

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Price Index Concepts and Measurement

Volume Author/Editor: W. Erwin Diewert, John S. Greenlees and Charles R. Hulten, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-14855-6

Volume URL: <http://www.nber.org/books/diew08-1>

Conference Date: June 28-29, 2004

Publication Date: December 2009

Chapter Title: Price and Real Output Measures for the Education Function of Government: Exploratory Estimates for Primary and Secondary Education

Chapter Author: Barbara M. Fraumeni, Marshall B. Reinsdorf, Brooks B. Robinson, Matthew P. Williams

Chapter URL: <http://www.nber.org/chapters/c5085>

Chapter pages in book: (373 - 403)

Price and Real Output Measures for the Education Function of Government

Exploratory Estimates for Primary and Secondary Education

Barbara M. Fraumeni, Marshall Reinsdorf,
Brooks B. Robinson, and Matthew P. Williams

9.1 Introduction

This chapter presents new measures of real output for government-provided education in the United States. The research refines the measures in our previous experimental work (see Fraumeni et al. 2004) and also takes important steps forward by calculating chain-type Fisher quantity indexes and implicit price deflators.

Measuring the education output of the government is difficult, even though education is a near-market activity. For services, defining nominal output measures can be problematic, and measuring real output is challenging (Griliches 1994). Education is a service with significant nonmarket inputs, notably student and parent time, and the outcome of education depends upon factors outside the control of providers of educational services, such as student ability, family, peer group, and neighborhood factors (Rivkin 2000). Accordingly, isolating the contribution of providers of educational

Barbara M. Fraumeni is associate dean of research, chair of the PhD in Public Policy program, and professor of public policy at the Muskie School of Public Service, University of Southern Maine. She was chief economist of the Bureau of Economic Analysis when this chapter was drafted. Marshall Reinsdorf is a senior research economist at the Bureau of Economic Analysis, U.S. Department of Commerce. Brooks B. Robinson, a supervisory economist, most recently served as chief of the Government Division, Bureau of Economic Analysis, U.S. Department of Commerce. He currently works for the U.S. Department of the Navy. Matthew P. Williams served as economist at the Office of the Chief Economist, Bureau of Economic Analysis, U.S. Department of Commerce when this chapter was drafted.

This paper represents views of the authors and is not an official position of the Bureau of Economic Analysis or the Department of Commerce. All of the authors were employees of the Bureau of Economic Analysis when this chapter was drafted. Marshall B. Reinsdorf is the only author currently at the Bureau of Economic Analysis. We thank Michael Christian and Barbara Silk, both former employees of the Bureau of Economic Analysis, for their assistance on this project.

services is not easy. In addition, not all benefits of education are measurable, because education has broader effects on the welfare of individuals and of society than just raising earnings, for example. As this research continues, the exploratory measures presented may be substantially altered and refined and will be expanded to include other levels and types of education.

The objective of the government is to educate all individuals of school age, including those least and most able. The cost of educating students will vary substantially across students, with the cost particularly high for special education students and those requiring supplemental help beyond that available in a typical classroom. A recent National Education Association (NEA) report indicates that the average U.S. cost per special education student is more than twice the average cost across all students. As well, the report notes that the number of special education students has risen 30 percent over the last ten years.¹ Educating these students is clearly more expensive than educating other types of students. Bringing about marginal improvements in their educational attainment is probably also more expensive than for more able students. Our current experimental output measures do not adjust for student composition, except to reflect the number of students in high school versus lower grades. Accordingly, given the growth in special education students and the associated higher costs, it is not surprising that the price measures presented in this chapter grow at a faster rate than the gross domestic product (GDP) or gross domestic purchases price indexes. In addition, to the extent that our measures do not capture all quality improvements occurring over time, quantity changes may be underestimated and price changes may be overestimated.

In 2001, the education function of government accounted for approximately 5 percent of nominal GDP, as measured by final expenditures, ranking it with health and income security as among the three largest government function categories.² The Bureau of Economic Analysis (BEA) began publishing functional tables with quantity and/or price indexes for government in 2004. However, these output quantity and price indexes are estimated with a cost-of-inputs-based approach, as is currently performed for total government: federal, state, and local.³ Such input-based approaches do not recognize changes in output resulting from intangible inputs or from varying relationships between inputs and outputs, such as those arising from qualitative changes, and they do not allow for a meaningful estimate of productivity change.

1. National Education Association (2004).

2. See tables 1.1.5, 3.15.5, and 3.16 of the Bureau of Economic Analysis (BEA) National Income and Product Accounts (NIPAs). As government-by-function tables (3.15 and 3.16) appear later than other NIPAs tables (last published in the October 2002 *Survey of Current Business* [BEA 2002a, 12–13]), the data cited in this chapter do not reflect results of the NIPAs comprehensive revision published in December 2003. See BEA (2002a, 2002b, 2003).

3. The 2003 comprehensive revision new NIPAs table family 3.10 presents an alternative breakout of consumption expenditures.

Output-based measures of government output are preferred to input-based measures of output but are difficult to develop and implement. In recent years, national income accountants in other countries have looked to volume indicators using an output approach to improve measures of government education output (Powell and Pritchard 2002; Konijn and Kleima 2000a, Australian Bureau of Statistics 2000b).⁴ The emphasis in all cases has been on real output—or volume—measures, rather than on price measures. These volume indicators, such as those based on number of pupils or hours spent in school, may or may not be quality adjusted. Others have suggested an outcome-based approach directly to adjust for quality change, such as those based on test scores or incremental earnings (O'Mahony and Stevens 2003, 2004; Jorgenson and Fraumeni 1992).⁵ A third approach is a housing value approach, such as those that look at differential prices paid for houses near borders of school districts with differential performance ratings (Black 1998).

Volume indicators using an output approach are commonly not really independent of input measures. For example, teacher experience and pupil-teacher quality adjustments both depend upon an input measure. Although measures are becoming less reliant on input measures, as real output is not set equal to real input, education output measures still frequently rely on input measures.

This exploratory chapter begins by presenting a simple education production function and by discussing the issue of outputs versus outcomes. It next summarizes and analyzes the progress made by other countries to measure the education output of government to set the stage for a description of the U.S. initial efforts.⁶ It then focuses on a few possible quality-adjusted volume indicators for the United States for primary and secondary public education.⁷ Subsequent research at the BEA will continue this line of investigation and will look at quality-adjusted volume indicators for public higher education, libraries, and other education, and at the other output-based approaches for all subcategories of the education function of government. The sample of possible quality-adjusted volume indicator alternatives to the BEA's current methodology is presented within the context of the literature, and empirical estimates are developed.

4. Following the international *System of National Accounts* (Commission of the European Communities et al. 1993), most countries use the term volume to refer to what U.S. economists typically call quantity. In this chapter, the terms are used interchangeably.

5. Currie and Thomas (1999) show the relationship between test scores and future educational attainment and labor market outcomes.

6. Recent attempts to measure the education output of government in the national accounts from the output side (as opposed to the input side) began outside the United States.

7. In this chapter, primary education refers to kindergarten through eighth-grade education, and secondary education refers to ninth- through twelfth-grade education.

9.2 Prior Research on Output-Based Measures of Education Services of Government

As part of a general movement in the international statistical community toward using an output-based approach to measuring government output in their national accounts, a number of countries have implemented or experimented with the output-based measures of real government educational services.⁸ The Office of National Statistics (ONS) in the United Kingdom has gone the furthest with this approach, with nearly 70 percent of government expenditure being measured using direct volume measures.⁹ New Zealand measures over 60 percent of government expenditure in a similar fashion. Australia, the Netherlands, and Italy have also followed suit, each measuring up to 50 percent of government expenditures using direct volume measures. Other countries, such as Canada, Italy, Germany, Norway, Finland, Sweden, and Israel, have also developed real output measures, either recently implementing them for a small portion of government expenditures or currently considering doing so in the near future. Education and health are the two functions of government most commonly measured with an output-based approach.

A topic of debate in these efforts is the extent to which output measures can be based on outcomes. Sherwood (1994) gives the example of a teacher who faces a class of poor students. If the students learn nothing, is the output of the school for that class zero? Sherwood suggests that whether a service output, such as education, should be quality adjusted with an outcome measure, as opposed to being derived from a pure transactions count approach, depends upon the particular service. He points out that the price that individuals pay for a market service, such as a football game, in part depends upon the expected outcome of the game.

A few countries have experimented with output measures that use data on outcomes to quality adjust a quantity index of student years of education. Test scores are one such measure, and the Atkinson Report suggests using real earnings growth in an experimental measure of education output.¹⁰

8. See in particular Jenkinson (2003) and Pritchard (2002b). See also other documents, such as those authored by the Australian Bureau of Statistics (ABS; 1998, 1999, 2000a, 2000b, 2001b, 2002a, and 2003), by the Office for National Statistics (ONS; 1997, 2000, and 2005)—including others by A. Pritchard of the ONS (Pritchard and Powell 2001; Powell and Pritchard 2002; Pritchard 2002a, 2003) and by Caplan, formerly of the ONS (Caplan 1998; Neuburger and Caplan 1998)—and by Algera (1998) of Statistics Netherlands (CBS), by Konijn, formerly of the CBS, and by Kleima of the CBS (Konijn and Kleima 2000a, 2000b).

9. This is according to an e-mail on November 25, 2003, from A. Pritchard of the ONS, unconfirmed by the ABS, who states that implementing current ABS research on direct volume measures for justice (police, courts, and prisons), taxation, and social security would bring coverage of government output (using real measures) to 90 percent, making the ABS the world leader.

10. See Atkinson Commission (2005) and ONS (2005). Note that this approach requires either a strong assumption that the change in the price of education in the labor market can

Refinements to the use outcomes data that remove the influence of non-school factors from the output measure have not yet been developed. Next, we review prior results on measuring the output for the education function of government employed by statistical agencies of the United Kingdom, Australia, the Netherlands, and other countries.

9.2.1 United Kingdom

In the United Kingdom, the ONS produces both an official and an experimental quality-adjusted volume measure of the education function of government. Both use full-time equivalent number of pupils as the volume indicator, under the assumption that hours in school per pupil are constant across time, although it is recognized that pupil hours would be preferred. Both exclude higher education.

The official volume indicator is quality adjusted for all education categories. A 0.25 percent quality-adjustment factor per year is utilized, because “there is significant evidence that educational standards have been rising over a number of years,” and “there is evidence that the quality of teaching is rising.”¹¹ This is justified by General Certificate of Secondary Education (GCSE) examination results, which show a pattern of increases in the average point scores of pupils over a period of eleven years. An index of the number of pupils enrolled in nursery schools, primary schools, secondary schools, further education, and special education is constructed, with weights proportional to the expenditure on education in the base period to form the official volume indicator.

Pritchard (2002a) introduced the idea of using a lesson quality adjustment. In the United Kingdom, government inspectors make assessments regarding the quality of lessons. Powell and Pritchard note that weights could be assigned to the three ratings categories for lessons: good/very good lessons, satisfactory lessons, and unsatisfactory/poor lessons. If these assessments were used to form a lesson quality adjustment, the rate of growth of the volume indicator would be raised over the period 1995 to 2000. However, Powell and Pritchard say that they would prefer a “more coherent basis for estimates,”¹² so this adjustment is not used as part of the official measure, although it is part of the unofficial measure.¹³

be measured by a general deflator or the availability of a customized deflator to separate out the price and volume components of the increment to earnings resulting from additional education.

11. Caplan (1998, 48). Caplan also indicates that there may be a declining proportion of students completing higher education courses; therefore, an upward quality adjustment for these students may not be justified. Ronald Ehrenberg of Cornell University indicated in a recent discussion that in his opinion, an upward quality adjustment may not be justified for U.S. higher education.

12. See Powell and Pritchard (2002, 8).

13. See Powell and Pritchard (2002) for a description of the official measure and Pritchard (2002a, 30), for the unofficial measure estimates.

The official and experimental ONS estimates show how sensitive results can be to methodology. From 1995 to 2000, the annual rate of growth of the volume indicator for the experimental estimates with the “quality of lessons received” adjustment is 1.74 percent.¹⁴ The comparable figure for the official index with the 0.25 percent quality adjustment is 0.88 percent.¹⁵

From 1995 to 2000, the annual rate of growth of the implicit price deflator for the official estimate is 3.76 percent. This reflects a typical pattern seen for the few countries, including the United States (see table 9.3 and figure 9.2), for which a price can be calculated based upon published information. In all these countries, the rates of growth in the prices account for at least two-thirds of the rates of growth of nominal expenditures, with the United Kingdom being on the high side for that time period at approximately 80 percent.¹⁶ In contrast, the rates of growth of the GDP price may account for less than half of the rates of growth of nominal GDP for the countries for which education prices can be calculated: Australia, the United Kingdom, the Netherlands, and the United States.¹⁷ However, to the extent that quality improvements occurring over time are not captured in the quality adjustments made to the education volume indicators, the price growth rates are overestimated, and the volume indicator growth rates are underestimated. Measuring the output of services is difficult, and measuring quality changes in the output of services is even more difficult. Accordingly, it is reasonable to assume that quality is imperfectly estimated in the education volume indicators of all countries.

9.2.2 Australia

The Australian Bureau of Statistics (ABS) examined a variety of approaches when researching possible volume indicators for education.¹⁸ These approaches included a volume indicator with and without quality adjustment and a modified incremental earnings approach. The quality adjustments considered include quality indicators such as class size, examination results, the quantity and quality of research publications, grants received, and the number of student research completions. In the official

14. See table 10 in A. Pritchard (2002a, 30).

15. See annex B in A. Pritchard (2002b, 11).

16. Rough estimates of a volume indicator for the United States covering higher education as well as primary and secondary education were calculated using quality-unadjusted enrollment for higher education to compare with the Australian, Netherlands, and U.K. estimates. From 1990 to 2001, the rate of growth of prices is about three-quarters the rate of growth of nominal expenditures.

17. The comparison is made here to GDP prices rather than to gross domestic purchases prices, as the former are available for all of the countries. Frequently, the term gross domestic final expenditures prices is used by other countries in place of the term gross domestic purchases prices used by the BEA.

18. The ABS does not distinguish between public and private education in its estimates; therefore, it is not possible to separate the government education function from the private education function.

index, class size was not adopted as a quality adjuster because of uncertainty about the relationship between class size and the quality of education received. Examination results are not adopted as a quality adjuster because of concern about the comparability of scores over time, particularly because of external factors, such as social capital, which can affect these scores.¹⁹ The modified incremental earnings approach, if ever adopted in the national accounts, would indirectly infer the “direction and size of long-term quality change” from a human capital model similar to that developed by Jorgenson and Fraumeni (1992), a model that has been implemented in a satellite account for Australia by the ABS.²⁰

The official volume indicator now used by the ABS does not quality adjust the output indicator. For primary and secondary education, the output index is an index of enrollment numbers converted to an equivalent full-time student unit (EFTSU) basis. For vocational education, module hours are used. For higher education, enrollments are converted to an EFTSU basis and weighted by the Higher Education Contribution Scheme (HECS) charges levied on students to form the final higher education index. A university research index depends upon the number of publications and student research completions. Other education services, such as preschool education, are still measured using input price indexes. All of the individual indexes are weighted together using cost shares.²¹

The new education output method results in growth rates that are higher and more stable than the previous input method. For years ending June 30, the average annual rate of growth of gross value added from 1993 to 1994 through 1999 to 2000 under the new method is 1.9 percent, compared to 1.5 percent under the previous method.²²

For years ending June 30, the 1994 to 2003 annual rate of growth of the implicit price deflator is 3.3 percent. This accounts for two-thirds of the rate of growth of nominal expenditures.

9.2.3 The Netherlands

Statistics Netherlands (CBS)²³ experimented with five possible volume indicators to replace their current input-based output index (Statistics Netherlands National Accounts Branch 2003; Konijn and Kleima 2000a). Edu-

19. Examination results were considered according to ABS (1998, 4) but were not used according to ABS (2002b, 3).

20. ABS (1998, 4), ABS (2002b, 3), and ABS (2001a).

21. ABS (2001b, 4–5) and ABS (1999, 13), the latter for the definition of a equivalent full-time student unit.

22. ABS (2002b, 3–4). David Bain of ABS provided the authors on April 30, 2004, with a worksheet containing education nominal and chain volume measures through the year ending June 30, 2003. (All ABS annual economic statistics are calculated from July 1 through June 30.) Growth rates for the nominal and implicit price deflator were calculated by the authors.

23. Both Statistics Netherlands and the Central Bureau of Statistics of Israel are referred to as CBS.

cation is divided into ten levels, from primary through university, and each level's output is measured using an appropriate index. Two indexes depend only on number of pupils. One is an unweighted index; the other weights number of pupils by expenditures per type. Three combination indexes use number of pupils for primary education along with number of pupils, number of pupils moving up, and/or number of graduates for other levels or types of education. In some cases, a two-year moving average of number of graduates is used to smooth the series. For several categories of secondary education and above, part-time students are counted as 0.5 in the pupil count.

Other quality adjustments are considered. For primary education, these include the composition of the pupil stock, the percentage of pupils that move up each year, and the scores of the level test, which is the National Institute for Educational Measurement (CITO) test. A quality adjustment for the composition of the pupil stock incorporates information on the share of students who have a lower level of education for two subcategories: those whose parents are Dutch and those whose parents are foreigners. This quality adjustment was not included in an estimate because of the uncertainty in the resource cost factors to serve these different students. A quality adjustment for pupils moving up was not incorporated into the estimate, because these adjustments would be almost constant over the years. The CITO test results changed little during 1995 to 2000, so they are not employed as a quality adjustment, but new tests may be a fruitful source of quality adjustment in later versions of a volume indicator. Pupils moving up was not used as an indicator for any of the university education volume indicators, because the financial grant period was shortened during this time period; accordingly, the study duration decreased. Pupils moving up was not used as an indicator for vocational colleges, because data were not available on this for the whole period 1990 to 1998.

The conclusion of Konijn and Kleima is that the best volume indicators are those that go beyond just tracking the number of pupils. Of the three combination indexes, the index that uses: the number of pupils for primary, secondary, and vocational education; a two-period moving average of the number of graduates of vocational colleges and universities; and pupils moving up as the quality adjuster for all other categories of education is their first choice.²⁴ Cost shares are used as weights. A volume indicator very similar to this index is now used in their national accounts.²⁵

Konijn and Kleima estimate volume indicators with the current input method and five alternative indexes; however, they indicate that current

24. Konijn and Kleima (2000a, 19, 25). The two-period moving average is used to mitigate the effect of a small absolute change looking large in relative terms, compared to a small population of graduates.

25. E-mail from Kleima on November 14, 2003.

input method estimates for 1996 to 1998 may not be reliable. In 1995, the volume of labor input to education was adjusted upward by 15 percent. Unrevised estimates for 1990 to 1995 show labor input to education almost constant.²⁶ The 1991 to 1995 annual growth rates of the volume indicators vary from 0.34 percent for the two pupil numbers indexes to 1.42 percent for the preferred combination index. The 1991 to 1995 annual growth rate of the current input method is 0.86 percent. The 1991 to 1998 annual growth rates of the volume indicators vary from 0.23 percent for the weighted pupil numbers index to 1.25 percent for the preferred combination index.²⁷ From 1991 to 1997, the annual rate of growth of the implicit price deflator is 2.41 percent, which is two-thirds the rate of growth of nominal expenditures.

Figure 9.1 shows the implicit price deflator for all levels of the government function of education for Australia, the Netherlands, and the United Kingdom. As the implicit price deflators for these countries are not available separately for the category primary and secondary education, the U.S. implicit price deflator is not shown. As there are at most only five overlap years, always including the base year 1996, it is difficult to make comparisons.²⁸

9.2.4 Other Countries

While a half dozen or so other countries have developed real output measures for a portion of government expenditures, in many cases, it is difficult to get a clear indication of the approaches used, much less a full description of the methodologies. Much of the work is still in developmental stages, and published explanations are hard to come by. Nevertheless, other approaches to measuring the output of the education function of government by other nations contribute valuable insight.

A number of countries use the number of pupils as a volume indicator, with quality-adjustment factors under consideration or already adopted. Among the countries that have adopted this approach are: Statistics Canada, the Federal Statistical Office of Germany, the National Institute of Statistics of Italy, and the Central Bureau of Statistics of Israel (CBS). Canada and Germany use or plan to use the number of pupils without a quality adjustment.²⁹ Italy uses the number of pupils for education-based functions, with some qualitative adjustments related to class size as captured by a congestion measure. In some cases, weights used in output aggregation also adjust for the use of equipment and teaching aids. For service-based functions in education, Italy uses the number of employees, the number of

26. Konijn and Kleima (2000a, 23).

27. Calculations based on estimates in Konijn and Kleima (2000a, 22).

28. The base year is set to 1996 in this chapter, because the U.S. data used in this chapter is the pre-NIPAs comprehensive revision data that has 1996 as its base year.

29. Jenkinson (2003, 4).

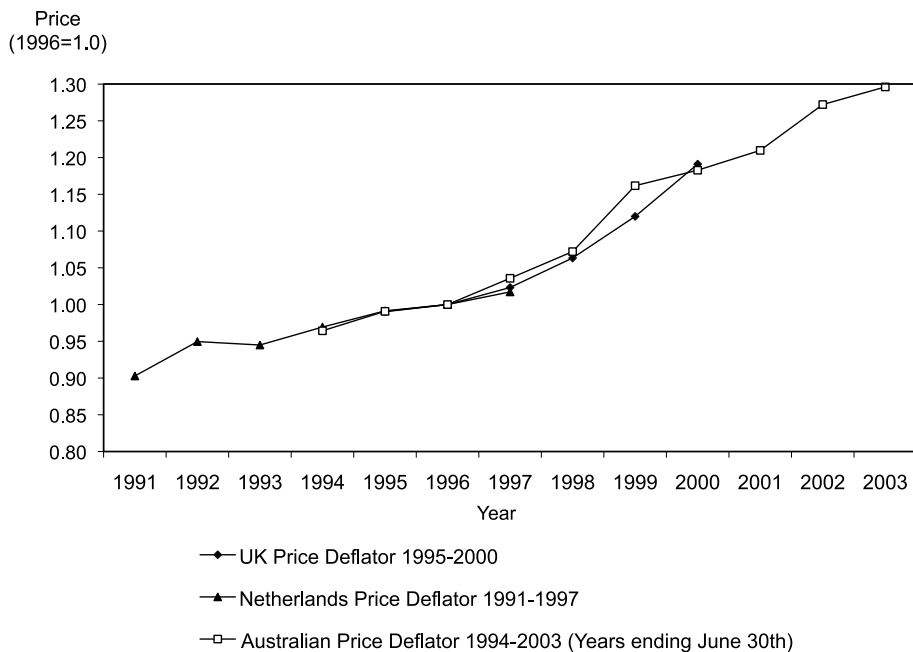


Fig. 9.1 The United Kingdom, the Netherlands, and Australia: All levels of education price deflator

users of the services, or the number of services provided.³⁰ Israel may be quality adjusting the number of pupils with a variety of indicators of quality change for higher education: the percentage of students succeeding in their studies each year; the number of students receiving diplomas or academic degrees; the percentage of students studying toward first, second, and third degrees; and the number of students in various study disciplines.³¹

Statistics Finland uses a variety of volume indicators. Teaching hours are the volume indicator for 99 percent of educational services produced by municipalities, which include services provided by vocational institutes and community colleges, as well as primary and secondary education institutions. The number of degrees completed, generally separated into graduate and postgraduate, measures university education output. Either the number of days of study or courses completed measures adult and continuing education, depending upon the university. The number of publications is used for the output of research, and the number of visitors is used for libraries.³²

30. Malizia (1998, 18–24).

31. Hadar, Madler, and Barzel (1998, 9–13).

32. Niemi (1998).

9.3 Experimental Estimates for the United States

9.3.1 Introduction

In this section, quality-adjusted volume indicators are presented that might serve as a basis for measurement of output of government educational services. Each begins with the number of pupils enrolled as the base index, then considers possible quality adjustments to this base index. The list of possible quality adjusters is not exhaustive, and improvements to these experimental estimates are still underway. Accordingly, these estimates should not be taken as an indication of what measure (if any) may be adopted in the future by the BEA. Estimates are presented for 1980 to 2001 for primary and secondary education.³³ Quality adjustments presented include adjustments by teaching staff composition indexes, pupil-teacher ratios, and high school dropout rates.

9.3.2 Defining a Production Function for Education

A difficult question in the development of an output volume measure for the education function of government is whether outcome ought to be distinguished from output. Outcome generally refers to the level of knowledge or skills possessed by those who have received education. Outcome can be affected by a host of factors other than schools themselves (e.g., ability, parental support, the quality of home life, and social capital in general). On the other hand, output generally refers only to the impact of schools on the level of knowledge and skills of students. For example, test scores or graduation rates are frequently used to quality adjust volume indicators for education, yet these are often affected by factors other than schools. Cipollone and Rosolia (2007), for example, find that conscription prospects and peer group outcomes affect graduation rates. Students' ability and prior preparation also affect current educational outcomes. Finally, families provide inputs into the learning process; so for example, students from families that do not speak English will generally require more educational services to achieve the same outcome as measured by test scores than will native speakers of English.

Some of the services that schools provide are in areas other than education itself, such as athletics and socialization, but for the sake of convenience, we will refer to our quality-adjustment factor for outcomes as learning. Let θ_{it} denote the average learning outcome by a student at education level i (primary, secondary, or higher education), and let q_{it} denote the number of students completing a year of education at level i in year t . Then, volume of learning in year t is:

33. The estimates do not incorporate revised data from the December 2003 NIPAs comprehensive revision.

$$(1) \quad Q_i = \sum_i \theta_{it} q_{it}$$

Changes in θ_{it} result both from changes in the educational services produced by schools using inputs of teachers, other staff, supplies, and capital stock and from changes in nonschool factors. Let teacher inputs be represented by the vector of \mathbf{T}_i , where the elements of \mathbf{T}_i are the numbers of teachers of each experience and education level teaching at school type i . Also, let \mathbf{A}_i represent administrative and support staff, let \mathbf{K}_i represent the capital stock, and let \mathbf{M}_i represent intermediate inputs, such as supplies. Finally, let the factors that influence outcomes but not output be e_i , an index of factors other than teachers or schools that influence student effort levels, and b_i , an index of student background and ability levels. Then, the outcome function at educational level i for learning per pupil is:

$$(2) \quad \theta_i = f_i(\mathbf{T}_i, \mathbf{A}_i, \mathbf{K}_i, \mathbf{M}_i, e_i, b_i, q_i).$$

If θ_i is measured by average test scores, $f_i(\cdot)$ equals the maximum score achievable by the school with inputs $\mathbf{T}_i, \mathbf{A}_i, \mathbf{K}_i, \mathbf{M}_i$, given the external factors e_i and b_i and student body size of q_i .

Over the relevant range for the arguments of equation (2), the average amount of learning by a student at level i is increasing in \mathbf{T}_i, e_i , and b_i . Also, $f_i(\cdot)$ is decreasing in q_i in the region where we expect schools to operate, meaning that the marginal effect of a rise in the student-teacher ratio is to reduce learning per student. It is also increasing in $\mathbf{A}_i, \mathbf{K}_i$, and \mathbf{M}_i in some local region (though in the case of \mathbf{A}_i , it is not always clear that schools are operating in that region). Finally, we assume that $f_i(0, 0, 0, 0, e_i, b_i, q_i) = 0$.

Because $f_i(\mathbf{T}_i, \mathbf{A}_i, \mathbf{K}_i, \mathbf{M}_i, e_i, b_i, q_i) > f_i(\mathbf{T}_i, \mathbf{A}_i, \mathbf{K}_i, \mathbf{M}_i, 0, 0, q_i)$ for $e_i > 0$ and $b_i > 0$, the average product of the inputs $\mathbf{T}_i, \mathbf{A}_i, \mathbf{K}_i, \mathbf{M}_i$ depends on the levels of e_i and b_i . To measure the per-student educational output produced by the inputs into production, we must therefore condition on some set of reference values of e_i and b_i . If e_i and b_i are constant over time, we can use their actual values as the reference values and treat the observed θ_i as a measure of the educational services produced by $\mathbf{T}_i, \mathbf{A}_i$, and \mathbf{K}_i . Otherwise, we must choose some level of reference values of e_i and b_i , such as their initial level, their final level, or some average in between these. Letting \hat{e}_i and \hat{b}_i denote these reference values, the conditional education production function is defined as:

$$(3) \quad \phi_i = (\mathbf{T}_{it}, \mathbf{A}_{it}, \mathbf{K}_{it}, \mathbf{M}_{it}, q_{it}, \hat{e}_i, \hat{b}_i) = q_{it} f_i(\mathbf{T}_{it}, \mathbf{A}_{it}, \mathbf{K}_{it}, \mathbf{M}_{it}, \hat{e}_i, \hat{b}_i, q_{it}).$$

Equation (3) can be used to measure the change in output of educational services from time t to time s as $\phi_i(\mathbf{T}_{is}, \mathbf{A}_{is}, \mathbf{K}_{is}, \mathbf{M}_{is}, q_{is}; \hat{e}_i, \hat{b}_i) - \phi_i(\mathbf{T}_{it}, \mathbf{A}_{it}, \mathbf{K}_{it}, \mathbf{M}_{it}, q_{it}; \hat{e}_i, \hat{b}_i)$. To estimate this change, the observed θ_{it} and θ_{is} must be adjusted for the effect of substituting \hat{e}_i and \hat{b}_i for the actual values e_{it} and b_{it} and of e_{is} and b_{is} .

We make no such an adjustment in this chapter, however. As a result, the change in outcome probably understates the growth in output of educa-

tional services in the recent past. In particular, increasing numbers of special education students and students whose parents do not speak English have probably had adverse effects on student outcomes.

The basic measures developed here lay the foundation for future research on adjustment of outcomes to reflect changes in nonschool factors. Furthermore, outcome is the appropriate variable for some important questions. Perhaps for this reason, in other industries where external factors heavily influence outcomes, the convention is to ignore the external factors and to accept outcomes as measures of output. In agriculture, for example, weather is a crucial determinant of the size of the harvest, and the spread across borders of disease-causing organisms can affect deliveries of animal products to industry customers. No provision is made for these effects in the calculation of real agricultural output for national accounts purposes.

9.3.3 Use of Input Quantity and Quality to Infer Changes in Output

Empirical research has shown that some input quantity and quality measures are linked to improved educational outcomes, as measured by test scores. These include pupil-teacher ratios and teaching staff composition measures, such as years of education and experience. When direct measures of educational outcomes are unavailable, counting the expected change in educational outcomes that would arise from changes in input quantities or quality in the output measure is better than assuming that output per student educated is constant. A common practice, therefore, is to quality adjust volume indicators by factors that measure the amount or quality of inputs that have been shown to have an important effect on output. The difficult part is to estimate the precise value of the change in output resulting from a given change in inputs. For example, if class sizes drop by 10 percent, does θ_{it} increase by 10 percent? Furthermore, if the quantitative impact of inputs on output is estimated at some point in time, changes in other factors, such as the composition of the student body, might alter the relationship.

9.3.4 Enrollment Data

The U.S. Census Bureau (Census) Current Population Survey student enrollment statistics are used in preference to other sources, such as the U.S. National Center for Education Statistics (NCES). The NCES enrollment data were incomplete in some years. While considered superior, the Census enrollment figures used are also imperfect. Over the time period we consider (1980 to 2001),³⁴ three adjustments to the data had to be made, as can be found in table 9.1: the data for 1981 to 1992 are revised to be consistent with the 1990 Census estimates; interpolation is used for 1980 to deal with the lack of a public/private breakdown of students; and estimates of students aged

34. See appendix B-1 in Williams (2003) for a note on the time series.

Table 9.1 Adjusted census enrollment figures (in thousands)

Year	Primary and secondary	Primary Grades K–8	Secondary Grades 9–12	College
2001	47,775	32,945	14,830	12,421
2000	46,982	32,551	14,431	12,008
1999	47,069	32,431	14,638	11,659
1998	46,551	32,252	14,299	11,984
1997	47,213	32,579	14,634	12,091
1996	45,618	31,506	14,113	12,014
1995	45,308	31,558	13,750	11,372
1994	44,948	31,409	13,539	11,694
1993	44,852	31,867	12,985	11,594
1992	43,878	31,201	12,677	11,765
1991	43,182	30,738	12,444	11,436
1990	42,605	30,446	12,159	11,166
1989	41,947	29,661	12,287	10,644
1988	41,649	29,281	12,368	10,624
1987	41,365	28,549	12,816	10,368
1986	40,755	27,805	12,950	9,803
1985	40,220	27,286	12,934	9,916
1984	40,140	27,282	12,857	9,886
1983	39,960	27,066	12,894	9,466
1982	40,304	27,232	13,072	9,547
1981	40,983	27,426	13,557	9,254
1980	40,548	27,088	13,460	8,785

thirty-five years and over are added in for years before 1994, because these students are not included in the Census enrollment figures.³⁵

9.3.5 Teaching Staff Composition

The U.S. Department of Education NCES “Monitoring School Quality: An Indicators Report” (2000) found that “students learn more from teachers with strong academic skills and classroom experience than they do from teachers with weak academic skills and less experience.”³⁶ Rivkin, Hanushek, and Kain’s (2001) analysis “identifies large differences in the quality of schools in a way that rules out the possibility that they are driven by non-school factors. . . . We conclude that the most significant [source of achievement variation] is . . . teacher quality.”³⁷ Hanushek (1998) states that the “differences in student achievement with a good versus a bad teacher can be more than 1 1/2 grade levels of achievement within a single school year.”³⁸ The NCES report identified thirteen indicators of school quality that recent research suggests are related to school learning; of these, four relate to the

35. See appendix B-2 in Williams (2003) for a full explanation of adjustments.

36. NCES (2000, i).

37. Rivkin, Hanushek, and Kain (2001, 32).

38. Hanushek (1998, 35).

quality of teachers: teacher academic skills, teacher assignment, teacher experience, and professional development.³⁹

Data produced by the NEA "Status of the American Public School Teacher" provide information on teacher educational attainment. Although educational attainment does not perfectly predict how well a person will teach, there is "broad agreement that teachers' academic skills are linked to student learning."⁴⁰ Students appear to learn more from teachers with strong academic training. For example, Darling-Hammond (2000) concludes, "The most consistent highly significant predictor of student achievement in reading and mathematics in each year tested is the proportion of well-qualified teachers in a state."⁴¹ Surveys by the NEA and NCES separate teachers with no degree, a bachelor's degree, a master's degree, a professional diploma, and a doctor's (PhD) degree. Indicating quality change, results show that from 1961 to 1996, the percentage of public elementary and secondary school teachers with a master's degree, specialist's degree, or doctor's degree almost doubled.⁴²

Independent of educational attainment, teacher assignment can directly affect student learning and the quality of education. Many teachers are currently teaching courses in disciplines other than those in which they have been formally trained, and the student achievement has suffered.⁴³ The NCES report states, "Given the apparent benefits students receive from being taught by well-qualified teachers, it is worth assessing the extent to which students are taught by teachers who are teaching without proper qualifications."⁴⁴ While teacher assignment is an important indicator of school quality, defining a teacher as qualified versus unqualified is difficult, and meaningful data are not available.

Studies show that students also learn more when taught by more experienced teachers. Rivkin, Hanushek, and Kain (2002) show that fourth- and fifth-grade students in Texas whose teachers had more than two years of experience increased their math and reading test scores by between 0.12 and 0.19 standard deviations more over the course of a year than those whose teachers had fewer than two years of experience. The NEA and NCES surveys report detailed information regarding teacher experience.

Even though experts would likely agree that professional development should enhance student learning, there is no concrete statistical evidence of such an association.⁴⁵ Conceptually, professional development opportuni-

39. NCES (2000, 4).

40. NCES (2000, 5).

41. See Darling-Hammond (2000, 27).

42. NCES (2003, 82).

43. NCES (2000, 12).

44. NCES (2000, 11).

45. NCES (2000, 14).

ties seem important to help retain quality teachers, but research is needed to document such a relationship.

Of the four indicators of school quality associated with teachers, teacher academic skills (educational attainment) and teacher experience offer the best hope of empirically capturing quality change. Using NEA and NCES survey data that are available for selected school years, the Government Division of the BEA computes a quality-adjusted constant-dollar estimate of labor compensation for education. Educational attainment and experience are taken into account to adjust average real compensation estimates to represent changes in the teaching staff composition. Specifically, annual estimates of the number of teachers cross-classified by experience categories and highest degree obtained are multiplied by 1996 average wages for these same groups, then divided by the total number of teachers in each year to derive an estimate of an annual real average wage.⁴⁶ This series, normalized to 1.0 in 1996, is an index of teaching staff composition. It is used in this chapter as a quality adjuster, under the assumption that differences in average wages paid reflect teacher quality differences.⁴⁷ Table 9.2 shows that although this index of teaching staff composition increased for the period as a whole and for the first subperiod, 1980 to 1990, it decreased during the 1990 to 2001 subperiod. This is probably a reflection of the significant changes in teacher experience shown between the 1990/1991 and 1999/2000 NCES surveys of teachers. This indicator of teaching staff composition change is applied to both primary and secondary education, as there is no evidence of a differing impact upon different grades.

9.3.6 Class Size

Does size matter? Intuition says it must. If class size did not matter, it would be perfectly logical to increase a second-grade class from thirty to sixty students—or to 120, for that matter. Supplemental, out-of-class tutoring would be just as effective when done with groups of ten students as with one-on-one instruction. Although intuition necessitates this conclusion, the measurable impact of class-size variation is debatable and tough to measure.

Finn (1998b) summarizes the findings of some pivotal studies on class size.⁴⁸ Glass and Smith's (1978) statistical meta-analysis of the findings of

46. The NEA and NCES provided the BEA with their survey data, cross-classified by experience and highest degree obtained categories. Experience categories include less than five years of experience, five to ten years of experience, eleven to fifteen years of experience, sixteen to twenty years of experience, twenty-one to twenty-five years of experience, and over twenty-five years of experience. Highest degree obtained categories include no degree, two-year degree, bachelor's degree, master's degree, and doctor's degree.

47. Experience-based adjustments to labor input indexes implicitly assume that wage differentials reflect actual relative marginal productivity differences (perhaps as determined by a merit pay system) as opposed to wage differentials primarily arising from seniority-based wage systems.

48. Finn (1998b).

Table 9.2 Annual rates of growth in prospective quality-adjustment factors (percentages)

	1980–2001	1980–1990	1990–2001
Teaching staff composition	0.13	0.49	–0.20
Pupil-teacher ratio	–0.77	–0.83	–0.71
High school dropout rate	–1.31	–1.52	–1.11
College enrollment rate	1.07	2.00	0.24

over eighty empirical studies show that “reduced class size can be expected to produce increased academic achievement.”⁴⁹ The Educational Research Service analyzed a much larger set of studies, finding mixed results.⁵⁰ One of Robinson’s conclusions is that the class-size effects are more apparent with early primary education. Tennessee’s Project STAR (Student-Teacher Achievement Ratio), a controlled scientific experiment that assigned over ten thousand students to small and large classes at random and then tracked their progress over four years, “provided educators with definitive answers about the impact of small classes in the primary grades.” Project STAR found that statistically significant differences existed among the students in the different size classes on every achievement measure for every year of the study.⁵¹ After returning to regular-size classes, the students of Project STAR were subsequently tracked by the Lasting Benefits Study. It found small but positive carryover effects through at least eighth grade.⁵² Finn’s study (1998b, 4) concludes that “small classes (17 pupils or below) are more effective academically than larger classes (22 and above) in the primary grades in all subject areas.” Class sizes seem especially important, as “teachers spend more time in direct instruction and less time in classroom management when the number of students is small” (4).

Ivor Pritchard (1999) also synthesized previous studies, concluding “the pattern of research findings points more and more clearly toward the beneficial effects of reducing class size.”⁵³ He noted Slavin’s (1989) findings that “reduced class size had a small positive effect on students that did not persist after their reduced class experience.”⁵⁴ Robinson and Wittebols (1986) found that the clearest evidence of the positive effects of smaller classes is in the primary grades. Ferguson (1991), using data on more than eight

49. Glass and Smith (1978, iv).

50. Robinson (1990).

51. Finn reaches this conclusion (1998b, 4). Mosteller (1995) and Krueger (1999) both support the conclusion that Project STAR results show that class size does matter, especially with younger and more economically disadvantaged children.

52. This is the conclusion of I. Pritchard (1999, 4), who cites Finn’s (1998a) citation of Nye et al. (1995).

53. I. Pritchard (1999, 1).

54. I. Pritchard (1999, 2) gives Slavin’s conclusion, citing Finn (1998a) as the source. Finn’s bibliography does not give a citation for Slavin (1989) as a sole author source. Finn’s bibliography includes a 1989 article by Slavin and Madden and a 1989 book edited by Slavin.

hundred districts and 2.4 million students in Texas, found that in grades one through seven, “district student achievement fell as the student/teacher ratio increased for every student above an 18 to 1 ratio.”⁵⁵ Krueger (1998), “in an external re-analysis of the Project STAR data, reconfirmed the original finding that ‘students in small classes scored higher on standardized tests than students in regular classes’ even when the data analysis took into account adjustments for school effects, attrition, rerandomization after kindergarten, nonrandom transitions, and variability in actual class size.”⁵⁶ Ivor Pritchard makes the following conclusions to his synthesis:

- Existing research shows that smaller classes in the early grades lead to higher achievement.
- Reducing class size from over twenty students to under twenty students moves the average student from the fiftieth percentile to the sixtieth percentile in achievement measures.
- Students, teachers, and parents all agreed that smaller classes increase the quality of classroom activity.

On the other side of the debate, Hanushek (1998) claims that in 277 independent studies, only 15 percent found a statistically significant correlation.⁵⁷ “The evidence about improvements in student achievement that can be attributed to smaller classes turns out to be meager and unconvincing.”⁵⁸ The results suggest that while some factors, such as teacher quality, do affect the output of education, class size does not. Using National Assessment of Educational Progress standardized tests data in conjunction with aggregate data on national pupil-teacher ratios over time, Hanushek concluded that smaller classes simply do not outperform larger classes on a consistent basis and that the data do not support the assertion that smaller classes ensure a higher level of output.

Hanushek (2002) suggests possible explanations for the lack of correlation between small classes and improved performance. One is that intraschool class sizes are not decided at random: schools put their lower-achieving students who need extra resources in smaller classes. Also, identification of exogenous determinants of class size is extremely difficult; accordingly, the generalizability of any findings may be jeopardized. As an example, he cites a study by Lazear (2001). Lazear looks at the probability that a student may impede his own learning or others’ learning and suggests that higher-quality teachers may be more capable of keeping students on track. This study raises the question in Hanushek’s mind of whether the probability of disruption should be considered an exogenous factor or dependent upon the teacher’s

55. As cited and quoted in I. Pritchard (1999, 2).

56. I. Pritchard (1999, 5).

57. Krueger (2002) disputes Hanushek’s conclusions after reviewing the same studies covered in Hanushek (1998).

58. Abstract in Hanushek (1998).

classroom management ability.⁵⁹ Except for a few scientifically controlled studies such as Project STAR, the bulk of the studies have no way to control for exogenous factors and simply compare achievement by class size. Other experiments (California, 1996; Indiana's Prime Time Project, 1994; Burke County, North Carolina, 1990; Wisconsin's Student Achievement Guarantee in Education Program, 1996) that systematically reduce class size across a school, district, or state may miss some of the benefits of having smaller classes, because they require hiring new, inexperienced teachers to accomplish the class-size reductions.⁶⁰

Actual class sizes are unavailable, but pupil-teacher ratios, which are available, are a proxy for class size.⁶¹ We therefore use pupil-teacher ratios for quality adjustment. Primary and secondary education pupil-teacher ratios have declined from 18.7 in 1980 to 15.9 in 2001.⁶² Table 9.2 shows the rate of decline in this ratio for the whole period and two subperiods. *Ceteris paribus*, this trend improves the quality of education, resulting in an increase in the output. Because of the controversy regarding the link between pupil-teacher ratios and the quality of education, we dampen, the effect of pupil-teacher ratios by raising them to the 0.1 power, a conservative assumption.⁶³ Letting ρ_{it} denote the student-teacher ratio in year t and w_{i0} denote a weight proportional to expenditures on educational level i (where the levels are primary and secondary), we can define a Laspeyres index of the educational services volume measure in equation (1):

$$(4) \quad Q^{\text{Laspeyres}} = \sum_i w_{i0} \left(\frac{q_{it}}{q_{i0}} \right) \left(\frac{\rho_{it}}{\rho_{i0}} \right)^{-0.1}.$$

With this quality adjustment, a 10 percent decrease in class size results in a 1 percent increase in the output measure. Pupil-teacher ratios are applied as a quality adjustment just for primary education (grades K–8), because an effect on primary education output has greater support in the literature than an effect on both primary and secondary education output.

9.3.7 High School Completion Factor

Two additional quality-adjustment factors that are worth considering are the percentage of the relevant population who complete high school and the percentage who go on to higher education. Two possible proxies for these

59. Hanushek (2002, 48–51).

60. I. Pritchard (1999, 9).

61. Pupil-teacher ratios are not the best measure of class size but are the best data available. See Hanushek (1998, 16) for reasons that the two measures differ, such as effect of special education teachers and aids on pupil-teacher ratios.

62. Eventually, it would be preferred to substitute pupil-teacher ratios for K–8, but these are not readily available, even through the NCES or other sources. The pupil-teacher ratios used come from table 65 in the NCES *Digest of Educational Statistics, 2002* (2003).

63. Krueger (1999) shows that a one-third reduction in class size over four years produced an average gain of 0.2 standard deviations in student achievement. See Hanushek (2002, 65).

factors were considered briefly: the high school dropout rate and college enrollment rates. Additional research is needed to identify and quantify these and other possible quality adjusters.

Research literature needs to be examined to answer two basic questions: To what extent are dropout rates determined by what schools do as opposed to other factors, such as social (including cultural) capital? And, are rising college enrollment rates primarily a sign of schools better preparing students for higher education (e.g., producing higher-quality students), or is this phenomenon mainly a function of changing labor market conditions? To give a sense of how important these potential quality adjustments might be, volume indicators are calculated with and without a dropout rate quality adjustment. The rates of growth of dropouts and college enrollments for recent high school graduates are shown in table 9.2.⁶⁴ The dropout rate quality adjustment is implemented at a 0.1 power, as dropout rates are taken to be an indicator of success for a portion of the high school population.⁶⁵ If the college enrollment quality adjustment is incorporated at a later date after further research, it also might be incorporated at a rate less than 1:1. Table 9.2 shows that the high school dropout rate reduction is larger in absolute value terms (if employed at a 1:1 rate instead of a 10:1 rate) than in any other possible quality-adjustment factor, where a decrease in the dropout rate would produce a higher adjustment than any other shown, with the exception of college enrollment rates for 1980 to 1990.⁶⁶ Over the 1980 to 2001 period, the increase in the college enrollment rate (again if employed at a 1:1 rate) would have the next-largest impact; however, in 1990 to 2001, this possible quality-adjustment factor would have a significantly smaller effect, as college enrollment rates peaked in 1997 at 67.0 percent before dropping to 61.7 percent in 2001.⁶⁷

9.3.8 Prices and Volume Indicators

Table 9.3 presents annual growth rates of a number of alternative prices and volume indicators for selected periods. These fall into three categories: (a) unweighted quality-unadjusted total enrollment; (b) quality-unadjusted enrollment, where the volume indicators are chain-type Fisher quantity indexes; and (c) quality-adjusted enrollment, where the volume indicators are chain-type Fisher quantity indexes. In all cases, the prices are implicit

64. Tables 108 and 183 in the NCES *Digest of Educational Statistics, 2002* (2003). A caveat to the dropout rate table states, "Because of changes in data collection procedures, data may not be comparable with figures for earlier years."

65. The high school dropout rate for persons aged sixteen to twenty-four years varies from a high of 14.1 percent in 1980 to a low of 10.7 percent in 2001. This rate is the average rate across public and private school students. See table 108 in NCES (2003).

66. As with the pupil-teacher ratio, the quality-adjustment factor for the dropout rate is the negative of the growth rates shown in table 9.2.

67. The economy may explain the drop in 2001 or even 2000, but the drop in 1998 and 1999 cannot be explained by an economic slowdown.

Table 9.3 Annual rates of growth in prices and quantities (volume indicators) for primary and secondary public education and gross domestic purchases (percentages)

	1980–2001		1980–1990		1990–2001	
	Price	Quantity	Price	Quantity	Price	Quantity
1 Quality-unadjusted enrollment, unweighted: Total	6.02	0.78	7.29	0.50	4.88	1.05
2 Primary growth rate	6.17	0.94	7.33	1.18	5.12	0.72
3 Secondary growth rate	5.83	0.46	7.55	-1.01	4.29	1.82
<i>Enrollment, chain-type Fisher quantity indexes, and implicit price deflators</i>						
4 Quality-unadjusted enrollment: Total	6.05	0.76	7.41	0.38	4.83	1.10
5 Primary contribution	4.19	0.65	4.88	0.80	3.57	0.51
6 Secondary contribution	1.90	0.14	2.52	-0.33	1.34	0.56
<i>Quality-adjusted enrollment: Totals</i>						
7 Adjusted by teaching staff composition index	5.91	0.89	6.88	0.88	5.04	0.90
8 Adjusted by 0.1 × pupil-teacher ratio	6.00	0.81	7.35	0.44	4.78	1.15
9 Adjusted by 0.1 × high school dropout rate	5.99	0.81	7.34	0.45	4.79	1.14
10 Adjusted by teaching staff composition index and 0.1 × pupil-teacher ratio	5.86	0.94	6.82	0.93	4.99	0.95
11 Adjusted by teaching staff composition index, 0.1 × pupil-teacher ratio, and 0.1 × high school dropout rate	5.80	0.99	6.75	1.00	4.95	0.99
12 Gross domestic purchases	3.05	3.29	4.16	3.35	2.06	3.23

Note: The sum of rows 5 and 6 may not equal the total in row 4 because of rounding.

price deflators. The third category of volume indicators is being used or is under consideration in the most countries. The first two categories of volume indicators are presented in this chapter mainly for purposes of comparison.

Each one of the methods used in table 9.3 can be criticized. An unweighted quality-unadjusted enrollment volume indicator assumes that all pupils in primary and secondary education receive the same quantity of education (e.g., that the output of schools is the same, regardless of whether they are educating a kindergartner or a twelfth grader). Also, it assumes that the quantity of education represented by a pupil year does not change over time. Clearly, these are simplifying assumptions. The growth rates shown for primary education versus secondary education are the unweighted growth rates for these subcategories; accordingly, they do not add up to the growth rate for the total. The methodology underlying the second and third categories, quality-unadjusted and quality-adjusted enrollment, where the volume indicators are chain-type Fisher quantity indexes, is preferred to the methodology underlying the first category, because under certain assumptions, including the assumption that public schools allocate their budget between primary and secondary education to maximize the output produced, cost shares used in Fisher indexes reflect relative marginal products of resources devoted to primary versus secondary education. The growth rates shown for primary education versus secondary education are Fisher index decompositions for these subcategories; accordingly, they do add up to the growth rate for the total. As is true for the first category of indicators, using a quality-unadjusted volume indicator assumes that the quantity of educational output per pupil year within either primary education or secondary education has not changed over time. This seems unlikely, even during the twenty-one-year-period examined.

The preferred approach uses a chain-type Fisher quantity index and includes adjustments for quality changes. However, the question of which quality indicators to include in the measure of quality change and of how to specify the equations for their effect are difficult to answer. Table 9.3 shows the prices and volume indicators implied by three possible indicators and by two possible combinations of these indicators. At this time, because further research needs to be performed on the use of high school completion as a quality indicator, the enrollment volume indicator, quality adjusted by an index of teaching staff composition and pupil-teacher ratios, and the implicit price derived from the volume indicators are favored. However, all measures are exploratory.

The second and third category of alternative volume indicators can be written as follows: Let $z_{p,y}$ represent enrollment in primary school in year y , and let $z_{s,y}$ represent enrollment in secondary education in year y . The enrollment growth rates for primary and secondary education are calculated as $GR(z_{p,1980,2001}) = (z_{p,2001}/z_{p,1980})^{1/21} - 1$ and $GR(z_{s,1980,2001}) = (z_{s,2001}/z_{s,1980})^{1/21} - 1$, respectively.

Let $GR(TSCI_{1980,2001})$ denote the growth rate of the teacher composition index, and let $GR(z_{p,1980,2001})$ and $GR(z_{s,1980,2001})$ denote the growth rates of primary and secondary school enrollment. Then, the growth rate of the volume indicator with a teaching staff composition adjustment for primary education is:

$$(5) \quad GR(z_{p,1980,2001}, TSCI_{1980,2001}) = GR(z_{p,1980,2001}) + GR(TSCI_{1980,2001}),$$

and the growth rate of the volume indicator with a teaching staff composition adjustment for secondary education is:

$$(6) \quad GR(z_{s,1980,2001}, TSCI_{1980,2001}) = GR(z_{s,1980,2001}) + GR(TSCI_{1980,2001}).$$

The growth rate of the volume indicator with a pupil-teacher ratio (PTR) adjustment for primary education is:

$$(7) \quad GR(z_{p,1980,2001}, PTR_{1980,2001}) = GR(z_{p,1980,2001}) - 0.1 GR(PTR_{1980,2001}),$$

and the growth rate for secondary education is $GR(z_{s,1980,2001})$ as calculated above, as the pupil-teacher adjustment is only applied to primary education. The growth rate of the pupil-teacher ratio is entered with a negative, as an increase in the ratio is associated with a decline in output quality, and a decrease is associated with a rise in output quality.

To adjust for changes in the dropout rate (DOR), the growth rate of the volume indicator adjusted for changes in school completion rates for secondary education as proxied by the changes in the dropout rate is:

$$(8) \quad GR(z_{s,1980,2001}, DOR_{1980,2001}) = GR(z_{s,1980,2001}) - 0.1 GR(DOR_{1980,2001}).$$

The growth rate in the dropout rate is entered with a negative, as an increase in the rate is associated with a decline in output quality, and a decrease is associated with a rise in output quality. The growth rate for primary education is $GR(z_{p,1980,2001})$ as calculated above, as the dropout rate adjustment is only applied to secondary education.

The growth rate of the primary education volume indicator adjusted for changes in teaching staff composition and the pupil-teacher ratio is:

$$(9) \quad GR(z_{p,1980,2001}, TSCI_{1980,2001}, PTR_{1980,2001}) \\ = GR(z_{p,1980,2001}) + GR(TSCI_{1980,2001}) - 0.1 GR(PTR_{1980,2001}).$$

The growth rate for secondary education is $GR(z_{s,1980,2001}, TSCI_{1980,2001})$ as calculated above, as the pupil-teacher adjustment is only applied to primary education.

The growth rate of the volume indicator adjusted for changes in teaching staff composition, the pupil-teacher ratio, and the high school dropout rate for secondary education is:

$$(10) \quad GR(z_{s,1980,2001}, TSCI_{1980,2001}, PTR_{1980,2001}, DOR_{1980,2001}) \\ = GR(z_{s,1980,2001}) + GR(TSCI_{1980,2001}) - 0.1 GR(PTR_{1980,2001}) \\ - 0.1 GR(DOR_{1980,2001}).$$

The growth rate for primary education is $GR(Z_{p,1980,2001}, TSCI_{1980,2001}, PTR_{1980,2001})$ as calculated previously, as the high school dropout rate adjustment is only applied to secondary education.

Quality-adjusted volume indicators are then calculated for primary and secondary education by applying the quality-adjusted growth rates to a 1996 base set equal to enrollment in 1996. Implicit price indexes are estimated by dividing nominal expenditures by the volume indicators. The resulting implicit price index is normalized to 1.0 in 1996. The final step is to calculate a chain-type Fisher quantity index with quality-adjusted enrollment and implicit prices for primary and secondary education as the inputs and to calculate the implicit price index associated with the chain-type Fisher quantity index.⁶⁸

Decomposing the Fisher chain-type indexes allow for estimation of the contribution of the subcomponents: primary and secondary education to growth in prices and quantities for the aggregate. The results of a decomposition for the quality-unadjusted estimates for the preferred indexes (that which uses teaching staff composition and the pupil-teacher ratio to adjust enrollment) is shown in the middle panel of table 9.3.

The growth rate of the decomposition of the chain-type Fisher quality-unadjusted volume index, c_i , is calculated as

$$(11) \quad GR(c_{iy}) = \bar{s}_{iy} \left(\frac{q_{iy+1}}{q_{iy}} - 1 \right),$$

for i = primary education or secondary education, where

$$(12) \quad \bar{s}_{iy} = \frac{F^P p_{iy} q_{iy} + p_{iy+1} q_{iy}}{F^P (\sum_j p_{jy} q_{jy}) + \sum_j p_{jy+1} q_{jy}}.$$

The variable F^P is a chain-type Fisher price index for year y ; $p_{iy+1} q_{iy}$ represents expenditures on education level i in year y , adjusted for price change between year y and year $y + 1$; and \bar{s}_{iy} may be interpreted as a weighted average of the expenditure share for education level i in year y and its hypothetical share at year $y + 1$ if only prices had changed. The quality-unadjusted chain-type Fisher quantity indexes for primary and secondary education are then calculated from the growth rates in the same manner as previously described.

The decomposition of the chain-type Fisher quality-unadjusted price index is calculated using equations (10) and (11), with the price relative substituted for the quantity relative in equation (10) and with chain-type Fisher quantity indexes, F^Q , substituted in for the Fisher price indexes in equation (11). The quality-unadjusted chain-type Fisher price indexes and implicit price indexes for primary and secondary education are then calculated in

68. For an explanation of how chain-type Fisher indexes are constructed and a discussion of their properties, see Young (1992), Diewert (1993), and Triplett (1992). Because of the properties of Fisher indexes, the implicit price indexes are Fisher price indexes.

a manner parallel to the quality-unadjusted chain-type Fisher quantity indexes and implicit price indexes, with appropriate normalization.

Table 9.3 shows that price change is always greater than quantity change for the periods listed, with the price change typically being in the ballpark of twice the U.S. gross domestic purchases price change. When making comparisons, it should be remembered that the price changes in table 9.3 are probably overstated and the quantity changes understated. This is because of quality improvements occurring over time that have not yet been, or perhaps never will be (due to lack of data), captured in the estimates and because of other factors leading to higher expenditures per pupil, such as the increase in the number of special education students. For example, has the quality of education received in high school increased, as evidenced by an increase in Advanced Placement courses? The comparison is made to gross domestic purchases prices rather than to GDP prices to exclude exports, which are included in GDP and excluded in gross domestic purchases, and to include imports, which are excluded in GDP and included in gross domestic purchases. Figure 9.2 plots the preferred price deflator (derived from the volume indicator that uses the teaching staff composition index and the pupil-teacher ratio to adjust enrollment) against the gross domestic pur-

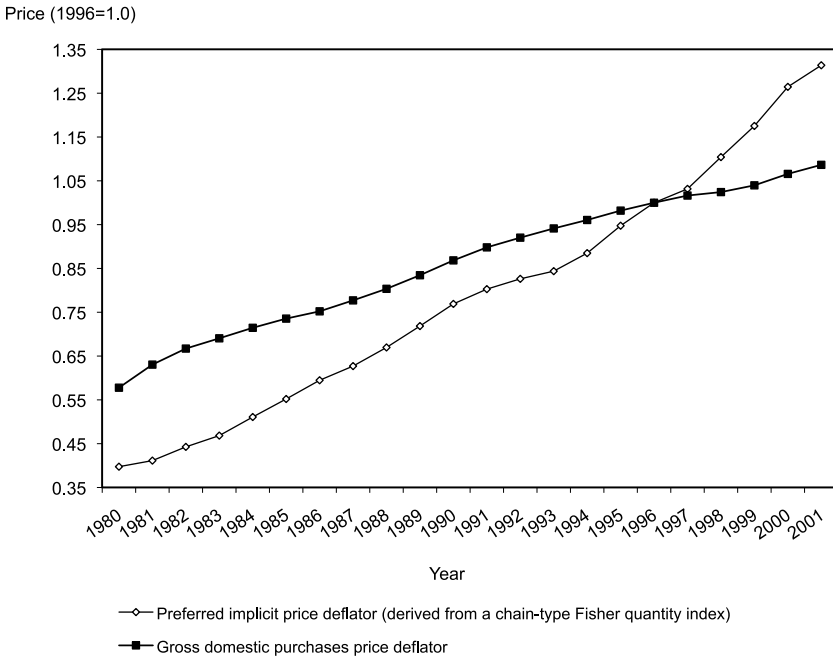


Fig. 9.2 Preferred implicit price deflator versus the gross domestic purchases price deflator

chases price deflator. Except for a brief period during the early nineties, the preferred price deflator rises at a rate faster than the gross domestic purchase price deflator. The decomposition of the price deflators derived from chain-type Fisher quality-unadjusted enrollment indexes in the middle panel of table 9.3 show that this is primarily because of the significantly higher contribution of primary education price change (4.19 percent versus 1.90 percent, 4.88 percent versus 2.52 percent, and 3.57 percent versus 1.34 percent). The rate of price change did moderate significantly in the last period, 1990 to 2001, compared to the first period, 1980 to 1990.

Enrollment data, which are the foundation for all volume indicators, show the influence of demographics. Noticeable is the decline in the population of high school students during 1980 to 1990, which ripples through all measures, but it is most apparent in the unweighted quality-unadjusted enrollment growth rates for secondary education in the top panel of table 9.3. Total enrollments nonetheless have increased during all three periods.

The difference between the top panel and the middle panel total growth rates reflect the fact that it is substantially more expensive to educate a secondary-school student than a primary-school student. The average expenditure per secondary student is estimated to be significantly higher than that per primary student.⁶⁹ On average, only either 30 percent or 31 percent of all primary and secondary students attend secondary school. Relative expenditures enter into the Fisher index calculation.

Looking at the middle panel of table 9.3, the total growth rates for the quality-unadjusted measures can be compared directly to the quality-adjusted enrollment volume indicators growth rates. Note that the change in the quantity index is offset by a change in the opposite direction in the price deflators.⁷⁰ This fact again highlights the sensitivity of the price results to quality adjustment of the quantity indexes. It is easiest to compare the quality-unadjusted estimates with those adjusted by the teaching staff composition index, as this difference, except for rounding, is exactly equal to the growth rate for the teaching staff composition index, shown in table 9.2. However, as the pupil-teacher ratio and high school dropout rate quality adjustments affect only one part of enrollments—not all enrollments, as with the teaching staff composition index—it is much more difficult to make a direct comparison. The impact of both are reduced, because the

69. It is difficult to estimate expenditure per student for primary versus secondary students, because expenditures may be reported on a school district basis, aggregated across primary and secondary schools, and because of different school formats (e.g., middle schools versus junior high schools). Our expenditure per-student estimates are based on *Digest of Educational Statistics* tables. See various issues of the NCES, *Digest of Educational Statistics*.

70. With Fisher indexes, the growth rates are related by the following equation:

$$(1 + n) = (1 + p) \cdot (1 + q),$$

where n is the nominal growth rate, p is the price growth rate, q is the quantity growth rate, and the growth rates are in decimal format (e.g., a 6.00 percent growth rate appears as .0600).

weights are less than one and because minus the pupil-teacher ratio and the dropout rate, both are entered at a 0.1 power. Accordingly, even though the absolute value of the rates of growth of the pupil-teacher ratio and the dropout rate are greater than that for the teaching staff composition index (see table 9.2), the volume indicators with the pupil-teacher ratio and the dropout rate adjustments grow at a slower rate for 1980 to 2001 than that with the teaching staff composition adjustment.⁷¹

These estimates show that quality adjusting a volume indicator can have a significant effect on estimated output and prices. The difference between the growth rates for the quality-unadjusted measure and the preferred quality-adjusted measure (that using the teaching staff composition index and the pupil-teacher ratio) is 0.18 percent, 0.55 percent, and -0.15 percent for 1980 to 2001, 1980 to 1990, and 1990 to 2001, respectively.⁷² The impact on output is greater than the impact on prices, as the rates of growth of quantities are much smaller than the rates of growth of prices. Chained BEA 2000 dollar estimates for primary and secondary education using an input cost-based output approach became available in October 2004. A comparison can be made between those estimates and the quality-adjusted output estimates presented here.⁷³

9.4 Conclusion

Given its goal of continuously improving the U.S. national accounts, the BEA is examining a number of possible changes to the way it measures the output of the government sector. This exploratory chapter looks at one possible methodology that might be adopted if a change is made. Focusing on prices particularly highlights that much additional research needs to be undertaken, both for primary and secondary education and for other components of the government education function (e.g., for higher education and libraries). For primary and secondary education, beyond looking at high school completion factors, additional research is needed. This includes research on trends in numbers of teaching specialists; and research on the number and sizes of special education classes, English as a second language (ESL) classes, and other special classes to interpret or modify the pupil-teacher ratios; research on the impact and growth of school-sponsored activities; and research on the composition of the student body, as it affects

71. The fact that the growth rates for the volume indicator with a pupil-teacher quality adjustment and for the volume indicator with a high school dropout rate quality adjustment are almost identical is coincidental. The product of the (higher) expenditure weight for primary school with the absolute value of the (lower) rate of growth for the pupil-teacher growth rate is equal to the product of the (lower) expenditure weight for secondary school with the absolute value of the (higher) rate of growth for the high school dropout rate.

72. Recall that changes in the experience distribution seem to be driving the decline in the teaching staff composition index over the 1990 to 2001 period. See table 9.2.

73. The relevant BEA category is titled "elementary and secondary education."

learning—these are just a few possible avenues of future work. As the title indicates, this chapter is exploratory.

References

- Algera, S. 1998. The measurement of price and volume changes of government output: The case of the Netherlands. Paper presented at the joint meeting of the Economic and Social Commission for Asia and the Pacific (ESCAP) and the Organization for Economic Cooperation and Development (OECD); System of National Accounts, 1993: Five Years on. 4–8 May, Bangkok.
- Atkinson Commission. 2005. *Atkinson Review: Final report; measurement of government output and productivity for the national accounts*. Hampshire, England: Palgrave Macmillan.
- Australian Bureau of Statistics (ABS). 1998. Measuring non-market sector output: Recent work by the Australian Bureau of Statistics. Paper presented at the OECD Meeting of National Accounts Experts, STD/NA(98)3. 22–25 September, Paris.
- . 1999. Non-market output: Recent work by the Australian Bureau of Statistics. Paper presented at the OECD Meeting of National Accounts Experts, STD/NA(99)41. 21–24 September, Paris.
- . 2000a. Australian national accounts: Concepts, sources and methods. ABS Publication no. 5216.0, December.
- . 2000b. Non-market output: Recent work by the Australian Bureau of Statistics. Paper presented at the OECD Meeting of National Accounts Experts, STD/NA/RD(2000)04. 26–29 September, Paris.
- . 2001a. Experimental estimates of human capital for Australia. Paper presented at the OECD Meeting of National Accounts Experts. 9–12 October, Paris.
- . 2001b. New chain volume estimates for the services sector. Australian national accounts: National income, expenditure and product. ABS Publication no. 5206, March.
- . 2002a. Australian national accounts: Non-profit institutions satellite account. ABS Publication no. 5256.0, November.
- . 2002b. New volume estimates for health and education services. *Year book Australia, 2002*. ABS Publication no. 1301.0, January.
- . 2003. Outputs and outcomes. *Measuring learning in Australia: A framework for education and training statistics, 2003*, chap. 9. ABS Publication no. 4213.0, January.
- Black, S. 1998. Measuring the value of better schools. *FRBNY Economic Policy Review* 4 (1): 87–94.
- Caplan, D. 1998. Measuring the output of non-market services. *Economic Trends* 539 (October): 45–49.
- Cipollone, P., and A. Rosolia. 2007. Social interactions in high school: Lessons from an earthquake. *American Economic Review* 97 (3): 948–65.
- Commission of the European Communities, International Monetary Fund, Organization for Economic Cooperation and Development, United Nations, and World Bank. 1993. *System of National Accounts, 1993*. New York: United Nations.
- Currie, J., and D. Thomas. 1999. Early test scores, socioeconomic status and future outcomes. NBER Working Paper no. 6943. Cambridge, MA: National Bureau of Economic Research, February.

- Darling-Hammond, L. 2000. Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives* 8 (1). Available at: <http://epaa.asu.edu/epaa/v8n1/>.
- Diewert, W. E. 1993. Fisher ideal output, input and productivity indexes revisited. In *Essays in index number theory*, vol. 1, ed. W. E. Diewert and A. O. Nakamura, 317–54. Amsterdam: North-Holland.
- Ferguson, R. F. 1991. Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation* 28 (2): 465–98.
- Finn, J. D. 1998a. *Class size and students at risk: What is known? What is next?* Washington, DC: U.S. Department of Education.
- . 1998b. Class size: What does research tell us? *Laboratory for student success: Spotlight on student success*, no. 207: 2–4.
- Fraumeni, B. M., M. B. Reinsdorf, B. B. Robinson, and M. P. Williams. 2004. Real output measures for the education function of government: A first look at primary and secondary education. Paper presented at the National Institute of Economic and Social Research Public Services Performance workshop, 2 March, London.
- Glass, G., and M. L. Smith. 1978. *Meta-analysis of the relationship of class size and student achievement*. San Francisco: Far West Laboratory for Educational Research.
- Griliches, Z. 1994. Productivity, R&D, and the data constraint. *American Economic Review* 84 (1): 1–23.
- Hadar, E., P. Mandler, and A. Barzel. 1998. Indicators for changes in output of non-market services. Paper presented at the OECD Meeting of National Accounts Experts, 22–25 September, Paris.
- Hanushek, E. A. 1998. The evidence on class size. University of Rochester, W. Allen Institute of Political Economy. Occasional Paper no. 98-1, February.
- . 2002. Publicly provided education. NBER Working Paper no. 8799. Cambridge, MA: National Bureau of Economic Research, February.
- Jenkinson, G. 2003. Measuring the volume of government outputs. Paper presented at the OECD Meeting of National Accounts Experts, 7–10 October, Paris.
- Jorgenson, D. W., and B. M. Fraumeni. 1992. The output of the education sector. In *Studies in income and wealth*, vol. 56, *Output measurement in the service sectors*, ed. Z. Griliches, 303–43. Chicago: University of Chicago Press.
- Konijn, P., and F. Kleima. 2000a. Volume measurement of education. Paper presented at the OECD Meeting of National Accounts Experts, STD/NA(2000)27, 26–29 September, Paris.
- . 2000b. Volume measurement of education. Statistics Netherlands. Occasional paper, November.
- Krueger, A. B. 1998. Experimental estimates of educational production functions. NBER Working Paper no. 6051. Cambridge, MA: National Bureau of Economic Research, March.
- . 1999. Experimental estimates of educational production functions. *Quarterly Journal of Economics* 114 (2): 497–532.
- . 2002. Economic considerations and class size. NBER Working Paper no. 8875. Cambridge, MA: National Bureau of Economic Research, April.
- Lazear, E. P. 2001. Educational production. *Quarterly Journal of Economics* 116 (3): 777–803.
- Malizia, R. 1998. The estimation of general government services at constant prices: Methodology and application proposal for Italy. Paper presented at the joint meeting of ESCAP and OECD; System of National Accounts, 1993: Five Years on, 4–8 May, Bangkok.
- Mosteller, F. 1995. The Tennessee study of class size in the early school grades. *The Future of Children: Critical Issues for Children and Youths* 5 (2): 113–27.

- National Center for Education Statistics. 2000. Monitoring school quality: An indicators report. U.S. Department of Education, Office of Educational Research and Improvement. December.
- . 2003. *Digest of education statistics, 2002*. Available at: <http://www.nces.ed.gov/programs/digest/>.
- National Education Association. 2004. Special education and the Individuals with Disabilities Act. Available at: <http://www.nea.org/specialed/>.
- Neuberger, H., and D. Caplan. 1998. The measurement of real public sector output in the national accounts. *Economic Trends* 531 (February): 29–35.
- Niemi, M. 1998. Measuring government sector output and productivity in Finland: Application of the Output Indicator Method. Paper presented at the OECD Meeting of National Accounts Experts. 22–25 September, Paris.
- Nye, B., B. D. Fulton, J. Boyd-Zaharias, and V. A. Cain. 1995. *The last benefits study: Eighth grade technical report*. Nashville: Center of Excellence for Research in Basic Skills, Tennessee State University.
- Office for National Statistics, United Kingdom. 1997. Use of productivity estimates in the United Kingdom output measure of gross domestic product. Paper presented at the joint meeting of ECE, OECD, and Eurostat on National Accounts, STD/NA(97)26. 3–6 June, Paris.
- . 2000. Development in the measurement of general government output. Paper presented at the OECD Meeting of National Accounts Experts. 26–29 September, Paris.
- Office for National Statistics, U.K. Center for the Measurement of Government Productivity, United Kingdom. 2005. Public service productivity: Education. Available at: http://www.statistics.gov.uk/articles/nojournal/Education_productivity_2007_main.pdf.
- O'Mahony, M., and P. Stevens. 2003. International comparisons of performance in the provision of public services: Outcome based measures for education. National Institute of Economic and Social Research (NIESR). November.
- . 2004. International comparisons of performance in the provision of public services: Outcome based measures for education. Paper presented at the National Institute of Economic and Social Research Public Services Performance workshop. 2 March, London.
- Powell, M., and A. Pritchard. 2002. Measuring government output: Mystical or misunderstood? Paper presented at the twenty-seventh general conference of the International Association for Research in Income and Wealth. 18–24 August, Djurhamn, Sweden.
- Pritchard, A. 2002a. Measuring productivity change in the provision of public services. *Economic Trends* 582 (May): 20–32.
- . 2002b. Measuring productivity change in the provision of public services. Paper presented at the NIESR conference on Productivity and Performance in the Provision of Public Services. 19 November, London.
- . 2003. Understanding government output and productivity. *Economic Trends* 596 (July): 27–40.
- Pritchard, A., and M. Powell. 2001. Direct measures of government output: A few conceptual and practical issues. Paper presented at the OECD Meeting of National Accounts Experts. 9–12 October, Paris.
- Pritchard, I. 1999. Reducing class size: What do we know? National Institute of Student Achievement, Curriculum and Assessment. Office of Educational Research and Improvement, U.S. Department of Education. Available at: <http://www.ed.gov/pubs/ReducingClass/index.html>.
- Rivkin, S. G. 2000. The estimation of productivity change in education. Amherst College. April.

- Rivkin, S. G., E. A. Hanushek, and J. F. Kain. 2001. Teachers, schools and academic achievement. Revised 1998 NBER Working Paper no. 6691. Cambridge, MA: National Bureau of Economic Research.
- . 2002. Teachers, schools and academic achievement. Revised 1998 NBER Working Paper no. 6691. Cambridge, MA: National Bureau of Economic Research, July.
- Robinson, G. E. 1990. Synthesis of research on effects of class size. *Educational Leadership* 47 (7): 80–90.
- Robinson, G. E., and J. H. Wittebols. 1986. Class size research: A related cluster analysis for decision-making. Arlington, VA: Education Research Service.
- Sherwood, M. K. 1994. Difficulties in the measurement of service outputs. *Monthly Labor Review* 117 (3): 11–19.
- Slavin, R. E., and N. A. Madden. 1989. What works for students at risk: A research synthesis. *Educational Leadership* 46 (5): 4–13.
- Statistics Netherlands, National Accounts Branch. 2003. Inventory of sources and methods for price and volume measures in the Dutch national accounts. BPA no. 02254-02-MNR.
- Triplett, J. E. 1992. Economic theory and BEA's alternative quantity and price indexes. *Survey of Current Business* (April): 49–52.
- U.S. Department of Commerce, Bureau of Economic Analysis. 2002a. Annual NIPA revision: Newly available tables. *Survey of Current Business* (October): 12–19.
- . 2002b. Table 3.15.1: Percent change from preceding period in real government consumption expenditures and gross investment by function. Available at: <http://www.bea.gov/national/nipaweb/TableView.asp?SelectedTable=307&Freq=Year&FirstYear=2006&LastYear=2007>.
- . 2003. Table 1.1.5: Gross domestic product. Available at: <http://www.bea.gov/national/nipaweb/TableView.asp?SelectedTable=58&Freq=Qtr&FirstYear=2006&LastYear=2008>.
- Williams, M. P. 2003. Real output measures for the education function of government: A more relevant approach. Draft, December.
- Young, A. H. 1992. Alternative measures of change in real output and prices. *Survey of Current Business* (April): 32–48.