

Measures of health care volume with applications to the Dutch health sector

Antonio G. Chessa¹

Abstract. This paper presents a measurement framework for production volume of health care. The framework consists of: (1) criteria for differentiating health care ‘products’, (2) conditions that volume measures should satisfy. The volume measures of the products are counting measures in mathematical sense: they need to be additive over time. A second additivity condition states that volume measures are not differentiated by provider or process characteristics. For instance, volumes for inpatient and outpatient care should be added for the same product. The notion of ‘health care package’ is proposed as counting unit, which is conceived of as a series of individual health care activities aimed at either recovery (cure) or the control of a patient’s health (care). Examples of valid and inadmissible counting measures are given.

Volume indices are calculated for Dutch hospital care and elderly care, and effects of different levels of product differentiation on volume growth are quantified. The average annual growth rate for hospital care drops from 3.5 to 2 percent during 1995-2007 when differentiating also by clinical and day treatments. The strong and weak points of the volume measures for hospital and elderly care are identified. Possibilities of improving the measures with new data for hospital care (DBC’s) and long-term care (ZZP’s) are discussed.

‘Quality’ adjusted volume measures result in this framework from a refinement of product differentiation. Adjusted volume indices are derived, which are illustrated with an example concerning pressure sores in nursing homes. The first results indicate a positive contribution to volume growth with several tenths of a percentage point. Some remarks are made about the applicability of the framework to other sectors.

Keywords: Volume index, counting measure, additivity, product differentiation, output, quality, cure, care, inpatient and outpatient care.

1. Introduction

In the past years, increasing efforts have been made by statistical agencies to develop explicit measures of production volume for service sectors (OECD, 2008). One of the recommendations of the Eurostat handbook on price and volume measurement is to develop volume indices based on measurements of quantities produced for different types of health services instead of deriving volume indices from price changes (Eurostat, 2001, p.117). A motivation for following the ‘volume approach’ is that differentiated product counts open possibilities for detailed analyses of labour productivity.

Finding suitable measures of production volume for service sectors seems to be more complex at first sight than for other sectors of economy. What types of ‘products’ do schools and health institutes deliver and what are the units of production to be counted? For cars, food products, computers, etc., numbers produced are counted, which are differentiated according to characteristics like brand name and model version. The product counts are aggregated in some way to an overall volume measure.

¹ Statistics Netherlands (CBS), Division of Macro-economic Statistics and Dissemination (MSP), Development and support department (MOO), P.O. Box 24500, 2492 HA The Hague, The Netherlands. E-mail: acsa@cbs.nl. The views expressed in this paper are those of the author and do not necessarily reflect the views of Statistics Netherlands.

Different volume measures have been proposed by statistical agencies for different types of services (e.g., see OECD (2008)). Numbers of patients, patient days and numbers of treatments are used for health care, while numbers of students and pupil hours are instances of measures used for education services. The range of proposed measures could suggest that there is ample freedom of choice. However, different measures may give quite different results, even for the same type of service. Is it possible to assess the suitability of a volume measure with respect to other measures? A set of guidelines for developing volume measures and to test their validity seems to be lacking in the literature.

This paper makes an attempt towards a framework for defining and quantifying health care production and volume, which is applied to hospital care and care for elderly patients. The measurement framework is described in Section 2. In Section 2.1, criteria are proposed for differentiating, or stratifying, health care production into different types. In Section 2.2, additivity conditions are set up that volume measures should satisfy and an abstract concept for counting units of production is described. Examples of admissible volume measures are given for different types of health care. Examples of indicators that do not satisfy the conditions of volume measures are also given.

In Section 3, the volume measurement framework is applied to hospital care in the Netherlands. The production volume index is calculated and compared to volume indices obtained for other levels of product differentiation. The volume measure is discussed in Section 3.2. The volume index for elderly care is presented and analysed in Section 4.

Data on health aspects ('quality indicators') are discussed in Section 5. Adjusted volume measures are derived and calculated for elderly care, based on prevalence scores of pressure sores. The adjusted volume measures are compared with the unadjusted versions.

In Section 6, new data are discussed that have become available recently for hospital care and long-term care. The new data follow from changes in the financing systems that were introduced in the past years. The data will be used to verify on which points the current volume measures can be improved.

The main findings will be summarised in Section 7. Points of further improvement for volume measurement will be outlined. Some remarks about the applicability of the measurement framework to other sectors, such as education, will also be made.

2. Measurement of health care volume

2.1 Differentiation of health care products

The central subject of a volume measurement framework should be the patient. There will be no production volume when health care providers do not have patients. Health care supply often starts when a client having certain symptoms visits a general practitioner. The client enters a trajectory, which may end with medication and recovery at home. But a client may also receive a doctor's referral for further examination and treatment at a health care institute, for instance, at a hospital or an institute for long-term care.²

The general idea behind volume measurement of health care production is as follows in this paper. We measure production volume by comparing a patient's health state during health care supply with some threshold. We count a unit of production when the threshold is exceeded; otherwise no production will be counted. The units to be counted are conceptualised in Section 2.2. Before defining counting measures of production volume, we characterise a space in which health care products can be positioned and counted.

² Health care refers here to the direct delivery or contact between the provider on the one hand and the patient on the other hand. Components not originating in care, such as promotion of healthy lifestyle, are left out.

In our volume measurement framework, health care products are characterised by an initial health state and an end state. Initial and end states can be interpreted in a broad sense. The states may correspond with the entire path of health care supply, where an initial health state may be described by a client's symptoms, while full rehabilitation may be the end state. Initial and end states may also be confined by health care supply at institutional level. For instance, the states could mark the initial and end state of a patient's health during hospital stay. The initial health state may be expressed in terms of a diagnosis, while the end state may be full recovery. As this paper focuses on hospital care and long-term care, we will characterise initial and end states for these types of care.

The dimensions of the state space of products are thus initial health state and end state, and, of course, also time. A product is represented as a point in a multi-dimensional space. The time component denotes the time at which the end state is reached, which coincides with the end of a treatment (e.g., hospital discharge). In this section, the focus is on characterising initial health states. Health states reached at the end of a treatment are often linked with 'outcomes' and 'quality' of health care (Castelli et al., 2007; OECD, 2008; Schreyer, 2008). We include final health states in our volume measurement framework in Section 5. A characterisation of health states is given below. These represent criteria for product differentiation.

D1. Purpose of health care process: cure or care

The purpose of health care activities is either: (a) the recovery of a patient's health, or (b) the control of a patient's health. Process type (a) usually acts on a shorter time scale, which eventually results in a patient's discharge from a health institute. Process type (b) mostly comprises long-term health care. The focus in elderly care, for instance, is on providing particular types of care and assistance repeatedly in time, like daily household and personal care. Process types (a) and (b) are in fact equivalent with cure and (maintenance or dependency) care. This distinction has implications for volume measurement, as will be argued in Section 2.2.

D2. Diagnosis (cure) or health indication (care)

Purpose of process D1 can be considered as a very rough, first subdivision of initial health states. Cure and care can be refined into specific diagnoses (cure) and health care indications (care). We do not intend treatment when we speak of diagnosis or health care indication. There may be different types of treatment for the same diagnosis.

D3. Consumer characteristics

Patient characteristics, such as age, could also be considered as health state indicators. Older patients may be more sensitive to additional complications or side effects than younger patients, also for the same diagnosis. We therefore suggest differentiating by patient age. Socio-economic environment of patients may also affect initial health state and could thus be used as an additional criterion for product differentiation. Certain diagnoses are sex specific, such as cervical cancer, which should thus also be used for differentiating health care products.

D4. Health state at the end of a treatment

A part of the indicators for end states should relate to initial health state indicators in order to measure progress in health state during treatment. Examples of indicators are the extent to which patients are cured or cared for diagnoses or health care indications. Full recovery from a diagnosis is an example of an end state. A clean house and a well-cared body are examples of end states in elderly care. Additional indicators may become relevant as end states, such as occurrence and severity of pressure sores in elderly care, incontinence, and complications after surgery.

We do take end states into account in volume measurement as treated in Section 2.2, but only with regard to satisfying a threshold for health states. For instance, we look at whether patients

have recovered or have been cared for ‘sufficiently’, but not at the extent to which a threshold is exceeded. Additional differentiation according to end states will be treated in Section 5. The above framework for product differentiation is applied in Section 3.1 to hospital health care and in Section 4.1 to health care for elderly patients. In sections 3 and 4, we will also quantify the effects of different levels of product differentiation on volume growth.

2.2 Counting measures of production volume

Once a characterisation of health care types is available, counting measures need to be developed in order to quantify the amount of production for each type. In other words, we need to define counting and (aggregate) volume measures on a state space. Different volume measures have been proposed in the literature, also for the same type of care (OECD, 2008). Examples are: number of days of care, patients treated and patients discharged. Of course, choices depend on available data, which may partly explain the range of different measures. Nevertheless, a legitimate question to ask is whether these measures are well-defined and whether they properly reflect product counts.

Attempts towards a comprehensive measurement framework for production volume have not yet been made. Below, we will first list conditions that volume measures should satisfy. We will show that certain volume indicators do not satisfy these conditions. Next, we will introduce an abstract concept of a counting measure of production volume, which will be discussed for cure and care. Volume measures should satisfy the following two conditions.

C1. Additivity over time

A volume measure for every differentiated product should be regarded as a counting measure in a mathematical sense. In abstract terms, a counting measure is a function that takes nonnegative integer values on subsets of some space (e.g., the real line or some multi-dimensional space).³ Space is equivalent with time for every differentiated product. The scope is to measure volume of production on a yearly or quarterly basis.

Counting measures are additive: the value of a measure is equal to the sum of the values for disjoint (‘non-overlapping’) subsets of some space. This property should also be imposed on volume measures in national accounts. Product units counted for a whole year should be equal to the sum of product counts for smaller time intervals within a year, such as months or quarters. As a consequence, additivity is an essential requirement for linking quarterly and annual volume statistics.

In more abstract terms, additivity can be denoted as follows. Let Q be a measure on subsets of some time interval T (e.g., a year). Let $T_i, i = 1, 2, \dots$, be a partition of T , so that the sets T_i do not overlap. The value of Q on T is then equal to the sum of Q over the sets T_i , that is:

$$(1) \quad Q(T) = \sum_i Q(T_i).$$

This is a necessary condition for volume measures, and should apply to all economic activity.

Additivity of volume measures may seem a trivial characteristic. However, this property excludes certain candidates of volume measures. Examples are number of patients in long-term care and number of students in education. Let us consider residents in homes for the elderly who receive care throughout the whole year. As the patients reside in elderly homes at any time of the year, the numbers of residents in the four quarters of a year are equal. But their sum is four times the

³ For technical details about measures see, for instance, Ash (1972).

number of residents at a yearly basis, which is in conflict with the number of residents for the whole year. The additivity property therefore does not apply in this case.

It almost sounds contradictory when saying that the number of patients is not an admissible volume measure, as the patient is the subject of the health care process. The problem with using patient counts in long-term care is that production processes in execution are in fact counted. However, we do not want to count processes when quantifying production volume, but the focus should be on the termination of a process (i.e., the end state) and the difference with the initial state. This idea is developed further in the concept of counting measure, which is described at the end of this section. Of course, patients should be taken as a point of departure in any counting measure. Production volume will be zero when there are no patients. The point is that patient counts are not sufficient; some additional structure should be imposed when developing counting measures.

C2. Additivity over provider or process characteristics

An important question when measuring health care volume is whether inpatient and outpatient treatments should be counted as equal units or not. As we stated at the end of condition C1, we intend to base counting measures of production volume on the completion of production processes, not on processes being carried out. The end result of a process is relevant to counting production, not the way in which a process is organised, what input factors are used. It makes no difference to our counting measures whether a patient with a certain diagnosis and age receives an inpatient or outpatient treatment, what drugs are administered, what kind of equipment and labour are used, whether the health provider is located nearby, etc. In other words, the way in which the transition from initial to final health state is realised is irrelevant to volume.⁴

For health care products that are differentiated according to purpose of a process, diagnosis or health indication, and consumer characteristics (D1-D3), the corresponding counting measures are additive over provider or process characteristics. We could denote this additivity property as follows. Let the product differentiation according to criteria D1-D3 give rise to a set H of product categories, and let differentiation according to provider characteristics result in a set S of different categories. Elements $J \in H$ can be denoted as vectors, such as (cure, appendicitis, age 50). Elements of the set S could be clinical and day treatment.

The volume measure $Q_J(T)$, for $J \in H$ and some time interval T , is equal to the sum of the volume measures $Q_{J,s}(T)$ over the set S :

$$(2) \quad Q_J(T) = \sum_{s \in S} Q_{J,s}(T).$$

The dimensions specified by the provider characteristics are thus summed out when counting production volume, for any $J \in H$. Product counts for the same diagnosis and patient age are added together for different provider characteristics. In the example of appendicitis this means that the counting measure of production volume, for a given patient age class, is equal to the sum of the numbers of completed clinical and day treatments. The distinction between clinical and day treatments is thus irrelevant. This additivity property also finds support in other contributions (OECD, 2008, pp.72-73; Schreyer, 2008).

⁴ If treatments in different institutes lead to different health states for the same diagnosis and patient age, then the numbers of discharged patients are not summed over the institutes. The additivity property only applies to discharges with the same initial and final health states. Discharges with different health states are considered here as different products. Differentiation of health care products with respect to final health states will be considered in Section 5.

Counting measures for cure and care

Health care packages as counting units

In Section 2.1 we stated that products are described by initial and final health states. The full set of health care activities performed between initial and final state represents our counting unit of production. This means that the counting units are not the individual health care activities in a health care package, unless a package consists of a single activity. In the case of cure, the activities in a package may encompass laboratory examinations, treatment of a diagnosis and activities aimed at a patient's recovery. In the case of care, a health care package contains activities for treating a health care indication or diagnosis. For example, a health care package in elderly care may consist of several hours of household assistance and personal care per week.

Suggestions have been made in the literature to use "complete treatments" as counting unit, which refer to the entire pathway that an individual takes through different health care institutes (see OECD (2008, p.72) and references, e.g. Berndt et al. (2001) and Aizcorbe et al. (2008)). The use of complete treatments has limitations, so that the OECD-handbook suggests "episodes of treatment provided by institutional units" as a "working definition of output". We recognise the limitations of complete treatments, but we also claim that health care products can be defined in a broad sense by making different choices about initial and end states.

Threshold

The health care activities in a package are intended to improve a patient's initial health state, which corresponds to the diagnosis or health indication received before treatment. A health care package should be counted only if a patient's health state exceeds some threshold. A threshold can be seen as a minimum health state that is sufficient for a patient's recovery or acceptable for being cared for. Thresholds also apply to other sectors of economy. For instance, produced cars or computers should be counted only when they function properly.

A health threshold may consist of different components for the same diagnosis or health indication. In the case of a hip or knee replacement, the threshold may be defined in terms of the possible occurrence of certain complications, such as infections or bleedings. The threshold could be set at 'no complications', that is, treatments for hip and knee replacement would be counted only when no complications arise during the whole treatment. In the case of elderly care, thresholds may include a clean house, a well-cared body and the occurrence of pressure sores. For instance, it could be decided to count health care packages only when patients have a clean house, a well-cared body and when no pressure sores occur, or when pressure sores do not include damage to bone and skin tissue.

The extent to which a threshold is exceeded could be used to further refine volume measures. This is often referred to as 'quality adjustments' of volume measures. We will illustrate how this problem could be dealt with in Section 5. Until then, we will count production volume as the number of health care packages, for which patients' health states satisfy some threshold, irrespective of the extent to which the threshold is exceeded. We denote this counting measure as *output volume measure*.

Examples of counting measures

In the case of cure, counting health care packages that satisfy a threshold is equivalent to counting patient discharges when patients are in medically acceptable conditions (or in the best state allowed in view of medical and health knowledge applied). A patient may undergo a series of admissions and discharges for the same diagnosis. Only the final discharge should be counted in the case of cure, since the emphasis is on a patient's recovery (obviously, the threshold was not reached before the final discharge).

A series of discharges could be counted all in the case of chronic diseases, such as varicose veins. The focus is on controlling a patient's health state, as full recovery is often not feasible.

During each admission and treatment, a patient's health state could be compared with thresholds for level of discomfort and complications, such as leg ulcers, swelling or skin discolouration.

Counting discharges does not make sense for the greatest part of long-term care, such as elderly care and care for disabled patients. In these cases, patients receive a health care package repeatedly in time. A package contains the health care activities for which patients received a health care indication. The number of health care packages should be counted that are supplied periodically to patients, for instance, one or two hours of household assistance and personal care on a daily basis. In this example, this amounts to counting the number of patient days.

Health care packages versus activities

In theory, counting measures based on health care packages are different from counting measures based on the individual health care activities in a package. In the latter case, we normally have a more detailed product differentiation and we would count the number of times that a health care activity is performed or the time spent on an activity. As we do not always have complete data on health care packages, it is interesting to know under what conditions the volume indices for the two counting measures give the same results.

In Appendix A, it is shown that the two volume indices are equivalent if: (1) the price or tariff of a health care package is expressed in terms of the tariffs of the underlying health care activities, (2) the number of times that an activity is carried out, or the time spent on an activity per package, is the same in two subsequent years for all the activities in a package, and (3) effects due to packages that fall in more than one calendar year are negligible ('censoring effects').

The results of Appendix A have important implications. If the volume indices for counting packages and activities give the same result, then we could use the volume index based on activities without making adjustments, when data on health care packages are not available. On the other hand, if the time spent on individual health care activities per package decreases from year to year, then the volume index based on the individual health care activities is smaller than the volume index based on counts of health care packages. The first volume index would thus underestimate volume growth. A nice illustration of this effect will be given for hospital care in the Netherlands in the next section. The effect occurs mainly because clinical treatments are substituted by day treatments. On average, less hospitalisation days are required for treating the same diagnosis.

3. Output volume for hospital care

3.1 Volume index

Data and volume measures

In this section, we summarise the method for calculating an output volume index for clinical and day treatments in hospitals. The method is based on data from the Hospital Discharge Register (HDR).⁵ We decided to focus on clinical and day treatments and not on other types of hospital care in this section in order to emphasise the effects of clinical and day treatments on output volume growth. Clinical and day treatments constitute about 70 percent of the entire hospital care. Outpatient (polyclinical) services are not part of the HDR and are not considered here. Descriptions of the complete volume method can be found in Chessa and Kleima (2006) and in Kleima et al. (2004).

The HDR contains the following data for every patient admitted:

⁵ The Dutch name of the database is Landelijke Medische Registratie (LMR). The data are managed by the Dutch organisation Prismant.

- Date of birth, which allows us to create age classes;
- Main diagnosis and subdiagnoses, according to 11,182 ICD-9 codes;
- The number of hospitalisation days;
- Clinical and day treatments are distinguished. Clinical treatment involves a stay of at least one night in the Netherlands, while day treatments are without a night stay;
- Patient admission and discharge.

The following choices were made for the volume measures:

- Products are differentiated according to main diagnosis and patient age. The main diagnoses are grouped according to a 3-digit ICD-9 classification, which results in approximately 1000 diagnosis groups. Patient ages are subdivided into 7 age groups (0, 1-14, 15-44, 45-59, 60-69, 70-79, 80 and older). This gives about 7000 diagnosis-age groups;
- No distinction is made between clinical and day treatments, so that additivity property C2 in Section 2.2 is satisfied;
- The counting measure used is the number of hospital discharges.

Expenditures per discharge are needed in order to calculate volume indices. However, this information is not available for each year. The information is assembled in “Cost of Disease” studies (CoD), which are performed once every five years (for more details, see Kleima et al. (2004)). This implies that the CoD-studies do not allow annual adjustments of the value shares of the diagnosis-age groups. The Dutch National Accounts use Laspeyres volume indices with annually adjusted value shares.

On the other hand, the CoD-studies show that expenditures are determined for about 85% by the price of hospitalisation. This motivated Statistics Netherlands to base the value shares of the diagnosis-age groups on the number of hospitalisation days. The advantage of this choice is that hospitalisation days are available in the HDR every year. The length of a day treatment is set equal to one hospitalisation day.

We now give the volume index for clinical and day treatments of hospital care in the Netherlands. Let H denote the set of diagnosis-age groups. Let $D_{J,t}$ denote the mean number of hospitalisation days and $Q_{J,t}$ the number of discharges for diagnosis-age group $J \in H$ in year t . The Laspeyres volume index for clinical and day treatments in year t , with respect to year $t - 1$, is:

$$(3) \quad \frac{\sum_{J \in H} D_{J,t-1} Q_{J,t}}{\sum_{J \in H} D_{J,t-1} Q_{J,t-1}}$$

Results

Figure 3.1 shows the (chained) volume index for the period 1995-2007. An interesting question at this point is to what extent the behaviour shown in Figure 3.1 changes when the product differentiation is modified. We consider two other product differentiations: (1) a distinction is made between clinical and day treatments, beside age and diagnosis, (2) no distinction is made at all, neither with respect to diagnosis and age. The first differentiation gives rise to $7000 \times 2 = 14,000$ product types. In the second approach, the total volume is simply equal to the total number of hospital discharges.

Before we compare the result in Figure 3.1 with product differentiation (1) and (2), we first give the expression for the volume index when differentiating output also with respect to clinical and day treatments. Let $D_{J,c,t}$ and $Q_{J,c,t}$ denote the mean number of hospitalisation days for clinical treatments and the number of clinical discharges, respectively, for diagnosis-age group $J \in H$ in year t . We denote the number of day treatments for diagnosis-age group $J \in H$ in year t by $Q_{J,d,t}$.

Since the duration of a day treatment is set to one hospitalisation day, we can write the volume index for this product differentiation as:

$$(4) \frac{\sum_{J \in H} D_{J,c,t-1} Q_{J,c,t} + \sum_{J \in H} Q_{J,d,t}}{\sum_{J \in H} D_{J,c,t-1} Q_{J,c,t-1} + \sum_{J \in H} Q_{J,d,t-1}}$$

The volume indices for the three types of product differentiation are shown in Figure 3.2, where the two alternative product differentiations are compared to the result in Figure 3.1. The results in Figure 3.2 show that the volume index as calculated in Figure 3.1 does not differ that much from the index for the total number of discharges. However, the difference becomes very large when output is also differentiated by clinical and day treatment.

Figure 3.1 Chained volume index for clinical and day treatments (1995 = 100).

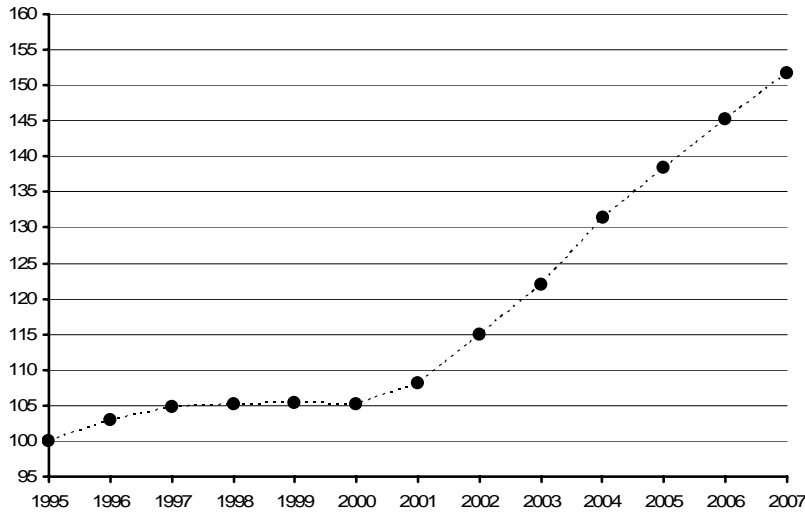
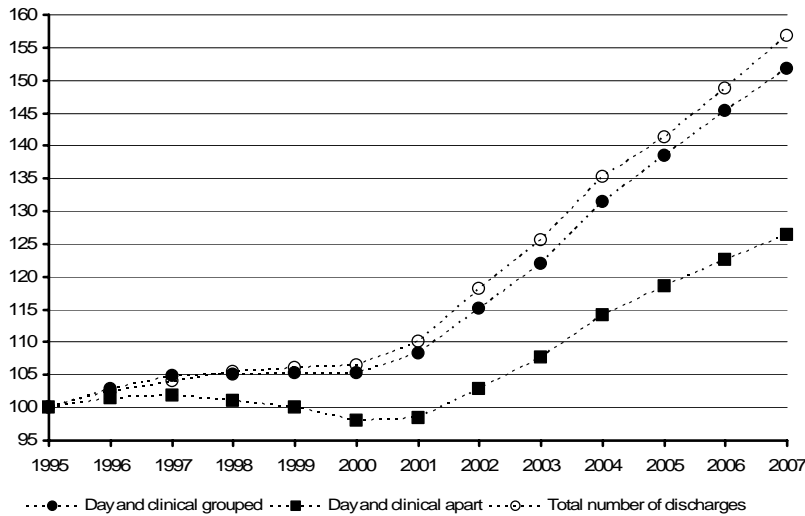


Figure 3.2 Chained volume indices for three levels of product differentiation (1995 = 100). The volume index for 'day and clinical grouped' is the index shown in Figure 3.1.



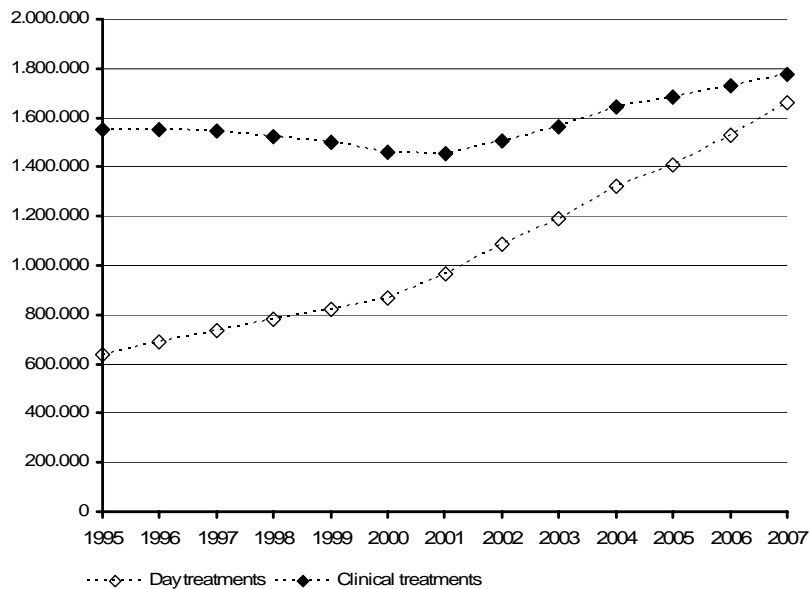
The average annual output growth for the three methods in Figure 3.2 is as follows: 3.5 percent for day and clinical treatments combined (Figure 3.1), 2 percent for output differentiated also by

clinical and day treatment and 3.8 percent for the total number of discharges. The average growth figures apply to the period 1995-2007.

A further differentiation of products according to clinical and day treatments gives a much smaller growth rate because of a combination of two factors. First, clinical and day treatments are given different hospitalisation day weights (see expression (4)). These ‘weights’ are equal in (3), while day treatments receive less weight than clinical treatments in (4). Second, day treatments show a much faster growth rate than clinical treatments (Figure 3.3). The different weight schemes thus give rise to quite a different behaviour of the volume index when differentiating output to clinical and day treatments as well.

As we already mentioned, the total number of discharges shows an average growth rate that differs only 0.3 percentage points from the volume index used at Statistics Netherlands for clinical and day treatments. The small difference implies that the overall contribution of product differentiation with respect to diagnosis and patient age to output volume growth is rather small. The contribution of product differentiation is slightly negative, since the total number of discharges increases faster than output volume (see the two upper lines in Figure 3.2). This could be caused by shifts from more severe types of care to less severe types, that is, to types of care with shorter hospital stays.

Figure 3.3 Number of clinical and day treatments in Dutch hospitals.



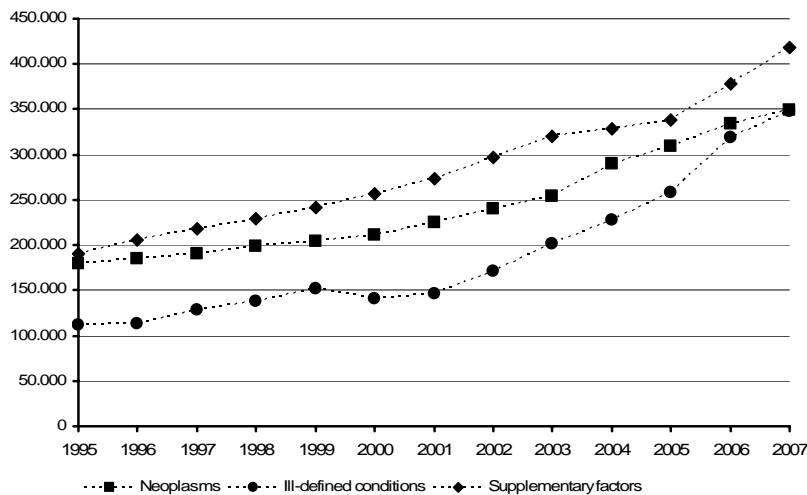
Although the focus of this section is on product differentiation and volume measurement, it is worth to spend a few notes on the development of output volume itself. Figures 3.1 and 3.2 clearly show two phases in the development of output volume. The discrete growth in the period 1995-2000 is followed by a much faster growth in the subsequent period. The change in growth behaviour in 2001 coincides with changes in the financing of health care. The Dutch Ministry of Public Health, Welfare and Sport made available additional financial means in order to reduce waiting lists. Health care institutes were allowed to deliver additional health care, which was not possible in preceding years. Until 2001, health care institutes could only deliver health care on the basis of a fixed budget.

Whether the changes in the financing system are the only or main reason for the rapid volume growth in the period 2001-2007 is disputable. The total number of discharges increased at a high rate throughout this whole period (6.1 percent per year on average), especially for day treatments. Shifts have taken place from clinical to day treatments. For example, appendicitis is

treated more often with laparoscopy than with traditional (clinical) surgery. The length of hospitalisation stay is reduced by day treatments, but their contribution to output volume is the same as for clinical treatments.

Furthermore, several ICD-groups ('chapters') have contributed substantially to output volume growth. Figure 3.4 shows the three ICD-9 groups with the highest increase of the number of discharges. Leaving these diagnoses out would decrease the growth rate to 4.6 percent for the number of discharges (2.6 percent instead of 3.8 percent for the period 1995-2007). Treatments of neoplasms, both benign and malignant, have nearly doubled in 12 years. A strong contribution to hospital discharges also comes from admission of patients with ill-defined conditions. These encompass a broad range of signs, such as mental and sensory disturbances, dizziness, palpitations, excessive sweating, etc. The rapid increase shown in Figure 3.4 could indicate that patients visit a doctor quicker than in past years. The 'supplementary factors' denote services for purposes other than remedying health states. Examples are encounters of persons with health services for specific procedures and aftercare, and special screening examinations. This category of services applies to persons without reported diagnosis (the so-called "V-codes" in ICD-terminology).

Figure 3.4 Numbers of discharges for three groups of diagnoses in Dutch hospitals.



3.2 Discussion

In this section, we evaluate the output volume method described in the previous subsection. We focus on the two core issues of Section 2: product differentiation and volume measurement.

The results for the output volume index have shown that volume growth changes substantially when output would be differentiated also with respect to clinical and day treatments. Figure 3.2 shows that clear guidelines for product differentiation are needed in order to create a basis according to which choices for output volume measurement can be motivated. The output volume index for clinical and day treatments in hospitals differentiates output according to diagnosis and patient age. Differentiation by clinical and day treatment is not carried out, since these are provider or process characteristics. In Section 2.2, we imposed additivity over provider characteristics on counting measures. The focus is on the completion of health care processes and the recovery of a patient's health, not on the way processes are organised.

For every diagnosis-age group, output volume is quantified by counting the number of hospital discharges. In Section 2.2, we suggested health care packages as counting units. We argued that counting packages and discharges is equivalent when: (1) patients require only one

hospital admission, (2) patients have a chronic disease and require long-term care. In the second case, patients may be admitted and discharged more than once for the same diagnosis. The purpose of the process (care) justifies counting all discharges.

There are situations where the distinction between cure and care is not always clear, such as for cancer treatments. If the disease can be cured, then only the final discharge should be counted. If chemotherapy is used as palliative treatment, then we would count all discharges, as the focus of the health care process is on controlling a patient's health state (care). In less evident situations it is not clear whether to count all discharges or only the final discharge.

The question whether a series of discharges should be counted as one or as separate contributions to output volume also arises for the 'supplementary factors', or the ICD V-codes. Figure 3.4 shows a rapid increase of the number of 'discharges' for this group. One can imagine that counting discharges can be justified in certain cases, such as screening. But difficulties may emerge for other types of services. Different encounters of persons with health services, for instance regarding procedures about the same type of aftercare, should be counted as a single contribution to output volume.

The output volume index is based on about 1000 diagnosis groups. As there are more than 10,000 ICD-codes, this means that an average diagnosis group contains slightly over 10 ICD-codes. It is therefore not excluded that certain diagnosis groups are still heterogeneous. It was decided at Statistics Netherlands to combine ICD-codes in order to avoid problems with 'empty' ICD-codes. If no patients would be discharged for a certain ICD-code in year $t - 1$ while discharges would be registered for the same code in year t , then it is not possible to calculate a volume index for that ICD-code. However, it is possible to handle the problem of entering and disappearing products. This problem was treated for elderly care and care for disabled patients at a later stage at Statistics Netherlands (Section 4).

Another limitation of the volume index follows from the set-up of the Hospital Discharge Register. The HDR does not include outpatient (polyclinic) visits of patients. Their contribution to the overall volume index for care by hospitals and specialists has to be calculated separately in the Dutch National Accounts. Additivity over provider characteristics cannot be applied, for instance, in cases where day treatments can also be applied in a polyclinic setting. In addition, if polyclinic services are part of a treatment, then these should be combined with other activities into a single health care package. Possibilities of doing this are offered by the new data on hospital care, the 'diagnosis treatment combinations' (DBC's), which are discussed in Section 6.1.

4. Output volume for elderly care

4.1 Volume index

Data and volume measures

In this section, we summarise the method for calculating output volume indices regarding care for patients in nursing homes, homes for the elderly and at patients' homes (home health care). We denote the ensemble of these types of care as 'elderly care' in the sequel, as practically all types of care apply to older patients. However, a small part of care concerns younger patients.⁶

The method is based on data from the Dutch health authority NZa. The NZa-data cover the following aspects:

⁶ This holds in particular for home health care. An example is antenatal (or prenatal) care at home. Another example is support given at home to children with mental disabilities. Of course, the bulk of this type of care is offered by institutes for patients with mental disabilities (not included in this paper, see Chessa (2008, 2009)). There is no longer a true one-to-one relation between type of health care institute and types of health care offered by institutes.

- The health care activities for which patients have a so-called ‘indication’;
- Health care activities can be classified as nursing home care, elderly home care or home health care. A subdivision of health care into inpatient care (nursing home care and elderly home care with stays including at least one night) and outpatient care (day care and home health care) can thus also be made;
- For almost every health care activity in each of the three types of care, the quantity of care and the costs per unit of quantity are available. Quantities are expressed in days of care for inpatient care and in hours for most types of outpatient care. The costs comprise personnel and material costs and cost of capital use. The unit costs, or ‘tariffs’, are based on maximum tariffs set in the ‘NZa policy rules’.

The health care indications are in fact the ‘diagnoses’ in long-term care. For example, a patient with home health care may have an indication for household and personal care, with a total of 10 hours per week, say.⁷ The NZa-data contain more than 150 health care activities.

Patient data are not explicitly available. Nevertheless, the NZa-data implicitly contain a rough subdivision of health care activities by patient age. For example, the data make a distinction between children and adolescents, adults and elderly for the same types of psychotherapy. Of course, the greatest part of health care applies to older patients, so that a subdivision by age is less relevant than for hospital care.

The following choices were made for our volume measures for elderly care:

- The individual health care activities are the products differentiated;
- The counting measures for the differentiated products are: (1) patient days for inpatient care and care in day-care centres, (2) hours of care and face-to-face contacts in psychotherapy for the rest of outpatient care.

The NZa-data do not contain information about which health care activities are bundled and assigned to patients. Knowing the set of health care activities for which patients received a health care indication would have allowed us to use bundles or packages of health care activities as counting units. At best, the counting measures for inpatient care could be viewed as health care packages. A day of care may consist of different health care activities, which a patient undergoes repeatedly in time. Counting health care packages is thus equivalent to counting patient days. The composition of a day of inpatient care is unknown, however.

Furthermore, we would have liked to apply the ‘additivity over provider characteristics’ to the counting measures, as was done for clinical and day treatments in hospital care. However, the data do not allow us to sum volumes for inpatient and outpatient care. In order to do so, we should at least know the equivalent of a day of inpatient care in terms of hours of outpatient care. For instance, how many hours of household and personal care are given to patients at their home on average per day? Knowing the answer would give us the opportunity to convert hours of outpatient care into days, which could then be summed with inpatient days. We were thus forced to treat inpatient and outpatient care as different types of output. We will return to this additivity topic in Section 4.2.

Because of the separate treatment of inpatient and outpatient care, the output volume index is similar to expression (4) for hospital care. The hospitalisation day weights in that expression are replaced here by the tariffs of the health care activities and the discharges by the counting measures for inpatient and outpatient care. Computational problems caused by entering and

⁷ These examples of care are considered as health care in the Dutch national accounts. This convention is questionable in many countries and international organisations, as the examples of care mentioned are part of social care.

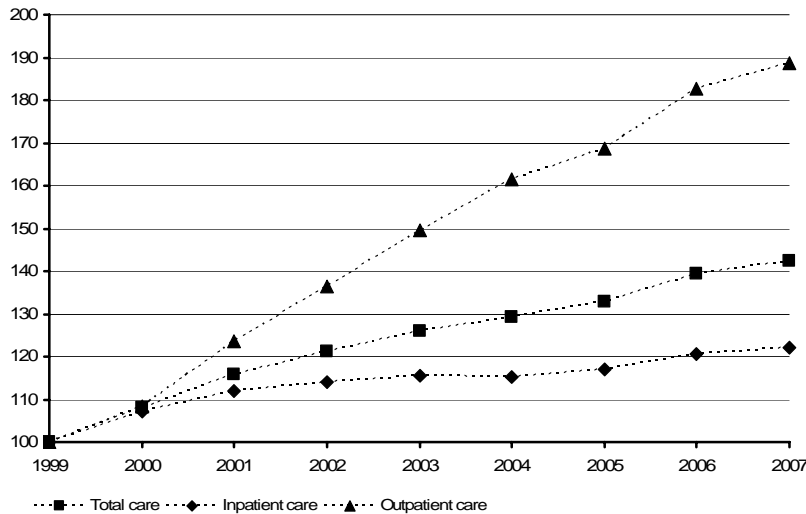
disappearing health care activities from year to year are treated in Chessa (2007) and in Chessa and Okkerse (2007).⁸

Results

Figure 4.1 shows the output volume indices for inpatient and outpatient care and for the entire elderly care in the period 1999-2007.⁹ The volume for outpatient care has increased at a much higher rate than for inpatient care. The respective average annual growth rates are 8.3 and 2.5 percent. Combining the output volume indices for inpatient and outpatient care gives an average annual growth rate for the entire care of 4.5 percent. The value share for inpatient care is larger than for outpatient care, so that the growth rate for the entire care is closer to the growth rate of inpatient care. We see a tendency in elderly care that is similar to that for clinical and day treatments in hospital care.

An interesting question within the scope of this study is to what extent the results of Figure 4.1 change when another level of product differentiation is chosen or a different measure for output volume. This question is a bit more difficult to answer than for hospital care, because different counting measures were used for different types of care and because quantities are missing for some types of care. Several types of care were therefore excluded from the subsequent analysis.

Figure 4.1 Volume indices for inpatient and outpatient elderly care and for both types of care combined (1999 = 100). Care financed with personal care budgets (PGB in Dutch) is not included.



In Figure 4.2, we compare an output volume index for long-term inpatient care with the total number of corresponding days of care in nursing homes and homes for the elderly combined, and with the number of patients in nursing and elderly homes. Care with respect to short stays is excluded, since patient numbers for this type of care are missing for several years. This part of care gives a large contribution to the output volume growth for inpatient care, which can be observed by comparing the output volume indices in figures 4.1 and 4.2.

Figure 4.2 shows that the number of patient days for long-term inpatient care has remained quite stable. However, the corresponding output volume has increased with 9 percent in 2007

⁸ The English-speaking reader may consult the report by Chessa (2009) on care for disabled patients. Entering and disappearing health care activities are handled by linking the price indices for these subsets of activities to the price index for those activities that are supplied in two successive years and for which the tariffs are known in both years.

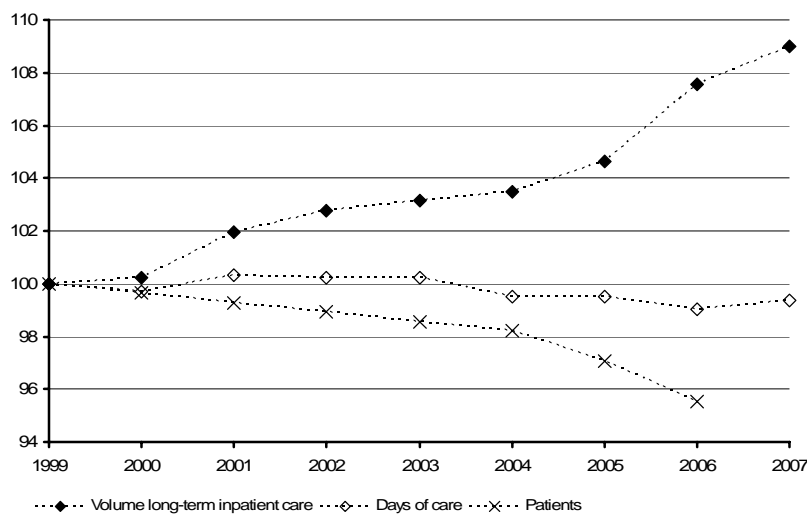
⁹ The NZa-data cover the period 1998-2007. We excluded the year 1998 from the results, because the first year for which we have data on patient numbers is 1999. These data will be compared with output volume later in this section.

with respect to 1999 (1.1 percent per year on average). Also in this case we thus observe a clear effect of product differentiation on output volume growth. A part of the difference between the two output volume indices can be ascribed to differences in the growth rates for days in nursing homes and days in homes for the elderly. Days in nursing homes increase while days in homes for the elderly decrease. As the value shares for nursing days are higher, the output volume index will be larger than the index for the total number of days of care.

Another contribution to output volume growth comes from the increase in the number of days of care for groups of patients that require special forms of care in nursing homes, such as patients with Korsakoff's syndrome, Huntington's chorea and rheumatism. This contribution is especially strong after 2004. Additional funds became available for nursing homes in order to give these patient groups even more adequate care.

Another important observation that follows from Figure 4.2 is that the number of patients decreased in the period 1999-2006. This means that, on average, the amount of care per patient increased, in terms of both output volume and days of care. It should be noted in this respect that days of care and patient numbers are based on different sources. In addition, patient numbers are sampled on one day of a year. These factors could contribute to the difference between the two indices in Figure 4.2.

Figure 4.2 Chained volume indices for long-term inpatient care, total number of days of care and patients for nursing and elderly home care combined (1999 = 100). Patient numbers refer to numbers on December 31 of every year. There are no patient data available for 2000-2003 and for 2007. Patient numbers for the period 2000-2003 were estimated by linear interpolation between the numbers for 1999 and 2004. Number of patients on December 31, 1999: 157,740.



It should be reminded that the number of patients is not a valid counting measure, since it does not satisfy the additivity over time property C1 in Section 2.2. The framework presented in Section 2 therefore proves its usefulness, as we can ascertain that the lowest curve in Figure 4.2 is not correct.

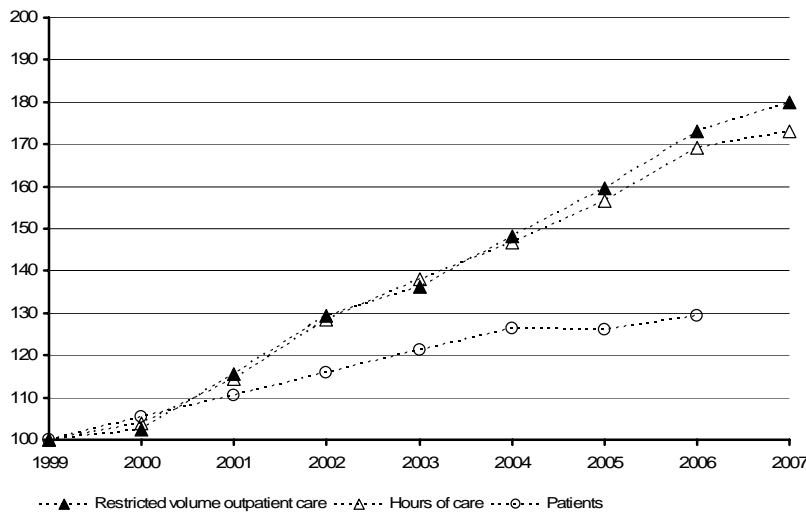
Figure 4.3 shows a similar graph for outpatient care. Also in this case we derived an output volume index for a reduced part of care, which we compared with the corresponding index for the total number of hours of care and the number of patients with outpatient care. We only considered health care activities for which the quantities are expressed as hours of care or face-to-face contacts of psychotherapy. Care in day care centres was omitted, because setting one day of care equivalent to 8 hours did not yield a reasonable index for the total number of hours. Finding a suitable conversion factor from day to hours would then become a subjective matter. The consequence of this omission is that the output volume index in Figure 4.3 is somewhat lower

than the complete volume index shown in Figure 4.1. We assume that one psychotherapeutic contact is equivalent with one hour of care.

The restricted output volume index does not differ much from the index for the number of hours. The respective average annual growth rates are 7.6 and 7.1 percent. In contrast with the output volume index for inpatient care in Figure 4.2, the greatest part of the output volume growth for outpatient care can thus be attributed to the increase of the number of hours of care. The contribution of product differentiation to output volume growth can be said to be limited for outpatient care for the whole period. The (positive) contribution is larger for the more recent years, which suggests that heavier types of care had an increased value share in the output volume index and/or showed increasing annual growth rates in the final years of the series.

Figure 4.3 also shows that both output volume and the total hours of care grow much faster than the number of patients with outpatient care. The average annual growth rate for the number of patients is 3.7 percent for the period 1999-2006. It should be noted that the output volume index in Figure 4.3 corresponds with a restricted part of outpatient care, while the patients counted also include patients with day care. This means that we must compare the increase in the number of patients with the output volume index in Figure 4.1. It follows that the output volume of outpatient care per patient has increased very fast. The strong increase of hours of care per patient could be partly caused by a shift from care given by volunteers to funded care.

Figure 4.3 Chained volume indices for a restricted part of outpatient care, total hours of care and patients with outpatient care (1999 = 100). Patients are counted that received outpatient care at any time of a year. Also in this case there are no patient data for 2000-2003 and for 2007. Patient numbers for the period 2000-2003 were estimated by linear interpolation between the numbers for 1999 and 2004. Number of patients in 1999: 380,930.



Also the results in this section and Figure 4.2 in particular show that clear guidelines for product differentiation are needed. Different levels of differentiation can give substantially different volume growth rates. Figures 4.2 and 4.3 confirm again the usefulness of a framework for volume measurement. Such a framework helps to identify invalid volume measures and implausible index numbers (based on patient counts in this case).

4.2 Discussion

In this section, we evaluate the output volume index for elderly care. As we did for hospital care in Section 3.2, the volume index will be evaluated with respect to the product differentiation and the counting measures chosen.

Product differentiation

The output volume index for elderly care uses a product differentiation based on individual health care activities, such as nursing, household and personal care. The level of health care activities is too detailed. Patients receive an indication for one or more health care activities. Patients can be treated adequately only when they receive each type of care specified in their indication. The health care indications or, equivalently, the packages of health care activities should be used as product differentiation.

Counting measures

As data about health care indications are not available in the NZa-data, we could not use health care packages as counting units. Instead, we count the time spent during a year for each health care activity in the packages. The two volume indices are compared in Appendix A. The volume indices give the same result if: (1) the tariffs of the health care packages are expressed in terms of the tariffs of the health care activities, and (2) the times spent on the health care activities in a single health care package remain the same every year for all health care activities and packages. In this case, we can thus use a volume index based on times spent for individual health care activities. This is an important result, as data about health care packages were not available.

Suppose now that a new technology is introduced in a subsequent year, which makes it possible to spend less time on providing some types of care.¹⁰ In Appendix A it is shown that a volume index based on time spent for health care activities underestimates a volume index based on health care packages. This result would clearly emerge for hospital care. However, substitution of clinical treatments by day treatments shortens length of stay, but does not affect output.

It should be noted that all the time spent on each health care activity is counted. It is therefore implicitly assumed that the thresholds regarding patients' health states are reached. The thresholds may involve cleanness of accommodation, personal care and occurrence and severity of pressure sores.

Additivity over inpatient and outpatient care

Unfortunately, we did not have information in order to apply this additivity property, so that inpatient and extramural outputs were treated as different types. Inpatient care is expressed in days, while the greatest part of outpatient care is expressed in hours of care. For a day of care in a home for the elderly, for example, we need to know how many hours of household assistance, personal care and other forms of support are provided. Otherwise, it is not possible to add inpatient and outpatient days.

Although hard conclusions about the effect of differentiation by inpatient and outpatient care cannot be drawn, we have tried to set up different scenarios in order to gain some insight into the possible range of this effect. As example, we consider long-term care in homes for the elderly and its counterpart in outpatient care. In the latter case, we consider household assistance, personal care (dressing and washing patients), support in organising daily life, and nursing.¹¹ Days of care in homes for the elderly are given by the NZa-data. For outpatient care, we used a factor in order to convert hours of care into days. An outpatient day may be any mixture of the four types of care mentioned above. We studied the effects of varying the total number of hours of care per outpatient day on the volume index when inpatient and outpatient care are not differentiated into different types. The results are shown in Figure 4.4

The 'differentiated' volume index in Figure 4.4 is obtained by weighting the year-to-year indices for inpatient and outpatient care according to their value shares. This is in fact the index that is obtained when differentiating output also by inpatient and outpatient care. The solid lines

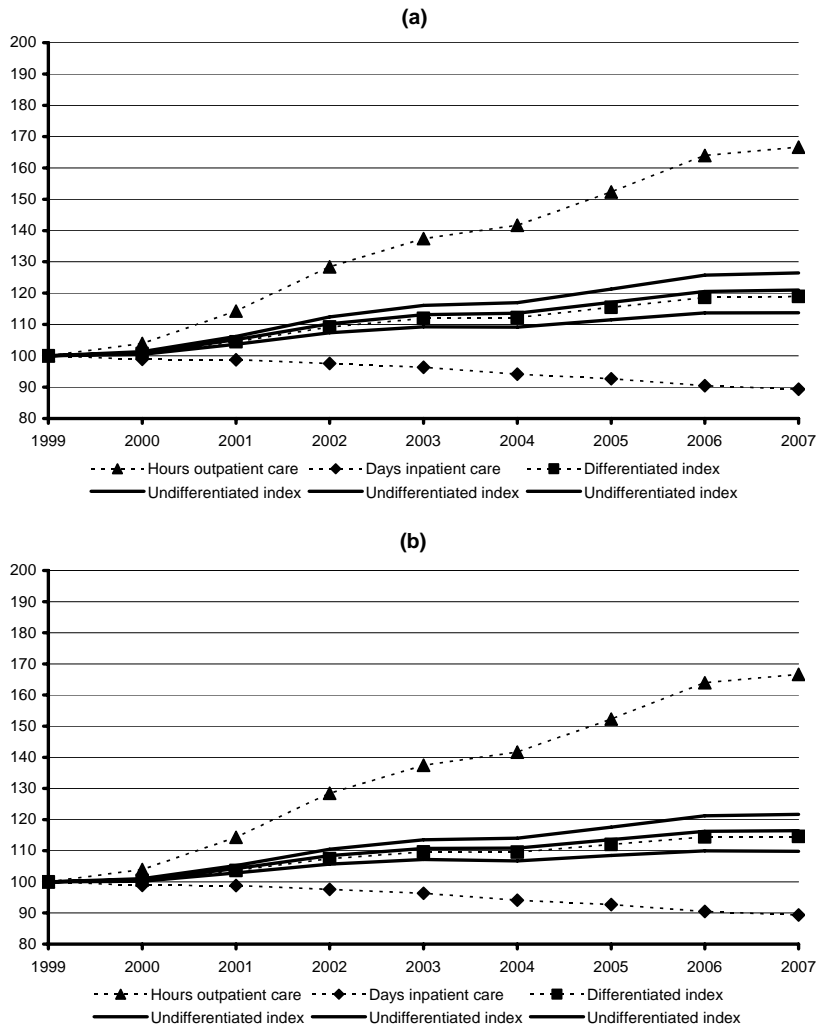
¹⁰ An example of a new technology in home care is a technical device for taking off support stockings.

¹¹ These types of care are indicated in Dutch as: huishoudelijke verzorging (HV), persoonlijke verzorging (PV), ondersteunende begeleiding (OB) and verpleging (VP), respectively.

do not distinguish by inpatient and outpatient care, so that the volume indices are obtained by simply adding inpatient and outpatient days and by calculating the ratio of the sums of these days for two successive years.

The number of hours of care per outpatient day is an uncertain factor in the calculation of the ‘undifferentiated’ index. Figure 4.4 shows the effect on the index for three different values for hours per day. The index becomes smaller when hours per day increase. There will be a smaller number of outpatient days, so that the share of outpatient days on the total number of inpatient and outpatient days decreases. This will move the undifferentiated index towards the index for inpatient days.

Figure 4.4 Chained volume indices for long-term care in homes for the elderly (inpatient) and a counterpart in outpatient care (1999 = 100). Volume indices are also shown for inpatient and outpatient care differentiated. The solid lines show volume indices for undifferentiated output as a function of number of hours of care per outpatient day: 1.5 hours/day (upper line), 2 hours/day (middle line), 3 hours/day (lower line). Also the percentage of total outpatient hours that is combined with elderly home care is varied: 90% in figure (a), 70% in (b).



An idea about the effect of product differentiation into inpatient and outpatient care can be obtained by comparing the differentiated and undifferentiated indices. The largest difference between the differentiated index and the three solid lines is almost 0.8 percentage point on

average per year, in both Figures 4.4-(a) and (b). This difference occurs for 1.5 hours of care per outpatient day. The smallest difference occurs when an outpatient day consists of two hours of care (0.2 percentage point). The results show that the differentiated index lies within the range of the undifferentiated output indices. The differentiated index could thus be considered as a suitable 'point estimate'.

The number of hours of care per outpatient day is not the only source of uncertainty. We also need to know how many hours of outpatient care are reserved for care without treatment (i.e., the equivalent of care in homes for the elderly). For instance, when patients receive a certain treatment at home in combination with household and personal care, then the hours for the latter two types of care should be left out. We have expressed the hours reserved for care without treatment as a percentage of the total hours of care in Figure 4.4 (90% in figure (a), 70% in figure (b)). This percentage is kept fixed over time in the two subfigures. The two percentages used hardly affect the differences between the three undifferentiated indices and the differentiated index. But the four volume indices decrease when the percentage decreases, as the number of outpatient days and their share in overall care become smaller.

While it is true that inpatient and outpatient days as used above are heterogeneous, the results of Figure 4.4 are nevertheless insightful. The effects of output differentiation into inpatient and outpatient care are shown for care without treatment. How big the effect could be for elderly care as a whole is difficult to say. This depends on the extent to which nursing home care can be substituted by outpatient care.

Although the proposed volume index does not fit well into the volume measurement framework of Section 2, the analyses presented in this section show that the volume index gives plausible results under certain conditions. These conditions are related to possible changes in the efficiency of health care provision from year to year, for instance by new technologies, and the composition of days of outpatient care in terms of time spent on health care activities.

Changes in the financing of health care are also taking place for long-term care. In Section 6.2, we will summarise the new financing system based on the so-called ZZP's ('packages by severity of care'). We will discuss whether these packages offer possibilities for improving product differentiation and counting measures.

5. Adjusted output volume measures

5.1 Refinement of the volume measurement framework

In Section 2.2, we introduced the concept of 'health care package' as a counting unit for volume measures. A health care package consists of individual health care activities and specifies times or numbers of times that activities are carried out within a single package. The tariffs of each activity in a package are also specified. Health care packages are differentiated according to the criteria D1-D3 described in Section 2.1: the purpose of a health care process (cure/care), diagnosis or health care indication, and consumer characteristics, such as patient age.

Counting measures that are differentiated according to criteria D1-D3 were called *output volume measures* in Section 2.2. Health care packages are counted when they reach some threshold regarding the health state of a patient. The extent to which a threshold is exceeded does not play a role in the definition of output given here. We could ask ourselves whether output volume should be differentiated further, by taking into account patients' health states and other factors related to health care provision. Output volume measures that are differentiated according to criteria D1-D3 could still be too coarse.

Should health care packages with the same diagnosis or health care indication, patient age, etc., be counted as separate products when they give rise to different health states or could they just be counted as equivalent products? Suppose that the same number of health care packages is

supplied in two subsequent years in elderly care, but that the frequency of severe pressure sores has decreased in the second year. It seems hard to defend that the volume of health care is the same in both years. Of course, patients could reach different health states because of differences in their physical condition. But different health states among patients could also be caused by differences in health care supply. In that case it is reasonable to expect that production volume increased in the second year, because the percentage of patients with less severe or no pressure sores increased.

In this section, the aim is not finding an exhaustive list of ‘quality factors’ that could be used for further product differentiation. We rather state that product differentiation criteria should also contain aspects about a patient’s health state and maybe also other aspects of health care provision. The focus in this section is on how data about health states could be incorporated into the volume measures presented in the previous sections. We propose the following working model.

Product differentiation and counting measures

The central idea is that product differentiation based on criteria D1-D3 is refined. These criteria are considered to be indicators of initial health states. Further refinement is based on final health states, which are classified as D4 in Section 2.1. The structure of the volume measurement framework of Section 2.2 is preserved. Health care packages are counted for every package type that is differentiated, also according to D4 now, such that additivity properties C1 and C2 are satisfied. For example, inpatient and outpatient health care packages for the same diagnosis and patient age, which reach the same health state, should be counted as equivalent packages.

Tariffs of health care packages

In a conventional national accounts setting, counts of different product types are combined into an overall volume measure by the tariffs of the product types. This also holds when combining product counts for different health states. In practice, however, tariffs are the same for different health states. In fact, we have ‘average’ tariffs for health care packages over final health states.

We claim that the ‘average’ tariffs should be re-valued for each health state. These ‘true’ tariffs are related in some way to the known, average tariff of a health care package. Variation in the tariffs of a package over health states could arise from different factors:

- More or less time may be actually spent on different health care activities than is established in the specification of a health care package;
- There could be differences in the tariffs of the individual health care activities, which may arise from heterogeneity in labour force and materials used;
- There could be other factors that are not specified in health care packages, which therefore are not elaborated in the tariffs.

Volume index

The working model described above is formalised in Appendix B, which is used to derive volume indices for different situations. The general expression for the Laspeyres volume index is given by (B6) in Appendix B. Volume index (B7) is a simplified version of (B6). Expression (B7) applies to the situation where the set of possible health states and their effect on the volume index are the same for every package type. We summarise that volume index below.

Let $Q_{J,t}$ denote the number of supplied health care packages of type J in year t , where the package types are the result of product differentiation according to criteria D1-D3. Let health states be grouped into a finite number of classes. For health care package J in year t , we introduce $Q_{J,k,t}$ to denote the number of health care packages that reach health state class k . We write this number as $Q_{J,k,t} = v_{k,t}Q_{J,t}$, where $v_{k,t}$ denotes the fraction, or percentage, of packages of type J that reach health state class k in year t .

Let $P_{J,t}$ denote the tariff of package type J in year t . As is stated above, the tariffs $P_{J,t}$ are in fact average tariffs and do not reflect heterogeneity in health care packages due to different health states. Let $P_{J,k,t}$ denote the tariff that corresponds with health state class k , which we write as $P_{J,k,t} = s_{k,t}P_{J,t}$. The factor $s_{k,t}$ denotes the extent to which the tariff with respect to health state class k deviates from the average $P_{J,t}$ in year t . If $s_{k,t}$ and $v_{k,t}$ are the same across package types for every health state k , then the Laspeyres volume index takes the following expression:

$$(5) \quad \frac{\sum_J P_{J,t-1} Q_{J,t}}{\sum_J P_{J,t-1} Q_{J,t-1}} \left\{ 1 + \sum_k s_{k,t-1} (v_{k,t} - v_{k,t-1}) \right\}.$$

This volume index, which is derived in Appendix B, consists of two parts:

- The ratio in (5) is a Laspeyres volume index for the package counts $Q_{J,t-1}$ and $Q_{J,t}$. This index does not differentiate according to health states. The numbers of packages in years $t-1$ and t are ‘weighted’ with the known (average) tariffs $P_{J,t-1}$;
- The second term is an adjustment factor, which contains two types of information about the health states: (1) the factor $s_{k,t-1}$ according to which the tariffs regarding health state class k deviate from the average tariffs of the package types J in year $t-1$, and (2) the fraction $v_{k,t}$ of health care packages that reach health state class k in year t .

From expression (5) it follows immediately that the adjustment term is equal to 1 when the fractions or percentages $v_{k,t}$ of packages falling in health state k remain the same in two successive years, that is, $v_{k,t} = v_{k,t-1}$ for all health state classes k . This means that differentiation of packages with respect to health state does not give a contribution to volume growth when the distribution of health care packages over the health state classes remains the same from year to year. This result holds irrespective of the values of the tariff factors $s_{k,t-1}$.

Castelli et al. (2007) also propose adjusted volume indices. One of the main differences with our approach is that they use levels of health states directly in volume indices, such as the amount of Quality Adjusted Life Years (QALYs) gained for a unit of output (i.e., package type J). We could use QALYs as health state indicator, but we would subdivide the range of QALYs into health state classes and count the number of health care packages in each class. This is a different approach from the one used in Castelli et al. (2007).

In practice, data are usually available for calculating the Laspeyres volume index for the package counts $Q_{J,t-1}$ and $Q_{J,t}$, or a proxy based on the amount of time spent on the individual health care activities in a package. Data about the two factors in the adjustment term are often unavailable, which is certainly true for the tariff factors $s_{k,t}$. Data on some aspects regarding health state have become available recently for Dutch health care (Halfens et al., 2007). A part of these data are used in the next subsection in order to derive the fractions $v_{k,t}$ for different years. These fractions are used to calculate the possible contribution from differentiation by health state to volume growth. The theoretical result that there is no contribution when $v_{k,t} = v_{k,t-1}$ for all health state classes k will be used as a benchmark.

5.2 Example for elderly care

Since 1998, a Dutch survey on measuring prevalence of health care related aspects (“Landelijke Prevalentiemeting Zorgproblemen”, LPZ) is carried out by the University of Maastricht (Halfens et al., 2007). In the LPZ, data about the prevalence and risk of developing pressure sores, incontinence, undernourishment and other indicators are collected at hospitals, nursing homes, homes for the elderly and in home health care. These health aspects are not contained in the data

used in sections 3 and 4. A part of the LPZ-data is tracked during different years. It is interesting to find out to what extent output volume indices may change when LPZ-data are incorporated.

Pressure sores constitute a major problem in health care. The prevention and treatment of pressure sores is estimated to cover (at least) one percent of the total Dutch health care budget (Severens et al., 2002). The prevalence of pressure sores is measured in the Netherlands since 1998. Prevalence shows quite a stable picture until 2004, but decreases afterwards. It is difficult to say what factors have contributed to this decrease. According to Halfens et al. (2007), health institutes may be giving more attention to pressure sores, which may be influenced by the increased focus on measuring quality indicators for health care.

In this study, we consider pressure sores as the only health aspect. This choice is made merely for illustrative purposes. The prevalence data of Halfens et al. (2007) are subdivided into two classes of pressure sores: one class contains pressure sores for stages of seriousness referred to as 2, 3 and 4, and the second class is the complementary group (stage 1 sores and no pressure sores).¹² We denote these two classes as health state classes $k = 1$ and $k = 2$ respectively, which means that $k = 2$ is the better health state. We only consider pressure sores for nursing home care. The data for elderly home care and home health care are based on small samples and are excluded from the results of Halfens et al. (2007).

In order to compute a volume index that takes into account the distribution of patient counts over the two classes, we make the following choices and assumptions:

- As we do not have data on health care packages (ZZP's), we use the output volume index used in Section 4 as a proxy. The output volume index is based on times spent and numbers of times that health care activities are carried out;
- The prevalence data are not differentiated according to the type of nursing home care. We assume that the prevalence scores are the same for all types of care;
- The prevalence scores are measured on one day of a year. We assume that these scores are representative for the whole year.

We consider the following volume index:

$$(6) \quad \frac{\sum_i p_{i,t-1} q_{i,t}}{\sum_i p_{i,t-1} q_{i,t-1}} \left\{ 1 + \sum_{k=1}^2 s_{k,t-1} (v_{k,t} - v_{k,t-1}) \right\},$$

where $q_{i,t}$ denotes the time spent, or the number of times that activity i is carried out in year t , and $p_{i,t}$ denotes the tariff of health care activity i in year t . The prevalence scores in year t are denoted by $v_{k,t}$ for health state classes $k = 1$ and $k = 2$. The prevalence of pressure sores in class $k = 1$ for nursing home care was about 16 percent in 2003 and in preceding years. This prevalence decreased to about 12 percent in 2004 and 2005, which improved further to 7 percent in 2007.

The only unknown factors in expression (6) are $s_{1,t}$ and $s_{2,t}$, which represent the extent to which the underlying, unknown tariffs of health care packages for health state classes $k = 1$ and 2 deviate from the (known) average tariff in year t . The variation in tariffs over health state classes could be due to different factors, as we discussed in the previous subsection.

We define $s_{1,t}$ and $s_{2,t}$ as follows. Let r denote the ratio of the tariff for health state class $k = 2$ and the tariff for health state 1, which we assume to be a constant ratio for all years t . As $k = 2$ is the better health state, we set $r \geq 1$. This means that the tariff for health state class 1 cannot be

¹² There are four "stages" of pressure sores, also called "pressure ulcers" in the literature: Stage 1 contains sores of a superficial nature, such as red skin. Stage 4 contains the most severe types of sores, which may also consist of damage, and even necrosis, of muscles and bone tissue.

greater than the tariff for class 2. For instance, a better health state could be associated with better prevention schemes. Such information is not specified in health care packages. The parameter r is introduced in order to capture missing information and heterogeneity of health care packages over health states. Expressions for $s_{1,t}$ and $s_{2,t}$ can be derived in terms of r :

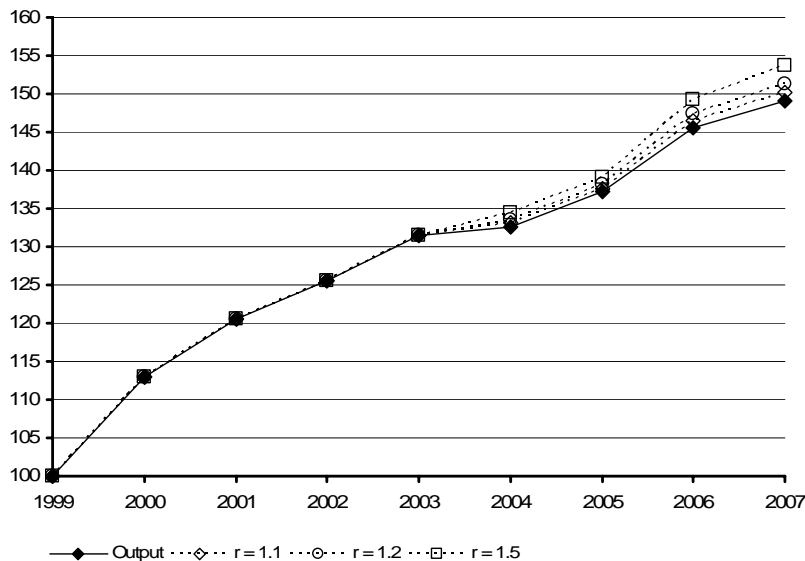
$$(7) \quad s_{1,t} = \frac{1}{v_{1,t} + rv_{2,t}},$$

$$(8) \quad s_{2,t} = \frac{r}{v_{1,t} + rv_{2,t}}.$$

If $r = 1$, then $s_{1,t} = s_{2,t} = 1$, so there is no heterogeneity of health care packages over health states. Patients may still end up in different health states in that case, for example, because of differences in physical condition. It should be emphasised that these differences do not affect the volume index in this example, since r only captures (missing) information about the health care process. Only the contribution from the health care process in reaching different health states is reflected by r .

At this stage, we do not have a method for determining a value or bounds for r . Therefore, we examined the effect of variations in r on the output volume index for nursing home care by performing a sensitivity analysis. The results are shown in Figure 5.1. The output volume index is not affected until 2004, since the prevalence scores are stable. There is no effect on the output volume index in 2005 either, since the prevalences are the same as in 2004. There is a positive contribution to output volume growth in 2004, 2006 and 2007, since the prevalence of the most serious pressure sores decreased in these years.

Figure 5.1 Chained output volume index for nursing home care in the Netherlands and adjusted volume indices based on prevalences of pressure sores in two health state classes, for different values of r (1999 = 100).



The effect of the prevalence of pressure sores on output volume growth is limited to several tenths of a percentage point in Figure 5.1. The average annual output volume growth, which is represented by the solid line in the figure, is 3.2 percent for the period 2003-2007. The average annual growth for $r = 1.1$ and $r = 1.2$ is 3.4 and 3.6 percent respectively. Also for $r = 1.5$ the

contribution to output volume growth is less than one percentage point. A value of $r = 1.5$ means that the tariff with respect to health state class 2 is valued 50 percent higher than for health state class 1.

The above example shows that volume adjustments based on prevalences of pressure sores have an effect on output volume growth. The contribution to output volume growth is greater than zero when $r > 1$. This is an important finding in itself, but also with respect to, for instance, the impact on productivity. The results in Figure 5.1 imply that productivity based on output volume gives a lower bound for productivity.

6. New data sources

In this section, we will review new data sources for hospital care and long-term care. New registration and financing systems have been introduced for these health care sectors, which are based on the Diagnosis Treatment Combinations (DBC's) for hospital care and the Packages by Severity of Care (ZZP's) for long-term care.¹³ We will investigate the possibilities of new data for improving volume measurement. The following questions will be addressed:

- Are DBC's and ZZP's suitable for product differentiation?
- Are DBC's and ZZP's suited as counting units?
- Is it possible to apply additivity over inpatient and outpatient care to counting measures based on DBC's and ZZP's?
- At present, we only have data on DBC's. Can we rely on the data? Put differently, how do the data compare to the HDR-data currently used, and to other data sources?

6.1 The DBC-system

6.1.1 Suitability of DBC's for volume measurement

The DBC-system was introduced as a new financing system for hospital care in 2005.¹⁴ From a financing point of view, care is split up in two parts. The so-called "A-segment" consists of care to which fixed tariffs apply. These tariffs are established by the Dutch Health Authority NZa. The tariffs of the "B-segment" result from negotiations between health care providers and health insurers. The idea behind the B-segment is to allow competition and to move towards 'market prices'. The Dutch government allows the size of the B-segment to increase every year, which covers 34 percent in 2009. The pricing mechanisms of both segments can be seen as attempts towards a different control of expenses. The expenses in hospital care have increased with almost eight percent per year on average in the period 2000-2006.

In principle, a DBC consists of all health care activities for a specific diagnosis. Activities encompass surgery, laboratory examinations and also polyclinic services. DBC's are suited for product differentiation, as they contain information about diagnoses and consumer characteristics like patient age. It should be possible to distinguish between cure and care type of DBC's on the basis of specifications about diagnoses.

In theory, DBC's are suited as counting units of production volume. A DBC is a health care package containing all activities between the start and the end of a treatment. Polyclinic services are part of a DBC, so that DBC's give the possibility to improve volume measurement of hospital care compared with the current method. In the case of chronic diseases, new DBC's are opened

¹³ The Dutch terms for the registration and financing systems are: Diagnose Behandelings Combinaties (DBC's) and ZorgZwaartePakketten (ZZP's).

¹⁴ The curative part of mental health care is also financed now according to the DBC-system.

for every hospital admission. Every DBC should be counted as a volumetric unit, at least, when the corresponding treatment proved to be successful. A point of attention is the fact that DBC's do not span more than one calendar year. If a treatment continues in the next year, then the current DBC will be closed and a new one will be opened in the next year. If the closure of a DBC does not correspond with the end of a treatment, then it should not be counted yet. The question is whether it is possible to trace the follow-up DBC in the data.

It is possible to apply the additivity over provider characteristics to counting measures based on DBC's. This holds for clinical and day treatments, but also for day treatments that can be performed in a polyclinic setting. In the latter case, we obtain an improvement over the current method, as polyclinic services are not contained in the HDR and are not differentiated according to diagnosis and patient age in other data sources.

In theory, there is thus room for improving the current volume measurement method with the use of DBC-data. The question that will be answered in the next subsection is whether the DBC-data that are currently available can be already used for volume calculations.

6.1.2 Data coverage

In order to gain insight into the completeness of the DBC-data, the data have been compared with other data sources on the following points:

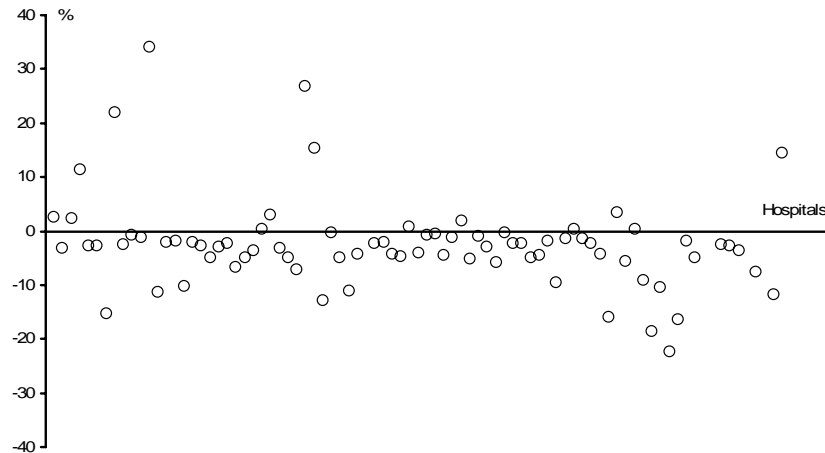
- The total number of hospitalisation days;
- The total number of DBC's;
- Production value.

Hospitalisation days recorded in the DBC-datasets for 2005 and 2006 have been compared with the data in the Hospital Discharge Register. The DBC-data show about 20 percent less days of hospitalisation than in the HDR-data, both for clinical treatments and day treatments. Days of hospitalisation have to be registered only for the system to accept a DBC as clinical or otherwise. The registration of the number of days in the system does not affect the tariff of a DBC. Given the big differences in hospitalisation days, it is likely that hospitals register the mere presence of hospitalisation days in DBC's rather than exact numbers. DBC's are primarily meant for reimbursement purposes. The results do not imply that the DBC-data are incomplete, but rather that hospitalisation days are not appropriate for assessing the completeness of the data.

Counting the number of DBC's offers a better way of assessing the completeness of the data. The number of DBC's was not compared with the number of hospital discharges from the HDR, since the HDR does not contain data about polyclinic services. It was decided to find another source in order to make direct comparisons. The annual reports of hospitals were used for this purpose. Numbers of completed DBC's registered in annual reports are probably taken from hospitals' own registration system instead of the DBC-system.

The DBC-dataset for 2007 contains 4.8 percent less DBC's than the annual reports. The differences per hospital can be much larger, as is shown in Figure 6.1. It is not clear yet what causes certain differences to be large. Values were also computed for the DBC's in the DBC-system and for the annual reports. Tariffs of DBC's were multiplied with the number of reported DBC's and summed over all DBC's. The total value for the DBC-system is about 12 percent smaller than the total value for the annual reports. The analyses that have been done so far show that we are not yet in a position to use the DBC-data for volume measurement. For more details about the analyses of the DBC-datasets, see Spanjaard (2009).

Figure 6.1 Percentual differences between the total numbers of completed DBC's in the DBC registration system and in the annual reports of 81 hospitals in 2007. A positive difference means that the number in the DBC-system is larger.



6.2 The ZZP-system

Institutes for elderly care, long-term mental health care and disabled care were financed on the basis of tariffs and health care delivered for single health care activities, such as nursing days and days of care in nursing homes and homes for the elderly, and household assistance, personal care, support and daily activities provided by home health care organisations. A new financing system based on Packages by Severity of Care (ZZP's) was introduced in January 2009. Clients receive health care indications specified in terms of quantities and type of health care activities, which are bundled in ZZP's. There are 10 ZZP's for elderly care, 13 ZZP's for long-term mental health care and 29 ZZP's for disabled care. Like DBC's, the ZZP's were introduced in order to obtain greater transparency and better control in the financing of care. Until the introduction of ZZP's, tariffs had to be established for several hundreds of single health care activities every year.

Data about ZZP's are not available yet. The first data are expected to be delivered by the NZa in the course of 2009. In spite of this, it is possible to make several comments about the suitability of ZZP's for volume measurement and the improvements that are expected over the current method. ZZP's allow differentiating by health care indication, which will result in a better product differentiation compared with the current method. At present, the single activities are the products differentiated. However, patients often undergo more than one health care activity. As a consequence, ZZP's also provide better counting units of production volume. The counting units are packages of health care activities provided to patients on a weekly basis.

The ZZP-system is introduced only for inpatient long-term care. Outpatient care is still financed on the basis of tariffs for single health care activities. The question is whether it is possible to apply the additivity over inpatient and outpatient care to counting measures. The health care activities in the inpatient ZZP's are specified in the same units as for outpatient care. The question is whether it is possible to find a feasible, unique partition of the quantities of health care activities for outpatient care into outpatient type ZZP's. If so, then we can apply the additivity property to the numbers of inpatient and outpatient ZZP's. We have to await production data on ZZP's in order to investigate the possibilities in this respect.

7. General discussion

Review of the measurement framework

In this paper, we presented a measurement framework for production volume in health care. The framework consists of two parts: criteria for product differentiation, and properties that volume measures should satisfy. The initial framework developed in Section 2 gives rise to measures of *output volume*. Counting measures for output volume, which are based on the notion of *health care package* as counting unit, are additive with respect to:

- (1) The time domain on which volume measures are defined;
- (2) Provider (or process) characteristics;
- (3) Health state reached at the end of a treatment.

The third additivity property means that health care packages are counted when patients reach a medically acceptable threshold regarding their health state. The extent to which a threshold is exceeded does not affect output volume; only the fact whether it is exceeded or not matters.

The measurement framework is extended with the inclusion of health state in Section 5.1. This is done by considering health state as an additional criterion for product differentiation. The measurement framework remains intact; the only implication is that additivity property (3) above is dropped. The extended framework leads in a natural way to an adjusted measure of output volume, where the adjustment term consists of two components:

- (a) The frequencies according to which health care packages are distributed over health states;
- (b) Relative ‘values’ of health care packages that belong to different health state classes.

It is important to stress that the relative values with respect to different health states in (b) only capture contributions from the health care process in reaching health states. Variations in health states across patients due to interpersonal differences are not included. The relative values can be interpreted as summary parameters that quantify the contribution of missing information about activities in health care packages. The parameter r in the case with pressure sores in Section 5.2 is an example of such a parameter. A method for determining the value of r has not been developed yet and is thus left for future research.

In the example of Section 5.2, we used pressure sores as the only health state indicator. However, more indicators could be proposed for certain diagnoses or health care indications. For instance, memory performance or some other behavioural measure could be proposed as health state indicator beside pressure sores for Korsakoff patients. An example of a health state class in this case could be an ‘above average’ memory performance combined with pressure sores in stage 1. The expressions for the volume indices derived in Appendix B hold for general (discretised) health state spaces.

Further research has to be done on which health state indicators to use. Beside the LPZ-study of Halfens et al. (2007), we should closely follow efforts made by others. For instance, the Dutch website www.kiesbeter.nl contains ‘quality charts’ of health care institutes, which distinguish between patient experiences and health aspects measured by health care personnel. Patient experiences involve attention given by personnel to patients and how well a patient’s body is cared. Health aspects involve the occurrence of pressure sores, symptoms of depression, undernourishment, pain after surgery, etc. Mattke et al. (2006) are developing a set of indicators with the intention of proposing these for international comparisons.

Evaluation of current measures

The volume measures used for hospital health care stand closer to the measurement framework than the measures used for long-term care. Counting discharges is the same as counting health

care packages in many cases, while the counting measures are additive over clinical and day treatments. The contribution of this property to volume growth is large: the average annual volume growth for clinical and day treatments is 3.5 percent over the period 1995-2007, which decreases with 1.5 percentage points when discharges would also be differentiated according to clinical and day treatments. A drawback of the method is that polyclinical services are treated as distinct products. The data used so far do not specify polyclinical services by diagnosis and the Hospital Discharge Register does not contain such data. However, polyclinical and day treatments should be added and not differentiated when they can be used for the same diagnosis. In this respect, the DBC-data offer possibilities for improving the current volume measures.

The volume measures for long-term care use times spent on health care activities, or the number of times that an activity is carried out, as counting measures. The health care activities are the 'products' differentiated. In addition, the counting measures are not additive over provider characteristics, as inpatient and outpatient care are treated as types of care that lead to different products. We were forced to make these choices because the data did not yet contain information about health care packages. In Section 4.2, it was argued that the volume indices for elderly care give accurate results under certain conditions. However, as ZZP's coincide with the notion of health care package, it is desirable to use the data on ZZP's in the future. This holds in particular when times spent on health care activities in a ZZP change from year to year, for instance because new technologies are introduced.

The volume measures used for health care at Statistics Netherlands are measures of output volume. The measures are not differentiated according to final health states of patients. Also in this case, data availability has been a problem. Data about health state indicators, such as the prevalence of pressure sores, have become available recently. A part of the data presented in Halfens et al. (2007) contains time series, so that these are suitable for inclusion into volume measures. The data in Halfens et al. (2007) show a shift to less serious pressure sores in nursing homes, which results in a positive contribution to volume growth (Section 5.2). Although the contribution seems limited to several tenths of a percentage point, data about health states should be included when these are available. Both volume and productivity growth can be calculated more accurately.

Applicability to other sectors: Education

An interesting question is whether the measurement framework of sections 2 and 5.1 could be applied to other sectors. We limit the discussion on this point to some hints about the applicability of the framework to education.

The criteria for product differentiation listed in sections 2.1 (D1-D3) and 5.1 ('health state') can be transferred to education services. Knowledge acquisition can be considered as the purpose of instruction. The types of care for which patients obtain indications now become school types (primary, secondary, higher education) and grades. An example of consumer characteristics is socio-economic environment of students. Instances of provider or process characteristics could be public versus private education and different instruction methods. 'Health state', which was proposed in Section 5.1 for further product differentiation, would now become 'knowledge state'. Knowledge state classes could be created on the basis of marks for exams.

Production volume of education should be differentiated by socio-economic background of students. Initial knowledge may vary with socio-economic background. It is probably difficult to acquire data about initial knowledge levels. If data about socio-economic background can be obtained, then these could be used as indicators for initial knowledge. Volume measures should capture the contribution of instruction to the change in knowledge state. If the final state is equal to the initial state, for instance, because the subject matter for a certain exam was known before instruction, then there is no contribution to volume.

The two additivity properties imposed on counting measures for health care in Section 2.2 should also be imposed on counting measures regarding production volume for education. A

consequence of additivity over time is that counting students is not a valid measure, like counting patients in long-term care is incorrect. On the other hand, also in the case of education, the subject that undergoes the production process should be the point of departure in the construction of counting measures. There will be no production volume when there are no students. The number of students must therefore be adjusted in some way in order to obtain a valid counting measure. For example, student months is a well-defined counting measure. A month could be conceived of as a counting unit for subject matter or 'instruction package'.

Additivity over provider characteristics should also be imposed. The objective is not to measure processes. Counting measures for each combination of knowledge state, school type, grade and consumer characteristics should be added over different instruction methods and public/private education.

The volume measurement framework presented in this paper should be applicable to other sectors as well. The starting point of setting up a state space of differentiated products, with dimensions based on initial and final states, should be transferable to other services and industries beside health care and education.

Main recommendations

The volume measures used for hospital care and long-term care at Statistics Netherlands have been analysed and discussed within the context of the measurement framework of sections 2 and 5.1. One of the main findings is that the method for hospital care stands closer to the framework than the methods used for long-term care. However, there is room for improvement with regard to both methods. We end this study by summarising the main recommendations that were made in this respect:

- The data on health care packages that will become available for long-term care can be used to improve the current volume measures, as a ZZP coincides with the concept of counting unit that we suggested. ZZP's are introduced only for inpatient care, so that the additivity over inpatient and outpatient care could still be left as an unresolved problem. Nevertheless, an improvement over the current methods for long-term care can be expected;
- We would also recommend the use of DBC's for hospital care. Polyclinic services are treated as separate products in the current method, a problem that will be resolved by the use of DBC's as counting units. However, more research has to be done on the data, as we do not yet feel confident about their use in volume measurement;
- The example of Section 5.2 on pressure sores in nursing homes shows that the inclusion of data on health states into volume measures may lead to an adjustment of volume growth. The contribution is small in the example, but should not be ignored. More research should be done on the availability of health state information and its inclusion into volume measures. In particular, a method for assessing the contribution of health care processes to reaching different health states should be developed. This refers to the parameter r in the example of Section 5.2.

Acknowledgements

This study was carried out with Eurostat Grant agreement theme 40, *Improvement of quality of national accounts*, no. 40100.2007.001-2007.319. The author is greatly indebted to Marcel Spanjaard for providing the results of his analyses of the DBC-data. Numerous other colleagues at Statistics Netherlands made suggestions for improving the manuscript, for which the author expresses his gratitude.

References

- Aizcorbe, A.M., Retus, B.A., and Smith, S. (2008). Toward a health care satellite account. BEA Briefing, May 2008.
- Ash, R.B. (1972). *Real Analysis and Probability*. Academic Press, Inc., San Diego.
- Berndt, E.R., Cutler, D.M., Frank, R.G., Griliches, Z., Newhouse, J.P., and Triplett, J.E. (2001). Price indexes for medical goods and services: An overview of measurement issues. In: Cutler, D.M. and Berndt, E.R. (Eds), *Medical Care Output and Productivity, Studies in Income and Wealth*, vol. 62, pp. 141-198. University of Chicago Press.
- Castelli, A., Dawson, D., Gravelle, H., and Street, A. (2007). Improving the measurement of health system output growth. *Health Economics*, **16**, 1091-1107.
- Chessa, A.G. (2007). Price and volume indices for elderly care: New methods and results based on NZa-data 1998-2006. Internal report, Statistics Netherlands. (In Dutch)
- Chessa, A.G. (2008). Care for disabled patients grows rapidly. In *Gezondheid en Zorg in Cijfers 2008*, pp. 51-62. Statistics Netherlands. (In Dutch)
- Chessa, A.G. (2009). A new index method for price and volume of care for disabled patients in the Netherlands. Internal report, Statistics Netherlands.
- Chessa, A.G., and Kleima, F.J. (2006). The Dutch experience in measuring health output and labour productivity. Discussion paper 06002, Statistics Netherlands. (Paper presented at the joint OECD/ONS/Government of Norway workshop *Measurement of non-market output in education and health*, London, October 3-5, 2006.)
- Chessa, A.G., and Okkerse, M. (2007). Elderly care grows rapidly. In *Gezondheid en Zorg in Cijfers 2007*, pp. 81-89. Statistics Netherlands. (In Dutch)
- Eurostat (2001). Handbook on price and volume measures in national accounts.
- Halfens, R.J.G., Meijers, J.M.M., Neyens, J.C.L., Offermans, M.P.W., Hoeymans, J., Schols, J.M.G.A., van Bokhorst-de van der Schueren, M.A.E., Joosten, K.F.M. and Hulst, J.M. (2007). *Report of prevalence measurements in Dutch health care*. Research institute Caphri, Department of Health Care and Nursing Sciences, University of Maastricht, The Netherlands, November 2007. (In Dutch)
- Kleima, F.J., Warns, P., and Opperdoes, E. (2004). Constructing a volume index for hospital services in the Netherlands. *Statistical Journal of the United Nations ECE*, **21**, 291-297.
- Mattke, S., Kelley, E., et al. (2006). Health care quality indicators project; initial indicators. Report OECD, *Health Working Papers*, no. 22.
- OECD (2008). *Towards Measuring the Volume of Health and Education Services*. OECD, Draft Handbook, October 2008.
- Schreyer, P. (2008). Output and outcome – Measuring the production of non-market services. Paper presented at the 30th General Conference of The International Association for Research in Income and Wealth, Portorož, Slovenia, August 24-30, 2008. The paper can be downloaded from the website <http://www.iariw.org/>
- Severens, J.L., Habraken, J.M., Duivenvoorden, S., and Frederiks, C. (2002). The cost of illness of pressure ulcers in the Netherlands. *Advances in Skin & Wound Care*, **15** (2), 72-77.
- Spanjaard, M. (2009). Can the Dutch DBC system be used for statistics? Internal report, Statistics Netherlands.

Appendix A

A comparison of output volume indices when counting health care activities or packages

As we argued in Section 2.2, our preferred counting measure for health care production is the number of ‘health care packages’ that patients receive. These numbers should be differentiated according to criteria D1-D3 (purpose of process, diagnosis/ health care indication and consumer characteristics). However, data about packages are not always available. Suppose that the data consist of the time spent on health care activities or the number of times that activities are carried out during a year. Health care packages consist of one or more activities.

In this appendix, we compare a volume index based on numbers of packages with a volume index based on quantities (time spent or frequency) for health care activities. We will show where differences arise between the two volume indices and we also derive conditions under which the volume indices give the same results.

We introduce the following notation. Let G be the set of individual health care activities, $q_{i,t}$ the total quantity delivered for activity $i \in G$ in year t and $p_{i,t}$ the tariff per unit of quantity for activity $i \in G$ in year t . We assemble the activities in G into health care packages as follows. A package J in year t consists of quantities $q_{i,J,t}$ for one or more activities $i \in G$. An elderly patient with a health care indication for package J , for example, may receive $q_{1,J,t}$ hours of household assistance (activity 1), $q_{2,J,t}$ hours of personal care (activity 2) and $q_{3,J,t}$ afternoons in a day-care centre (activity 3) per week, say, in year t .

The quantities $q_{i,J,t}$ for the activities i may vary from year to year for the same package. For instance, the amount of time spent for certain activities may decrease in subsequent years by the introduction of new technologies. In practice, not all the health care activities will be part of a package, so that $q_{i,J,t} = 0$ for some activities $i \in G$ and health care packages J .

The volume index based on numbers of packages is derived below by making the following simplifications:

- Health care packages fall within one calendar year. In other words, beginning and end of a process both fall in the same year;
- The sets of health care activities and packages are the same in every year. We thus exclude situations with disappearing and entering activities and packages;
- The tariffs of the health care activities do not change within a year.

The expressions that are derived below can be generalised when these three restrictions are dropped. The formulas become more complex, so we make the simplifications merely for reasons of convenience and clarity of exposition.

We denote the full set of health care packages by H . The number of times that package J is delivered in year t is denoted by $Q_{J,t}$, while $P_{J,t}$ is the tariff of package J in year t . This tariff is related to the tariffs of the health care activities as follows:

$$(A1) \quad P_{J,t} = \sum_{i \in G} p_{i,t} q_{i,J,t},$$

for all packages $J \in H$ and all years t . The tariff of package J is thus obtained by multiplying the quantities of the individual activities in the package by their tariffs and by summing these monetary values over the types of activities.

The total quantity $q_{i,t}$ delivered for activity $i \in G$ in year t results by summing the quantities delivered for activity i per health care package type over all the package types, that is:

$$(A2) \quad \sum_{J \in H} q_{i,J,t} Q_{J,t} = q_{i,t},$$

which holds for all activities $i \in G$ and all years t .

The Laspeyres volume index in year t , with respect to year $t - 1$, which adopts packages as counting units, is defined as

$$(A3) \quad \frac{\sum_{J \in H} P_{J,t-1} Q_{J,t}}{\sum_{J \in H} P_{J,t-1} Q_{J,t-1}}.$$

Substitution of expression (A1) for year $t - 1$ into the denominator of the volume index gives

$$(A4) \quad \sum_{J \in H} \sum_{i \in G} p_{i,t-1} q_{i,J,t-1} Q_{J,t-1},$$

which can be rewritten, by reversing the order of summation, as

$$(A5) \quad \sum_{i \in G} p_{i,t-1} \sum_{J \in H} q_{i,J,t-1} Q_{J,t-1}.$$

The inner sum is equal to expression (A2) for year $t - 1$, so that (A5) simplifies to

$$(A6) \quad \sum_{i \in G} p_{i,t-1} q_{i,t-1}.$$

The health care packages are valued according to the tariffs of the individual activities, so that the total value of the health care packages delivered in year $t - 1$ is equal to the total value of the quantities delivered for the health care activities.

The numerator of volume index (A3) has a somewhat more complex form. By following the steps for the denominator that lead to expression (A5), we obtain for the numerator

$$(A7) \quad \sum_{i \in G} p_{i,t-1} \sum_{J \in H} q_{i,J,t-1} Q_{J,t}.$$

The inner sum of this expression consists of terms $q_{i,J,t-1} Q_{J,t}$. This sum represents the quantity of activity $i \in G$ that would be delivered in year $t - 1$ with the numbers of health care packages supplied in year t . If $q_{i,J,t-1} = q_{i,J,t}$ for all $i \in G$ and $J \in H$, then, by making use of expression (A2), it follows that the volume index simplifies to

$$(A8) \quad \frac{\sum_{i \in G} p_{i,t-1} q_{i,t}}{\sum_{i \in G} p_{i,t-1} q_{i,t-1}},$$

which is the Laspeyres volume index for the quantities $q_{i,t}$ and $q_{i,t-1}$ of the health care activities. Calculation of (A8) thus gives the desired volume index. If the quantities of the activities do not change over time for all the packages, then volume indices (A3) and (A8) give the same result.

Suppose now that, for some reason, $q_{i,J,t} < q_{i,J,t-1}$ for some activities i and packages J , while $q_{i,J,t} = q_{i,J,t-1}$ otherwise. Situations where $q_{i,J,t} < q_{i,J,t-1}$ may result from the introduction of new technologies in year t , which may reduce the time required for performing some activity. This case implies that volume index (A8) underestimates the desired volume index (A3). It is easily verified that numerator (A7) of volume index (A3) is greater than the numerator of (A8). The opposite holds if $q_{i,J,t} > q_{i,J,t-1}$ for some health care activities i and packages J , and $q_{i,J,t} = q_{i,J,t-1}$ otherwise. The same findings can be obtained for other types of volume indices, such as the Paasche and Fisher index.

The above findings also hold when the sets G and H of health care activities and packages change over time, due to new and disappearing activities and packages, and when the tariffs $p_{i,t}$ of the health care activities vary within a calendar year. It is not difficult to prove that volume indices (A3) and (A8) give the same results when the quantities $q_{i,J,t}$ per package remain constant in time and that volume index (A8) underestimates (A3) when $q_{i,J,t} < q_{i,J,t-1}$, for some activities i and packages J . These results do not necessarily hold when the health care activities of a package cover more than one year together ('censoring effects'). In general, it can thus be stated that volume indices (A3) and (A8) are equal when the censoring effects for health care packages are negligible.

Appendix B

Adjusted output volume measures

In order to illustrate how ‘quality’ aspects, such as heterogeneity of final health states, can be incorporated into volume measures, we make use of the same notation as in Appendix A. We start from expression (A3), which is a Laspeyres volume index based on counts of health care packages.

Product differentiation by health state is represented here as follows. For every health care package $J \in H$, we introduce a set K_J of (discretised) health states. Let $Q_{J,k,t}$ denote the number of supplied health care packages of type $J \in H$ that belong to health state class $k \in K_J$ in year t , and let $P_{J,k,t}$ denote the tariff of these packages in year t . A Laspeyres volume index similar to (A3), which now also differentiates according to health state class, can be defined as follows:

$$(B1) \quad \frac{\sum_{J \in H} \sum_{k \in K_J} P_{J,k,t-1} Q_{J,k,t}}{\sum_{J \in H} \sum_{k \in K_J} P_{J,k,t-1} Q_{J,k,t-1}}.$$

The number of packages $Q_{J,t}$ in Appendix A is equal to the sum of the health care packages over the health state classes in K_J , for all package types $J \in H$ and every year t :

$$(B2) \quad Q_{J,t} = \sum_{k \in K_J} Q_{J,k,t}.$$

The tariff $P_{J,t}$ in Appendix A is in fact an average over the health state classes in K_J , since the following equality must hold, for all $J \in H$ and t :

$$(B3) \quad P_{J,t} Q_{J,t} = \sum_{k \in K_J} P_{J,k,t} Q_{J,k,t}.$$

We now express the tariffs $P_{J,k,t}$ in terms of the average tariff $P_{J,t}$ for package J , such that $P_{J,k,t} = s_{J,k,t} P_{J,t}$. The factor $s_{J,k,t}$ thus expresses the extent to which the tariff for health state class $k \in K_J$ deviates from the average tariff.¹⁵ We also introduce $v_{J,k,t}$, which denotes the fraction of health care packages of type $J \in H$ that fall in class $k \in K_J$ in year t . In other words, we have $Q_{J,k,t} = v_{J,k,t} Q_{J,t}$. Equalities (B2) and (B3) can now be rewritten as:

$$(B4) \quad \sum_{k \in K_J} v_{J,k,t} = 1$$

and

$$(B5) \quad \sum_{k \in K_J} s_{J,k,t} v_{J,k,t} = 1.$$

Volume index (B1) can now be rewritten in the following form:

¹⁵ We allow the tariffs of health care packages to vary over health states. If we would use the same tariff $P_{J,t}$ for different health states, then it can be easily verified that (B1) is equal to (A3). In other words, differentiated counts of health care packages by health state would then not lead to an adjustment of the undifferentiated index (A3).

$$(B6) \quad \frac{\sum_{J \in H} P_{J,t-1} Q_{J,t} \sum_{k \in K_J} s_{J,k,t-1} v_{J,k,t}}{\sum_{J \in H} P_{J,t-1} Q_{J,t-1}}.$$

The denominator of this volume index is simplified by substituting expression (B3) into the denominator of (B1).

The form of (B6) is similar to (A3), that is, to the volume index that does not differentiate according to health state class. The difference between the two volume indices is represented by the sum over the set K_J of health state classes in the numerator of (B6). This term can be seen as a ('quality') adjustment for the number $Q_{J,t}$ of health care packages of type J in year t . The information needed to make the adjustment consists of the distribution of the packages over the health state classes of K_J , in terms of the fractions $v_{J,k,t}$, and the corresponding 'price relatives' $s_{J,k,t-1}$ with respect to the average tariff $P_{J,t-1}$ of package J in year $t - 1$. We conclude this appendix by discussing some special cases of volume index (B6).

Situations without 'quality' effects

If the adjustment term in the numerator of (B6) is equal to 1 for all health care packages J , then volume index (B6) is equal to volume index (A3). There will be no effects on the volume index due to the distribution of health care packages over different health state classes. There are cases where the adjustment term is equal to 1, also when different health state classes are nonempty for the same package J . For instance, this occurs when the fractions $v_{J,k,t}$ do not change over time for every health state class $k \in K_J$, for each package J . From equality (B5) it follows that the adjustment term is equal to 1.

The result just stated is important from a practical point of view. Data about quality indicators are still limited. Distributions of health care packages over health state classes are hard to find. Data are usually available at an aggregate level, while tariffs are averaged over health state classes. When health experts believe that distributions of supplied health care packages do not change significantly over time, the adjustment term could be set to 1 so that volume index (A3) would be suggested.

The adjustment term is close or equal to 1 also when the tariffs for the health state classes are close or equal to the average tariff, for every package type. No matter how health care packages are distributed over health state classes, health care packages belonging to different health state classes are counted together when there is no difference among the tariffs for the health state classes.

Identical adjustments for packages

Volume index (B6) can be simplified somewhat in cases where the sets K_J of health state classes are the same for all packages J . Let us consider situations where $s_{J,k,t}$ and $v_{J,k,t}$ have the same values for every package type $J \in H$, for each $k \in K_J$ in any year t . The set K_J can be denoted by K , since the subscript for package type J is now irrelevant. Equal values for $s_{J,k,t}$ over package types means that the tariffs $P_{J,k,t}$ deviate from the average tariff $P_{J,t}$ by the same percentage for every package J . Equal values for $v_{J,k,t}$ means that different treatment or package types J give rise to the same distribution of supplied packages over health state classes. The frequencies of health states thus give the same distribution, irrespective of diagnosis or health care indication, patient age, etc.

Volume index (B6) can be written as follows:

$$(B7) \quad \frac{\sum_{J \in H} P_{J,t-1} Q_{J,t}}{\sum_{J \in H} P_{J,t-1} Q_{J,t-1}} \left\{ 1 + \sum_{k \in K} s_{k,t-1} (v_{k,t} - v_{k,t-1}) \right\}.$$

We have dropped the subscript for package type J from the terms in the sum over the set of health state classes K . Furthermore, we made use of equality (B5) in order to obtain the form within brackets in (B7). Expression (B7) is of interest in cases where data about health states are available only at aggregate level, that is, not for the individual package types.

In addition, if data about supplied numbers of health care packages are not available, but only the time or number of times that health care activities are performed, then we could substitute volume index (A8) into (B7) and use the following volume index:

$$(B8) \quad \frac{\sum_{i \in G} P_{i,t-1} q_{i,t}}{\sum_{i \in G} P_{i,t-1} q_{i,t-1}} \left\{ 1 + \sum_{k \in K} s_{k,t-1} (v_{k,t} - v_{k,t-1}) \right\}.$$

It is legitimate to use this volume index when the times $q_{i,t}$ reserved for the health care activities i in a health care package do not change from year to year.

Volume index under threshold

We conclude this appendix with some remarks about the way in which the expressions derived above are affected by the inclusion of thresholds for patients' health states. Thresholds can be accommodated in an easy way in expressions (B1)-(B7). Setting a threshold for health care package J is considered here as counting only those packages that belong to the health states of some subset L_J of K_J . The states k that do not belong to L_J , in other words $k \in K_J \setminus L_J$, do not satisfy the threshold.

Expressions (B1)-(B7) keep the same form when $Q_{J,t}$ is (re-)defined as the number of packages of type J that are counted only in the health state classes L_J . The tariff $P_{J,t}$ then applies to the average tariff with respect to L_J . It is useful to note that volume indices (B6) and (B7) consist of two parts: one part gives the contribution to the volume index by merely satisfying the threshold, while the second part takes into account the distribution of health care packages over those health state classes that satisfy the threshold. With regard to expression (B8), it should be noted that only those quantities $q_{i,t-1}$ and $q_{i,t}$ of health care activities should be counted, which correspond with packages that satisfy the threshold.

Price index under threshold

Health state thresholds thus give rise to a volume index that is restricted to health state sets $L_J \subseteq K_J$, with $J \in H$. The index of value change, however, applies to the complete sets K_J . The price index, which is the ratio of the value and volume index, can be written as

$$(B9) \quad \frac{1 + \left(\sum_{J \in H} \sum_{k \in K_J \setminus L_J} P_{J,k,t} Q_{J,k,t} / \sum_{J \in H} \sum_{k \in L_J} P_{J,k,t} Q_{J,k,t} \right) \sum_{J \in H} \sum_{k \in L_J} P_{J,k,t} Q_{J,k,t}}{1 + \left(\sum_{J \in H} \sum_{k \in K_J \setminus L_J} P_{J,k,t-1} Q_{J,k,t-1} / \sum_{J \in H} \sum_{k \in L_J} P_{J,k,t-1} Q_{J,k,t-1} \right) \sum_{J \in H} \sum_{k \in L_J} P_{J,k,t-1} Q_{J,k,t-1}}.$$

The second term is a Paasche price index with respect to the sets L_J . The first term contains the total values, in years $t - 1$ and t , of health care packages that do not satisfy the threshold. As can be seen in (B9), these values end up in a kind of adjustment term that affects the price index, but not the volume index. Health care packages that do not satisfy health state thresholds are not counted and do not contribute to the volume index. The total production value generated by these packages translates into an overall price effect.