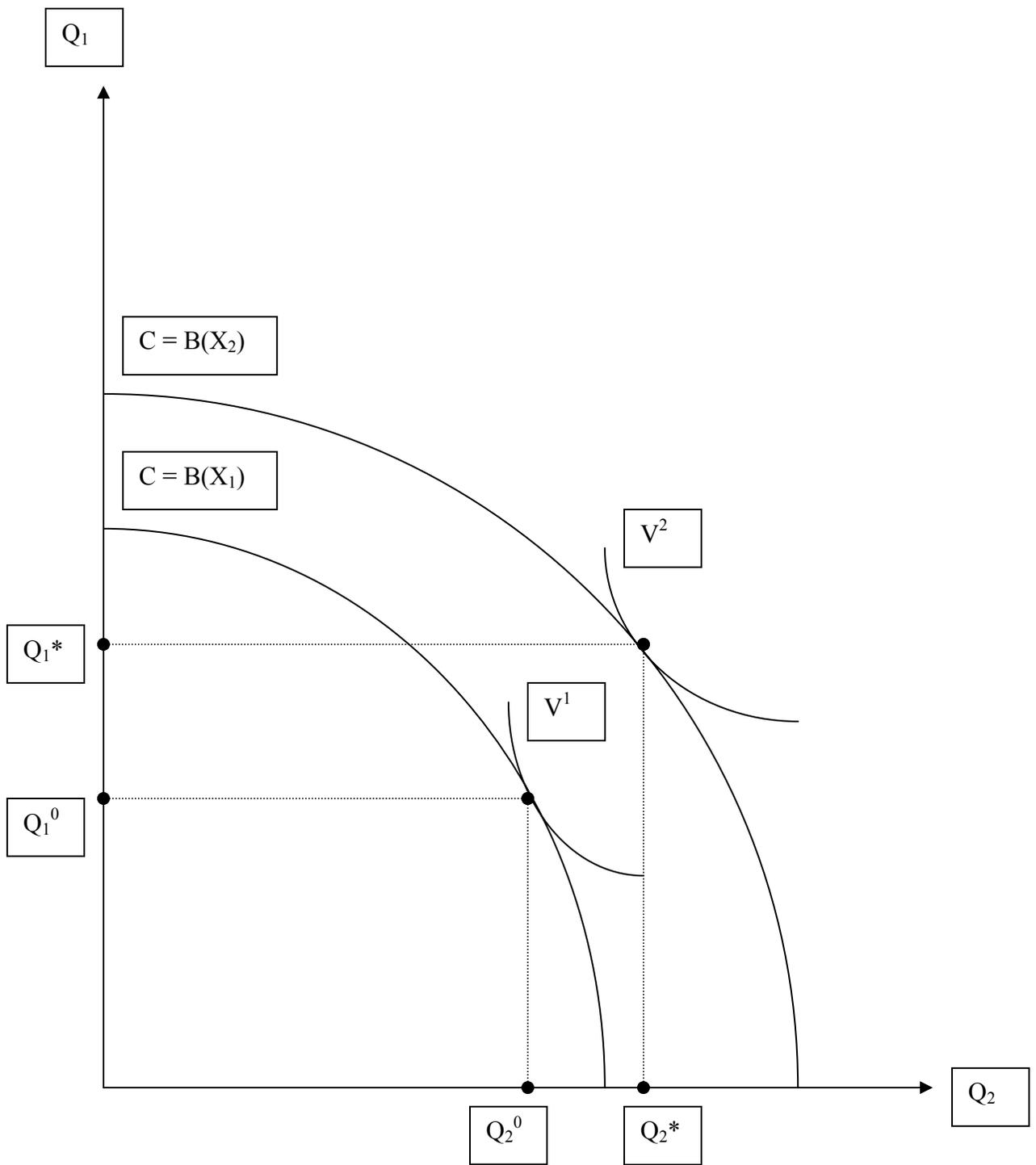


FIGURE 1: Two school districts with identical budgets choose different output combinations.



**FIGURE 2:** Variations in demographics change the school budget and the output mix.

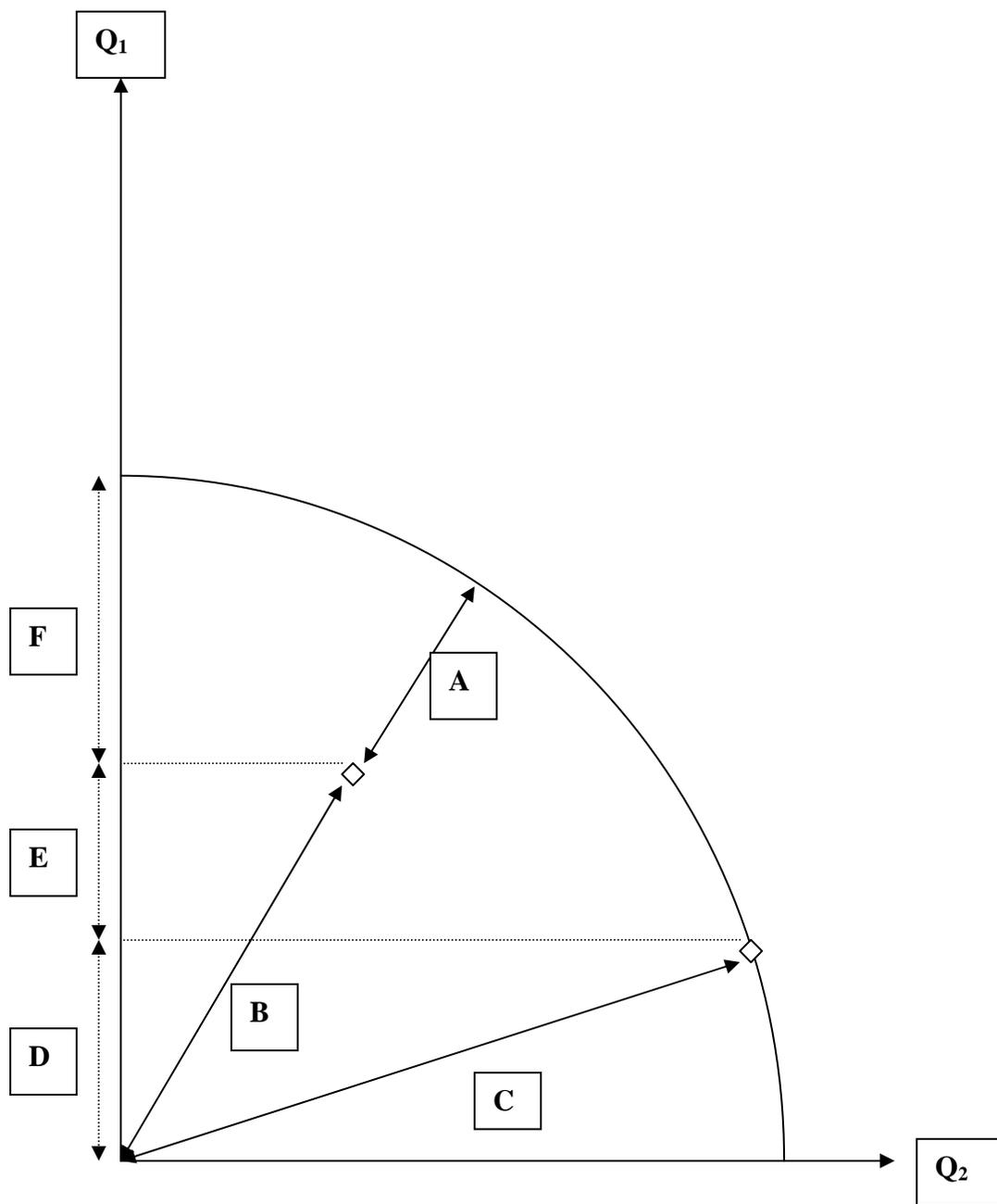
**TABLE 1**

Direction of Change in the Single-Output School Efficiency Index (I)

for a Change in Demographics (X)

By Scale and Demand Characteristics

	<u>Scale Characteristics</u>		
	Diseconomies	Constant Returns	Economies
<u>Demand Response</u>	<u>(<math>e_{MX} &lt; 1</math>)</u>	<u>(<math>e_{MX} = 1</math>)</u>	<u>(<math>e_{MX} &gt; 1</math>)</u>
Elastic ( $e_{QX} > 1$ )	+	+	?
Unitary ( $e_{QX} = 1$ )	+	0	-
Inelastic ( $e_{QX} < 1$ )	?	-	-



**FIGURE 3:** An inefficient school district appears most efficient if only  $Q_1$  is observed.

