Scale adjusted volume measurement of educational services

10

Jeroen van den Tillaart

Discussion paper (201023)

Statistics Netherlands

Explanation of symbols

.	= data not available
*	= provisional figure
**	= revised provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2008–2009	= 2008 to 2009 inclusive
2008/2009	= average of 2008 up to and including 2009
2008/'09	= crop year, financial year, school year etc. beginning in 2008 and ending in 2009
2006/'07–2008/'09	= crop year, financial year, etc. 2006/'07 to 2008/'09 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Scale Adjusted Volume Measurement of Educational Services

Jeroen van den Tillaart[1]

Abstract

This exploratory paper proposes a new method to measure the output of government-provided secondary education in the Netherlands. In this method transferred knowledge and skills, not pupil-hours or pupil-years, constitute the principal output measure of education. The legal possibilities to transfer between different types of theoretical education offer an opportunity to construct a ratio scale which reflects the transfer of knowledge and skills between these types of education. Counts of (successfully completed) pupil-years should be adjusted in confirmation with this scale. Because transferred knowledge and skills per pupil-year – and not pupil-years per se – are used as output measure, a transformation of traditional unit cost weights is needed.

Calculations show that when transferred knowledge and skills are taken as output measure, in recent years output volume is higher than when plain pupil-years are taken as principal output measure. This is the result of two causes: the proportion of pupils who successfully finish their school year is growing, and proportionally more pupils succeed in finishing a higher level of secondary education.

Keywords: (cost) weights, education, quality adjustment, output measure, volume index.

---

[1] Statistics Netherlands, Henri Faasdreef 312, 2492 JP The Hague, Netherlands; j.vandentillaart@cbs.nl. This paper represents views of the author which do not necessarily coincide with the official position of Statistics Netherlands.

1. Introduction

This exploratory paper presents a new method to measure the output of government-provided secondary education in the Netherlands. The research which led to this article was funded by a grant from Eurostat and has been carried out between 2009 and 2010. The most important aim of this research was to refine existing output measures for secondary education and to look for well-founded ways to adjust the measured output for quality. The method proposed here is still in development and is as yet not introduced in the Dutch national accounts.

Measuring non-market education output is known to be hard. Defining outputs, units of production and determining weights to aggregate between different types of services is challenging. Isolating the marginal benefits of education is problematic, because education is a service with considerable inputs that lay beyond the control of the supplier of the service, such as student and parent time. The outcome of education not only depends on the service delivered, but on student ability, family, peer group, and neighborhood factors as well (Rivkin in Fraumeni et al. 2008: 1-2). Furthermore, some benefits of education such as social well-being, good citizenship, and life skills are difficult or even impossible to measure. Nonetheless one can discern at least three general approaches that are used to determine change in output volume of non-market services such as education:

(1) Input methods: According to this approach, which was chiefly used between the early sixties to the mid-nineties (Smith & Street 2007: 46), the change in the volume of inputs is representative for the change in the volume of output. The advantage of this approach is that it avoids the need to measure and value education output. However, it is not certain that more or better inputs automatically lead to more or better output. When the change in output does not correspond with the change in inputs this method wrongly estimates change in volume. Moreover, technological improvements which lead to reductions in expenditure may appear to reduce output, when in fact only inputs may have been reduced. In short, using the assumption that output equals inputs makes it impossible to analyze changes in productivity.

(2) Pseudo output methods: 'Pseudo' output price indices are derived by adjusting input price indices for observed productivity growth in other related fields of production, or by basing growth on observed output price indices of similar products (SNA 2008: 309). The most important drawback of this approach is that only by direct measurement of the output one can verify the accuracy of the assumed adjustments.

(3) Output volume methods or direct volume measurement approach: This approach is recommended for the measurement of individual services such as health and education (Atkinson 2005; Eurostat 2001; OECD 2009; SNA 2008; Stiglitz et al. 2009). The output volume indicator is a weighted measure of outputs of the various non-market goods and services produced. The measure should fully reflect changes in both quantity and quality (SNA 2008: 309). When it does, it can be used to analyze productivity.

The Eurostat Handbook (2001: 1, 17-18) states that from January 1st 2006 onwards all member states from the European Union (EU) are required to use direct volume measures for individual non-market output in the National Accounts. This requirement led to major research efforts to the development of new output methods. In the Netherlands output methods for primary and secondary education were implemented in 2001 (see Konijn & Kleima 2000), and for tertiary education in 2006 (see Van den Berg 2007).

Most of the volume indicators developed thus far are based on the number of pupils as a proxy for the number of pupil-hours taught, in conformation with Eurostat and OECD recommendations. Because simple number of pupils (or pupil-hours) is too undifferentiated an output measure to be meaningful, many countries are looking for more accurate measures of volume growth of education output. Without taking the quality of the service delivered into account, an output measure may still over- or underestimate growth in the sector. After all, the reliability of output-based measures depends largely on the question whether they are detailed enough 'to avoid conflating true volume changes with compositional effects' (Stiglitz et al. 2009: 100). For example vocational education may be much more expensive than theoretical education. Changes in the distribution of pupils over these types of education should be accounted for. A first step to do this is to stratify pupil numbers to type of education and to weigh each with their specific unit costs. This type of adjustment, which is often referred to as 'implicit quality adjustment', is strongly recommended by the OECD (2009: 36).

But the quest to more 'explicit' forms of quality adjustment has started as well. One can think of quality adjustments on the basis of exam results or other kinds of attainment rates, input quality indicators such as teaching experience, school inspection reports, and class size. The diversity of the indicators and variety of approaches used to integrate these in the volume indicators are characteristic for the difficulties involved with this type of quality adjustment.

In this paper a new method is proposed for measuring secondary education output in the Netherlands. It is not based on pupil-hours as preferable output units, but on units of transferred skills and knowledge. The fundamental idea behind this view is that not the activity of teaching expressed by pupil-hours itself, but the amount of knowledge and skills transferred by schools should be the principal measure of output. If it is accepted that the transfer of knowledge and skills is the core business of schools, this output concept enables us to capture more closely the production of schools. Next to that, it enables us to account for the changes in educational output caused by changes in the distribution of pupils over different types of secondary (theoretical) education. The discussion of the implications of taking transferred knowledge and skills as principal measure of output will be the main topic of this article.

The paper begins with an overview of the most commonly used output measures and quality adjustments. Next it discusses the rationale behind these methods and some of the problems connected with them. In section three the notion of educational output as transferred knowledge and skills is introduced and elaborated. Thereafter the weights and Laspeyres volume index formula are discussed which are needed to aggregate different units of transferred knowledge and skills. This more theoretical part of the paper is followed by a discussion of results using the newly developed method in the context of Dutch secondary education, before finishing with concluding remarks and future steps.

2. Prior research on measurement of educational outputs

In recent years much attention has been given to the development of volume indicators based on direct output measurement. For the European Union (EU) an important stimulus was a requirement dated from 2002 that made direct output measurement for individual government services compulsory from 2006 onwards. Since then most EU-members have implemented output measures. Outside the European Union research institutes and statistical agencies have looked for volume

indicators using an output approach as well (See for example Douglas 2006: 8; Fraumeni et al. 2008; Statistics New Zealand 2010; Trewein 2002: 779).

Because education is an individual service, numbers of pupils form the basic counting measure in all these output methods (see OECD 2008): No pupils, no production. Many countries, the Netherlands included (Konijn & Kleima 2000), take pupil numbers as principal *counting measure*, and continue taking pupil-hours or pupil-years as *output measure*.[2] This is in agreement with the recommendations of Eurostat and the OECD. Note, that taking pupil-hours or years as output measure means that the volume indicator is closely connected to demographic developments. There is a strong correlation between the production of secondary education and the population between 12 to 18 years. If only pupil-years were taken as volume indicator production would grow or decline in proportion to the population of pupils in secondary education. Of course the output volume of schools is largely, but not fully, determined by demographic developments. Therefore, there is a strong need for quality adjustment of the output measure.

A first step to construct a more satisfying output measure for education services involves differentiating pupils by type of education, and weighting the pupil numbers of each stratum by their relative value. Suppose that pupils who follow a vocational technical education cost more to educate than pupils who attend a theoretical oriented education. It is then regarded necessary to stratify pupil counts to account for changes in the composition of the pupil population. In this way, not only pupil numbers, but the distribution of the pupils over the two types of education as well, determines output volume. Most statistical agencies using output measures make this kind of stratifications (see OECD 2009: 46-50). The stratified pupil counts (q_i) are generally weighted by accompanying costs per type of pupil-year (p_i). The level of differentiation in types of education services may vary and is often limited by the availability of data on cost weights.

The index formula V_L that describes changes in outputs for types of education i, $i \in I$ from year $t\text{-}1$ to t then can be written by means of a standard Laspeyres volume index:

$$(1) \qquad V_L = \frac{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t}}{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t-1}}$$

Eurostat and the OECD call for more explicit ways to adjust for quality as well. The statistical agencies of different countries have experimented with a wide range of possible quality adjustments. Italy uses a model which adjusts output with respect to average class size (Collesi et al. 2007). In the United Kingdom different test results are used to derive a quality adjustment factor that expresses the contribution that each year of schooling makes to the average test scores achieved. In Scotland Briggs (2008) and Murray (2007) examined the use of attainment rates differentiated into three levels under hold of a threshold. The three levels are

---

[2] An output measure is strictly speaking a volume measure. Therefore it is more correct to use the concepts 'pupil hours' or 'pupil-years' instead of 'number of pupils'. In this article 'pupil hours' refers to the actual number of hours taught. The concept 'pupil-year', in contrast, refers to a full-time equivalent (fte) of teaching given to a pupil in the period of one school year. The number of teaching hours that constitutes one fte of teaching (or one pupil-year) may be varying over the years. This means that the concepts 'pupil-hours' and 'pupil-years' are not used as synonyms and thus cannot be used interchangeably.

weighted subsequently with future earnings as a proxy to adjust for quality. In a recent publication from Statistics Denmark PISA-scores and drop-out rates are put forward as possible candidates to adjust education output for quality (Deveci 2009). Fraumeni et al. (2008: 20), finally, explore using input quality measures such as pupil-teacher ratio, teaching staff composition, when direct measures of educational outcomes are unavailable. In addition they propose to use information about the percentage of pupils that completes high school and the percentage of pupils that enrolls in higher education.

Unfortunately most of these quality adjustments lack the robustness and reproducibility that is necessary for them to be implemented into the national accounts. The majority of all research to quality adjustment still bears an exploratory character. For the purpose of the national accounts most countries end up using stratified pupil numbers as volume indicator. Therefore output measures of education still remain closely tied to demography, and at a fair distance from the increase of competences of pupils (Stiglitz et al. 2009: 47). In the next section a new output measure for education services is proposed that may be helpful to determine education outputs more in line with the characterization given by Stiglitz et al.

3. Education output reconsidered

As already discussed, finding suitable measures of production volume for non-market services such as education is complicated. First one has to determine what education output is, what the units of production are that should be counted, and what types of 'products' are delivered. The Eurostat Handbook defines education output as '[t]he quantity of teaching (that is, the transfer of knowledge, successfully or not) received by the students, adjusted to allow for the qualities of the services provided, for each type of education' (2001: 114), or as 'the amount of teaching consumed by a pupil' (2001: 32). And the OECD (2009: 28) defines education output as 'the organized communication of knowledge', to be discerned from 'learning as an outcome (the change in behavior, information, etc.).' Following this depiction of education output, the units of production to be counted are pupil-hours or pupil-years, differentiated sufficiently to correct for compositional effects using cost weights.

The SNA (2008: 309) emphasizes that one of the pitfalls that should be avoided when measuring the volume of output of non-market individual services is that the measurement should not be restricted to reflect the inputs or activities of the unit producing the service. Activities constitute intermediate, not output variables. Only in the absence of such adequate output indicators, activities should be used as indicator (ibid.). For instance, with regard to the measurement of the health system it is recommended to count complete treatments, and not the individual activities that constitute the treatment. In the light of these remarks it is unclear why it is often thought that pupil-hours or pupil-years should constitute the principal output measure. As De Haan (2009: 12) notes correctly, hours or years taught is a process or input measure rather than an output measure. Therefore it might be useful here to reconsider the concept of education output.

Transfer of knowledge and skills
As a first step in determining education output one should ask what the principal goal of educational institutions is: Do they produce daycare services for children? Is it their task to give guidance to young people to become well informed citizens? Or is it their principal task to reproduce knowledge? In the paper at hand the

principal goal of educational institutions is determined as transferring skills and knowledge. Skills and knowledge can be taken very broadly as to include life skills and social skills next to knowledge transfer of subjects like history, economy and mathematics. This definition of education output is in no sense new (see Deveci 2009; Fraumeni et al. 2008; De Haan 2009; Struik 2008).

Following this definition of education output the ideal units of production to be counted should be units of transferred knowledge and skills during a certain amount of time, and not pupil-hours or pupil-years. As mentioned earlier, pupil-hours or years are a process or input measure rather than an output measure and should be avoided when possible. A unit of transferred knowledge can be thought of as a package of transferred knowledge and skills, for example 'a year of knowledge and skills transfer at a certain type of education', or 'the total amount of knowledge and skills transfer needed to obtain the diploma of a certain kind of education'.

The activities and processes that constitute a package of transferred knowledge and skills are intended to improve a pupil's (initial) state of knowledge and skills. The end state of knowledge and skills at the moment when primary school is finished, is assumed to be equal to the initial state of knowledge and skills of a pupil who enters secondary education. When the pupil leaves secondary education with a diploma s/he has improved on his/her initial level of skills and knowledge at least enough to exceed the minimal requirements to obtain the diploma. Each package of knowledge and skills transfer can thus be characterized by referring to an initial and an end state of transferred knowledge and skills. Educational products can be defined by varying the characterization of begin and end states.

In this paper academic or school years delineate the time dimension of the units. This seems a natural choice, because in general in secondary education the progress to a higher year is marked by some form of evaluation of the progress of the pupil, which defines whether or not a pupil exceeded the minimal demands for starting at the next level. Taking a longer period, for example the number of years that are needed to obtain a diploma would create a rather long time lag between the start of the production process and the end of it. It would take a long time before quality improvements would be reflected in an increase of number of graduates. And none of the teaching to pupils who do not graduate counts as output. Taking a period (much) smaller than a year, such as hours of knowledge and skill transfer would make it practically impossible to determine what information and skills are conveyed.

In general it is advisable to use the most refined product differentiation for which all relevant data is available. In this way one implicitly compensates for changes in the compositional mix of pupils, their characteristics and those of the packages of knowledge and skills transferred to them. The topic of weighting and aggregating the different homogenized categories of units of knowledge into a single volume indicator will be discussed in section 4.

Threshold
In the method proposed here, a unit of knowledge and skills transfer should only be counted as production when it is successfully transferred. A unit of knowledge and skills is then successfully transferred when some threshold, which can be seen as the minimum of skills and knowledge necessary to transfer to the next class, is exceeded. It is assumed here that it does not matter to what degree the pupil exceeds the threshold. Grades or exam results are not taken into account. A pupil who does not exceed the threshold – that is, a pupil who has to repeat a year or who

drops out from secondary education all together – does not give rise to any production in principle. An analogy with health can make this choice clear.

Both the OECD (2009: 71) and the Eurostat (2001: 117) handbooks propose to count complete treatments when it comes to health, and not the individual activities that comprise such treatments. The rationale behind this is that activity indicators often do not lead to reasonable productivity numbers (Eurostat 2001). For instance, when the number of operations declines drastically, due to the availability of newly developed medicines, taking the number of operations as an indicator would imply an unjustified decrease of output. Counting complete treatments, whether they are treatments by operation or by medicine would lead to the right output. Moreover, it becomes clear from this example that the complete treatment can be defined only by recourse to the health state of the patient. One can only define two different sets of activities as a complete treatment for the same illness by referring to a certain begin state (diagnosis) and some end state of a patient's health, at the beginning and at the end of the treatment. Once it has been established what generally constitutes a complete treatment one can choose whether to count failed treatments as production or not. Taking only successful treatments into account has the advantage that the volume indicator shows an increase (/decrease) when the success rate increases (/decreases).

The same applies to education. The begin state and end state of knowledge and skills on the part of the pupil play a role in determining the education product that s/he receives. After one has established what generally constitutes a particular educational product, one can choose again whether to count pupils who failed to pass their exams or other requirements as production or not. Because motivation, talent, pupil time, and parent time all play an important role in the success of the pupil one might be hesitating in attributing success or failure to the education service itself. Yet thresholds apply to other sectors of economy as well. It is often assumed that a not functioning car, mobile phone or microchip at the end of the production line should not be counted as a unit of production. Moreover, taking only the successful transfer of knowledge and skills as output has the advantage, that changes of success rates become visible in the volume indicator. If a school manages to increase the success rate of their pupils, this will be reflected as an increase in production, whereas this otherwise would not be the case.

In the method proposed here, a school only then delivers a unit of production when it succeeds in realizing the (minimal) change in knowledge and skills aimed for. This means that emphasis is put on the responsibility of schools to ensure their pupils successfully complete their school year.

Attainment

Education output is defined as the amount of transferred knowledge and skills. Production units are characterized as units or packages of successfully transferred knowledge and skills. Ideally the transferred amount of knowledge and skills to be measured precisely equals the difference in level of knowledge and skills between the beginning and end of the school year on the part of the pupil, and only in so far this difference is caused by the education service. Because it is impossible to measure this for each pupil individually, as a second best solution it is assumed that the amount of transferred knowledge and skills in some school year equals the difference between minimal end state of knowledge and skills (threshold) and the required initial state of knowledge and skills: For class *t* at least class *t-1* level knowledge and skills are required, and the threshold for success is passing the exams of class *t*. The difference in level of knowledge and skills between the two is equal to the required amount of knowledge and skills transfer in class *t*. Further it is assumed that the amount of knowledge and skills that someone learns outside of

school will be more or less proportional to that what someone learns at school during the year.

Let $k_{i,t}$ denote an average package of knowledge and skills transfer for type of education i in year t. The 'size' of such a yearly package of knowledge transfer expresses the *attainment rate* for the type of education at hand. Let $q_{i,t}$ from now on denote the number of successfully completed pupil-years for type of education $i \in I$ in year t. The total amount of transferred knowledge and skills – the output – for a particular type of education i in year t, $K_{i,t}$, can now be written as:

$$(2) \qquad \mathrm{K}_{i,t} = k_{i,t} \cdot q_{i,t}$$

In the following it is assumed that for all classes of a certain type of education i the increase of knowledge and skills, which can be ascribed to the educational intervention, is the same: the gain of knowledge and skills is linear throughout secondary education.

In the next section the outlined concept of education output will be used to construct the weights needed to aggregate different types of education, and to formulate an output volume index formula.

4. Weighting transferred knowledge and skills

If we want to express the change in output volume for different types of education the OECD (2009) advocates using cost weights. Following this recommendation a Laspeyres volume index that describes the change in outputs between year t and year *t-1* would look like:

$$(3) \qquad V_L = \frac{\sum_i \widetilde{p}_{i,t-1} \cdot K_{i,t}}{\sum_i \widetilde{p}_{i,t-1} \cdot K_{i,t-1}},$$

in which $K_{i,t}$ represents the total amount of knowledge and skills transfer, and $\widetilde{p}_{i,t}$ the accompanying cost weights for each type of education $i \in I$ in year t. The obvious choice for $\widetilde{p}_{i,t}$ would be the costs per unit of transferred knowledge and skills when aggregating different amounts of knowledge and skills. These weights can be determined by dividing the total production value of a certain type of education i in year t by the total number of transferred units of knowledge:

$$(4) \qquad \widetilde{p}_{i,t} = \frac{p_{i,t} \cdot q_{i,t}}{k_{i,t} \cdot q_{i,t}} = \frac{p_{i,t}}{k_{i,t}},$$

in which $p_{i,t}$ represents the costs per pupil, $k_{i,t}$ the attainment rate, and $q_{i,t}$ the number of pupils for each type of education $i \in I$ in year t. Notice however, that this interpretation of $\widetilde{p}_{i,t}$ would result in expression (1) in cases when the amounts of knowledge transfer $k_{i,t}$ do not alter between given years t and *t-1*.

A simple example shows that aggregating changes in transferred units of knowledge for different types of education with weights $\widetilde{p}_{i,t} = p_{i,t} / k_{i,t}$ yields unexpected results (see table 1 and 2 below). In the example the change of output

(based on pupil numbers and attainment rates) of three existing types of secondary education from the Netherlands (vmbo-gt, havo and vwo, see section 5) is determined using index formula (3) and weights following (4). For the sake of simplicity it is assumed that the attainment rates used here express absolute amounts of knowledge and skills transfer and that these can be added up together: The amount of transferred knowledge and skills of three years of havo (3 x 1,33) is assumed to equal four years of vmbo-gt (4 x 1,00). The results are compared with the change in total amount of transferred knowledge and skills:

Table 1 and 2: Weighting with costs per unit of knowledge transfer results in unexpected values for the volume index.

	costs year 1 (p)	attainment (k)		number of pupils (q)		output (q x k)	
		year 1	year 2	year 1	year 2	year 1	year 2
vwo	100	1,67	1,67	100	50	167	83
havo	100	1,33	1,33	100	50	133	67
vmbo-gt	100	1,00	1,00	100	200	100	200
total				300	300	400	350
index (using (4))	1,00						
growth of output	0,88						

	costs year 1 (p)	attainment (k)		number of pupils (q)		output (q x k)	
		year 1	year 2	year 1	year 2	year 1	year 2
vwo	100	1,67	1,67	100	200	167	333
havo	100	1,33	1,33	100	50	133	67
vmbo-gt	100	1,00	1,00	100	50	100	50
total				300	300	400	450
index (using (4))	1,00						
growth of output	1,13						

From the example it becomes clear that the change in output volume is overestimated in the first table and underestimated in the second table. Overestimated in the first table, because more people attend a type of education with a lower attainment rate; underestimated in the second table, because more pupils attend a type of education with a higher attainment rate in the year 2. Weighting transferred units of knowledge with unit costs clearly does not lead to the correct index for change in the volume of outputs.

Whereas it is often assumed that for non-market services, such as education and health, 'the relative valuation by the consumer equals the relative costs incurred by producers' (OECD 2009: 24), this seems not to be the case here. The catch is that fixed costs are in place for three different types of education, each characterized by different attainment rates. Therefore the choice for the consumer is not between a cheaper or a more expensive product, but the choice is between different products, with different benefits for the same amount of money.

Attainment rates, not costs, determine the valuation by the consumer. A benefit maximizing pupil who can choose to reach havo-3 level of education in three years by attending havo, or in four years by attending four years of vmbo-gt, will take the shortest route – of course assuming her / his talents allow so. First, this is more time-efficient. Second, suppose education would not be subsidized taking the faster route would be cheaper for the consumer. Equally, in general a government that has to give account to its tax payers will stimulate students to exploit their talents,[3] while taking the most cost-efficient route through the school

[3] In Europe this goal is explicitly laid down in the Lisbon agreement from 2002.

system. Because costs per pupil per year are fixed, cost- and time-efficiency are dependent on attainment rates.

More generally, costs per unit of transferred knowledge are determined by both costs per year of attendance ($p_{i,t}$) and attainment rates ($k_{i,t}$). Now suppose that costs per pupil-year $p_{i,t}$ do not differ between the given types of education. In such a situation a consumer can minimize the costs per unit of knowledge transfer by attending a type of education with a higher attainment rate. The marginal benefits for the consumer alter with the difference in attainment rate of the type of education s/he chooses. Put otherwise, the marginal benefits assigned to a certain kind of education $i \in I$ in comparison to a reference or base type of education $0 \in I$ can be expressed as the ratio of their attainment rates $k_{i,t} / k_{0,t}$. Therefore, if we want to use weights that express the relative valuation by the consumer as the OECD (2009) recommends, we should not use $p_{i,t} / k_{i,t}$ as weights, but $p_{i,t} / k_{i,t}$ multiplied by $k_{i,t} / k_{0,t}$. This means that the cost weights as expressed in (4) need to be transformed into:

$$(5) \qquad \tilde{p}_{i,t} = \frac{p_{i,t}}{k_{i,t}} \cdot \frac{k_{i,t}}{k_{0,t}} = p_{i,t} / k_{0,t}$$

Using (3) and (5) our volume index formula now becomes:

$$(6) \qquad V_L = \frac{\sum_i (p_{i,t-1} / k_{0,t-1}) \cdot q_{i,t} \cdot k_{i,t}}{\sum_i (p_{i,t-1} / k_{0,t-1}) \cdot q_{i,t-1} \cdot k_{i,t-1}}$$

If we introduce a $\lambda_{i,t}$ for the ratio of attainment rates $k_{i,t} / k_{0,t}$ we can rewrite (6) as follows:

$$(7) \qquad V_L = \frac{\sum_i p_{i,t-1} \cdot q_{i,t} \cdot \lambda_{i,t}}{\sum_i p_{i,t-1} \cdot q_{i,t-1} \cdot \lambda_{i,t-1}} \cdot \left[\frac{k_{0,t}}{k_{0,t-1}} \right]$$

The introduction of $\lambda_{i,t}$, expressing the *attainment ratio* of a type of education with regard to some base type of education, has the advantage that no information regarding the absolute amount of knowledge and skills transfer is needed to calculate the ratio in the left part of the expression. It is sufficient to use the information about possibilities to transfer between levels of education to determine all lambdas. Because all attainment ratios $\lambda_{i,t}$ are defined relative to the same base type of attainment rate $k_{0,t}$, they can be said to form a scale. Therefore the output volume method presented in this paper may be called a 'scale adjusted' output method.

The right hand term of equation (7) expresses possible changes in the attainment rate of the reference type of education between year *t* and *t-1*. If this term changes, all lambdas change consecutively, which means that the possibilities to transfer between types of education alter. In practice this is not very likely to happen anywhere in the near future. For sake of simplicity in what follows it is stipulated that $k_{0,t} / k_{0,t-1} = 1$. The method presented in this paper allows though – at least to some extent –, that changes in the amount of knowledge and skills transferred may occur at one or more types of education, without the necessity to

change all attainment ratios accordingly. These changes may be interpreted as changes in the quality of education and are subject of investigation in appendix A. It will be shown that changes in the right hand term of expression (7) can be seen as a special kind of quality adjustment.

Earlier Struik (2008) and Struik & Van Mulligen (2008) explored taking the transfer of knowledge and skills as output concept. Although their contributions were valuable for the development of the method proposed in this paper, there is a major disadvantage sticking to their proposal. Their experimental 'vovo-method' uses attainment ratios similar to the ones used here to weigh between types of secondary education, but cost information is not used at all to determine weights. All types of education discussed in the paper were seen as the same product, only differing in attainment ratio. This is not totally appropriate. A short example can make this clear.

In section 5 a distinction is made between a type of education called 'vmbo-gt' and 'lwoo-gt'. Both types have the same level. The difference between them is that lwoo-gt is apt for pupils who have more learning difficulties, such as concentration problems or behavioral disturbances, and who require more attention. With a class size of 16 at maximum, the average lwoo-gt class size is much smaller than that of a vmbo-gt class. Outside class teaching, more personal attention is given to the lwoo-pupils. Not surprisingly, these differences are reflected in the costs per pupil per year. In 2007 costs for a lwoo-gt pupil were estimated at 12.644 euros per year, while for vmbo-gt pupils at 7.166 euros per year.[4]

It becomes easily visible that renouncing to take costs into account (in the case of the vovo-volume index) leads to problems here. Suppose, for example, there is a sudden change in the characteristics of the pupil population, due to which relatively more pupils will attend lwoo-gt. Then the value index of secondary education rises more quickly than the volume index based solely on the transfer of knowledge and skills. In this case costs per pupil seem to have risen, while in fact not costs, but only the composition of the pupil population has changed (see table 3).

Table 3: Weighting without costs results in an unexpected value for the price index.

	costs year 1 (p)	attainment (k)		number of pupils (q)		output (q x k)	
		year 1	year 2	year 1	year 2	year 1	year 2
lwoo-gt	150	1,00	1,00	100	150	100	150
vmbo-gt	100	1,00	1,00	100	50	100	50
value index	1,10						
vovo volume index	1,00						
scale adjusted volume index	1,10						
vovo price index	1,10						
scale adjusted price index	1,00						

In sum, if the assumption suggested by the OECD (2008: 24) is indeed justified, that – for types of education with similar attainment ratios – 'on average, the relative valuation by the consumer equals the relative cost incurred by producers', then it makes sense to treat types of education with different price tags as different products. When types of education differ in attainment ratios, this

---

[4] It is up to the reader to decide to what extent costs of education services can be seen as an indicator for teaching efforts.

should be reflected as well in the weights used to construct an overall volume index.

5. Estimates of educational output for Dutch secondary education

The Dutch system of secondary education
Secondary education in the Netherlands can be split up in 'secondary general education' and 'secondary vocational education'. Within the first category there are around 700 secondary schools, almost all of which are publicly run. Secondary 'general' education encompasses schools providing pre-university education (vwo; 6 years, age 12-18), senior general secondary education (havo; 5 years; age 12-17), pre-vocational secondary education (vmbo; 4 years, age 12-16) and practical training (pro; age 12-18). The vmbo type of education can be split further in the theoretically oriented junior general secondary education (mavo / vmbo-gt) and the practically oriented pre vocational education (vmbo-bk). All four types of secondary education are for children aged twelve and over and all begin with a period of basic secondary education. Pupils who are unable to obtain a vmbo qualification, even with long-term extra help, can receive practical training. This type of education is not divided in classes of consecutively higher level, but still prepares pupils for entering the labor market. Vmbo with additional long-term extra help is known as 'lwoo'.

Pupils who finish vwo can go to university level education. Pupils with a havo diploma may go to higher professional education, or they can choose to start in vwo-5 to obtain the vwo-diploma in two additional years. The different vmbo (or lwoo) diplomas give access to different levels of (upper) secondary vocational education (mbo; ½ - 4 years, depending on level). The vmbo-gt type gives access to more higher levels of mbo than vmbo-bk. A vmbo-gt diploma gives a pupil also the opportunity to start in havo-4. The havo diploma can then be obtained in two added years. The legal possibilities to transfer between type and level of education will play an important role in this paper in determining the average relative amounts of knowledge and skills transfer between vmbo-gt, havo and vwo.

'Secondary vocational education' is provided for at around 40 regional training centers in the Netherlands. They provide vocational education in four sectors: Engineering & Technology, Economics, Health & Social Care, and Agriculture, Natural Environment & Food Technology. Vocational education offers participants from the age of 16 a choice of 700 vocational courses, four training levels and two different routes in which courses can be followed. There is a full-time college-based route that includes work placements and there is the part-time work-based route, which combines part-time education with an apprenticeship in a company.

A final part of secondary education is adult education (vavo), one part of which consists of part time secondary general education for adults. Sometimes this type of secondary education is referred to as 'second chance secondary education'. Another part of vavo offers adults courses focusing on social sufficiency, literacy, citizenship, and the mastery of Dutch as a second language. This last part of vavo is not taken further into account in this paper.

Data
In this section, results obtained by the application of the newly developed method on data from the Dutch system of secondary education are summarized. The data

on pupil numbers come from the so-called 'education-matrix'. This register contains yearly pupil and student counts for all types of publicly funded education.[5] Additionally it contains information about all transfers of pupils and students that took place between types and levels of education from the one year to the other. Data is available from school year 1997/'98 to 2007/'08 (preliminary) and is updated annually. Within this period there are 1495 different transitions registered that pupils took from one type and level of education to another. Most important transfers are those to the next class of the same type of education, or finishing secondary education with a diploma. A relative small, but not insignificant amount of pupils stays down a class or steps out of publicly funded education altogether. Furthermore, small groups of pupils are registered taking all kinds of imaginable transitions, like transferring from class 1 to class 3, or transferring between sectors (from agrarian to administrative education for instance), or transitioning to a lower class and/or level.

All transfers in and between lwoo-gt, vmbo-gt, havo, and vwo can be assigned an attainment ratio which expresses the relative amount of transferred knowledge and skills. First the attainment ratio for the type of education at hand is determined. Next it is established how much knowledge and skills is transferred during the school year to the pupils in consideration. If a pupil has to repeat a class no production is delivered. In the method presented it is assumed that there was no transfer of knowledge of skills in such cases.[6] The same holds for pupils who leave secondary general education altogether without a diploma. In contrast, when the pupil finishes the class successfully the next year, the relative amount of knowledge transfer ($\lambda_{i,t}$) is 1 on lwoo-gt/vmbo-gt (base type of education), 4/3 on havo, or 5/3 at vwo. This follows from the transfer possibilities inherent in the Dutch system of secondary education. A vmbo-gt or lwoo-gt diploma (4 years) gives access to havo-4, and thus equals three successful completed years of havo. Equally, five years of havo equals four years of vwo. A pupil who transfers between vmbo-4 and havo-4 has learned 1 during the fourth year of his/her education. S/he started at level 3 x 1 = 3, and needs 3 x 4/3 to start at havo-4 level. The difference thus is 4 - 3 = 1. A pupil who transfers from 3 vwo to 4 havo, started at 2 x 5/3 and ends at level 3 x 4/3. S/he must have learned at least 2/3 during his/her third year of studies.

It is not possible to ascribe attainment ratios to the other types of secondary education that can be mutually compared. For special secondary education no detailed information exists. Only in the small numbers of cases where diplomas for vmbo-gt havo or vwo are registered for special education an attainment ratio is added. For practical training no attainment ratios are possible, because this type of education is not characterized by consecutively higher level classes. When pupils leave special education without diploma this is not seen as a loss of production. The principal goal for most pupils on schools for special education or practical training is not an increase of knowledge and vocational skills, but an increase of social and life skills.

---

[5] Privately funded secondary education is very small in pupil numbers. In 2007 1.909 pupils aged under 18 attended privately funded schools specialized in preparing pupils to pass the exams needed to obtain a diploma, the so-called 'zelfstandige exameninstellingen' (Inspectie van het Onderwijs 2009: 17). Additionally, in 2007 903 pupils were registered at 34 privately funded schools for primary or secondary education, so called 'B3-institutions' (Van der Meer & Van der Ploeg 2007).

[6] See section 3.

Pupils who finished vmbo or did some years of practical training can transfer to mbo. It is not possible to ascribe attainment ratios to mbo types of education either. There are over 700 study programs to be found on mbo which are not comparable with regard to contents. For instance, how should one compare the knowledge and skills transfer from an electrician to that of a hairdresser or a laboratory assistant? Further within mbo there exist many levels of education and lengths of study programs vary considerably. There is no detailed information available about repeaters in mbo. In the method proposed all student-years, successfully finished or not, count as production. For the group of students who disappear from secondary education without diploma the choice is made to count half of them as successful production and half of them as failed production. This choice is made because one may expect that not all dropouts are pupils who did not successfully complete their classes. Some of them may have decided to quit school and start with a paid job. A large number of students on mbo is over 18 years old,[7] and therefore no longer obliged to go to school. A sensitivity analysis is made to get hold of the consequences of this choice (see Appendix C).

The pupil numbers, corrected for repeaters, dropouts, and attainment ratios are then categorized into eight types of secondary education. Additionally, pupil-years for part time educations are counted with a weight of 0,5. The distinction into eight types of education is based on data available at Statistics Netherlands on yearly expenditures per type of education. Yearly expenditures are largely based on the annual reports of the board of commissions of the schools for secondary education. While funding of schools by the government is based on the total number of pupils, the yearly expenditures per type of education *per pupil* used here are determined by dividing the total expenditures by the number of students who successfully finish the school year. Following the Eurostat handbook (2001:115), in case of vocational training at (v)mbo level only government expenditures for teaching are taken into account and not the (mainly by private companies made) expenditures for 'on the job' training.

Results

Figure 1 shows the chained scale adjusted volume index for the period 1998-2007. It is interesting to look closer at the effects of the introduction of attainment ratios, because this constitutes the most important difference with volume indicators based on pupil-years. For that reason in figure 1 the comparison is made between the chained index of pupil-years, the chained index of pupil-years, corrected for dropouts and non-promoted pupils, and the chained scale adjusted index. Remember that the scale adjusted index is always already adjusted for dropouts and non-promoted students.

Figure 1 shows that the scale adjusted index and the index corrected for dropouts and non-promoted pupils do not differ much from each other. The differences become larger when plain pupil-years are used. Between 1999 and 2002 the index for pupil-years lays above the other two indices. This results from a growing group of dropouts and non-promoted pupils, which pulls the index corrected for non-promoted pupils downwards. From 2002 onwards this group is declining again, which translates in a faster rise of the indices that are corrected for non-promotion and dropouts, than for the index that plainly takes pupil-years as volume indicator. For the total period 1998-2007 it means that the corrected index grows with an average of 0,3 percent per year faster than the not corrected index.

---

[7] In 2007/'08 74 percent of all mbo students was 18 or older.

Figure 1: three chained volume indices for secondary education.
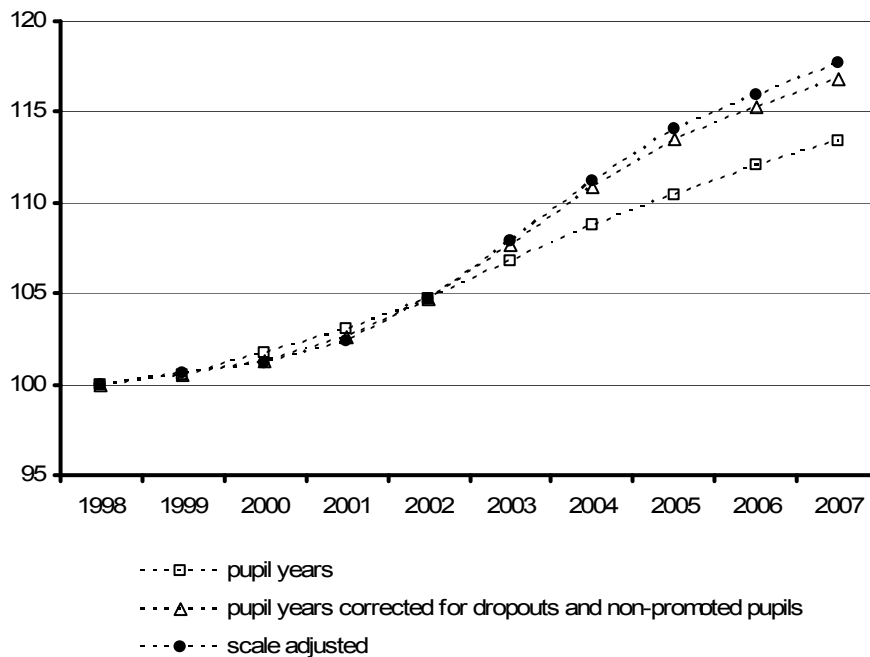
When the index corrected for dropouts and non-promoted pupils is compared with the one which is also adjusted for attainment, there is a positive effect on average annual output of 0,1 percent that results from a higher average transfer of skills and knowledge (attainment ratio) per pupil. This means that proportionately more pupils successfully move through havo and vwo.

In figure 2 the effects of the introduction of the attainment ratios are made more visible. The chained index labeled 'pupils years' represents the index for all the produced pupil-years in secondary education between 1998 and 2007 to which an attainment ratio could be assigned. This means that only pupils who successfully transferred in or between lwoo-gt, vmbo-gt, havo, and vwo are counted. When adjusted for attainment ratios, the second, scale adjusted index is obtained. Figure 2 shows more clearly that between 1998 and 2001 nothing happened with regard to the distribution of pupils over types of education with different attainment ratios. In contrast, from 2001 onwards the average pupil-year is characterized by an ever increasing attainment ratio.

Figure 2: The effects of the introduction of attainment ratios on the volume indicator.

··△·· pupil years ··●·· scale adjusted
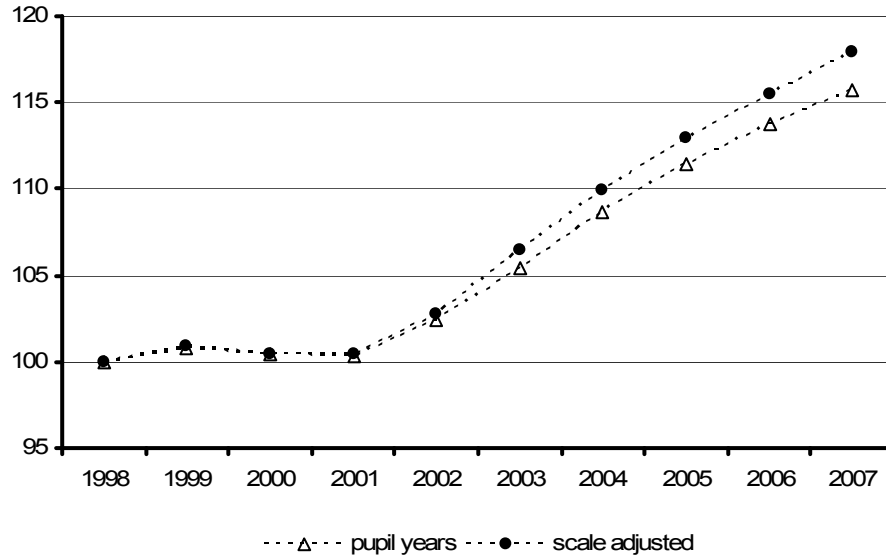
This increase of the average amount of transferred knowledge and skills per pupil per year goes along with a downward pull on the price index. At first sight this may seem awkward, because the distribution of pupils over different types of education itself should not have an influence on the price index. After all, that is the reason why unit costs are used as weights to aggregate different types of education together. Differences in the distribution of pupils over the types of education should be registered as a volume effect, not as a price effect.

But, taking the price index into closer consideration, it becomes clear that the introduction of attainment ratios may cause the price index to decrease when the volume index increases, even when the costs per year for the different types of education do not alter. First notice that the value index (*W*) of secondary education is determined by the change in registered number of pupils, weighted by their accompanying unit costs:

$$(8) \qquad W = \frac{\sum_{i \in I} p_{i,t} \cdot q_{i,t}}{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t-1}}$$

The price index that can be derived from expressions (7) and (8) shows that the changes in the price index may occur as a result of a change in the average amount of knowledge and skills transferred only:

$$(9) \qquad P_P = \frac{\sum_{i \in I} p_{i,t} \cdot q_{i,t}}{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t-1}} \cdot \left[\frac{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t} \cdot \lambda_{i,t}}{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t-1} \cdot \lambda_{i,t-1}} \right]^{-1}$$

The change in the average amount of knowledge and skills transferred per pupil per year is discounted for in the bracketed and inverted term of expression (9). If this average amount grows, *ceteris paribus*, the numerator of the right hand term will increase, and as a result the Paasche price index will decrease. This

18

characteristic of the price index can be shown even more clearly by turning to a situation in which just a single type of education j would exist. Then the price index would express the change in costs multiplied by the inverse of the change in knowledge and skills transfer between the years t and $t-1$:

$$(10) \qquad P^j = \frac{p_t}{p_{t-1}} \cdot \left[\frac{\lambda_t}{\lambda_{t-1}} \right]^{-1}$$

The property that the price index can alter even though for all types of education the unit costs stay equal over time is distinctive for the scale adjusted method outlined in this paper. It results from the observation made earlier, that, when taking the amount of transferred knowledge and skills transferred as output concept, the relative valuation of education services cannot be expressed exclusively in monetary terms, when it is evident that changing expenditures do not (fully) reflect changes in the amount of transferred knowledge and skills. The transformation of weights by means of attainment ratios (see section 4) can be interpreted as an attempt to restore the balance. Simultaneously, when the total output of educational services alters, while the total value (expressed in monetary terms as in expression (8)) stays equal, the average price must have changed as well. For example, when proportionally more people attend havo and vwo in relation to vmbo-gt, *ceteris paribus*, more output is delivered for the same amount of money. The price index should be pulled downwards as a result. Exactly this holds for Dutch secondary education since 2001.
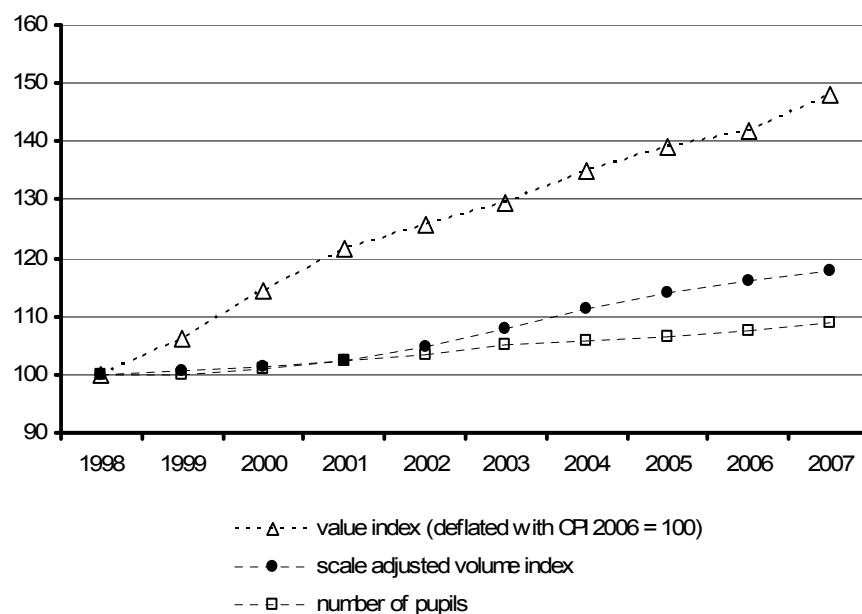
The total expenditures of secondary education services have risen with almost 80 percent between 1998 and 2007. Corrected for inflation using the CPI (2006 = 100), this becomes 49 percent. The volume of secondary education has only risen with 17,3 percent, meaning that prices have risen faster than volume. This implies that secondary education services have become more expensive between 1998 and 2007. The chained price index would be 1,0 percent higher in 2007 when no adjustment for attainment was carried out, but only for dropouts and non-promoted pupils. When only pupil-years were counted and weighted with costs, the CPI deflated prices would seem to have increased 4,3 percent faster compared to the prices following the scale adjusted method.

Figure 3: Value (deflated with CPI 2006 = 100), volume, and price index for secondary education, based on the scale adjusted method. Small differences with earlier published official statistics of Statistics Netherlands may occur, due to a different categorization of special secondary education, practical training and lwoo types of education.

- – –△– – – value index (deflated with CPI 2006 = 100)
- – –◇– – – price index
- · · –●– · · volume index (scale adjusted)
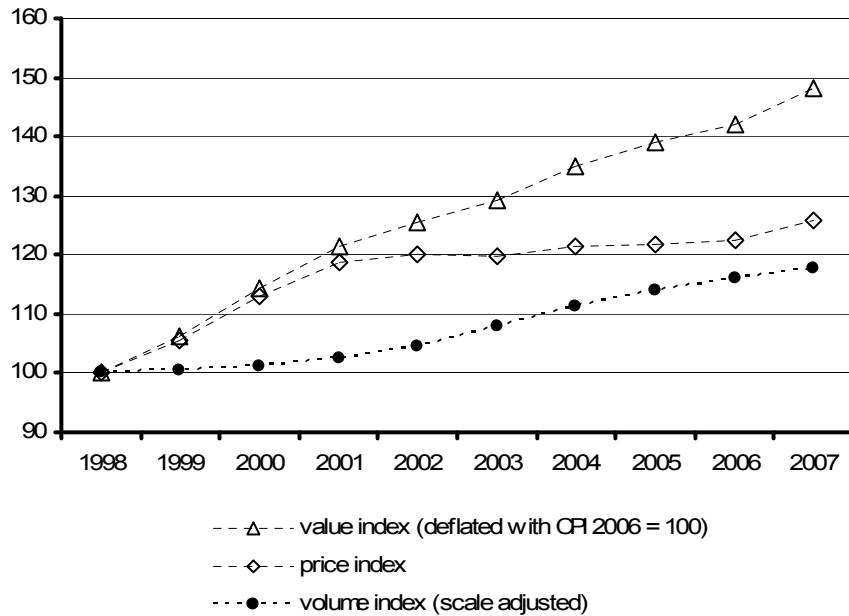
The value index of secondary education rises rather quickly. Demographic developments only played a modest role in this. Figure 4 shows the growth of the population of pupils attending secondary education between 1998 and 2007 compared with the scale adjusted volume index and the value index. Compared to the 49 percent rise of expenditures, the rise in pupil numbers of just about 9 percent is relatively small.

Figure 4: Chained indices for number of pupils, the value of secondary education and scale adjusted volume.

- · · –△– · · value index (deflated with CPI 2006 = 100)
- – –●– – scale adjusted volume index
- – –□– – number of pupils

An important factor that can explain the rising expenditures in secondary education are the investments made in buildings and inventory such as furniture, computers, and teaching materials (Elbers & Kaashoek 2010). Especially in the period 1998-2001 this made the expenditures per pupil to increase quickly, as the price index shows.

Another effect that played a role in the increasing expenditures of education can be called the 'absolute participation effect' (Elbers & Kaashoek 2010), which expresses changes in the percentage of age groups in the total population who attend secondary education. Especially the relative size of the group of 17 to 21 year olds in secondary vocational education (mbo) has increased. Simultaneously, the numbers of dropouts for secondary vocational education dropped significantly from approximately 73 thousand in 2001 to 58 thousand in 2007. Furthermore Elbers and Kaashoek notice an effect in the employee/pupil ratio. According to them this ratio has been rising in recent years due to growing pupil numbers who attend lwoo-types of education and due to efforts to lessen the number of dropouts without a 'starting qualification' from secondary vocational education. Indeed the Dutch Ministry of Education, Culture and Science has committed itself as a part of the Lisbon agreements in 2002 and 2003 to diminish the number of dropouts and to raise the percentage of graduates in higher secondary education. Much efforts are made to reduce the number of dropouts without a starting qualification.[8] One way of accomplishing these goals is by stimulating the flow through secondary education by enabling pupils to switch more easily between different types of general or vocational education. For example, from August 1st 2010 onwards it becomes possible in some circumstances for talented pupils who finished vmbo-gt class 3 to transfer to havo class 4.[9]

6. Final remarks and next steps

In this exploratory paper a new method to measure the output of government-provided secondary education in the Netherlands is proposed. As opposed to the current method in the Dutch national accounts for secondary education, in which outputs are defined as pupil-years, outputs are defined as transferred knowledge and skills. While the absolute amounts of knowledge and skills transferred may be difficult to measure directly, it has been argued in this paper that the relative differences between transfer of knowledge and skills between various types of secondary education can be used as a second best solution. The legal transfer possibilities between different types of education offer a solid basis to determine the relative differences in transfer of knowledge and skills.

Taking transferred knowledge and skills as an output measure means that traditional cost weights no longer suffice for aggregating different types of education in cases where costs can not be assumed reflecting the amount of

---

[8] Recent numbers from Statistics Netherlands show that the number of people between 15-25 years who do not attend any type of education and do not possess a starting qualification dropped from 15 percent in 2001 To 9 percent in 2009. See http://www.cbs.nl/nl-NL/menu/themas/arbeid-sociale-zekerheid/publicaties/artikelen/archief/2010/2010-3093-wm.htm.

[9] Because this holds only for a small group of talented pupils, no adjustments have to be made with regard to the method proposed here. The attainment ratio for this group of pupils who receive the education product 'from vmbo-gt-3 to havo-4' would be 2 (starting level havo-4 minus starting level vmbo-gt-3).

transferred knowledge and skills. A transformation of costs weights is needed to restore the balance between costs and amount of transferred knowledge and skills. For the Dutch situation the attainment ratios between certain types of secondary education could be used. It does not suffice to use attainment ratios only, as was made clear with regard to the vovo-method (see Struik 2008). Costs still play an important role in determining the weights, because they are particularly apt to express the input-efforts that are needed to deliver a unit of production.

For (parts of) educational systems in which there exist no transfer possibilities between different types of education, the same framework could be applied when there is data available on the number of courses that pupils attend (for example a pupil who finishes 7 courses a year learns 7/6 more than one who finishes 6 courses a year). The only requirement is that there is a solid scale needed to work from.

The main disadvantage of the framework is, that it rests on the assumption that different trajectories through secondary education which in the end result in the same diploma (for example a vwo diploma), can be characterized by the same amount of knowledge and skills transfer. After all, one still may object that the skills and knowledge transferred are the same. With regard to the threshold model of outputs this is to say that one objects that reaching the threshold to acquire a diploma is not a well-suited indicator for knowledge and skills transferred.

The advantages of our framework outweigh this disadvantage though. Firstly, to see educational outputs as knowledge and skills transferred and not as pupil-years lies conceptually closer to the core activity of schools. After all, the most important goal of a school is not to provide pupils with some daily occupation, but to transfer knowledge and skills to them. Secondly, if, *ceteris paribus*, more knowledge and skills are transferred by types of education that can be characterized by attainment ratios, the scale adjusted volume indicator will show a rise. The method currently used in the national accounts does not show such a rise, because it fails to take into account the amount of knowledge and skills transferred all together. If the number of pupil-years does not alter for these types of education, neither does the volume index. Thirdly, the scale adjusted method does not necessarily require much more data than currently used in the Dutch national accounts. Of course, more detailed cost information and pupil counts would lead to a better approximation of output volume. Such more detailed information is already available at Statistics Netherlands.

Finally, the scale adjusted method proposed here removes the biggest disadvantage of the experimental vovo-method proposed by Struik (2008), by taking both costs and attainment ratios into account when determining the weights needed to construct an overall volume index for secondary education.

When data is available, the framework presented here can be refined by introducing 'quality adjustments' in the form of a more extended product differentiation. In appendix B the example of an extended product differentiation in pupils with and without dyslexia is given. When information that may hold clues about the quality of education is only partial, sensitivity analysis can be used to get hold of the possible effects of such quality indicators (see Appendix C).

Next steps
Further research should be carried out to discern the relevant pupil characteristics that may influence the expenditures of the average school. While much information about pupil characteristics in the Netherlands can be obtained via the Social Statistics Database (SSB), it is not certain though whether all the relevant cost information can be obtained.

Future research could aim for discerning transfer possibilities within vocational secondary education. The introduction of the 'education number'[10] which allows to track more closely the paths pupils take through secondary education could prove to be a help to do this. Because vocational secondary education constitutes a big part of secondary education, the method would benefit greatly.

---

[10] In Dutch 'onderwijsnummer'.

References

Atkinson, T. (2005). *Atkinson Review: Final report. Measurement of Government Output and Productivity for the National Accounts*. Palgrave Macmillan: Basingstoke.

Ayoubkhani, D. (2009). *Methods Development: Decomposing Quality Change in the Output of Publicly Funded School Education*. UK Centre for the Measurement of Government Activity: Newport.

Berg, P. van den (2007). *Onderzoek ontwikkeling van een outputindicator voor het tertiair onderwijs*. Mimeo. Statistics Netherlands: Voorburg.

Briggs, J. (2008). *Developing a quality-adjusted output measure for the Scottish education system*. Updated version of a paper originally developed for an OECD workshop on Measuring Education and Health Volume in June 2007.

Chessa, A. (2009). *Measures of health care volume with applications to the Dutch health sector*. Statistics Netherlands: The Hague.

Collesi, D., Guerrucci, D., Versace D. & Zannoni, S. (2007). *The use of class size and the Italian method*. Presented at the OECD Workshop on measuring Education and Health Volume, June 2007.

Deveci, N. (2009). *Productivity and Quality of the Public Sector*. Statistics Denmark: Copenhagen.

Douglas, J. (2006). *Measurement of Public Sector Output and Productivity*. Policy Perspectives Paper 06/09. New Zealand Treasury: Wellington.

Elbers, H. & Kaashoek, R. (2010). *A realistic look at educational expenditure*. Statistics Netherlands: The Hague. http://www.cbs.nl/en-gb/menu/themas/overheid-politiek/publicaties/artikelen/archief/ 2010/2010-educational-expenditures-pub.htm.

Eurostat (2001). *Handbook on price and volume measurement in national accounts*. Office for Official Publications of the European Communities: Luxembourg.

Fraumeni and Jorgenson (1992). 'The output of the Education Sector', in Zvi Griliches (ed.) *Output Measurement in the Service Sectors*. National Bureau of Economic Research Studies in Income and Wealth, Vol 56. University of Chicago Press: Chicago.

Haan, J. de (2009). *On the Measurement of Secondary Education Output in the National Accounts*. Statistics Netherlands: The Hague.

Inspectie van het Onderwijs (2009). *Zelfstandige exameninstellingen voortgezet onderwijs: leerlingenpopulatie en examenresultaten*. http://www.onderwijsinspectie.nl/binaries/content/assets/Actueel_publicaties/2009/ Zelfstandige+exameninstellingen+voortgezet+onderwijs.pdf

Konijn, P. & Kleima, F. (2000). *Volume Measurement of Education*. Mimeo. Statistics Nethedrlands: Voorburg.

Meer, M. van der & Ploeg, S. van der (2007). *Particuliere B3-scholen 2007-2008*. Regioplan: Amsterdam.

Murray, R. (2007). *Developing a quality-adjusted output measure for the Scottish Education System*. Presented at the OECD Workshop on measuring Education and Health Volume, June 2007.

Organisation for Economic Co-operation and Development (OECD) (2009). *Towards Measuring the Volume Output of Education and Health Services*. Statistics Directorate, Committee on Statistics, OECD: Paris.

Office for National Statistics (ONS) (2007). *Methods for public service productivity: quality adjusting school education output.* ONS, September 2007.

Rivkin, S. (2000). 'The Estimation of Productivity Change in Education.' Amherst College. April.

Smith, P. & Street, A. 'The measurement of non-market output in education and health', in *Economic & labour market review*, Vol 1, no.6, 2007.

SNA (2008). *System of National Accounts*. European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank.

Statistics New Zealand (2010). *Measuring government sector productivity in New Zealand: a feasibility study*. Wellington: Statistics New Zealand.

Stiglitz, J., Sen A. & Fitoussi, J. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress.

Struik, M. (2008). *A Volume Measure for Secondary Education Output* (in Dutch). Mimeo. Statistics Netherlands: The Hague.

Struik, M. & Van Mulligen, P.H. (2008). *Measuring Output of Secondary Education in the Netherlands*. Mimeo. Statistics Netherlands: The Hague.

Trewein, D. (2002). *Year Book Australia 2002*. Number 84. Australian Bureau of Statistics: Canberra.

Appendix A: 'Explicit' quality adjustments for secondary education

In paragraph 4 the question of quality change was addressed shortly. In this appendix the issue is discussed in somewhat more detail. It is not within the scope of this paper to discuss to full extent what 'quality improvement' of education might mean. Nevertheless, one might think of quality improvement induced by an increase in the added benefits for society or economy, for example due to the introduction of information and computer science in secondary education. Or one might think of the added value of technological or pedagogical progress which enables teachers to transfer knowledge and skills more easily to their weaker pupils, or which enables teachers to transfer knowledge and skills which are better adapted to the demands of society.

One must make a sharp analytical distinction between indicators that refer to the quality of inputs (technological improvements, teacher experience, class size) and those which refer to the quality of outputs and outcomes (grades, well being of pupils, or even future earnings). If one uses input quality indicators to adjust an output index for change in quality, one runs the risk of taking the effects of the input quality indicator twice into account, namely in cases when the output indicator itself is already co-determined by the input factor under scrutiny. Therefore, when an output volume method is used, one should only adjust outputs for quality using indicators that refer explicitly to the output (or outcomes) of the service addressed.

It is important to stress that the quality adjustments here are *independent from* the possibilities that exist within the system of secondary education to transfer between types of education. Put otherwise, the quality adjustments discussed here have no influence on attainment ratios $\lambda_{i,t}$. Whether a pupil finished a qualitative good or qualitative bad havo education, s/he still has the (legal) possibility to transfer to the fifth class of vwo.[11]

The quality change for some type of education i between year t and $t-1$ can be described by means of a function $f_i(x_{t,t-1})$ of the form:

(A1) $\qquad f_i(x_{t,t-1}) = (\mu_{i,t} / \mu_{i,t-1}) \cdot \varepsilon$

One could propose to use data on class size, average grades or exam scores for $\mu_{i,t}$, if one could show that those variables have an effect on the transfer of knowledge and skills. Still it remains hard to determine the precise effect of such quality indicators on the total volume index. This is represented by the unknown parameter ε.

Note that if $f_i(x_{t,t-1})$ is the same for all types of secondary education, the function describes an overall quality change within secondary education. In practice it will be difficult to find the correct value for parameter ε though. Examples of similar kinds of (explicit) quality adjustments can be found in Ayoubkhani (2008), Collessi et al. (2007), and Deveci (2009).

---

[11] This may be interpreted as a further motivation to stipulate that right hand term of expression (7), namely $k_{0,t} / k_{0,t-1}$, equals 1. (See also end of section 4.)

Appendix B: Quality adjustments for made-to-measure pupil-years[12]

In this section it is illustrated how quality aspects that concern the characteristics of pupils and the educational intervention offered to them can be incorporated into the volume measurement framework. For example a pupil can have dyslexia, and therefore be in need for some extra support in language and grammar courses. Another example could be that a pupil has a deficit in knowledge of the language spoken at school. Additional language courses may be given to this pupil. The quality adjustment introduced here can be seen as an alternative approach to differentiate between the educational services delivered, for instance to differentiate between pupils without dyslexia who attend class 4 of havo and those who are diagnosed with dyslexia and attend class 4 of havo. Conditions or learning (dis)abilities like dyslexia, deficit in spoken language at school, or dyscalculia can occur at all types of education. It is assumed that the pupils under scrutiny here, and who attend a certain type of education i despite a certain learning disability, all have the capabilities needed to successfully finish some education type i, but only when they get support in a specific part of their curriculum. (The alternative for them would be to go to another school type, for instance one with a lower attainment ratio, or possibly to a type of school that offers a more practical instead of theoretical education. In general this is supposed to be a solution less fit for the pupil in question.)

Some extra notation conventions are needed first. For each education type $i \in I$, a set of L learning (dis)abilities is introduced.[13] Let $q_{i,l,t}$ denote the number of pupil-years of type $i \in I$ that belong to learning ability class $l \in L$ in year t. Let $p_{i,l,t}$ denote the costs of these educational services for this class of learning abilities in year t. A Laspeyres volume index similar to (7) can be written as follows:

$$
(B1) \qquad \frac{\displaystyle\sum_{i\in I}\sum_{l\in L} p_{i,l,t-1} \cdot q_{i,l,t} \cdot \lambda_{i,t}}{\displaystyle\sum_{i\in I}\sum_{l\in L} p_{i,l,t-1} \cdot q_{i,l,t-1} \cdot \lambda_{i,t-1}}
$$

The sum over all pupil-years differentiated to all kinds of learning (dis)abilities in year t equals the total number of pupil-years $q_{i,t}$ taught at some type of education i in year t:

$$
(B2) \qquad q_{i,t} = \sum_{l\in L} q_{i,l,t}
$$

The costs $p_{i,t}$ per pupil for a school type i in year t can be seen as an aggregated average price[14] over the different classes of learning (dis)abilities L, since the following equality must hold for all $i \in I$ and t:

---

[12] This appendix draws heavily on Chessa (2009).

[13] Here the example of a learning disability is taken. Of course, one could also think of making quality corrections for schools that offer high-gifted children possibilities for extra knowledge and skill transfer.

[14] This contains the crux of the argument made in this appendix: if there is more detailed cost and consumer information about a specific group of pupils available (for example the extra costs for the groups of pupils with dyslexia at a certain type of education), then, using this information, it is possible to explicate that certain price and volume effects result from changes in the distribution of learning (dis)abilities within the population under scrutiny.

(B3) $$p_{i,t} \cdot q_{i,t} = \sum_{l \in L} p_{i,l,t} \cdot q_{i,l,t}$$

Now it is possible to express the costs of $p_{i,l,t}$ in terms of the average costs $p_{i,t}$ for school type i, such that $p_{i,l,t} = s_{i,l,t} \cdot p_{i,t}$. The factor $s_{i,l,t}$ thus expresses the extent to which the costs to educate a pupil with a certain learning (dis)ability deviate from the average costs to educate a pupil. Next $v_{i,l,t}$ is introduced, which denotes the fraction of pupils with learning (dis)ability l within i in year t. Therefore, it holds that $q_{i,l,t} = v_{i,l,t} \cdot q_{i,t}$. Further notice the following two properties:

(B5) $$\sum_{l \in L} v_{i,l,t} = 1$$

(B6) $$\sum_{l \in L} s_{i,l,t} \cdot v_{i,l,t} = 1$$

Using the just introduced factors $s_{i,l,t}$ and $v_{i,l,t}$, the volume index (B1) can now be rewritten as follows:

(B7) $$\frac{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t} \cdot \lambda_{i,t}}{\sum_{i \in I} p_{i,t-1} \cdot q_{i,t-1} \cdot \lambda_{i,t-1}} \cdot \left\{ \sum_{l \in L} s_{i,l,t-1} \cdot v_{i,l,t} \right\}$$

The form of the left term in (B7) is similar to the left term in expression (7).[15] The difference between (B7) and (7) lies in the part of the equation that is bracketed. This term can be interpreted as the 'quality adjustment' factor.

In the case of full data availability there is no need to use (B7), and one could simply use expression (5). A more refined product differentiation with equally refined cost weights would yield the same results. In that case, each element $i \in I$ could be stratified further into categories that denote types of education as well as learning abilities. Often such information is not or not fully available. In such cases the approach to quality adjustment developed here offers the possibility to perform sensitivity analyses. This may give some indication about the effects of possible quality adjustments and the robustness of our volume index. An example can make this clear.

Suppose there is data available on the number of people with dyslexia on schools for secondary general education (i.e. vwo, havo, vmbo, lwoo/pro, and special secondary education). Suppose that the data is available from 2001 onwards. Assume further (for the sake of argument in this purely fictional example) that the property of being diagnosed with dyslexia is evenly distributed over the types of education that make up secondary general education. Between 2001 and 2003 five percent of the pupil population is dyslexic. In 2004 this increases to six percent and stabilizes at seven percent from 2005 onwards.

---

[15] It is still assumed that for the right hand term of expression (7) holds that $k_{0,t} / k_{0,t-1} = 1$.

Figure B1: Chained volume indices for secondary general education in the Netherlands, after discerning the pupil population into pupils with or without dyslexia, and using various assumptions about extra costs for pupils with dyslexia.
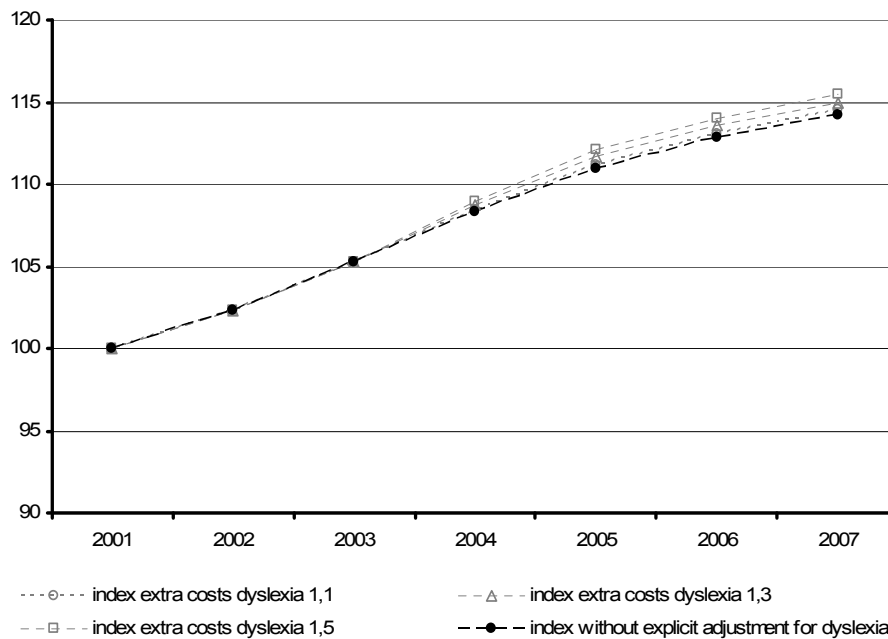
Different assumptions about the additional costs involved to educate this category of pupils will lead to different indices. In figure B1 some possible solutions are given. The lowest black dotted line represents the index without (quality) adjustment for dyslexia. The dotted lines above represent the indices under the assumption that costs per promoted pupil with dyslexia are 1,5 times (upper dotted line), 1,3 times, and 1,1 times higher than the aggregated average costs $p_{i,t}$.

In this example there is no change in the proportion of pupils diagnosed with dyslexia between 2001 and 2003. Therefore, there is no difference to be noticed in the volume index under the different assumptions made. Only when the proportion of pupils with dyslexia alters, changes in the volume indices become visible. This happens between 2003 and 2005 when the proportion of students with dyslexia is growing. Under the assumptions chosen here, these pupils carry proportionally more weight and consequently the volume indicator rises more quickly.

The example shows that volume adjustments based on such information may have an effect on volume growth. The sensitivity analysis may give rise to the idea that the volume indicator (not for dyslexia corrected) underestimates production. Other quality adjustments may show the opposite effect. For example one might think of certain technological innovations or a rise in the proportion of pupils with Dutch as native language.

Appendix C: Sensitivity analysis for dropouts at mbo level

As mentioned in section 5 half of the dropouts in secondary vocational education is counted as production. The other half is seen as 'failed' production. This choice is made because one may expect that a part of the pupils who finished their school year successfully decide to quit school because they prefer a paid job over pursuing a diploma. Pupils who would not have been promoted to the subsequent class may have started a job as well. In the model proposed in this paper the latter category should not be counted as production.

The choice to count fifty percent of the dropouts as production should be elucidated. Due to a lack of information about the study results of dropouts in the last year of their studies a sensitivity analysis is carried out. Various assumptions are tested about the proportion of successful pupils that leave secondary vocation education.

Figure C1: Sensitivity analysis for dropouts at mbo level. Dropouts treated as full production (1), no production (0), and intermediate steps.
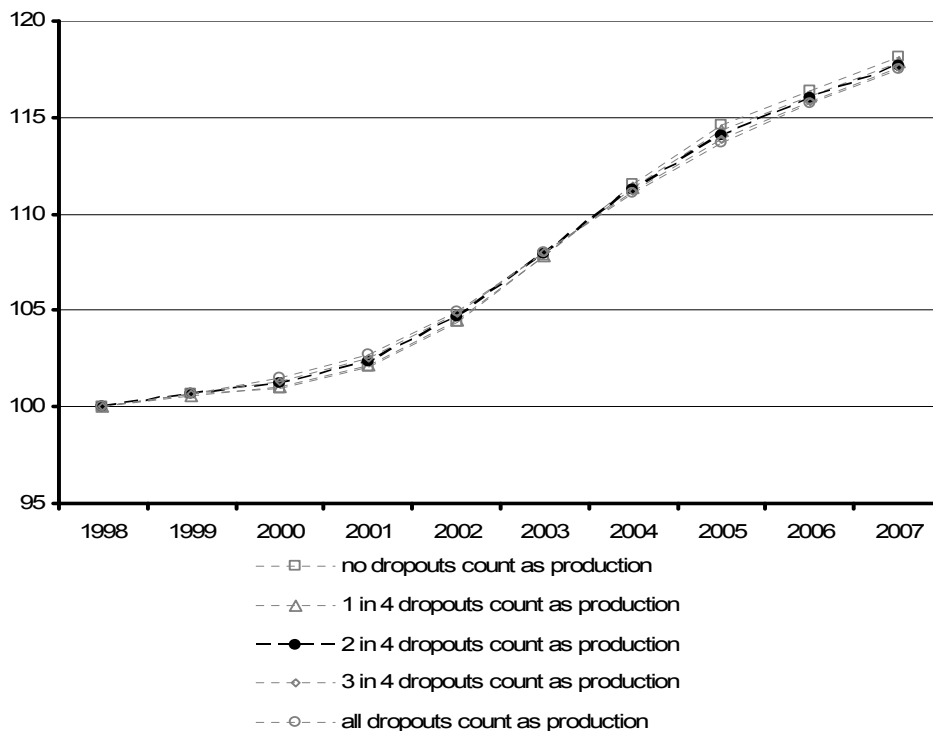

Figure C1 shows five different chained volume indices for vocational secondary education in the Netherlands between 1998 and 2007. The assumptions made vary between 'all dropouts are production' to 'none of the dropouts count as production.' As noted before, in the method proposed here a middle course is taken. The sensitivity analysis shows that whatever choice is made, the chained volume indices stay rather close to each other. If all dropouts are counted as production the volume index increases on average with only 0,2 percent per year faster than when none of the dropouts are seen as production between 1998 and 2007. The biggest differences between the two indices are found in 2001 and 2005. These points mark the maximum (73.251) and minimum (49.327) number of dropouts in secondary vocational education between 1998 and 2007.