4 Financial arrangements in the experiment and design of the research project

In the first section of this chapter, the financial arrangements in the experiment are described in more detail. It is shown which conditions had to be fulfilled by each project and which aspects were at the discretion of the projects. Differences between the regional projects are briefly described. The details of these differences are given in Appendix 4.1. Section 4.2 is concerned with the research design.

4.1 Financial arrangements

4.1.1 Introduction

The ‘Experiment Specialistenhonorering’ (ESH) consisted of five projects. In Chapter 3, the goals of the experiment were described. In short, these goals were: the possibility for medical specialists to work in peace; the harmonisation of the interests of the specialists and the hospital; efficiency; cost control; quality of care; and, accessibility of care. Some of the participants in the projects believed that the abolition of the FFS system would improve efficiency, since under FFS the efficient treatment of patients can have negative financial consequences for specialists. Others, especially some of the specialists, stated that they expected no differences in efficiency, since under the old system specialists had not been influenced by financial considerations.

Each project was - within certain (financial) boundaries - free to create an alternative payment system for medical specialists. The goals of the different experimental projects were reflected in the arrangements that were made. For example, in some projects harmonisation of the interests of the specialists and the hospital was considered very important. These projects were more likely to select a system for determining specialist budgets that resembled the budgeting system of the hospital. In other projects, the priority was that specialist budgets would not be influenced by the intensity of treatment. The underlying notion was that specialists would then be free to concentrate upon purely medical considerations, and not be distracted by financial considerations. These projects saw this as a way to improve efficiency. In such projects, the number of new patients was an important determinant of the specialist budget, since the
participants believed that this variable could not be influenced by choices that the specialists made (apart from giving good quality care).

In this section, the blueprints of the five projects are described. The details of the financial outlines for each project are given in the Appendix 4. There, the financial arrangements are stressed; less attention is given to the organisation of the projects.  

The arrangements of the five projects have a number of common elements, which are described in Section 4.1.2. Section 4.1.3 contains a summary of the different aspects that make up the financial arrangements of the projects. In Section 4.1.4, an overview of the plans is given and concluding remarks are made.

4.1.2 Common elements and boundaries

The project had three common elements: external financing; efficiency and quality projects; and billing according to the FFS-system. These common elements and the conditions are described in this section.

Conditions

Before the local agreements were signed, intensive talks among the three parties (specialists, hospital managers and health insurers) took place. These agreements had to satisfy the general requirements of the Zfr, as described in Chapter 3. In short, the criteria were that: the project had to stimulate suitable use of care; there would be no unintended shift of patient streams to other hospitals or private clinics; and, the proposals had to fit within the ‘acceptable financial level of 1994’, as given by the Ministry of Welfare, Public health and Culture (as it was then called). The starting point for developing proposals was that, in the new payment system, the direct link should be severed between the income of medical specialists and the number of diagnostic and therapeutic services they performed. The proposals had to pay attention to efficiency projects in order to stimulate suitable use of care. The local talks led to five concrete project proposals, which were subsequently given the status of experiment for a period of three years.

52 For a more detailed description of the organisation, see SEO and Ipso Facto (1998).
External financing

An important part of the project proposals dealt with the financial arrangements. In order to understand the financial structure of the plans, three elements are important: the distinction between external and internal financing; the internal determination of the starting budget of a partnership of medical specialists; and, the way in which budgets for a partnership change in the years after the start of the experiment. The projects were free to determine the internal starting budget per partnership and the adjustment of this budget over time. These aspects will be discussed in more detail in the next section.

Each project had an external and an internal financial framework. The external budget was the sum that the project was guaranteed to receive. For the internal budget there was no guarantee. The total external budget and the total internal budget could differ from each other, depending upon the way a project determined the internal budget. If the internal budget was higher than the external budget, a solution for this difference had to be found within the project. For example, in the project ‘Ziekenhuizen Noord-Limburg’, there was a deficit in the first year. This deficit could be closed in the following years, because the participants had come to strict agreements about the adjustment of the partnership budgets from year to year.

The starting budget for the medical specialists was determined externally by the COTG (see Section 3.1) for every hospital in the experiment. The budget in the initial year 1995 was based on the revenues of the partnerships of specialists in a base year, to be chosen by the project. In every project, one base year had to be chosen for the whole project: 1992, 1993 or 1994. So it was not allowed for every partnership within a particular project to select the year that was most favourable for that partnership. The volume of specialist care in the chosen base year was expressed in the tariffs of 1 April 1994. This meant that the general tariff reduction of 1993, as well as the differentiated tariff reduction of 1994, were influencing the external COTG-budget, as mentioned in Chapter 3. The revenues were, depending upon the chosen base year, corrected for the autonomous growth of volume in 1992, 1993 and/or 1994 (respectively, 0.75%, 0.75% and 1.0%), and for the 1995 trend in wages and prices. A productivity discount of 1.04% could be reinvested by the projects and was therefore not deducted from
the starting budget. Lastly, the number of 'vacancies for medical specialists' was added to the budget, expressed in full-time equivalents and taking into account the date of taking up the appointment. For each full-time equivalent, the starting budget was raised by 350,000 Dutch guilders.

The external budget of a project in the year following the start-year was the budget determined by the COTG, corrected for the FOZ percentage of growth, and for the growth of wages and prices considered relevant for the tariffs of specialist care. The FOZ growth percentage is the approved growth for health care costs that is determined by the government.

The budget that was available for the projects for the initial year 1995, as determined by the COTG, was supplemented with the subsidy of the Zfr. In Chapter 3, it was explained that this subsidy was necessary to ensure the participation of specialists in the project, since the 1993 and 1994 tariff reductions were reflected in the COTG-budget. The subsidy of 15 million Dutch guilders consisted of a fixed part of 5 million guilders and a variable part of 10 million guilders. Every project received 1 million guilders from the fixed part; the variable part was distributed according to the volume of specialist care in 1992 (expressed in tariffs of 1994). The amount of subsidy was adjusted each year to the rise in prices (following the Dutch Central Bureau of Statistics price-index of total consumption of families). The COTG-budget and the subsidy together formed the external financial framework for the experiment.

Efficiency projects

Another common element of the projects is that attention was not only given to the payment system for specialists, but also to elements of 'suitable use of care'. In practice, the effect was that all experimenting hospitals have developed projects designed to enhance quality and efficiency. The basic goals of these projects were: better allocation of care by stimulating transmural care; and, improved efficiency in the hospital. The latter can result in providing more or

\[53\] The number of places for medical specialist that was approved by the government but not filled in.

\[54\] This is reported in a Ministry of VWS publication, called Financieel Overzicht Zorg (FOZ, Financial Review of the Care Sector).

\[55\] The subsidy was officially destined to finance the implementation costs of the experiment, but, in practice, functioned more as a (partial) compensation for the tariff reductions.
better care at the same cost, or in offering the same amount and quality of care at reduced cost.

Billing

A final common element of the five projects was that all specialists continued to bill their services in the traditional way. This means that, in sending out their bills, the specialists continued to use the FFS system, even though their income was determined in another way, i.e. budgeted. These bills were presented to a central point for each project or, at the least were centrally administered. A possible difference between the sum of the 'FFS bills' and the external budget led to an adjustment of the tariff for a day of hospitalisation, subject to the approval of the COTG.

Provided that the FFS-system was abolished and the conditions of the Zfr and the Ministry were satisfied, the projects were free to shape the payment systems of medical specialists according to their own ideas. They could choose the degree of variability of the fee of medical specialists. The aspects that could be determined by the projects themselves are the subject of the next section.

4.1.3 Individual aspects of the financial agreements

This section summarises the aspects of the financial arrangements on which the projects were allowed to make their own decisions. Each project had to come to an agreement about the subjects described in Table 4.1.
Table 4.1 Individual aspects of financial arrangements

<table>
<thead>
<tr>
<th>aspect</th>
<th>possible solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  external base-year</td>
<td>1992, 1993 or 1994*</td>
</tr>
<tr>
<td>2  determination of internal starting budget per partnership</td>
<td>division of external budget</td>
</tr>
<tr>
<td></td>
<td>or determination independent of external budget</td>
</tr>
<tr>
<td>3  production parameters</td>
<td>following FB-system: Functionally-oriented Budgeting system of the hospital</td>
</tr>
<tr>
<td></td>
<td>or oriented to number of new patients</td>
</tr>
<tr>
<td>4  influence of production growth upon partnership budget in the following year</td>
<td>a partly or 100%</td>
</tr>
<tr>
<td></td>
<td>b possible use of bandwidth</td>
</tr>
<tr>
<td></td>
<td>c role of production agreements</td>
</tr>
</tbody>
</table>

* These possibilities were determined by the government. For the other aspects, the projects could have chosen other possibilities than those mentioned here.

Firstly, a base-year had to be chosen for the external budget, as described in 4.1.2 above.

**Internal starting budget of a partnership**

Secondly, a typical characteristic of the experiment was the freedom of the projects to determine the starting budgets per partnership of specialists in the start-year: the second aspect of the financial structure of the experiment. One possibility was to divide the external budget over the partnerships. Another possibility was to determine the internal starting budget independently from the external budget. Most projects that chose the last possibility chose a base-year for the internal budget, to which adjustments were made for developments between the base-year and the start of the experiment.

In most projects, the starting budget per partnership was based upon the volume of specialist care in a base-year, expressed in tariffs of 1 April 1994, and in most cases indexed to the level of 1995. This meant that the existing distribution of income between specialities was preserved. It would have been possible to redistribute income over the different specialities, but none of the projects wanted to do that at the time. Since there were large differences in income between specialities that could only partly be explained by differences in
workload and ‘disutility’\textsuperscript{56}, income redistribution was a very controversial subject. The projects considered it best not to burden the start of the experiment with this problem. In Alkmaar, it was decided, during the project, to harmonise the incomes of specialists. There, the specialists experienced the ‘unexplainable’ differences between the specialities as a barrier to the necessary cooperation.

Production parameters

Thirdly, it was necessary to choose the parameters in which the production was expressed. As stated above, some projects geared the definition of the production parameters to the Functionally-oriented Budgeting (FB) -system of the hospital. This meant that first visits, clinical admissions and day-treatments were important production parameters. The goal was to harmonise financial incentives between the specialists and the hospital. Other projects tried to find a reliable indicator for the number of new patients of a partnership in a year. They considered the number of new patients the most useful measurement of production, because this figure does not depend upon the intensity of treatment and therefore leaves specialists free to concentrate on purely medical considerations. Since the number of new patients, or the number of first visits to the outpatient department, was difficult to define, agreements were made about the exact definition of this concept.

Adjustment of internal partnership budgets over time

A fourth point of agreement at the level of the individual projects concerns the adjustment of the partnership budgets in the years after the start of the experiment. The production of a partnership can change from year to year. Under the old FFS system, this had a direct impact upon the revenues of a partnership. In designing the new payment system, it had to be determined if and how this would affect the budget. A growth of production as expressed in the production parameters did not necessarily have to lead to an equally large growth in the budget. A direct link between production and budgets in the next year has the advantage that it is an incentive to work hard. On the other hand, it means that financial considerations can still play a role in determining the specialists’ behaviour. In some projects, it was agreed that a growth in production would only partly be translated into a growth in the budget in the

\textsuperscript{56}For example: the number of urgent calls outside office hours.
following year. In other projects, a threshold value was determined. Only the growth of production above this threshold or below this threshold was translated into a rise or fall in the budget. Moreover, it was possible that the adjustment of the budgets over time was not determined by actual production but by production-agreements. It was also possible for a project to agree that increases in production would not lead to increases in budgets.

4.1.4 Overview of project plans

There are a number of important differences between the project plans, which will be summarised in this section. Details are given in Appendix 4.1. The main differences concern three aspects: the starting-point; the adjustment-system; and, the definition of the production parameters.

Starting-point

Firstly, not all projects had chosen the same base-year for the calculation of the total external starting budget by the COTG. Also, different choices were made in determining the internal starting budget per partnership. These choices are summarised in Table 4.2.

<table>
<thead>
<tr>
<th>project</th>
<th>external</th>
<th>internal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lievensberg (Bergen op Zoom)</td>
<td>1992</td>
<td>share of 1994 revenues in starting budget</td>
</tr>
<tr>
<td>Scheperziekenhuis (Emmen)</td>
<td>1993</td>
<td>1989 revenues plus corrections</td>
</tr>
<tr>
<td>Rijnmond</td>
<td>1994</td>
<td>1992 revenues in tariffs of 1-4-1994</td>
</tr>
<tr>
<td>Ziekenhuizen Noord-Limburg (Venlo/Venray)</td>
<td>1992</td>
<td>1989 revenues plus corrections</td>
</tr>
</tbody>
</table>

There were some differences in the way the internal budgets per partnership were determined. The only project that actually divided the starting budget was the project in Bergen op Zoom. There, the partnerships received their share in the 1994 revenues as a share of the total starting budget. In the other four projects, the determination of the partnership budgets was not directly connected to the external COTG-starting budget. Naturally, during the negotiations about the starting-level, it was known whether the COTG-budget was sufficient to finance
the sum of the partnership budgets, with the exception of the project Rijnmond (TPR). In the Rijnmond project, a number of developments combined to delay the final determination of the external budget\(^5\). In the end, a general reduction of 3.8% had to be applied to the formerly-determined internal budgets. As mentioned earlier, none of the projects attempted redistribution between specialties at the start of the project, though there were no objections from the government.

**Adjustment-system**

The projects differed in the way the (internal) budgets were adjusted from year to year. In Table 4.3 the conditions for a change of the budgets are given. For the sake of convenience, we leave out adjustments to developments in wages and prices.

**Table 4.3 Adjustment of partnership budgets**

<table>
<thead>
<tr>
<th>project</th>
<th>budget was adjusted:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCA (Alkmaar)</td>
<td>only if the production changed more than 10% (in absolute values) compared with 1994</td>
</tr>
<tr>
<td>Lievensberg (Bergen op Zoom)</td>
<td>if production changed, but with a limit of 5%</td>
</tr>
<tr>
<td>Scheperziekenhuis (Emmen)</td>
<td>partly, if the production-agreements changed</td>
</tr>
<tr>
<td>Rijnmond</td>
<td>partly, if the production-agreements changed</td>
</tr>
<tr>
<td>Ziekenhuizen Noord-Limburg (Venlo/Venray)</td>
<td>if the production-agreements changed, but not automatically</td>
</tr>
</tbody>
</table>

The projects differed considerably in the way the partnership budgets were adjusted. In Alkmaar, even before the income harmonisation, the specialists had a more or less fixed income. The budget changed only if the production compared with 1994 changed by more than 10%. Smaller changes were left without adjustment. Therefore, most incentives for production increase were removed from the system, and a decrease in production did not have negative financial consequences.

\(^5\) It took until the end of 1995 for all specialists in the Rijnmond project to sign the agreement, and it took a long time for all partnerships to hand in their annual accounts. Furthermore, the determination of internal budgets was delayed because increases in COTG-tariffs for internists sparked off a renewed discussion about the distribution of the partnership budgets.
The project in Bergen op Zoom was the only one with a direct link between the partnership budget and the production. There was no mitigating influence, in the sense that the first part of a rise in production was not translated into extra budget, or in the sense that a rise in production was only partly expressed in the budget. In Bergen op Zoom, a change in production automatically made itself felt in the partnership budget of the following year, unless the change in production was more than 5%. In that case, the change in budget was limited to 5%.

In Venlo/Venray, the change in the production-agreements was considered, not the change in actual production. But even a change in the production agreement brought no automatic change in the budget. This had to do with the agreement to limit the growth on the hospital level.

The adjustment systems of the projects in the Rijnmond and in Emmen were more or less comparable. They followed the FB-system to a certain degree. In both projects, there was a fixed component and a semi-fixed component that changed with adherence to the hospital. A growth in production agreements was in both projects translated into the budget, but not for 100%. In the Rijnmond, 50% of the budget was variable, and in Emmen 30%. The share of the fixed component in Emmen was 40% and in the Rijnmond 25%. The half-fixed components were, respectively, 30% and 25% for Emmen and the Rijnmond.

All projects have made exceptions to the adjustment systems described above, for example, for non-gate disciplines (e.g. radiotherapy) or for partnerships who were expecting expansion at the start of the experiment. Separate agreements for the adjustment of budgets for non-gate specialities were made in many projects, since the work of these disciplines is largely determined by requests that other specialities make. The agreements could also cover the way in which an expected need to take extra specialists into the partnerships during the experiment period would be treated.

By the ‘adherence’ of a hospital we mean – somewhat simplified – the number of people that use that hospital when they need hospital care. ‘Adherence’ is a measure of the number of patients who are treated in a certain hospital (compared with the number of patients who are treated in other hospitals). For example, when there is a decrease in the number of patients being treated in an academic hospital in the same region as a general hospital, and more patients are being treated in the general hospital, then the adherence of the general hospital is rising.
Definition of production

Thirdly, the projects differed in the way production was defined. In Table 4.4 an overview is given.

Table 4.4 The definition of production parameters

<table>
<thead>
<tr>
<th>project</th>
<th>definition of production parameters:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCA (Alkmaar)</td>
<td>admissions, day-treatments and new visitors to the outpatient department</td>
</tr>
<tr>
<td>Lievensberg (Bergen op Zoom)</td>
<td>new visitors to the outpatient department</td>
</tr>
<tr>
<td>Scheperezienhuis (Emmen)</td>
<td>admissions, day-treatments, new visitors and repeat visitors to the outpatient department</td>
</tr>
<tr>
<td>Rijnmond</td>
<td>admissions, day-treatments, new visitors to the outpatient department and patient-days</td>
</tr>
<tr>
<td>Ziekenhuizen Noord-Limburg (Venlo/Venray)</td>
<td>new visitors to the outpatient department and clinical cards</td>
</tr>
</tbody>
</table>

The project in Bergen op Zoom was using the most pure 'front-door criterion': production was expressed in the number of new patients of a partnership at the outpatient department. The intensity of the treatment of these patients did not play a role in the definition. In Emmen, the intensity of treatment did play a role, though it should be remembered that only 30% of the changes in production were translated into the budget. The budget of a partnership was adjusted on the basis of the number of admissions, day-treatments, new visitors and repeat visitors to the outpatient department in Emmen. Not all of these parameters are pure 'front-door' parameters in the sense that specialists cannot influence them. To a somewhat lesser degree, this also applies to the other projects in the Rijnmond, Alkmaar and Venlo/Venray. Of these three projects, Venlo/Venray used the most 'pure' front-door criterion. The other two projects had given precedence to fitting in with the FB-system of the hospital.

In Alkmaar, Emmen and the Rijnmond, weights were determined for the production parameters for each partnership or discipline. These weights expressed the workload that was associated with different parts of production. This was a way to take into account the fact that some partnerships experienced a larger workload with respect to, for example, an admitted patient than other partnerships. In Bergen op Zoom and Venlo/Venray, no explicit weights were defined, but implicitly they were taken into account: the value of the production
parameter - the partnership budget divided by the number of new patients of that partnership - was an expression of the workload to a certain degree.  

All projects have struggled with the problem of defining the ‘first visit’ to the outpatient department. In most cases, the projects tried to choose a definition that would fit in with the existing COTG-guidelines. In Bergen op Zoom, for example, the number of short cards, year cards and first visits of privately-insured patients was counted. In Alkmaar, the same happened for the gate-disciplines, but for the non-gate disciplines another definition was used that was determined by the Steering Committee. A provision was included in the agreement about production changes caused by a changed interpretation of definitions in Alkmaar. However, this was not sufficient to avoid a discussion about this subject.

The most important conclusion of this section is that there were significant differences between the projects in the way the production was defined and in the conditions under which production changes influenced the budget. A hypothesis to be tested is that, in these different projects, these differences led to different effects of the experiment with the payment system on the treatment of patients.

4.2 Design of the research project

In Section 4.2.1, the research questions are worked out in more detail. Section 4.2.2 gives an overview of the design of the research project. In Section 4.2.3 the practical decisions with respect to the research design are described. In the following sections, some relevant data sources are discussed in detail. Finally, in Section 4.2.7, detailed hypotheses are described that are tested in the following chapters.

4.2.1 Elaboration of the research questions

In Chapter 1 the following research questions were formulated:

59 For two partnerships of different disciplines, with a comparable volume and a comparable budget but with a different number of new patients per year, different parameter values were used.
1. What was the content of the agreements made in the different projects?
2. Did the experiment have an effect upon the treatment of patients?
3. If so, what was this effect?

The first question was answered in Section 4.1. To answer the second and third question we need hypotheses about the possible effects of the experiment. The hypotheses that can be derived from the theory and empirical findings were described in Chapter 2. We concluded that the change in financial incentives might affect many aspects of the diagnosis and treatment of patients. Physicians might stop performing unnecessary services or start referring patients more often for necessary services. They might decrease production in terms of elements that do not play a role in the new measurement system. By using care more economically, they might try to have more leisure and, at the same time, increase production in terms of the units in which it is measured in the new system. The degree to which actual changes were likely to take place depended upon the possibilities. For example, if a physician already works very efficiently, there would be no opportunity to increase efficiency. It is clear that the participants in the experiment believed there were opportunities for improvement in economical use of care, since such improvements were among the main goals of the experiment.

Because it was not possible to analyse all possible changes, we selected a number of important changes to analyse. First of all, inpatient treatment is an expensive part of the activities in a hospital. If the experiment encouraged doctors to treat more patients in day-treatment, or to decrease the length of stay in the hospital, that would have important consequences for hospital costs, especially in the long run. Therefore, we studied the probability that a new patient was admitted within a certain period and we studied the duration of stay of admitted patients.

Second, waiting times were analysed. At the time of formulating the detailed research questions, we considered the waiting period a subject of interest for the evaluation study for the following reasons. In the Netherlands, in the mid 1990s, for a number of specialities, there were quite long waiting lists for a first consultation with a specialist as well as for an admission. At that time, not much

60 In the short run, it may be difficult to realise savings, because these are most likely achieved by dismissing employees and/or closing wards.
policy attention was given to these waiting lists. The emphasis of policy was on cost control. As a research question, it was interesting, since the experiment could be expected to influence the waiting period in several ways. On the one hand, more economical use of care could decrease the waiting period. On the other hand, specialists choosing more leisure could increase the waiting period. Therefore, the waiting periods before patients could have a clinical intervention were analysed.

Thirdly, the substitution of a hospital stay by day-care or short stay could be expected to have consequences for the outpatient department. Substituting day-treatment for clinical admission might mean that patients had to return to the outpatient department more often for check-ups by the specialist and follow-up treatment. On the other hand, we expected an increase of referrals to the GP and others. Sending patients back to their GP sooner for check-ups after treatment in the hospital might lead to a decrease in the probability of repeat visits. Also, handing over check-ups of chronic patients to the GP might lead to fewer visits to the outpatient department. For these reasons, we studied the probability of repeat visits for new patients at the outpatient department. New patients were distinguished from chronic patients because, with the latter, the mechanism that leads to repeat visits is different.

So, the following detailed research questions were formulated with regard to the treatment of patients.

1. Did the experiment have an effect upon the probability of admission and, if so, what effect?
2. Did the experiment have an effect upon the duration of stay and if so, what effect?
3. Did the experiment have an effect upon the waiting period and if so, what effect?
4. Did the experiment have an effect upon the probability of repeat visits and if so, what effect?

**4.2.2 Overview of the design**

An important approach to answer the research questions is to measure changes over time within the experimental projects. Data were collected about the
situation before the experiment (pre-experiment measurement) and about the situation shortly before the end of the experiment (post-experiment measurement), and about developments between these points in time. A comparison of the pre-experiment and the post-experiment measurement gives an indication of the effects of the experiment. However, this comparison is not sufficient to determine the effects of the experiment, since it is not known to what extent changes over time can be ascribed to the experiment. Changes can also be caused by national developments or by local or regional developments that are independent from the experiment. For example, there might have been autonomous trends in the tendency to treat patients clinically (as compared with treatment at the outpatient department). Also, technical developments or government policy (not concerning the experiment) might have influenced the diagnosing and treatment of patients.

An example can clarify this further (see Figure 4.1). Suppose that the average probability of admission for a new patient had decreased in the experimenting hospital A between pre-experiment and the post-experiment measurement (for comparable groups of patients). This could be an effect of the experiment, since the experiment was destined to stimulate the efficient treatment of patients. But this could also be a more or less autonomous trend, caused by medical-technical developments. For example, less evasive surgery can make it possible to treat more patients in day-treatment instead of admitting patients. Suppose we compare experimenting hospital A with non-experimenting hospital B. If we found that the probability of admission had decreased less in hospital B, we might conclude that the experiment had promoted the decrease in the probability of admission. However, if we found that the probability of admission had decreased even more in hospital B, we might conclude that factors other than the experiment were responsible and that the experiment might even have been holding back the decrease in the probability of admission. Therefore, we need more information than just the changes over time within the experimental projects.
One possibility to control for external developments (outside the experiment) was to include a control group in the design. Therefore, at the start of the experiment, a control group was selected of hospitals that were not experimenting with the payment system for specialists. Comparison of the developments over time in the experimenting hospitals and the control hospitals would make it possible to isolate the effect of the experiment from other developments. For the pre-experiment measurement, data were collected in experimenting hospitals as well as in non-experimenting control hospitals.

However, within an unexpectedly short period, ‘lump-sum’ agreements were made in practically all hospitals in the Netherlands. These financial agreements were comparable to those in the experimenting hospitals. The Ministry of VWS was happy with this development, since it appeared to solve long-standing difficulties with the payment of medical specialists, and actively stimulated it. These other projects were called the ‘local initiatives’. In a sense, the experiment with alternative payment systems for medical specialists was no longer an
experiment, since in other hospitals more or less the same principles were adopted. Therefore, during the study it turned out to be impossible from a practical point of view to use a control group. These developments inevitably led to a discussion about the usefulness of continuing the research project. It was, however, eventually decided to proceed with the study of the experiment, since it was expected to yield useful insights anyway. It was still important to find out what the effects of the alternative payment systems would be, since in a few years a decision would have to be made about the optimal payment system for specialists in the longer run. The local initiatives were seen as a short-term solution.

The loss of the control group meant that a possibility was lost to distinguish between developments caused by the experiment and developments caused by other factors. Nevertheless, a positive factor was that there was a lot of variation in alternative payment systems within the experiment group. This gave us other possibilities to determine the effect of the change in payment systems. The different payment systems that were in use during the experiment could be quantified in the form of new 'experimental tariffs' that differed between experimental projects. These tariffs could be included in an econometric analysis. The pre-experiment tariffs and the different sets of new 'tariffs' of the experimental projects were among the factors used to explain a number of different aspects of treatment before and during the experiment. The lack of a control group was found to be less problematic when it became clear that the influence of the different payment systems could be tested empirically in such an econometric analysis. Furthermore, the lack of a control group could partly be compensated by intensifying the qualitatively-oriented analysis of the experiment (on the basis of interviews, group discussions and surveys). This made it easier to interpret changes that took place between the first and the second round of data collection.

The approach chosen to answer the research questions from 4.2.1 was basically as follows:

61 Though it might be argued that these hospitals differed from the original experiment hospitals, in that they were on the whole motivated to a larger degree by financial considerations and to a lesser degree by the desire to experiment with a new system.
1. Data collection took place along the classical lines of an initial pre-experiment monitoring process and a consecutive post-experiment monitoring.
2. Continuous data collection also took place during the experiment.
3. The data from different sources were combined.
4. New 'tariffs' for the experiment situation were calculated.
5. A statistical analysis of the collected (patient-) data was executed.

These five steps are described below. The research design is shown in Figure 4.2.

**Research method**

1. *Data collection at the start and at the end of the experiment*

Data were collected at two points in time: at the start of the experiment and after a period of approximately two and a half years. The following data were collected:

a) quantitative data on the level of the individual patient on illness-episodes for approximately 1000 patients per hospital;
b) interviews with specialists, managers, health insurers and other participants;
c) opinions of patients and hospital personnel collected in postal surveys;
d) information from group-discussions with primary care representatives.

The patient data (see a) above) were meant to give a detailed insight into the diagnosing and treatment of certain complaints and diseases. This information could be related to characteristics of the patients, and the relation between different aspects of treatment could be analysed.
Figure 4.2 Overview of data collection and analysis

start of the experiment

- patient-data
- interviews with specialists, management and insurers
- surveys of patients and personnel
- group discussions on primary care

continuous data collection:
- continuous contact
- production-data

analysis:
- combination of data
- calculation of tariffs
- econometric analysis

end of the experiment

- patient-data
- interviews with specialists, management and insurers
- surveys of patients and personnel
- group discussions on primary care
The interviews with participants (see b) above) in the experiment were included because we wanted to know what their goals and expectations were at the start of the experiment, and how they experienced the experiment. The motivations and experiences of the participants were important factors in determining the outcome of the experiment.

We also asked the opinions of patients and hospital personnel (see c) above), because we expected them to be affected by the experiment in some respects. The quality of care as perceived by the patient could change, as well as the waiting period for patients. Since the distribution of labour between the hospital and other parts of the health sector could change, and diagnosing and treatment within the hospital could change, it was perfectly possible that the workload and work satisfaction of the hospital personnel would be affected.

Group-discussions (see d) above) with primary care representatives were included in the design because the distribution of labour and the pattern of cooperation between the primary and the secondary care sector were expected to change. First, within efficiency and quality projects, GPs and specialists could come to new arrangements about the optimal place where certain activities had to be performed. For example, they could decide that the monitoring of some chronic illnesses could best be done by the GP, or that minor surgery or some diagnostic services could be switched from the specialist to the GP. Secondly, a possible financial incentive for specialists to do monitoring of patients themselves was removed. These two factors could lead to a reduction of referrals from GPs to secondary care, or it could lead to specialists referring patients back to primary care at an earlier stage.

The complete data collection is shown in Figure 5.1. The other elements of this figure are discussed below.

2. Continuous data-collection

To answer the research questions of the original study commissioned by the Zfr, it was necessary to collect data continuously during the experiment, not just at the start and at the end of the experiment. This was because we were interested to learn about the dynamic process by which changes were made, rather than

---

62 The quality of specialist care as perceived by the patient might change for a number of reasons, e.g. the specialist taking more time for a consultation or the patient being referred back to the GP at an earlier stage.
just which changes were made. Two types of data were collected continuously. First, there was continuous contact with participants in the projects in the form of visits, phone calls, interviews etc. The purpose of this contact was to monitor the process of implementing the experiment with alternative payment systems. Secondly, production data were collected during the entire period. Data were collected in each hospital per quarter per specialty (e.g. number of first visits to the outpatient department, number of repeat visits, number of admissions, number of patient-days, etc.).

3. **Relation between the collected types of data**

The patient-data are important, since the effects we wanted to analyse could only be traced accurately at this level. Information about the diagnosing and treatment of individual patients with certain complaints or diseases was collected in detail for a certain period in time. During that period every aspect of treatment in the hospital was laid down making it possible to analyse the relation between different parts of the treatment: for example, substitution between day-treatment and clinical admission. The possibility to relate the content of diagnosing and treatment to the characteristics of the patient was also important. Relevant characteristics of the patient that were collected are: age, sex, type of health insurance (private or public) and complaint/diagnosis. Because the patient-data were collected from the medical patient-files, they were not affected by possible registration-effects of the new payment system and definition-changes that influenced the financial administration. So the patient-data are detailed and reliable, but they do not show the broad picture of everything that takes place in a hospital. This is because a selection of complaints and diagnoses was made which will be explained more fully in Section 4.2.4. To get an overview of all activities in a hospital for all periods of time, the aggregated production-data are necessary. These data were collected for all specialties for the period 1994-1997 (as far as possible). These data, however, may suffer from definition-changes or differences in definition\(^6\) and, for this type of data, a link with patient characteristics is impossible. The data from interviews and group-discussions and from continuous contacts (the qualitative data) are an important source of information for interpreting the quantitative data (the
patient-data and the production-data). However, the qualitative data are not based upon ‘hard facts’, as they are subjective. Therefore, all these data-sources are used to complement each other.

In this thesis, most attention is given to the quantitative data on the patient level. The aggregated production data are used as a background against which to compare the patient-data. The qualitative data are used to interpret the econometric results from the analysis on the patient data. For example, the qualitative data can be used to explain estimated changes in the admission probability in a project between the start and the end of the experiment that are independent from the changes in financial parameters.

4. **Calculation of tariffs for the experiment situation**

Before the start of the experiment tariffs were determined by the COTG. In Chapter 3, a description of the tariff system was given. These tariffs were the same for all hospitals, but they differed between publicly- and privately-insured patients. At the end of the experiment, lump-sum budgets for the specialists were in place. So it might be said that there were no tariffs as such. However, it was still the case that, in exchange for these lump-sum budgets, a certain amount of production was expected. This production was measured and monitored using different production parameters for different projects, as described in Section 4.1. This means that it is possible to calculate an amount of money that corresponds to an element of production, e.g. a new patient. We called this amount of money the *parameter revenue*, in order to distinguish it from a traditional tariff. The name ‘parameter revenue’ seemed logical, since, in the different projects, each different production-parameter had its own revenue. The larger this parameter revenue is, the greater is the financial incentive to the doctor to have a new patient. First, a higher parameter revenue leads to a greater probability of a larger budget in the following year, depending on the exact arrangements for adjustment of the budget in a project. Secondly, the higher the parameter revenues, the larger is their contribution to meeting the arranged production targets.

---

63 Differences in definition do not just play a role in comparing hospitals. They also play a role in comparing information from the hospital-administration with that from the annual report for the same hospital.

64 A number of assumptions had to be made (see Appendix 4.3).
The parameter revenues at the end of the experiment were the same for publicly- and privately-insured patients. Between the different experimental projects and specialties, the parameter revenues differed during the experiment depending on the financial arrangements in the projects. We calculated parameter revenues for all the production parameters that played a role in the different experimental projects:

1. first visit of a new patient at the outpatient department;
2. repeat visit at the outpatient department;
3. clinical admission (for more than one day);
4. day-treatment;
5. a patient-day in the hospital (for each day during an admission).

Here an example is given for a project in which the calculation is simple. A general description of the calculation and more examples are given in Appendix 4.3. Suppose that only the number of new patients is of importance in the experimental system, as in the Ziekenhuis Lievensberg. This means that the only way for the specialist to earn more money in the next year is to help more new patients. Given the number of new patients, the numbers of admissions, day-treatments, repeat visits etc do not influence the budget for the specialists. In that case, the calculation is performed as follows. The lump-sum budget for a specialty is divided by the number of new patients in that hospital for that specialty in the relevant year. In the Ziekenhuis Lievensberg, a high parameter revenue results for new patients and a zero parameter revenue for all other possible production-parameters, like admissions. The revenues for the follow-up treatment can be considered as included in the revenue for new patients. The system was designed that way to keep physicians from doing not strictly necessary follow-up treatment just in order to add to their income.

In the Ziekenhuis Lievensberg, the parameter revenue can be seen as a pure financial incentive, since the budget in the following year increases in line with the number of new patients. The only complication is that there is a limit of 5% growth. A growth of over 5% is not reflected in the next year's budget. For most experimental projects, the calculation of the parameter revenues is more complicated, inter alia because different forms of production have to be made comparable (for example, new patients, day-treatment and clinical admissions). For the details, we refer to Appendix 4.3. In most cases the parameter revenue can not be seen as a pure financial incentive: for example, because the budget for
the following year depends on production arrangements and not on actual production. That does not alter the fact that the parameter revenues reflect the value of additional production in terms of achieving the arranged production level.

As is shown in the Appendix, for the same specialty and the same form of production, differences in ‘tariffs’ can be very large between hospitals in the experiment. For new patients who visit the cardiologist, the highest calculated experimental tariff is over 900 guilders in the Lievensberg in Bergen op Zoom and the lowest is 48 guilders in the Scheperziekenhuis in Emmen. It is important to note that this is not the difference in the income of cardiologists between the projects. The reasons for the large differences are differences in the extent to which the budget is variable and differences in production parameters and their weights (see Appendix 4.3).

The calculated parameter revenues and the COTG-tariffs played an important role in the econometric analysis.

5. **Econometric analysis**

An important element of the study was econometric analysis of the collected patient-data. This analysis was carried out at the level of the individual patient. As described above, it was possible to translate most financial agreements in the experiment into financial parameters that could be used in the econometric analysis. In this way, we could determine the effect of the alternative payment systems econometrically by using as explanatory variables not only the financial parameters (the tariffs and parameter revenues), but controlling for as many interfering factors as possible. These other factors were patient-characteristics, hospital-characteristics, information about complaints and diseases of patients and autonomous trends. Statistical testing could show whether there was a significant influence of the ‘new agreed tariffs’ on the treatment of patients.
By statistically determining the influence of the parameter revenues, the effect of the experiment could for a large part be disentangled from other influences. However, this influence could not be completely isolated, since not all aspects of the experiment could be caught in the calculated parameter revenues (see also Figure 4.3).

First, the parameter revenues reflected most of the financial arrangements, but not all, since some aspects of the adjustment system were impossible to model. In Alkmaar, adjustment took place if the threshold of 10% growth was crossed, and in Bergen op Zoom it took place as long as the threshold of 5% growth was not yet reached. This means that, for an individual patient in the econometric analysis, it can not be determined if the calculated parameter revenues reflect the 'extra financial gain' from treating this patient. This depends upon the level of production that was reached before the patient was treated. However, it can be said that the parameter revenues reflect the extra production that treating this patient generates (according to the definition of the project). Therefore, the parameter revenues indicate to what extent treating this patient contributes to reaching the desired or arranged production level. But it is not certain how large the 'pure financial incentive' is.
In Bergen op Zoom, it was expected that the threshold of 5% would not be crossed in the normal course of affairs. In that case, the parameter revenue does not just reflect the ‘production incentive’ but also the financial incentive. In Emmen and the Rijnmond, the parameter revenue also reflects the financial incentive, since the fact that the production growth is only partially translated into the budgets, is reflected in the parameter revenues.\(^65\) In Alkmaar, there was no direct link between the parameter revenue and the budget, since production changes smaller than 10% did not influence the budget.\(^66\) In Venlo, there was an arrangement that the specialist budget would not increase except under exceptional circumstances. In these circumstances, the parameter revenue reflects the production definition, but not the ‘financial gain’ of treating a patient.

Secondly, the organisational arrangements could not be modelled directly. For example, arrangements regarding transmural care could not be translated into variables to be used in the econometric analysis. The financial arrangements of a project may reflect the priorities of a project and, in this sense, the parameter revenues may also give an indication of the organisational arrangements. However, there is no direct variable to reflect this in the econometric analysis, and therefore the effect of organisational arrangements that are part of the experiment partly ends up in the variable that indicates all unexplained differences over time in a project. This factor may also contain regional developments that are independent from the experiment. No distinction can be made between these independent developments and elements of the experiment that are not reflected by the parameter revenues.

In short, the econometric analyses makes it possible to determine the effect of changes in financial incentives because of the experiment, but not all other changes that are part of the experiment can be isolated in the quantitative analysis. Information from the qualitative analysis is used in the attempt to explain such ‘combined influences’.

---

\(^{65}\) Only 30% and 50% respectively, of the production growth in Emmen and the Rijnmond is translated into the budget.

\(^{66}\) There was no link at all after the income harmonisation in Alkmaar.
Arrangements and design

4.2.3 Practical decisions

In advance of the study, a number of practical decisions had to be made. These decisions concerned the selection of hospitals to be included in the study, and the timing of the two data-collection rounds.

Selection of hospitals

Five projects participated in the ESH. Three of these projects involved more than one hospital. The largest project was TPR in the Rijnmond with 13 hospitals. It was considered unnecessarily complicated and expensive to include all experimenting hospitals in the study. Therefore, a selection of hospitals to be included in the study had to be made. In the Rijnmond, not one but two hospitals were chosen to participate in the study because of the size of the project. In all other projects one hospital participated in the study. The necessary choices are summarised in Table 4.5.

Table 4.5 Selection of hospitals for the study

<table>
<thead>
<tr>
<th>project name</th>
<th>number of hospitals in the project</th>
<th>hospital(s) or location participating in the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project ‘Noord-Holland Noord’ (TVV agreement)</td>
<td>2</td>
<td>Medisch Centrum Alkmaar (MCA) in Alkmaar</td>
</tr>
<tr>
<td>Ziekenhuis Lievensberg</td>
<td>1</td>
<td>Ziekenhuis Lievensberg in Bergen op Zoom</td>
</tr>
<tr>
<td>Scheperziekenhuis Emmen</td>
<td>1</td>
<td>Scheperziekenhuis in Emmen</td>
</tr>
<tr>
<td>Regionaal project Rijnmond (TPR agreement)</td>
<td>13</td>
<td>St. Clara Ziekenhuis in Rotterdam, IJssellandziekenhuis in Capelle aan den IJssel</td>
</tr>
<tr>
<td>Regional project ‘Ziekenhuizen Noord-Limburg’ (‘Experiment Zorgvernieuwing’)</td>
<td>1 (on 2 locations)</td>
<td>st. Maartens Gasthuis in Venlo</td>
</tr>
</tbody>
</table>

In the projects in Emmen and Bergen op Zoom, no selection had to be made, because only one hospital participated in each of these projects: respectively, the Scheperziekenhuis and Ziekenhuis Lievensberg. Both of these hospitals took part in the study.

At the time when hospitals had to be selected for the study, two hospitals participated in the project Noord-Holland Noord: the MCA and the Gemini
Ziekenhuis in Den Helder. The two hospitals decided together that only the MCA would participate in the study.

In the project of the Ziekenhuizen Noord-Limburg, also called the project Venlo/Venray, the situation was somewhat different. The Ziekenhuizen Noord-Limburg are the st. Maartens Gasthuis in Venlo and the st. Elisabeth Gasthuis in Venray, which operate as one hospital. Legally, the Ziekenhuizen Noord-Limburg are one entity. However, because there were two locations, the project could not be considered one hospital for the purposes of the study. A selection had to be made. The study took place in the location with the largest production capacity, i.e. the st. Maartens Gasthuis in Venlo.

In the Rijnmond, the two hospitals that participated in the study were the St. Clara Ziekenhuis in Rotterdam and the IJssellandziekenhuis in Capelle aan den IJssel. This selection was made by the project (TPR) itself.

All in all, there were six experimenting hospitals that took part in the study.

**Timing of data-collection**

The first round of data collection ideally had to take place shortly before the start of the experiment. This timing was kept to as far as possible. It was not implemented completely, because in some projects a final decision about the experiment was only made at a late stage (in the Spring of 1995) and the experiment started retroactively from 1 January 1995. In such cases, it was not possible to collect data before the official start of the experiment. It was, however, possible to collect data before the start, in practice, of the experiment, which means that the findings would be able to show the pre-experiment situation.

At the start of the study, there was some discussion about the timing of the second (and last) round of data collection. On the one hand, the experiment constituted an important change for the participants. Therefore, it was felt that the period between the two data-collection rounds should not be too short. The people concerned in the projects needed time to adjust. This would increase the probability of finding clear effects of the experiment. On the other hand, if we waited too long, developments could occur that would make the results of the
study obsolete. There was an additional argument for a shorter period. In principle, the experiment was supposed to last three years. At the start of the experiment, there was uncertainty about the situation that would prevail after the experiment. One of the possibilities was that the old FFS system would be reintroduced. In that case, participants might adjust their actions to the incentives of the FFS system even before the end of the experiment. This was another disadvantage of a long period between the data collection rounds. Therefore, it was decided at the start of the study to plan the second round of data collection two years after the first round.

Later on, during the research project, the Zfr, the organisation that had commissioned the study, decided to postpone the second round, after consultations with the participants in the projects. This decision was taken because of the start of the local initiatives that were mentioned above, with, as a result, the loss of the control group. Moreover, the participants in the projects felt that two years was too short a period to find all the effects they anticipated. The data collection was postponed. In the end, the period between the two rounds was approximately two and a half years, allowing the study a somewhat longer ‘observation period’.

4.2.4 Patient-data

We collected data for approximately 1000 selected patients per hospital. This was done at the start of the experiment and after two and a half years. These data were meant to give a detailed insight into the diagnosing and treatment of certain complaints and diseases. This process was completely mapped for a certain period of time (the observation period). We collected information not only about the activities of the specialty from where the patients for the sample were selected, but also about the other specialties that participated in the treatment of the individual patient.

Patients can come to the hospital with a great variety of complaints and diseases. Drawing a random sample of all patients in the hospital would result in a very

---

67 As described above, part of the design and purpose of the study became obsolete when practically all other Dutch hospitals made comparable financial agreements. The speed with which that happened was totally unforeseen at the time the experiment and the study started.
heterogeneous group. It would be difficult to draw firm conclusions about such a group, because there would be too little information about the influence of the particular complaint upon the diagnosis and treatment process. Therefore, we decided to select more homogeneous groups of patients with certain complaints and diseases.

The selection process is summarised in Table 4.6. The selection was made along the following lines. First of all, a distinction was made between clinical patients, patients at the outpatient department and patients in day-treatment. This was because we wanted to be able to analyse different kinds of treatments and processes. A sub-sample of clinical patients was taken to make sure that there would be enough clinical admissions to the hospital to analyse. The largest part of the sample of patients at the outpatient department consisted of patients who came for the first time to this department with a certain medical problem. These patients were selected at the start of their treatment-period in the hospital so that information could be collected from the start about all choices that specialists made about determining their diagnosis and about their treatment.

Within these three groups, a further selection was made on the basis of the specialty. The next step was a selection within the clinical group for the selected specialties based upon specific discharge diagnoses. Within the ‘outpatient department group’, the further selection was based on the type of complaints suffered by new patients and on diagnoses for existing patients who visited the hospital for monitoring of their treatment and/or condition. Day-treatment patients were further selected according to the type of intervention.
Table 4.6 The different steps in the selection process for patient data

<table>
<thead>
<tr>
<th>step 1</th>
<th>step 2</th>
<th>step 3</th>
<th>number of pat. to be selected per hospital</th>
<th>observation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>clinical patients</td>
<td>7 specialties</td>
<td>discharge diagnoses</td>
<td>300</td>
<td>5 months before first admission in the selection period and 3 months after</td>
</tr>
<tr>
<td>selection at</td>
<td>8 specialties</td>
<td>new patients: complaints</td>
<td>450</td>
<td>3 months from the first visit</td>
</tr>
<tr>
<td>outpatient</td>
<td>3 specialties</td>
<td>monitored patients:</td>
<td>150</td>
<td>4 visits back from ‘selection visit’</td>
</tr>
<tr>
<td>day-treatment</td>
<td>6 specialties</td>
<td>types of intervention</td>
<td>100</td>
<td>like clinical patients</td>
</tr>
<tr>
<td>patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The selection of complaints, diagnoses and interventions that were included in the sample was based upon a number of criteria. The general idea underlying these criteria was that the selected groups together should be as representative as possible for total specialist care in the hospital. This had the following consequences for the criteria. First, groups that were not often found in a hospital were not suitable.\(^{68}\) Secondly, it was considered best to avoid selecting especially those groups where the largest possible efficiency gains were expected. The study would then give an overly positive view of the total increase in economical use that could be realised. Thirdly, groups in which diagnosing and treatment were a totally standardised process were not very interesting. After all, no changes can be expected when everything is fixed and the specialist has no choices to make. Therefore, one of the criteria for selecting patients with certain complaints was that there would be no complete certainty about the way to diagnose and treat these complaints. Chest pain and back problems are included, amongst other reasons, because research shows a large variability of treatment (see Section 2.2.3).

Fourthly, the different specialties had to be reasonably covered, though this did not mean that every individual specialty had to be represented.

\(^{68}\) It is possible that with rare diseases there is more room for PID (see list of abbreviations). But, in that case, it is also likely that the doctor is more uncertain and there is less opportunity for standardisation and efficiency.
In the outpatient department group, 600 patients per hospital were selected, most of whom made a first visit to a specialist at the outpatient department with a new complaint or probability diagnosis. Naturally, this does not mean that the whole treatment had to take place only at the outpatient department. It is possible that, in some cases, the specialist would decide to admit a patient from this sample to the hospital. If so, then data about this admission were collected as well, as long as it took place within the observation period.

A sample of new patients with certain complaints could not be taken from the existing administration of the specialists and the hospital, because the necessary information was not computerised. Therefore 'pre-selection forms' were used, that were designed especially for the study of the experiment. The thirteen groups of complaints of new patients that are given in Table 4.7 below were selected for eight specialties. It was not the intention to select all thirteen groups in each hospital. Some of the groups were 'standby categories' to be used when it was not possible to find sufficient numbers of patients of the originally intended category in a certain period. In each hospital, it was intended to select nine categories of complaints of new patients at the outpatient department. In addition to those new patients, three categories of patients were selected who had been to the hospital before with their illness and who were still returning for monitoring.

From the clinically treated patients, seven groups of 45 patients each per hospital were selected on the basis of the discharge diagnosis and the specialty that discharged the patients (about 300 patients per hospital). These patients were selected at the end of the disease-episode. In Table 4.7, eight clinical groups are shown. One of the groups is a reserve group. Apart from the clinical patient-groups, three groups of patients in day-treatment were also selected in each hospital (about 100 patients per hospital). The selection differed between hospitals: in each hospital the three most frequent interventions were selected.

For new patients at the outpatient department, a different observation period was used than the one for clinical and day-treatment patients. The new patients were followed for three months after their first visit to the outpatient department. This means that the observations are truncated: treatment after three months is not part of the observation. The clinical and day-treatment patients were selected at a stage when an important part of their treatment was already finished. Their observation period ranged from five months before the
moment of selection to three months after. This was done so that the ‘post-clinical’ treatment could also be observed.

Table 4.7 shows the different complaints and diagnoses that were selected for new patients at the outpatient department and for clinically-treated patients.

Table 4.7 Selection of new patients at the outpatient department and clinical patients for collecting the patient-data

<table>
<thead>
<tr>
<th>specialties and groups of patients with various complaints/diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>new patients at the outpatient department</strong></td>
</tr>
<tr>
<td>cardiology</td>
</tr>
<tr>
<td>surgery</td>
</tr>
<tr>
<td>surgery</td>
</tr>
<tr>
<td>surgery</td>
</tr>
<tr>
<td>gynaecology</td>
</tr>
<tr>
<td>internal medicine</td>
</tr>
<tr>
<td>otolaryngology</td>
</tr>
<tr>
<td>otolaryngology</td>
</tr>
<tr>
<td>neurology</td>
</tr>
<tr>
<td>neurology</td>
</tr>
<tr>
<td>orthopaedics</td>
</tr>
<tr>
<td>orthopaedics</td>
</tr>
<tr>
<td>urology</td>
</tr>
<tr>
<td><strong>clinical patients</strong></td>
</tr>
<tr>
<td>cardiology</td>
</tr>
<tr>
<td>surgery</td>
</tr>
<tr>
<td>surgery</td>
</tr>
<tr>
<td>gynaecology</td>
</tr>
<tr>
<td>internal medicine</td>
</tr>
<tr>
<td>otolaryngology</td>
</tr>
<tr>
<td>neurology</td>
</tr>
<tr>
<td>orthopaedics</td>
</tr>
</tbody>
</table>

Data were collected about the following subjects for each patient in the sample for the duration of the observation period:

- specialties that treat the patient;
- referral to the specialist;
- visits to the outpatient department, admissions to the hospital and day-treatments (dates, specialties and reasons);
- diagnostic activities;
- number and type of services that were performed;
- waiting period for surgery and other elements of treatment;
- complaint(s) and final diagnoses (main and secondary);
- referral from the hospital to another hospital or another organisation (e.g. nursing home) or back to the GP;
- state of affairs at the end of the observation period;
- characteristics of the patient (age, sex, type of health insurance, some elements of the medical history).

Econometric analysis of the patient data

The variables explained by econometric analysis are: the probability of admission for patients who came to the outpatient department with a certain complaint for the first time; the duration of stay in the hospital for selected clinical patients with certain diseases; the waiting period before an intervention takes place for selected clinical patients; and, the probability of repeat visits for new patients who visited the outpatient department with a certain complaint.

Van Vliet (1988)\textsuperscript{69} carried out a literature survey of the factors that can influence hospital utilisation. This yielded the following categories for a comprehensive model:

- consumer/patient characteristics (health status, insurance coverage, hospitalisation related factors, socio-economic status and regional factors);
- regional supply of health care facilities;
- hospital characteristics;
- physician characteristics.

Naturally, in our case, the financial incentives from the different payment systems have to be added to the set of exogenous variables. Considering the research question, we are primarily interested in the effect of the financial

\textsuperscript{69} This reference may seem dated, but is used anyway, since Van Vliet's literature survey was extensive and the relevant exogenous factors will not have changed in the meantime.
variables and in a possible additional effect of the experiment that can not be captured in the old or new ‘tariffs’. The other exogenous factors are important merely because they enable us to estimate unbiased coefficients for the financial factors. Therefore, we do not use such a detailed set of explanatory variables as Van Vliet does. Sometimes variables within one category of explanatory factors can be joined together without loss of information for our purposes.

In the following paragraphs, we discuss the sets of exogenous variables used in our model: patient characteristics; hospital variables; physician characteristics; financial incentives; and, other factors.

Patient characteristics

Naturally, it is very important to correct for the health status of the patient. We had several variables at our disposal concerning the health status. The complaint or disease has an important influence on all aspects of the diagnostic process and the treatment. The characteristics of the patient are also important control variables for the analysis. Age, sex and type of insurance are included in the analysis. We would have liked to have more information about the patients, e.g. their income, educational level and household situation. However, such information is not to be found in medical patient files. In some cases the marital status or the nationality of the patient was known, but there were too many ‘missings’ to use these variables in the analysis. Van Vliet concludes that the socio-economic variables appear to play a minor role in explaining hospital utilisation. 70 Hopefully, the lack of these variables is not problematic in the analysis, but we keep it in mind in interpreting the estimation results. Some information was collected about the history of the complaint or disease. It was known whether the complaint had been treated earlier in the same hospital. In that case, it seemed likely that there was prior information about the complaint. This can influence the diagnostic process. Another possible explanatory variable is whether there was already a (probability) diagnosis when the patient came to the hospital (e.g. determined by the GP). Lastly, it was important to know something about the state of treatment at the end of the observation period. If the treatment in the hospital had not yet finished, it would mean that part of the treatment was not observed and could not be analysed. This circumstance can yield different results than for patients with finished treatment. This concerns
only the data for the new patients at the outpatient department where the observation stopped three months after the first visit.

Further patient characteristics that can be important according to Van Vliet are the day of the week on which the patient is admitted (important for the length of stay), population density and degree of urbanisation. The factors 'degree of urbanisation' and 'admitted in the weekend' were included in the econometric analysis.\footnote{Van Vliet (1988), p. 58.}

**Hospital variables**

The factor 'hospital' was treated as a composite variable (dummy). Since there are only six hospitals in the analysis, there is not enough variation to distinguish a number of individual characteristics of hospitals, nor are we interested in the effect of each separate characteristic. The hospital where treatment takes place is important in our analysis for two reasons. First, hospitals have their own traditions, habits, environment and working methods. Therefore, differences between hospitals could be expected even in the pre-experiment period. Second, this variable is also included in combination with the dummy for the post-experiment measurement. In this case, the hospital variable gives information about differences between project plans that can not be translated into precise variables, and about developments outside the experiment that took place in a region or a hospital in the experiment period. An example of the latter type of information is a reduction of hospital beds, a new hospital building or an increase in the demand for care in the environment of the hospital. An example of the former type of information is a threshold, above or below which no adjustment of partnership budgets for the next year took place.\footnote{We considered population density correlated too much with degree of urbanisation to include it as a separate factor.}

**Physician characteristics**

The treatment of individual patients may also be explained from doctor characteristics. We did not use those characteristics for several reasons. First of all, it would have been much more difficult to collect the necessary data, since

\footnote{It is impossible to make a separate variable that includes this factor, because it is unknown if and when the threshold was reached in relation to the collected individual data.}
individual physicians would have had to give up their privacy. Second, the inter-doctor variation is not really the phenomenon that we are interested in. It does not matter if this factor ends up in the error term, as long as the other coefficients (especially those of the parameters revenues) are not biased. We do not expect this inter-doctor variation to bias the coefficients since it can be seen as 'white noise', considering that we have already included hospital-dummies.

Financial incentives

The parameter revenues are the essential variables for answering the main research question. If diagnosing and treatment are influenced by financial incentives, we would expect to find a significant effect of these variables on aspects of treatment. However, there may be intervening factors that serve to reduce the expected effect, such as a lack of goal harmonisation between specialists and the hospital management.  

The period is naturally of importance because, as stated earlier, not all aspects of the experiment could be modelled and because there could be influential developments over time that were independent from the experiment. In addition to the interaction between the hospital and the period (start or end of the experiment), other interactions are modelled where appropriate. For example, the effect of a certain diagnosis on the factor to be explained may vary with the age of the patient.

Other factors

Van Vliet mentions regional supply variables as a possible explanatory factor. With respect to these variables he concludes:

"The only consistent finding relates to the strong positive impact of regional hospital bed supply on both admission rate and length of stay, but this consistency is restricted to macro-studies only."  

This might imply that these variables are not very important for our micro-analysis. We have not included regional hospital bed supply or other regional

---

73 An example is the situation in Bergen op Zoom, where the hospital management tried to slow down the decrease in the number of clinical admissions because they were worried about the hospital budget.

supply factors as separate variables. In as far as these are important, they end up in the estimated coefficients for the hospital dummies in our analysis.

Another possible exogenous factor is the medical-technological development. We did not expect a large effect in a period of two and a half years. However, in some cases it may have been important. Since we can not make this operational in the econometric analysis, it ends up in the trend term.

In short, the different elements of the way specialists treat patients were explained from the following broad categories of factors:

- characteristics of the patient;
- aspects of the history of the illness;
- the complaint or disease the patient is suffering from;
- the hospital where the patient is treated;
- the parameter revenues ('tariffs');
- the period of treatment (pre-experiment or post-experiment).

4.2.5 Interviews and group discussions

As mentioned above, both at the start and at the end of the experiment interviews were held with participants in the experiment, and group discussions took place within a group of GPs and other representatives of the primary care sector. Halfway through the experimental period, at each hospital, group discussions were organised with the participants in the experiment in order to monitor the progress of the experiment. Each of the approaches adopted is described in turn below. Apart from these methods, there were also continuous contacts with participants in the projects to monitor the process of change.

Interviews

In each project, two rounds of 15 individual interviews were carried out with important participants in the experiment and persons who could be expected to notice possible effects of the experiment in their work. Interviews were held with:

- the management of the hospital;
- managers in charge of nursing and/or laboratories;
- the health insurer;
- medical staff (at least eight medical specialists).

In selecting the eight specialists, it was taken into account that surgical and non-surgical specialties should be represented as well as gate- and follow-up specialties, and both supporters and opponents of the project. In any case, internists, cardiologists, surgeons, neurologists, anaesthetists, gynaecologists and radiologists were represented in each hospital.

The goal of the interviews at the start of the experiment was to collect information about the knowledge that the interviewed persons had about the experiment, their expectations, their opinions about the experiment, and what they considered to be weak points or possible problems. Questions were asked about the expectations regarding the contents of care, and regarding the financial and organisational aspects.

In the second round of interviews, at the end of 1997, the subjects of conversation were: the developments during the project; the extent to which the expectations were fulfilled and the goals realised; and, the plans and prospects for the period after the end of the experiment.

**Group discussions with primary care representatives**

A group discussion for representatives of the primary care sector was organised in each project, parallel with the interviews. The reason for organising these group discussions was that the experiment could also influence other parts of health care, such as the primary care sector (see 4.2.2). The following primary care representatives were invited for the group discussions:

- GPs;
- midwives;
- a home-care representative;
- a physiotherapist.

In the different regions, between six and twenty persons were present at the discussions, depending mostly upon the interest of the GPs.

At the start of the experiment, some of the subjects of discussion were: cooperation within the primary care sector and between primary care and secondary care; knowledge about the local experimental project, and opinions about the project and its possible weak points. Another subject was the expectations about the effects of the experiment on health care in the primary and secondary sector (especially changes in the division of labour between
primary and secondary care; quality of care; and, the relation between primary and secondary care).

At the end of the experiment, the group discussions concerned the experiences with the experiment in the primary care sector, and the extent to which the expectations had been fulfilled.

*Group discussions among participants in the experiment*

Somewhere in the middle of the experiment period, group discussions were organised with the participants in the experiment. The goal of these discussions was to keep up with the developments in the projects. Specialists, management of the hospital and health insurers could have discussions with each other about the developments, as a result of which differences in point of view and possible problems with the experiment could become clear.

4.2.6 *Surveys among patients and hospital personnel*

In planning the design of the research, we considered the possibility that the experiment would have effects on patients and on hospital personnel. Since a number of goals of the experiment was to improve efficiency, cost control and quality of care, the experiment could influence the way patients were treated, their perception of the care, the way in which hospital personnel carried out their responsibilities, and the labour satisfaction of the personnel. Therefore, surveys among these groups were included in the design. These surveys were conducted by mail.

*Patient-survey*

In each hospital, about 500 patients were selected at the start and at the end of the experiment (a total of more than 3000 patients in each period). The patients were aged 18 years and older, and had been treated by specialists shortly before the survey. The sample was stratified to ensure that enough clinical patients were selected.

Questions were asked about the following subjects:

- the waiting period before an appointment at the outpatient department could be made, and before admission could take place;
- opinion about the treatment by the specialist (concerning, among other things, available time and explanation about treatment);
- the perceived quality of care;
- necessary absence from work or inability to attend to normal daily activities.

**Personnel survey**

For this survey, employees of the hospital were selected who worked together with specialists.\(^{75}\) Per hospital and per period 300 surveys were mailed.

Questions were asked about:

- cooperation between the personnel and medical specialists;
- the perceived quality of care;
- workload;
- labour satisfaction.

Generally speaking, the results of the patient and personnel surveys showed few striking side effects of the experiment. It seemed that, in most respects, patients and personnel did not experience large changes because of the experiment. However, waiting times for admission are an exception: according to the patient survey, they increased. No further separate attention is given to the surveys in this thesis. The results are only mentioned when they are directly relevant to the subject under discussion (for example, waiting times).

**4.2.7 Detailed hypotheses**

In this chapter, we have not only formulated detailed research questions, but have also explained the calculation of pre- and post-experiment ‘tariffs’, i.e. the ‘parameter revenues’. On the basis of these building blocks detailed hypotheses can be formulated.

Table 4.8 summarises the changes in the parameter revenues that were caused by the experiment in the six hospitals.\(^{76}\) The revenue from a first visit increased in

\(^{75}\) The sample consisted of senior nursing officers, nurses, operating-room assistants, radiology assistants, nurses outside the nursing units of the hospital, and other personnel (e.g. medical secretaries).
practically all hospitals, except the Scheperziekenhuis in Emmen. This is because, in Emmen, only a modest part of the budget was variable and in the experiment period there were still four positive tariffs. In Bergen op Zoom and Venlo/Venray, there was only a positive tariff for first visits of new patients. Concerning the tariff for a repeat visit, Emmen was also an exception. In all other projects, repeat visits were no longer rewarded. Therefore, the tariff for repeat visits decreased in all hospitals, except Emmen, where it increased. For day-treatment the picture is mixed: an increase of the tariff in four hospitals and a decrease in two (the two where only new patients were counted). The tariff for clinical admissions also increased in four hospitals. It remained the same (zero) in Bergen op Zoom and Venlo/Venray. The tariff for a patient-day decreased in every hospital; only in the Rijnmond was it still positive during the experiment.

With regard to these tariff-changes the following hypotheses were formulated on the basis of theoretical considerations and common sense (see Table 4.9). The table gives the expected partial effects of a change in just one parameter revenue.

Table 4.8 Changes in parameter revenues caused by the experiment

<table>
<thead>
<tr>
<th>project</th>
<th>first visit</th>
<th>repeat visit</th>
<th>day treatment</th>
<th>admission</th>
<th>patient-day</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCA (Alkmaar)</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Lievensberg (Bergen op Zoom)</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>=</td>
<td>↓</td>
</tr>
<tr>
<td>Scheperziekenhuis (Emmen)</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Rijnmond St. Clara Ziekenhuis</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Rijnmond IJssellandziekenhuis</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Ziekenhuizen Noord-Limburg (Venlo/Venray)</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
<td>=</td>
<td>↓</td>
</tr>
</tbody>
</table>

The table is based upon average tariffs for the different specialties and the different insurance types of the patients at the start and near the end of the experiment.
Table 4.9 A priori expected effect of parameter revenues upon treatment

<table>
<thead>
<tr>
<th>parameter revenues</th>
<th>prob. of admission</th>
<th>duration of clinical stay</th>
<th>prob. of repeat visit</th>
<th>waiting period</th>
</tr>
</thead>
<tbody>
<tr>
<td>first visit</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>repeat visit</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>day-treatment</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>admission</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>patient-day</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Firstly, an increase in the tariff for a first visit to the outpatient department can be seen as an incentive to help as many new patients as possible in a certain period. This could be done by treating the patients as efficiently as possible (at least from the point of view of hospital care), which could be realised by: sending patients back to the primary care sector as soon as possible; sending patients ‘forward’ to the tertiary care sector (e.g. nursing homes); day-treatment instead of clinical admissions, where possible; and, treatment at the outpatient department instead of day-treatment, where possible.

When there are still opportunities to make the treatment more economical and these opportunities are used, a decrease in the probability of admission is to be expected and a decrease of the waiting period. Through a substitution of day-treatment for clinical admissions more patients can be treated in a certain amount of time and with a certain amount of beds. However, the effect upon the duration of the clinical stay and the probability of repeat visits is not a priori clear. A more economical treatment might mean that patients, given their characteristics, have a shorter stay in hospital than before. But the characteristics might change because of a stricter selection of which patients to admit and which not. In the latter case, the condition of admitted patients is likely to be more serious on average than before and this might lead to a longer duration of stay. So there are conflicting influences. The same goes for the probability of repeat visits. On the one hand, sending patients back to their GP sooner for check-ups after treatment in the hospital might lead to a decrease in the probability of repeat visits. On the other hand, substituting day-treatment for clinical admission might mean that patients have to return to the outpatient department more often for check-ups and follow-up treatment by the specialist.
Secondly, an increase in the tariff for a repeat visit to the outpatient department makes it more attractive to do check-ups in the hospital and less attractive to send the patient back to the primary care sector. Therefore an increase in the probability of repeat visits is expected. For the other analysed aspects of treatment, no clear influence is expected.

Thirdly, an increase in the tariff for day-treatment, given the tariff for clinical admissions, makes it more attractive to treat patients in day-treatment and less attractive to admit them. Therefore a decrease in the probability of admission is expected. In connection with this decrease, an increase in the duration of the clinical stay may be expected, since it is more likely that day-treatment can be substituted for short stays. An increase in the number of day-treatments is expected to cause an increase in the number of repeat visits to the outpatient department. For this reason, an increase in the probability of repeat visits is expected. With regard to the waiting period, a decrease is expected because of the substitution of day-treatment for short-stay admission.

Fourthly, for an increase in the tariff for clinical admissions, given the tariff for day-treatment, exactly the opposite effect is expected to that of an increase in the tariff for day-treatment, as described above. So, the effect in practice will depend upon the relative changes in the tariffs for day-treatment and admission.

Fifth, when the tariff for a patient-day increases, an increase in the duration of clinical stay is expected. This increase may lead to a decrease in the probability of repeat visits. When patients stay longer in the hospital, it may be less necessary to return to the outpatient department for check-ups, etc. No effect upon the probability of admission is expected. A longer stay in the hospital is expected to cause an increase in the waiting period, because more beds are occupied.