GAMING THE PAYMENT SYSTEM

EVIDENCE FROM GERIATRIC REHABILITATION

S.A.S. Verbunt

Gaming the Payment System

Evidence from Geriatric Rehabilitation

by

Sabine Antoinette Stephanie Verbunt

in partial fulfilment of the requirements for the degree of

Master of Science
in Engineering and Policy Analysis

at the Delft University of Technology,
to be defended publicly on Wednesday May 8, 2024 at 1:30 PM.

Graduation committee		
Chair:	Dr. ir. M. Kroesen	TU Delft
First supervisor:	Dr. N. van der Linden	TU Delft
Second supervisor:	Dr. ir. M. Kroesen	TU Delft
External supervisor:	Dr. T.O. Michielsen	CZ

TU Delft

CZ

Executive summary

Considering the ageing population and growing shortages in healthcare staff, Dutch policymakers are looking for ways to make high-quality and accessible healthcare available against affordable costs. Payment systems should be designed so that healthcare providers are incentivised to provide high-quality care whilst using the available resources as efficiently as possible. This also applies to the payment system in geriatric rehabilitation (GR), which is short-term, multidisciplinary care for vulnerable elderly with the aim of returning home. Each year, there are around 45,000 to 50,000 GR trajectories in the Netherlands, resulting in total healthcare costs of around 700–800 million euro.

The reimbursement scheme in GR contains stepwise tariffs depending on both the length of stay and treatment intensity: after a pre-determined number of inpatient days or treatment hours, the total reimbursement jumps to a higher level and then stays constant until the next threshold is reached. These discontinuous reimbursements may distort physicians' behaviour in a way that is detrimental to the efficiency of care provision. This thesis examines whether these behavioural distortions are present regarding GR and, if so, which factors are associated with stronger distortions on a GR provider level. For this goal, GR claims data from health insurer CZ covering the years 2017–2023 are used. These data are virtually representative of all claims at Dutch health insurers in terms of main diagnosis groups. Furthermore, this thesis qualitatively assesses the possible consequences of the proposed changes in the payment system of GR, including the abolition of the current stepwise tariffs, using insights from GR claims data and economic theory.

Behavioural distortions can be observed in the claims histogram of the relevant treatment duration variable as a 'bunch', which is excess mass around a threshold. This thesis is the first to analyse behavioural distortions due to discontinuous reimbursements for two aspects of the treatment duration at the same time. Visual inspection indicates that behavioural distortions are only present for the length of stay. For the treatment intensity, no distortions are visible at all. These insights are helpful for policymakers to better understand how similar financial incentives may affect the behavioural responses of healthcare providers differently. Subsequently, the applicability of the bunching approach is studied to estimate the magnitude of the distortions regarding the length of stay. The claims histogram shows that the distribution of the length of stay is not smooth due to weekly peaks. These peaks are also present in weeks with no jumps in the reimbursement. As a result, the central assumption of the bunching approach requiring smoothness outside the area affected by behavioural distortions does not plausibly hold.

A multiple linear regression analysis is performed to explain the variation in the strength of behavioural distortions regarding the length of stay on a GR provider level. The dependent variable is defined by taking the difference between the number of claims for a GR provider in the two days after and two days before a threshold. This is done for all thresholds, and the sum of the differences is scaled with the total number of claims to correct the fact that every GR provider has a varying number of claims submitted at CZ. Regarding the independent variables, data about the fraction of years in which the budget ceiling of a GR provider is exceeded, the share of female patients, the mean age of the patient group and the share of patient admissions on the different days in the week, are used. Including data about the budget ceiling, which is the maximum amount that healthcare providers can claim from health insurers in a year, has not been done before. Regression results indicate that a stronger behavioural distortion of a GR provider is positively correlated with a larger fraction of years in which the budget ceiling is exceeded. This result is statistically significant at a 5% level. Information about this relationship is valuable for health insurers to improve contracts with healthcare providers. Regression results also indicate that a larger share of patient admissions on Wednesday, Thursday and Friday are statistically significantly correlated with stronger behavioural responses of GR providers. These findings are consistent with previous research indicating that patients who reach the threshold day during the weekend are less likely to be discharged than patients who reach the threshold day in the days before the weekend.

The proposed changes to the current payment system in GR include abolishing the reimbursement scheme with stepwise tariffs and, instead, introducing a modular payment system. This new system consists of different modules in which the length of stay and treatment intensity are reimbursed with constant tariffs per inpatient day and treatment hour. Using insights from economic theory and claims data from CZ about GR trajectories, a qualitative re-evaluation of the changed incentives points out that the modular payment system in GR could result in a situation that is unfavourable in terms of efficiency as long as there are no waiting lists and enough healthcare personnel available. That is because the incentive for overtreatment is stronger in the new system compared to the old one, which increases the chances of gaming the payment system. Therefore, it might be better to adapt the current payment system in GR so that the criteria for an optimal payment system in the eyes of the Dutch Healthcare Authority are better met. Suggestions for adjustments to the current payment system include moving the location of the tariff thresholds, no longer registering the indirect patient-related time and abolishing the distinction between the different diagnoses.

Although this thesis shows how financial incentives in the reimbursement scheme of GR affect physicians' behaviour concerning the two aspects of the treatment duration differently, it remains unclear why physicians take the opportunity to game the payment system regarding the length of stay but not regarding the treatment intensity. Future research should shed light on this issue by, for example, conducting interviews with elderly care physicians and the management of healthcare institutions. Besides, a positive coefficient is estimated for the correlation between the strength of behavioural distortion and the fraction of years in which the budget ceiling is exceeded on a GR provider level. However, this finding does not reveal anything about the direction of causality. Follow-up studies can conduct experiments to estimate the causal effect(s), providing additional valuable insights for health insurers, policymakers and other stakeholders in the healthcare sector.

To conclude, the modular payment system will come into effect nationally from January 1, 2030. The five years before, 2025 to 2030, serve as a transition period in which an increasing number of GR providers will adopt the new payment system. Monitoring providers' performances during that period is crucial because GR providers will most likely extend the treatment duration in the modular payment system. It would also be wise if the Dutch Healthcare Authority thinks of alternatives when things do not turn out as expected. Considering the insights from this thesis, further developing the modules or maintaining the current payment system with some adjustments would be good alternatives. An advisable development of the modules is to implement decreasing — instead of constant — tariffs to limit the financial incentives for overtreatment. A pivotal adjustment to the current payment system in GR is to move the location of the tariff thresholds from whole weeks to (partly) reduce the existing behavioural distortions regarding the length of stay. Both options will help to manage future healthcare costs in GR.

Contents

1 Introduction

Dutch elderly care is under pressure. The demand for this type of care is increasing rapidly, mainly due to the ageing population. While there were 2.5 million Dutch people 65 years or older in 2010, this number has grown to 3.6 million in 2023 and is even predicted to rise to 4.8 million in 2040 (Centraal Bureau voor de Statistiek [CBS], 2023). More and more elderly need care, whilst at the same time, there is a growing shortage of healthcare staff (ABF Research, 2023). This problem becomes even more urgent in the future: there are currently three people of working age for every person over 65 years, but this ratio is expected to decline to two in the coming decades (CBS StatLine, 2023a).

Projections of elderly care spending indicate that the current Dutch elderly healthcare system is unsustainable in the long run (Rijksoverheid, 2023). Policymakers and other stakeholders in the healthcare sector are, therefore, looking for ways to keep elderly care accessible and affordable without compromising on quality. Several national initiatives, such as the Integrated Healthcare Agreement in 2022, attempt to contribute to this challenge. A key pillar within these initiatives is devoted to designing payment systems with the right financial incentives for healthcare providers so that they provide high-quality care while using the available resources as efficiently as possible.

The payment system in geriatric rehabilitation (GR), which is temporary, multidisciplinary care for frail people with multi-morbidity, is currently based on stepwise increasing reimbursements that depend on the treatment duration. More precisely, the reimbursement scheme follows a discontinuous step function: once the treatment duration has crossed a threshold, the reimbursement increases once and then stays flat until the next threshold is reached. As the distances between the thresholds are unequal, the reimbursement scheme resembles a staircase tariff pattern with unequal step lengths. For GR, the treatment duration is determined by both the treatment intensity (as measured in the number of treatment hours) and the length of stay (as measured in the number of inpatient days).

Reimbursement schemes with staircase tariff patterns, such as in GR, may distort physicians' behaviour in two ways. On the one hand, physicians may — possibly medically unnecessary — prolong treatment to cross a threshold to reach a higher reimbursement. On the other hand, physicians may stop treatment once a higher reimbursement is reached — while it might not be in the patient's interest. The goal of this thesis is to examine whether these behavioural distortions are present for GR providers and, if so, which factors are associated with stronger distortions. Furthermore, this thesis aims to analyse the possible consequences of the proposed changes in the payment system in GR, including the abolition of the current stepwise reimbursements, on the efficiency of care provision.

Using claims data from Dutch health insurance company CZ about GR treatments in the years 2017–2023, I first determine whether there are distortions in physicians' behaviour present and assess the applicability of the bunching approach, as explained by Kleven (2016), which enables estimation of the magnitude of the behavioural distortions. Next, I perform a multiple linear regression analysis with the GR claims data to explain the variation in the strength of responses between GR providers regarding the length of stay. Finally, I perform a qualitative assessment to evaluate the proposed transition to a new payment system in GR from an efficiency perspective and to provide recommendations on how the current payment system in GR can be adjusted so that that it better meets the criteria as drawn up by the Dutch Healthcare Authority (Nederlandse Zorgautoriteit [NZa], 2024a).

The contribution of this thesis is three-fold. First, I add to the understanding of whether stepwise reimbursements lead to distortions in physicians' behaviour within a sizeable field of elderly care in the Netherlands. By analysing the length of stay and treatment intensity at the same time, I provide insights into how similar financial incentives in the reimbursement scheme may affect physicians' behaviour differently. These insights are valuable for policymakers when designing payment systems in healthcare. Second, I provide evidence on the correlation between exceeding the budget ceiling, which indicates whether a healthcare provider exceeds the pre-determined maximum amount that a provider can claim from a health insurer, and the strength of behavioural distortions of a GR provider. The budget ceiling has not been analysed before in this context, while knowledge about this correlation is paramount to improving contracts between health insurers and providers. Third, I am the first to extensively assess the possible consequences of the new payment system in GR in light of the efficiency of care provision using insights from economic theory and GR claims data. Moreover, I give recommendations on how the current payment system in GR can be improved. This information can help policymakers to make informed decisions regarding sustainable payment systems in healthcare.

The remainder of this thesis is structured as follows. In Section 2, I explain the organisational structure of GR in the Netherlands. Then, in Section 3, I describe the payment system of GR in detail. Additionally, I discuss the positioning, including the contribution, of this thesis in the existing literature on payment systems with discontinuous tariffs. In Section 4, I summarise the GR claims data from CZ belonging to the years 2017–2023. Next, I outline the bunching approach, multiple linear regression analysis and the qualitative assessment in Section 5. Then, I present and evaluate the results in Section 6. Finally, I state the conclusion and discussion, including limitations and recommendations, in Section 7.

2 Organisational structure

2.1 Definition

GR is defined in Article 2.5c(1) of the Health Insurance Decree (*Besluit Zorgverzekering*) as "integrated and multidisciplinary rehabilitation care provided by elderly care physicians, in relation to vulnerability, complex multi-morbidity and limited learning and training ability, aimed at reducing the functional limitations of the insured individual such that return to the home situation is possible". From this definition, multiple elements can be distinguished:

- **Integrated and multidisciplinary rehabilitation care provided by elderly care physicians**. An integrated and multidisciplinary approach applies to GR in which several practitioners work closely together to provide appropriate care. The elderly care physician is the leader of the team. The composition of the team is further dependent on the rehabilitation goals of the patient, which are described in the treatment plan. In general, the team consists of (specialised) nursing staff, physiotherapists and occupational therapists. Depending on the patient's needs, speech therapists, dietitians, psychologists, music therapists or social workers may also be part of the team. Besides a list of the practitioners involved, the treatment plan includes a provisional discharge date, which is evaluated and possibly adjusted during (bi-)weekly team meetings.

- **In relation to vulnerability, complex multi-morbidity and limited learning and training ability**. The patient group of GR is characterised by vulnerability due to, for example, restricted mobility. Multi-morbidity refers to the situation when a person has multiple diseases, disorders, limitations or handicaps at the same time. Complex multi-morbidity occurs when there is a loss of well-being, and it is difficult to distinguish how the various morbidities separately contribute to this loss (Zorginstituut Nederland [ZIN], n.d.). The limited learning and training ability means that a patient has difficulty coping with physical or mental strain. As a result, the GR treatment must be adapted to the individual's recovery situation and training pace.

- **Aimed at reducing the functional limitations of the insured individual such that return to the home situation is possible**. A GR treatment aims to restore functional and cognitive capabilities so that patients can return home. In other words, patients should be able to perform basic activities of daily living independently again, such as going to the toilet or going up the stairs without assistance.

The other paragraphs of Article 2.5c of the Health Insurance Decree specify some features of a GR trajectory. First, Article 2.5c(2a) states that a GR trajectory is only possible within a week after hospital discharge. However, since 2020, it is also possible that patients get admitted from home, so without prior hospital admission (Verenso, 2020). Next, Article 2.5c(2b) defines that a GR trajectory always starts as an inpatient trajectory, meaning that a patient sleeps in a GR facility. Later in the treatment process, a patient may receive outpatient or ambulatory GR at home or at the GR facility. Hence, an ambulatory GR trajectory immediately from the start is not reimbursed nowadays, but there are currently experiments running that make this possible (ActiZ, 2022). Finally, Article 2.5c(3) formulates that the duration of a GR trajectory is set to a maximum of six months. In extraordinary circumstances, health insurers may allow a longer period after approval.

2.2 Providers

GR providers are mostly large healthcare institutions (>1000 employees) with multiple facilities that provide care and housing for primarily elderly people. They mainly offer GR within specialised rehabilitation departments of their nursing homes. However, it is also possible that they provide GR at separate rehabilitation facilities or specific units within hospitals. Due to the comprehensive provision of care, GR usually represents only a small part of the revenues of these healthcare institutions (SiRM, 2023). Almost all have the legal form of foundation and, hence, are not-for-profit oriented. In total, there are around 140 GR providers in the Netherlands (Vektis, 2021).

Regarding care provision and patient group, GR providers differ considerably from medical specialist rehabilitation providers. The latter group focuses on medically stable patients who can tolerate high-intensity therapy. Furthermore, medical specialist rehabilitation is not offered under the supervision of an elderly care physician, but under the supervision of a physiatrist in rehabilitation units within hospitals, specialised rehabilitation centres or independent treatment centres. Consequently, there is little overlap of patients between GR and medical specialist rehabilitation providers in practice (Studio GRZ, 2014).

There are five main diagnosis groups that GR providers distinguish: CVA (Cerebrovascular Accident, or informally: stroke), elective orthopaedic surgery, trauma orthopaedic surgery, amputations and other disorders. The corresponding treatment shares, as shown in Table 2.1, are in descending order: 37.8% for other disorders, 29.8% for trauma orthopaedic surgery, 16.5% for CVA, 13.3% for elective orthopaedic surgery and 2.6% for amputations (ZIN, 2020).

Several sub-diagnoses can be identified within these main diagnosis groups, as shown in Table 2.1. More than half (51.1%) of all patients has a CVA (16.5%), hip fracture (14.8%), other organ disorder (11.3%) or respiratory disorder (8.5%) sub-diagnosis (ZIN, 2020).

Table 2.1

Geriatric rehabilitation diagnosis groups with corresponding treatment share

Main diagnosis group	Share	Sub-diagnosis	Share
CVA	16.5	CVA	16.5
Elective orthopaedics	13.3	Hip new prosthesis	6.6
		Hip revision	1.6
		Knee new prosthesis	4.0
		Knee revision	0.5
		Other elective orthopaedics	0.8
Trauma orthopaedics	29.8	Upper extremity trauma	1.7
		Lower extremity trauma (excluding hip)	6.0
		Hip fracture	14.8
		Traumatic vertebral fracture	1.0
		Intracranial trauma	0.9
		Other trauma	5.5
Amputations	2.6	Amputation upper leg (and higher)	0.6
		Amputation lower leg/feet/toes	1.9
		Upper extremity amputation	0.0
Other disorders	37.8	Heart diseases	3.9
		Blood vessels	1.4
		Respiratory disorders	8.5
		Oncological disorders	4.0
		Other organ disorders	11.3
		Other upper extremity disorders	0.3
		Other lower extremity disorders	1.2
		Spine disorder	1.3
		Rheumatic diseases	0.2
		Other musculoskeletal disorders	1.5
		Other brain disorders	1.2
		Neuromuscular disorders	0.5
		Other neurological disorders	2.4

Note. Treatment shares in %. Source: National Health Care Institute (ZIN, 2020).

3 Payment system

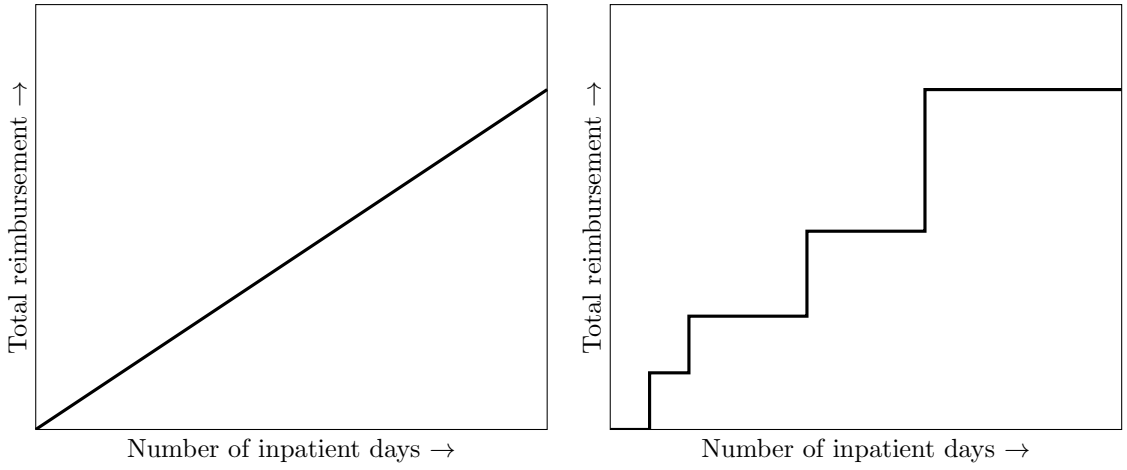
3.1 Transition to the Health Insurance Act

Originally, GR was funded under the Exceptional Medical Expenses Act (*Algemene Wet Bijzondere Ziektekosten*). The payment system under this act entailed that there were fixed budgets for GR providers, which were spent on GR patients by charging a fixed tariff per inpatient day. There was no separate reimbursement for the treatment intensity, which is the time spent by the different practitioners on each patient, because this was discounted in the fixed day tariff. In other words, the old GR payment system did not have any financial incentives to encourage shorter lengths of stay or to intensify treatment by providing more treatment hours.

From the 1st of January 2013, the funding of GR transferred to the Health Insurance Act (*Zorgverzekeringswet*) to improve the quality and efficiency of GR treatments (Parliamentary Documents II, 2010/11, 30597, no. 185). Better quality would be achieved because the new payment system would allow for more tailored care by reimbursing the treatment intensity next to the length of stay.[1] Greater efficiency would be achieved because the new payment system would reduce lengths of stay by not working with total reimbursement based on a fixed day tariff times the number of inpatient days, but instead, with pre-determined reimbursements based on subgroups of inpatient days. This difference between the two reimbursement schemes is shown in Figure 3.1 for a given treatment intensity.

Figure 3.1

Illustrative reimbursement scheme for the length of stay in geriatric rehabilitation under the Exceptional Medical Expenses Act (left) and the Health Insurance Act (right) for a given treatment intensity



[1] Incorporating the treatment intensity into the payment system is quite special because this only applies to medical specialist rehabilitation next to GR in the Netherlands. Aligning the payment systems of these two forms of rehabilitation was an important principle when developing the current payment system in GR (NZa, 2012).

The transition of the funding of GR to the Health Insurance Act turned out well. The mean treatment intensity increased by 37% and the median length of stay decreased by seven days, as measured in 2015 (Bouwstra et al., 2017). Whether these outcomes can entirely be attributed to the changed payment system is unclear because only a correlation and not a causal relationship has been shown. Consequently, it cannot be ruled out that other factors that changed due to the transition played a role as well. In addition, there could have been a shift in the patient population (Bouwstra et al., 2017). For example, patients who would be more likely to sustain high treatment intensities and return home after GR could have been selected after the transition. Or, conversely, patients who would benefit from long-term and low-intensity rehabilitation could have been excluded from a GR trajectory after the transition and may have ended up in a regular nursing department instead.

Total healthcare costs of GR decreased in the years following the transition: from 800 million in 2013 to 707 million euro in 2017 (ActiZ, 2019; ZIN, 2023). However, starting from the year 2017, the costs for GR have shown an increasing trend, as presented in Table 3.1. Even in the COVID-19 years 2020 and 2021, the costs for GR rose due to compensation for providers for lost income. Nevertheless, the rise in healthcare costs is still relatively modest as the annual cost growth factor for GR corresponds to only 2.0% in the years 2017–2022, which is lower than the average of 5.1% for the entire Dutch healthcare sector in that period (CBS StatLine, 2023c).

Table 3.1

Total healthcare costs for geriatric rehabilitation in the Netherlands

Costs	Year					
	2017	2018	2019	2020	2021	2022
Regular	707	733	749	652	722	780
Supplementary due to COVID-19	-	-	-	123	78	-
Total	707	733	749	775	800	780

Note. Unit of costs is in million euro. Source: National Health Care Institute (ZIN, 2023).

Part of the increase in healthcare costs can plausibly be explained by a growing number of GR trajectories, as shown in Table 3.2. The annual growth factor of the number of GR trajectories for the period 2017–2022 was 0.6%. The other part of the increase was then likely due to a rise in the average reimbursement per GR trajectory.

Similar to the dip in healthcare costs, the total number of trajectories in 2020 and 2021 was substantially lower than in the years before, as shown in Table 3.2. This drop is likely due to the COVID-19 pandemic, which caused significantly fewer people to stay in the hospital (CBS StatLine, 2023), which subsequently led to less outflow to GR facilities.

Table 3.2

Total number of geriatric rehabilitation trajectories in the Netherlands

	Year						
	2017	2018	2019	2020	2021	2022	2023[a]
Number of trajectories	46,224	49,534	46,443	41,131	43,903	47,919	41,780

Note. Source: Dutch Healthcare Authority (NZa, 2024b).
[a]Not complete yet.

3.2 Diagnosis treatment combinations

The payment system in GR under the Health Insurance Act is based on diagnosis-treatment combination (DTC) care products (in short: DTCs). These are packages of care activities and procedures required to treat a patient with a specific diagnosis. Hence, healthcare providers do not claim every provided service separately from the health insurer but get paid one price for the entire care pathway.

DTCs are labelled with codes that do not change over the years. However, the price for each DTC does vary over the years as healthcare providers make one- or multi-year price agreements with health insurers. Depending on whether the DTC falls within the regulated segment or not, a maximum price applies. That is, the Dutch Healthcare Authority sets a maximum price for each DTC in the regulated segment every year (NZa, 2024b). The prices are based on average costs, and are indexed annually and periodically recalibrated. The maximum prices form the upper bounds in the negotiating process for the price agreements. There is even a max-max price for GR, which implies that a provider may charge a maximum of 10% on the maximum price if this has been agreed on with the health insurer. In practice, health insurers reimburse, on average, between 97 and 99% of the maximum prices that apply for GR in the years 2018–2022 (NZa, 2024c).

A DTC belonging to a GR trajectory is opened on the first day that patient-related contact takes place in the context of admission, observation or examination. It is closed on the 42nd day after the date of the last registered care activity, provided that the maximum lead time of 120 days is not exceeded. In other words, it is closed by default on the 120th day after opening unless it is closed earlier based on the 42-day rule, which is only relevant if the last care activity is registered on day 77 or earlier. Since the duration of a GR treatment is set to a maximum of six months, multiple DTCs can belong to a single care pathway of a GR patient. However, a GR provider can claim the reimbursement immediately after the closure of a DTC and, hence, does not have to wait until the entire treatment is finished.

3.3 Structure of diagnosis-treatment combinations

All 36 DTCs within the regulated segment of GR have a nine-digit code starting with 9984180 (see for a complete overview of all DTCs: Table A.1 in Appendix A). Then, two digits are added to the end, depending on the specific combination of the following four components:

- **Trajectory**. A clinical (30 DTCs) or ambulatory (6 DTCs) trajectory are the only possibilities within the DTC structure of GR. A clinical trajectory entails that care is initially provided on an inpatient basis but can be followed by treatment on an outpatient basis. In contrast, within an ambulatory trajectory, care is solely provided on an outpatient basis. An ambulatory trajectory is only possible as a follow-up trajectory.

- **Diagnosis**. While multiple main diagnosis groups and sub-diagnoses are specified within GR (see Table 2.1), the DTC structure only distinguishes two diagnoses: CVA and other (18 DTCs each). The category 'other' contains thus all possible main diagnosis groups (elective orthopaedic surgery, trauma orthopaedic surgery, amputations and other disorders) except CVA. This classification originates from the time when GR was funded under the Exceptional Medical Expenses Act, where this distinction was also used (NZa, 2012).

- **Length of stay**. The number of inpatient days determines the length of stay. An inpatient day is a calendar day that is part of a nursing period, including overnight stays. The total length of stay runs from admission to discharge, where the admission day (provided it took place before 8 p.m.) and the discharge day are both calendar days to be registered. If a patient goes back to the home situation for a trial leave of a maximum of three days, these days may not be registered as inpatient days. Trial leave often happens when it is expected that, if all goes well, the patient will not return to the GR facility.

  Logically, ambulatory trajectories do not depend on the length of stay. For clinical trajectories, the length of stay can be divided into subgroups of inpatient days, independently of the diagnosis: 1–14, 15–28, 29–56, 57–91 and 92–120 days. Hence, the thresholds are located exactly after 2, 4, 8 and 13 weeks. Or, to put it differently, new reimbursement tariffs start on days 15, 29, 57 and 92. These boundaries originate from the frequency distribution of the length of stays in GR in 2011 (NZa, 2012).

- **Treatment intensity**. The number of treatment hours determines the treatment intensity. Both direct and indirect patient-related time are recorded in blocks of five minutes. Direct time includes face-to-face contact with the patient, including meetings in the patient's presence. Nurses only record direct patient-related time for specialist activities, such as infusion or respiratory support. Other nursing activities are not registered separately as these are discounted in the reimbursements for the length of stay. Indirect time encompasses time spent on the patient without the patient's presence, such as multidisciplinary consultations, writing a report or analysing test results. Given the multidisciplinary nature of GR, the recorded time of the elderly care physician, paramedical and behaviour support practitioners count equally. Non-patient-related treatment time, such as time spent on travel and training, may not be registered.

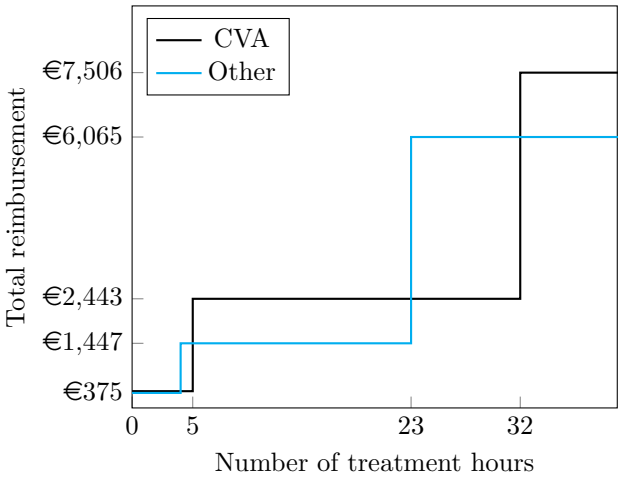
  Similar to the length of stay, the treatment intensity can be divided into different subgroups of treatment hours. The boundaries of these subgroups are based on historical treatment hours in which a classification of various treatment intensities was used (NZa, 2012). There are no fixed subgroups regarding treatment intensity because these depend on three factors: trajectory, diagnosis and length of stay. To illustrate, the four subgroups belonging to the clinical CVA trajectory with a length of stay between 1 and 14 inpatient days are 0–7, 7–13, 13–45 and >45 treatment hours, while the two subgroups belonging to the clinical CVA trajectory with a length of stay between 57 and 91 inpatient days are 0–75 and >75 treatment hours. If the length of stay is between 92 and 120 days, there are even no subgroups for treatment intensities at all, no matter what the diagnosis is. A complete overview of all treatment intensity subgroups for different trajectories, diagnoses and lengths of stay is given in Table A.1 in Appendix A.

3.4 Current reimbursement scheme

The DTC structure results in a reimbursement scheme with stepwise tariffs for the two possible ambulatory trajectories of CVA and other diagnoses, as shown in Figure 3.2. Since ambulatory trajectories only depend on the treatment intensity and not on the length of stay, a staircase pattern with various step lengths is easily visible. Figure 3.2 also shows that the reimbursement for CVA is, on average, higher than for the other diagnoses. This difference in tariffs originates from the time that GR was funded under the Exceptional Medical Expenses Act, where a surcharge for CVA trajectories applied (NZa, 2012).

Figure 3.2

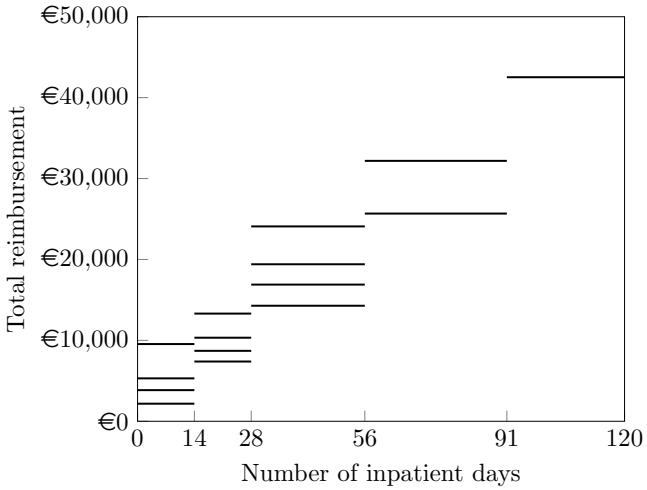
Reimbursement scheme of ambulatory trajectories in geriatric rehabilitation in prices 2023



The reimbursement scheme of clinical CVA trajectories in GR is shown in Figure 3.3. A similar scheme can be made for the clinical trajectories belonging to the other diagnoses (see Figure B.1 in Appendix B). The horizontal lines indicate the maximum price for different treatment intensities. It is more difficult to recognise a staircase pattern in Figure 3.3 compared to Figure 3.2 because the clinical trajectories depend on both the length of stay and the treatment intensity.

Figure 3.3

Reimbursement scheme of clinical CVA trajectories in geriatric rehabilitation in prices 2023

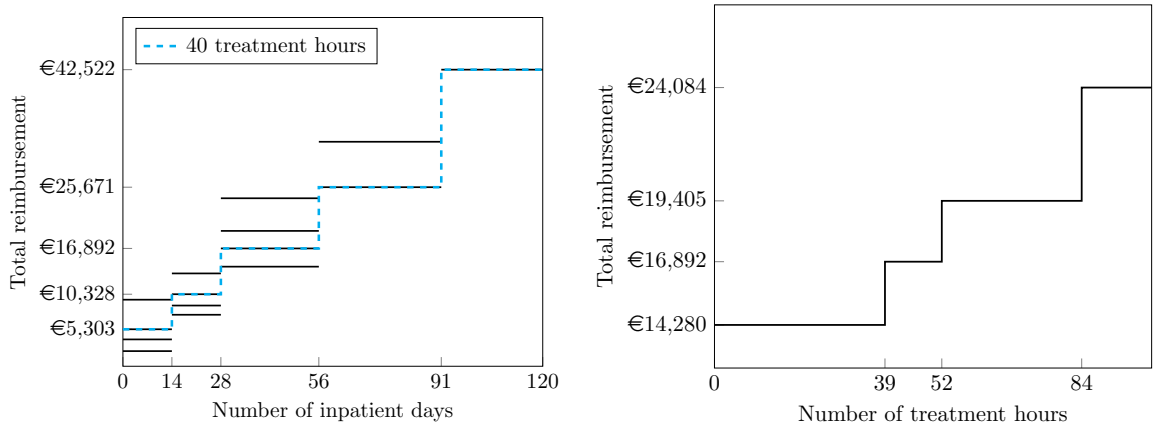


Note. The horizontal lines represent different treatment intensities.

To show the staircase tariff patterns in the reimbursement scheme for the clinical trajectories in GR, either the treatment intensity or the length of stay must be fixed. When the treatment intensity is fixed to, for example, 40 treatment hours, the discontinuous step function for the length of stay becomes visible again, as shown on the left in Figure 3.4 for the CVA diagnosis. When the length of stay is fixed to, for example, 28–56 days, then the discontinuous step function for the treatment intensity becomes visible, as shown on the right in Figure 3.4, again for the CVA diagnosis. The same trick can be applied to see the staircase patterns for other fixed treatment intensities or lengths of stay, or for the clinical trajectories belonging to the other diagnoses in GR.

Figure 3.4

Reimbursement scheme of clinical CVA trajectories in geriatric rehabilitation for a fixed treatment intensity of 40 treatment hours (left) or fixed length of stay of 29–56 days (right) in prices 2023



3.5 Proposed future payment system

In February 2024, the Dutch Healthcare Authority announced that the current payment system in GR will be renewed (NZa, 2024a). The initial plan was to change the payment system for different forms of temporary medical stays but not for GR. However, GR was later added to the plan because the Dutch Healthcare Authority concluded — in consultation with health professionals and other stakeholders in the healthcare sector — that multiple payment systems next to each other lead to complexity, sub-optimal use of available resources and unnecessary administrative burdens.

The new payment system in GR will be based on different modules. There is a module for residence that reimburses the length of stay per inpatient day (similar to the payment system in GR under the Exceptional Medical Expenses Act). For intensive nursing and care for complex target groups, there is a separate module for residence. Besides residential modules, there are modules for treatment intensity with a distinction between the time spent by the medical, paramedical or behavioural support

practitioner. The reimbursement unit for the treatment modules is per hour with the direct patient-related time recorded per five minutes. Hence, in contrast to the current payment system in GR, there will be no fixed subgroups of the length of stay or treatment intensity, no distinction between diagnoses, and no recorded indirect patient-related time anymore.

The modular payment system will be implemented at some GR providers in the Netherlands from January 1, 2025. The first years serve as a transition period to further develop the modules so that they better fit the practice. Then, as of January 1, 2030, the new system will be implemented nationally.

3.6 Economic perspective on different payment systems

According to Ellis and McGuire (1986), a physician serves as the key decision-maker who selects the services to be provided to a patient, and trades off the benefits to the patient against the benefits to the healthcare provider. Hence, a physician functions as an intermediary representing the interests of the patient and the healthcare provider. Financial incentives in payment systems may result in conflicting interests between these two. In other words, a treatment that is optimal for the provider might not coincide with one that is optimal for the patient.

When the payment system in GR was still funded under the Exceptional Medical Expenses Act, the patient and the provider shared a common interest in more treatment: the patient for more rehabilitation benefits and the provider for more profits (under the condition that the day tariffs were higher than the associated costs). Both interests were thus perfectly compatible. These interests will stay aligned under the modular payment system in GR, provided that the tariffs per treatment hour and in-patient day are higher than the corresponding costs. However, conflicting interests arise in the current payment system in GR. While the preference of a patient for more treatment remains unchanged,[2] the preference of a provider is no longer solely focused on more treatment because extending treatment is not necessarily associated with higher profits anymore. That is because the marginal revenue in the current payment system in GR is everywhere zero except at the thresholds, while the marginal costs are always positive. So, reasonably assuming that the marginal costs of treatment do not discontinuously change at the thresholds, providers' profits generally decrease when more treatment is provided. Moreover, as there are a finite number of steps with an absolute limit on the total reimbursement in the last step, there is no incentive for physicians to pursue the next threshold endlessly.

[2]While most patients likely prefer more treatment if there are rehabilitation benefits to gain, this does not apply to all patients because some value their time more than the potential added rehabilitation benefits. Since this was also the case in the old payment system, I conclude here that a patient's preference remains unchanged in that regard.

In practice, elderly care physicians likely base their decisions regarding the treatment duration in GR on a combination of the interests of the patient and the healthcare provider. If physicians predominantly value the patients' interests, they do not take the jumps in reimbursements into account but only base their decisions on rehabilitation considerations. In contrast, if physicians mainly value the financial interests of their employer, it is optimal for them to set the treatment duration such that the marginal revenue is positive, which is precisely at the thresholds. Highly valuing financial interests may lead to two types of behavioural distortions. On the one hand, this may result in over-provision of care when physicians decide to extend treatment such that a threshold will be crossed. This type of overtreatment may be financially inefficient but not necessarily detrimental to the quality of care because it may not be disadvantageous for a patient to get a little more treatment than would be necessary from a rehabilitation perspective. On the other hand, this may result in under-provision of care when physicians decide to stop treatment after a threshold has been crossed. This type of undertreatment may be beneficial for efficiency reasons but may harm the quality of care when a patient does not receive adequate treatment.

Even when completely ignoring patients' interests, physicians probably face optimisation frictions that prevent them from setting the treatment duration precisely at the thresholds. These frictions include personnel hours constraints, inattention, inertia, incomplete information and uncertainty in forecasting, among other things. As a result, the treatment duration is not expected to be exactly at a threshold but rather spread out over a range around a threshold.

3.7 Contribution to the literature

Determining the treatment duration based on financial considerations rather than clinical ones is known in the literature as strategic discharge behaviour (Eliason et al., 2018). Sometimes, it is also referred to as 'upcoding', but this is strictly speaking incorrect. Upcoding is a fraudulent medical billing practice in which a more expensive service is charged than is performed. Since GR providers actually provide the care for which they claim the reimbursement from health insurers, it is not a criminal act but rather improper or unethical behaviour.

Some research has been done on strategic discharge behaviour in the context of discontinuous reimbursement schemes. Notable studies are those of Eliason et al. (2018) and Einav et al. (2018) who examined the introduction of a discontinuous reimbursement scheme for the length of stay in long-term care hospitals in the United States of America. Both studies find that healthcare providers

respond to a large, one-off increase in tariffs by delaying discharges until the higher tariff is reached and then discharging a substantial proportion of patients immediately after. Similar behaviour has been found by Douven et al. (2015) in a Dutch mental healthcare setting, and by Gaspar and Koolman (2022) and Gaspar et al. (2023) in respectively an outpatient setting and inpatient setting, both for Dutch medical specialist rehabilitation. In contrast to these studies, Pletscher (2016) concludes that behavioural distortions regarding the length of stay are only present when the differences between tariffs are sufficiently large for psychiatric care in Switzerland, while Pott et al. (2021) found no significant behavioural distortions at all for psychiatric care in Germany.

This thesis adds to the literature by giving new insights into the possible existence of strategic discharge behaviour within GR in the Netherlands. While all related studies focus on only one aspect of the treatment duration, this thesis examines the length of stay as well as the treatment intensity. The main advantage of considering both aspects is that it can give insights into how similar financial incentives may have different effects on physicians' behaviour. This information is helpful for policymakers when developing optimal payment systems in healthcare.

Related studies have already examined the variation in behavioural distortions between healthcare providers. Factors that are found to be relevant in explaining differences are the specific admission day in the week[3] and patient characteristics (Gaspar et al., 2023), the profit orientation of a provider (Eliason et al., 2018) and whether a provider is in financial distress or not (Gaspar and Koolman, 2022). This thesis also examines differences in the strength of providers' responses by including data about the budget ceiling, which is the maximum amount that a healthcare provider can claim from the health insurer in a year. This factor has not been explored before in the context of discontinuous tariffs.[4] Insights into the relationship between exceeding the budget ceiling and behavioural responses are valuable for health insurers to improve contracts with healthcare providers.

Finally, while many healthcare systems worldwide move away from a fee-for-service type towards more bundling of care payment systems, an opposite movement will take place for Dutch GR shortly. Although it is too early yet to assess the effects of this change quantitatively, this thesis provides qualitative insights into the possible consequences of this transition for the efficiency of care provision. These insights can act as a starting point for new research and may be useful when comparing the actual outcomes in a few years with what is expected based on the qualitative assessment.

[3]Gaspar et al. (2023) show that fluctuations in the day of the week a patient was admitted (and therefore due to be discharged) are relevant. For example, a patient that reaches the threshold day during the weekend is less likely to be discharged that day than someone who reaches the threshold day on a Friday.

[4]A thorough literature search on this factor in relation to the topic of this thesis yielded no relevant results.

4 Data

4.1 Health insurer CZ

Anyone who lives or works in the Netherlands is obliged to conclude a contract for basic health insurance (Article 2 of the Health Insurance Act). There are currently eleven health insurers in the Netherlands (Zorgverzekeraars Nederland, n.d.). Health insurance company CZ (*Centraal Ziekenfonds*) is a not-for-profit organisation operating on behalf of three brands: CZ, Nationale-Nederlanden and OHRA. It is the second largest health insurer in the Netherlands with four million policyholders in 2023, corresponding to a market share of 23% (Vektis, 2023).

Historically, CZ has a strong presence in the southern regions of the Netherlands. However, there has been an acceptance obligation since 2006, implying that health insurers must accept all applicants seeking basic insurance (Article 3 of the Health Insurance Act). Hence, health insurers are not allowed to select applicants based on residential location or other factors, such as health condition, age or financial position. There is also a ban on premium differentiation for basic insurance, which means that health insurers are obliged to charge the same premium for everyone's basic insurance.

To ensure that all policyholders have access to all care covered by their basic insurance within a reasonable time and travel distance, health insurers must purchase sufficient care from healthcare providers across the Netherlands. This is stipulated in the legal duty of care for health insurers (Article 11 of the Health Insurance Act). Generally, Dutch health insurers have contracts with all large healthcare institutions in the Netherlands.

CZ has a care purchasing policy for GR that is uniform across all three brands. This policy formulates CZ's vision, principles and requirements in the procurement process regarding GR (CZ, 2023). These are, in short, the following. GR providers can choose to conclude a contract for one or two years. The specific agreements differ depending on providers' performances, regional circumstances and potentially planned experiments. Performance is evaluated by CZ's value model, which connects quality with efficiency, resulting in an A (best), B or C (worst) score. Furthermore, each contract specifies the budget ceiling. In the event of approaching this ceiling, supplementary agreements may be negotiated (e.g. raising the ceiling) or the provider may implement an admission stop for new patients. Providers must pay back the overclaimed amount if they exceed the budget ceiling. Finally, as a general rule, CZ requires that GR providers treat at least 60 patients with CVA, 70 with orthopaedic surgery (elective or trauma) or 60 with other diagnoses per year.

4.2 Descriptive statistics

4.2.1 Data availability

In correspondence with the Health Insurance Act and CZ's privacy statement, personal data from insured people may, in principle, be kept for seven years (CZ, n.d.). The GR claims data at CZ available for this thesis cover the period 2017 up to and including 2023. However, as providers can submit GR claims to CZ within twelve months after the closure of the DTC, the data concerning the year 2023 are not complete yet.

4.2.2 Number of providers

The total number of GR providers contracted with CZ fluctuated between 132 and 145 over the years, as shown in Table 4.1. There were no waiting lists for GR in the years 2017–2023. Hence, the insured people of CZ who needed GR in that period, received this care within 48 hours and 30 minutes of travel time from their place of residence.

Table 4.1

Overview of the total number of geriatric rehabilitation providers contracted with CZ

	Year						
	2017	2018	2019	2020	2021	2022	2023
Number of providers	145	139	137	140	136	135	132

4.2.3 Breakdown by trajectory

An overview of the total number of GR claims at CZ belonging to the clinical and ambulatory trajectories is presented in Table 4.2. Based on the market share of CZ that varied between 20–23% in the years 2017–2023 (Vektis, 2023), the total number of GR claims at CZ is slightly higher each year than one would expect based on the total number of GR claims in the Netherlands (see Table 3.2). A plausible explanation for this difference is that CZ's insured population is, on average, relatively older.

In the COVID-19 years 2020 and 2021, respectively 12 and 9% fewer people stayed in the Dutch hospitals compared to the pre-COVID-19 year 2019 (CBS StatLine, 2023b). This is largely in line with the total number of GR claims at CZ as displayed in Table 4.2, which shows a decline of respectively 12% and 6% in 2020 and 2021 compared to 2019.

Based on the total number of claims at CZ (Table 4.2) and the total number of contracted GR providers (Table 4.1) in each year, the mean number of claims per GR provider varied between 73 to 87 over the years, excluding 2023 because not all claims for that year have yet been submitted to CZ.

Since the total number of ambulatory GR claims only comprises 0.5% of the total number of GR claims at CZ over the years, ambulatory trajectories are not considered in the remainder of this thesis.

Table 4.2

Overview of the number of geriatric rehabilitation claims at CZ belonging to different trajectories

Trajectory	Year							
	2017	2018	2019	2020	2021	2022	2023[a]	2017–2023
Clinical	11,767	11,953	11,586	10,206	10,967	11,457	8966	76,902
Ambulatory	75	71	82	57	44	53	35	417
Total (clin. + amb.)	11,842	12,024	11,668	10,263	11,011	11,510	9001	77,319

[a]Not complete yet.

4.2.4 Characteristics of clinical claims

The GR claims at CZ belonging to clinical trajectories are virtually representative of the claims in the Netherlands based on the ratio between CVA and other diagnoses: 16.8% versus 83.2%, respectively (see Table C.1 in Appendix C). A more detailed breakdown by main diagnosis groups also indicates that these GR claims at CZ are representative (see again Table C.1 in Appendix C).

Table 4.3 shows the summary statistics of the 76,902 GR claims at CZ belonging to clinical trajectories from the years 2017–2023. There is considerable variation in the length of stay: the mean is 39 inpatient days with a standard deviation of 26. There is also substantial variation in the treatment intensity: the mean is 33 treatment hours with a standard deviation of 28. Next, the mean age is 79 years with a standard deviation of 10 years. Moreover, of all patients, 92% is 65 years or older, and 60% is women. Finally, the mean reimbursement per claim is 13,194 euro with a standard deviation of 8306 euro. Hence, the total costs of GR for CZ add up to over one billion euro for the years 2017–2023.

Table 4.3

Summary statistics of geriatric rehabilitation claims at CZ belonging to clinical trajectories

Mean length of stay (inpatient days)	39
Standard deviation length of stay (inpatient days)	26
Mean treatment intensity (treatment hours)	33
Standard deviation treatment intensity (treatment hours)	28
Mean age (years)	79
Standard deviation age (years)	10
Older than 65 years (%)	92
Female (%)	60
Mean reimbursement (€)	13,194
Standard deviation reimbursement (€)	8306

Note. Statistics are calculated based on 76,902 claims at CZ from the years 2017–2023.

A closer look at the length of stay shows that the mean length of stay varies significantly for the five main diagnosis groups, as presented in Table 4.4. The difference between the highest mean (57 for amputations) and the lowest mean (29 for elective orthopaedic surgery) is 27 inpatient days, which corresponds to almost four weeks. The standard deviations for the various main diagnosis groups are also relatively large, varying from 20 to 34 inpatient days.

Table 4.4

The length of stay for the different main diagnosis groups in geriatric rehabilitation

	Main diagnosis group				
	CVA	Elective orth. surg.	Trauma orth. surg.	Amputations	Other
Mean length of stay[a]	44	29	41	57	37
SD length of stay[a]	29	20	25	34	25

Note. Statistics are calculated based on 76,902 claims at CZ belonging to clinical trajectories from the years 2017–2023.
[a]Mean and standard deviation (SD) are measured in the number of inpatient days.

4.2.5 Role of the admission day and weekend

Patients are not admitted to a GR facility evenly throughout the week, as shown on the first row of Table 4.5. As the workweek progresses, relatively more people get admitted to a GR facility. An explanation for this phenomenon is that hospitals, where most GR patients come from, want to discharge their patients before the weekend because less staff is available on Saturday and Sunday. As a result, there are relatively few admissions (3.4%) to a GR facility on the weekend days.

The weekend effect is also present for the discharge day of GR patients, as shown on the second row of Table 4.5. Relatively many discharges (22.0%) happen on Friday, which is likely because patients often go on trial leave during the weekend and do not return afterwards. However, in contrast to the admission day, a larger share (12.2%) of the GR patients are discharged on weekend days. A reason for this could be that staff constraints play less of a role because the personnel planning in a nursing home may be more predictable since many patients stay there for a longer period compared to the hospital.

Table 4.5

Shares of geriatric rehabilitation patients admitted and discharged on a particular day in the week

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Admission day	18.1	19.2	19.2	19.6	20.5	2.7	0.7
Discharge day	15.2	15.9	17.0	17.6	22.0	7.2	5.0

Note. Shares (%) are calculated based on 76,902 claims at CZ belonging to clinical trajectories from the years 2017–2023.

Table 4.6 zooms in on the percentage of GR patients discharged on a particular day given the admission day. The relatively high values on the diagonal show that many patients are discharged after whole weeks. Also, the weekend effect can still be seen with a relatively high share of discharges ($> 20\%$) on Friday.

Table 4.6

Shares of discharge days of geriatric rehabilitation patient given the admission day

Admission day	Discharge day						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Monday	18.2	15.7	16.0	16.0	21.5	7.3	5.2
Tuesday	15.3	19.7	16.4	16.4	20.6	6.6	5.1
Wednesday	13.8	16.1	20.8	17.7	20.3	6.8	4.5
Thursday	13.9	13.9	17.6	21.0	21.5	7.1	4.9
Friday	15.1	14.5	14.7	17.3	25.9	7.7	4.9
Saturday	14.5	15.3	15.6	15.1	21.9	11.9	5.8
Sunday	13.1	15.8	16.1	16.1	22.7	9.9	6.2

Note. Shares (%) are calculated based on 76,902 claims at CZ belonging to clinical trajectories from the years 2017–2023.

4.2.6 Budget ceiling

Table 4.7 shows the number of GR providers exceeding the budget ceiling, including any interim agreed increases between the provider and the health insurer. In total, there were 167 different GR providers in the years 2017–2023. However, some GR providers merged, stopped or started somewhere during that period. This resulted in a varying number between 132 and 144 GR providers throughout the years, as shown in Table 4.1. Hence, not all GR providers were active in the full seven years, making a correction for the number of active years necessary.

Exactly 35% (58 providers) of all GR providers never exceeded the budget ceiling and around 56% (94 providers) exceeded the budget ceiling only in 25% or less of the active years, as shown in Table 4.7. Only 11% (18 providers) of all GR providers exceeded the budget ceiling in more than half of the active years. Note that the interval 0–25% includes GR providers that exceeded the budget ceiling once in the seven years ($\frac{1}{7} = 14\%$). Similarly, the interval 25–50% includes GR providers that exceeded the budget ceiling twice in the seven years ($\frac{2}{7} = 29\%$).

Table 4.7

Overview of the number of geriatric rehabilitation providers exceeding the budget ceiling

	Fraction of years in which budget ceiling exceeded (Y)				
	0	$0 < Y \leq 0.25$	$0.25 < Y \leq 0.5$	$0.5 < Y \leq 0.75$	$0.75 < Y \leq 1$
Number of providers	58	36	55	16	2

5 Methodology

5.1 Bunching approach

With the growing availability of administrative data, it is commonly observed that individuals or firms tend to cluster or self-select to a specific value in the range of a variable. Examples are firms reporting profits just below thresholds above which higher taxes apply or individuals targeting earnings below thresholds above which they would lose specific allowances. When this occurs, the histogram of the relevant variable (e.g. profits or earnings in the examples) shows a visible 'bunch', which is excess mass that would not be otherwise predicted by the surrounding bins. Often, a bunch is preceded or followed by missing mass, depending on the situation.

The economic literature about bunches focuses on identifying a behavioural response to non-linear incentives (Kleven, 2016). The classic bunching setup considers a policy threshold (e.g. a change in tariffs) at some continuous running variable (e.g. treatment duration), causing individuals or firms to 'manipulate' the running variable to be below or above a threshold. The bunching approach shows great similarities with regression discontinuity and regression kink designs, as explained by, for example, Imbens and Lemieux (2008) and Card et al. (2015), respectively. However, the key difference lies in whether the running variable can be manipulated or not (Kleven, 2016). If this variable is not subject to manipulation whenever discontinuous jumps in incentives are observed, then a regression discontinuity/kink design is applicable; otherwise, the bunching approach. In the context of this thesis, physicians can clearly influence the treatment duration for a GR trajectory. In that respect, the bunching approach is the most suitable method to apply.

The main goal of the bunching approach is to estimate the magnitude of the excess mass (bunches). This goal can be achieved in this thesis by comparing the empirical distribution of GR claims with the distribution of GR claims that would have been as if there was no bunching behaviour around a threshold. This latter 'counterfactual distribution' can be estimated using the following procedure. A polynomial function of order q is fit to the bin counts of the empirical distribution of GR claims, excluding observations in a range $[z_L, z_U]$ around a threshold point z^*, by estimating the following regression:

$$c_j = \sum_{l=0}^{q} \beta_l (z_j)^l + \sum_{k=z_L}^{z_U} \gamma_k \mathbb{1}[z_j = k] + \epsilon_j, \tag{1}$$

where c_j is the number of claims in bin j, z_j is the treatment duration level (either the number of inpatient days for the length of stay or the number of treatment hours for the treatment intensity)

in bin j, z_L and z_U are, respectively, the lower and upper bound of the window affected by bunching responses around the threshold (the 'bunching window'), γ_k is an intercept shifter for each of the bins in the bunching window, $\mathbb{1}[\cdot]$ is an indicator function and ϵ_j is the residual term. Hence, by including indicator functions for each bin in $[z_L, z_U]$, the polynomial part of (1) is estimated without considering the data affected by bunching behaviour around the threshold z^*.

The counterfactual claims distribution is then estimated as the predicted values $\hat{c}_j$ from (1), omitting the contribution of the indicator functions belonging to the claims falling in the bunching window $[z_L, z_U]$:

$$\hat{c}_j = \sum_{l=0}^{q} \hat{\beta}_l (z_j)^l. \tag{2}$$

The excess mass $\hat{b}$ is finally obtained by taking the difference between the observed and counterfactual bin counts in the part of the bunching window where the excess mass is located (i.e. $[z^*, z_U]$):

$$\hat{b} = \sum_{j=z^*}^{z_U} (c_j - \hat{c}_j). \tag{3}$$

Two assumptions must hold to apply the bunching approach (Bachas et al., 2021): 1) manipulation is one-sided and bounded, and 2) the counterfactual distribution is well-behaved. The first assumption implies that policy changes at thresholds incentivise people or firms to manipulate only in one direction (e.g. from below to above a threshold or the other way around) within a small window around a threshold. For example, a sudden increase in tax at a threshold will not incentivise anyone to shift earnings from just below to above the tax threshold. In addition, if people's earnings are far from the tax threshold, they are not likely to be incentivised. The second assumption implies that the counterfactual distribution close to the threshold can be estimated by fitting a distribution using only observations sufficiently far from the threshold that are not manipulated. If this second assumption holds, the polynomial function can be estimated using the observations outside the bunching window, and then extrapolated to the bunching window to provide estimates of where the observations would have been located without manipulation.

5.2 Multiple linear regression analysis

To examine the variation in the strength of behavioural distortions regarding the length of stay on a GR provider level, I perform a linear regression analysis with multiple independent variables. In the case that the two assumptions of the bunching approach plausibly hold, I define the dependent

variable, representing the strength of the behavioural distortion, as $\hat{b}$ in (3) but then calculated for each GR provider separately with the number of claims per provider added up for the active years.

When either one or none of the assumptions plausibly hold, I define the strength of the behavioural distortion of a GR provider as follows. First, I add up the number of claims for the active years for each GR provider p. Then, I take the number of claims in the R days following a threshold minus the number of claims in the R days preceding it. I do this for every threshold, located on days 15, 29, 57 and 92. Lastly, I scale the strength based on the total number of claims in the $[-R, R]$ windows for all four thresholds. Mathematically, this can be expressed as:

$$\begin{aligned} s_p = \Big(& \sum_{k=15}^{15+R-1} n_{k,p} - \sum_{k=15-R}^{14} n_{k,p} + \sum_{k=29}^{29+R-1} n_{k,p} - \sum_{k=29-R}^{28} n_{k,p} + \sum_{k=57}^{57+R-1} n_{k,p} - \sum_{k=57-R}^{56} n_{k,p} \\ & + \sum_{k=92}^{92+R-1} n_{k,p} - \sum_{k=92-R}^{91} n_{k,p} \Big) \div \Big(\sum_{k=15}^{15+R-1} n_{k,p} + \sum_{k=15-R}^{14} n_{k,p} + \sum_{k=29}^{29+R-1} n_{k,p} + \sum_{k=29-R}^{28} n_{k,p} \\ & + \sum_{k=57}^{57+R-1} n_{k,p} + \sum_{k=57-R}^{56} n_{k,p} + \sum_{k=92}^{92+R-1} n_{k,p} + \sum_{k=92-R}^{91} n_{k,p} \Big), \end{aligned} \tag{4}$$

where s_p is the strength of the behavioural distortion of GR provider p, and $n_{k,p}$ is the number of GR claims at a length of stay of k inpatient days for provider p. If there are no GR claims in any of the $[-R, R]$ windows around the thresholds for a GR provider, this provider is removed because dividing by zero is undefined. Overall, a higher value for s indicates a stronger behavioural distortion.

For the independent variables, I include data about the fraction of active years in which a GR provider at CZ exceeded the budget ceiling, which has not been done before. Hence, it is unknown yet what the sign (i.e. positive or negative) of the relationship between exceeding the budget ceiling and the behavioural distortion of a GR provider is. On the one hand, the sign could be negative along the line, for example, that it does not pay off to strong strategic behaviour when the budget ceiling is likely to be exceeded because the overclaimed amount will not be reimbursed. On the other hand, the sign could be positive along the line, for example, that it may be profitable to show strong strategic discharge behaviour to exceed the budget ceiling in order to improve the negotiating position for the next year(s). I also include data for each GR provider about the share of female patients, the average age of the patient group and the share of patient admissions on each day in the week to control for their impact, as shown in previous research (see Section 3.7). While the relevance of profit orientation and financial distress has also been proven, these are not considered as there is no variation in these factors between GR providers.

In mathematical terms, this regression setup on the level of a GR provider translates to the following equation:

$$\begin{aligned} s_p =& \beta_0 + \beta_1 BudgetCeiling_p + \beta_2 Female_p + \beta_3 Age_p + \beta_4 Monday_p + \beta_5 Tuesday_p + \\ & \beta_6 Wednesday_p + \beta_7 Thursday_p + \beta_8 Friday_p + \beta_9 Saturday_p + \epsilon_p, \end{aligned} \tag{5}$$

where for each GR provider p, $BudgetCeiling_p$ is the fraction of active years in which the budget ceiling is exceeded, $Female_p$ is the share of female patients, Age_p is the average age of the patient group, $Monday_p$ to $Saturday_p$ are the shares of patients that are admitted on each day of the week (excluding Sunday), and ϵ_p is the error term. To avoid multicollinearity issues, the share of patients admitted on the last day of the week, Sunday, is omitted from (5). In addition, no interaction terms are included as these could not be substantiated by arguments from economic theory.

To evaluate the fit of (5), I look at the R-squared, which measures the proportion of the variance in s_p explained by the independent variables. The R-squared tends to increase as more independent variables are added to the model, even if those variables do not significantly improve the model's explanatory power. Therefore, to penalise the addition of (unnecessary) variables, I also assess the adjusted R-squared which accounts for the number of independent variables. Finally, I perform an F-test to evaluate whether (5) provides a better fit than a model with no independent variables and, hence, to check the overall significance of (5).

5.2.1 Sensitivity analysis

The coefficients β_l for $l \in \{0, 1, \ldots, 9\}$ in (5) are particularly sensitive to the specification of s_p. This metric largely depends on the choice of the (symmetric) window $[-R, R]$. This window should span the entire area with excess or missing mass. A window that is too large contains days that are not affected by behavioural distortions. Likewise, a window that is too small leaves out days that are affected by behavioural distortions.

A common way to select the $[-R, R]$ window is by visual inspection. Sometimes, determining this window is obvious from the data but in many cases, it is more difficult because behavioural distortions are often diffuse. As visual inspection is thus vulnerable to the researcher's discretion, I choose the most logical value for R from my perspective, but vary this value with -1 and $+1$ to check the sensitivity of the estimated β_l's for this choice.

5.3 Qualitative assessment

The modular payment system with registration per hour, as described in Section 3.5, turned out best after a comparative analysis performed by the Dutch Healthcare Authority (NZa, 2024a). The assessment framework used contained the following five criteria with corresponding weights, which have been drawn up by the Dutch Healthcare Authority in consultation with other stakeholders in the healthcare sector:

- **Criterion 1: Easy to scale up and down (40%).** Healthcare providers must be able to provide the care that is needed for a heterogeneous patient group, no more and no less. Reimbursing flexibility and customisation of care tailored to patients' needs is, therefore, considered essential.
- **Criterion 2: Low administrative burden (20%).** Healthcare professionals should spend as much time as possible on providing care and as little as possible on secondary tasks, such as registering patient-related time.
- **Criterion 3: Right incentives (20%).** Perverse incentives for under- and overtreatment and risk selection must be prevented as much as possible. In addition, there must be an incentive to promote flow in the chain: from the hospital to the temporary stay facility and from the temporary stay facility to the home situation.
- **Criterion 4: Limited patients (10%)**. The payment system should be appropriate for groups with a limited number of patients. This criterion partly overlaps with the first criterion. Nevertheless, it is included because the involved stakeholders considered it desirable that healthcare providers are given enough resources to provide suitable care for small groups of patients.
- **Criterion 5: Simplicity (10%).** The payment system must be simple with a limited number of homogeneous components for the treatment intensity and the length of stay.

This thesis focuses on the impact of financial incentives on efficiency, and thus, relates best to the third criterion. While the Dutch Healthcare Authority, together with the other involved stakeholders, assessed this criterion of minor importance with a weight of only 20%, I motivate the importance and re-evaluate the assessment of this criterion extensively using insights from economic theory and GR claims data. Furthermore, I give recommendations on how the current GR payment system can be adjusted to meet the five criteria better.

6 Results

6.1 Applicability of bunching approach

To apply the bunching approach, the two assumptions as described in Section 5.1 must hold in practice. The plausibility of the 'manipulation within a window' assumption can usually be strongly argued from an economic perspective. This assumption likely holds in the context of this thesis because physicians will only extend and never shorten the treatment duration to cross a threshold. This is because the reimbursement scheme in GR contains discontinuous increasing (and not decreasing) tariffs. In addition, it is not plausible that an optimal treatment duration that is so far from a threshold (e.g. outside the bunching window) will be manipulated, reasonably assuming that elderly care physicians also base their decisions on medical considerations. The second 'regularity' assumption holds when the distribution of claims outside the presumed bunching window closely follows a smooth finite-degree distribution. Therefore, visual inspection of the histograms of the relevant treatment duration variables, which are the treatment intensity and length of stay for GR, is needed to assess whether this assumption holds. This is done in the following two paragraphs of this section.

6.1.1 Treatment intensity

The distribution of the treatment intensity of the GR claims at CZ belonging to the clinical trajectories with the CVA and other diagnoses for the years 2017–2023 are displayed in Figure 6.1 and Figure 6.2, respectively. The red vertical lines denote the tariff thresholds. As the location of a threshold differs depending on the specific subgroup of length of stay, there are four subfigures belonging to each diagnosis group (i.e. CVA and other). There is no subfigure belonging to a length of stay of more than 91 inpatient days because there are no tariff thresholds at all for this subgroup. The horizontal axis is further truncated at 100 treatment hours.

There is little to no bunching behaviour visible in Figure 6.1 and Figure 6.2. In other words, there are no peaks with excess claims observable immediately after the thresholds and holes with missing claims just before the thresholds. Varying the bin size, which is currently set to one hour, does not alter this conclusion. Consequently, applying the bunching approach to estimate the magnitude of the excess mass does not make any sense when no behavioural distortions are visible.

Figure 6.1

Distribution of the treatment intensity of geriatric rehabilitation claims at CZ belonging to clinical trajectories with CVA diagnosis from the years 2017–2023

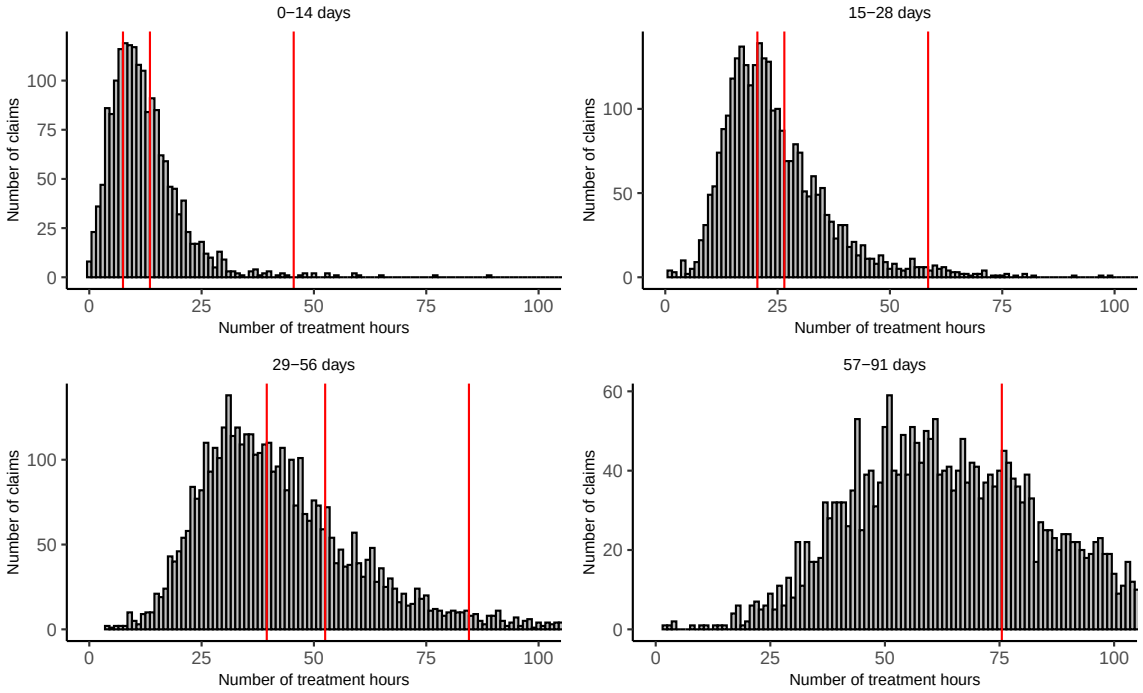
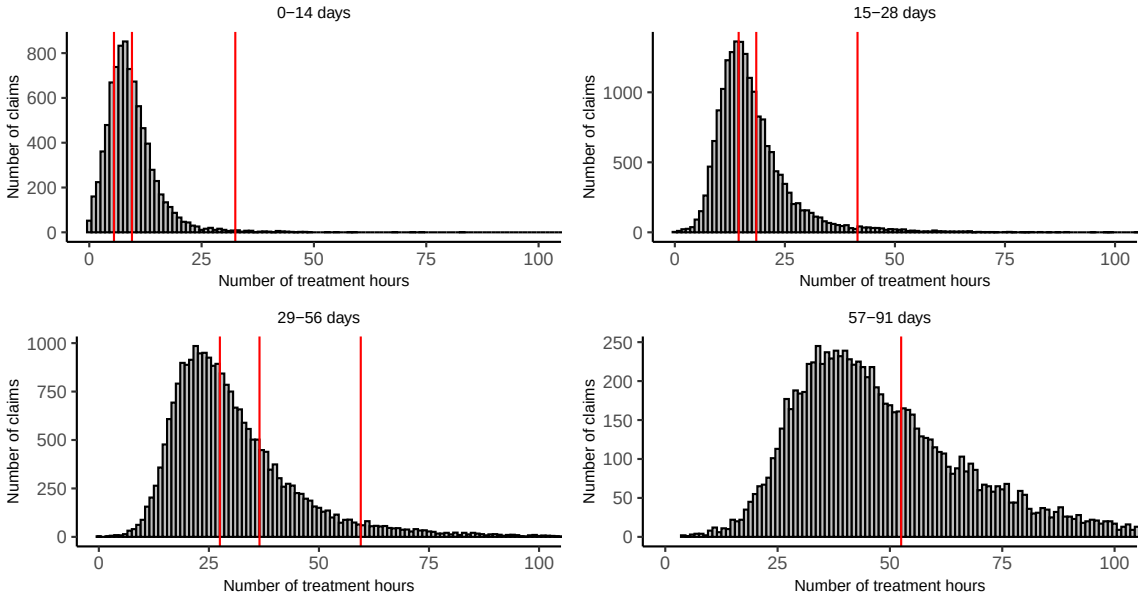


Figure 6.2

Distribution of the treatment intensity of geriatric rehabilitation claims at CZ belonging to clinical trajectories with other diagnoses from the years 2017–2023



The absence of strategic discharge behaviour regarding the treatment intensity may be surprising. Despite the (bi-)weekly multidisciplinary meetings where the planning for the treatment intensity for the coming week(s) is made, 'manipulating' the treatment hours may be more difficult in practice than it seems. Some reasons for the absence of bunches could be the following (in arbitrary order). First, the total available time of the different practitioners involved may be (too) limited to reach the desired treatment hours. Second, optimisation of the treatment intensity may be complicated due to the many practitioners involved. For example, if a meeting has six attendees who all participate for half an hour, the treatment intensity immediately increases by three hours. Third, the financial incentives may not be strong enough. As shown in Figure 3.3, the steps in total reimbursement are much higher between the subgroups of the length of stay than between the subgroups of the treatment intensity. Fourth, the thresholds are set too far from what is optimal from a rehabilitation perspective. As shown in Figure 6.1 and Figure 6.2, the peaks of the distributions often occur before the thresholds. Fifth, there might be inattention to the fact that the thresholds belonging to the various treatment intensities are located differently depending on the subgroup of lengths of stays and specific diagnosis. This is in contrast to the locations of the thresholds for the length of stay, which are always located at the same place, independently of the treatment intensity and diagnosis.

6.1.2 Length of stay

The distribution of the length of stay per GR claim at CZ belonging to the clinical trajectories for the years 2017–2023 is shown in Figure 6.3. The red vertical lines denote the tariff thresholds. As the location of a threshold does not depend on the diagnosis (CVA or other) or treatment intensity, no subfigures are shown.

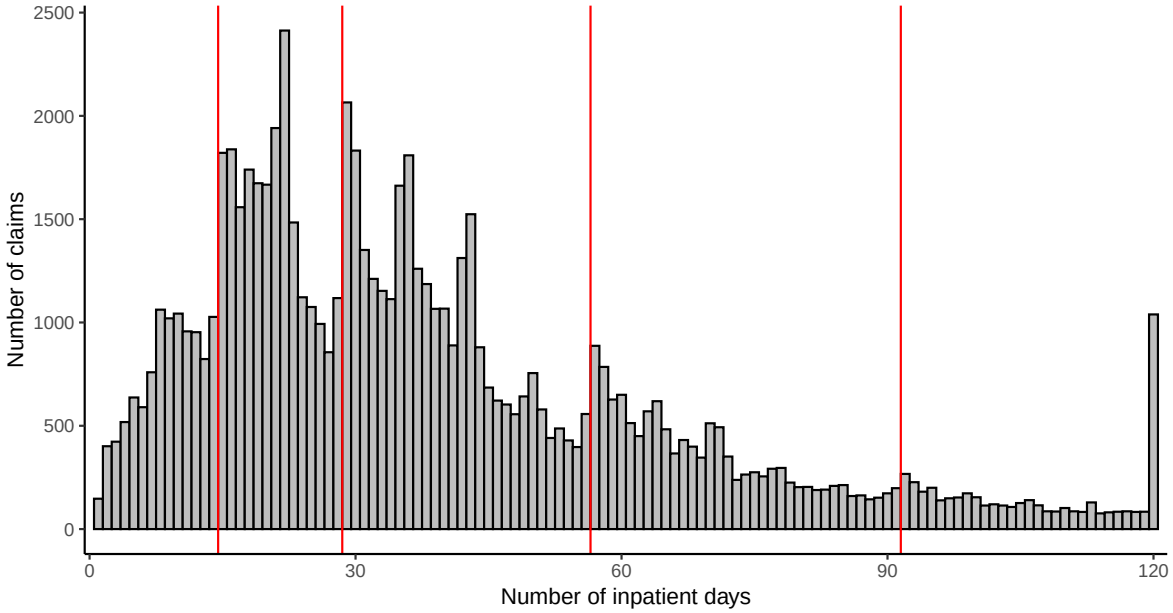
Strategic discharge behaviour around the thresholds is visible in Figure 6.3. Regarding the first two thresholds (located on days 15 and 29), the number of claims on the days just after the threshold is substantially higher than on the days just before. This pattern is also visible for the last two thresholds (located on days 57 and 92), but less strongly.

Besides the peaks in claims immediately after the thresholds, there are conspicuous peaks visible around every seven days. An explanation for these weekly peaks is that physicians may think in weeks of treatment rather than in days, especially if the treatment is longer than a couple of days. These peaks disrupt the smoothness of the counterfactual distribution because a high-order polynomial is needed to fit the data outside the bunching window. Correcting for the weekly peaks with, for example, dummy

variables will result in a smoother distribution but will also remove the peaks around the thresholds. A further complication to the smoothness of the counterfactual distribution is the peak on day 120, which corresponds to claims with the maximum lead time. On this day, the DTC must be closed and a new DTC can be opened again for which the inpatient day starts from zero again. Finally, there are relatively many claims in the third week (i.e. the week after the first threshold). A reason for this could be that many patients only need a relatively short length of stay to achieve their rehabilitation goals. All these peaks, except those at the tariff thresholds, violate that second assumption of the bunching approach that the counterfactual must be 'well-behaved'. As both assumptions are necessary conditions, the bunching approach is not applicable here to estimate the magnitude of excess mass regarding the length of stay in GR.

Figure 6.3

Distribution of the length of stay of geriatric rehabilitation claims at CZ belonging to clinical trajectories from the years 2017–2023



6.2 Multiple linear regression results

6.2.1 Different years

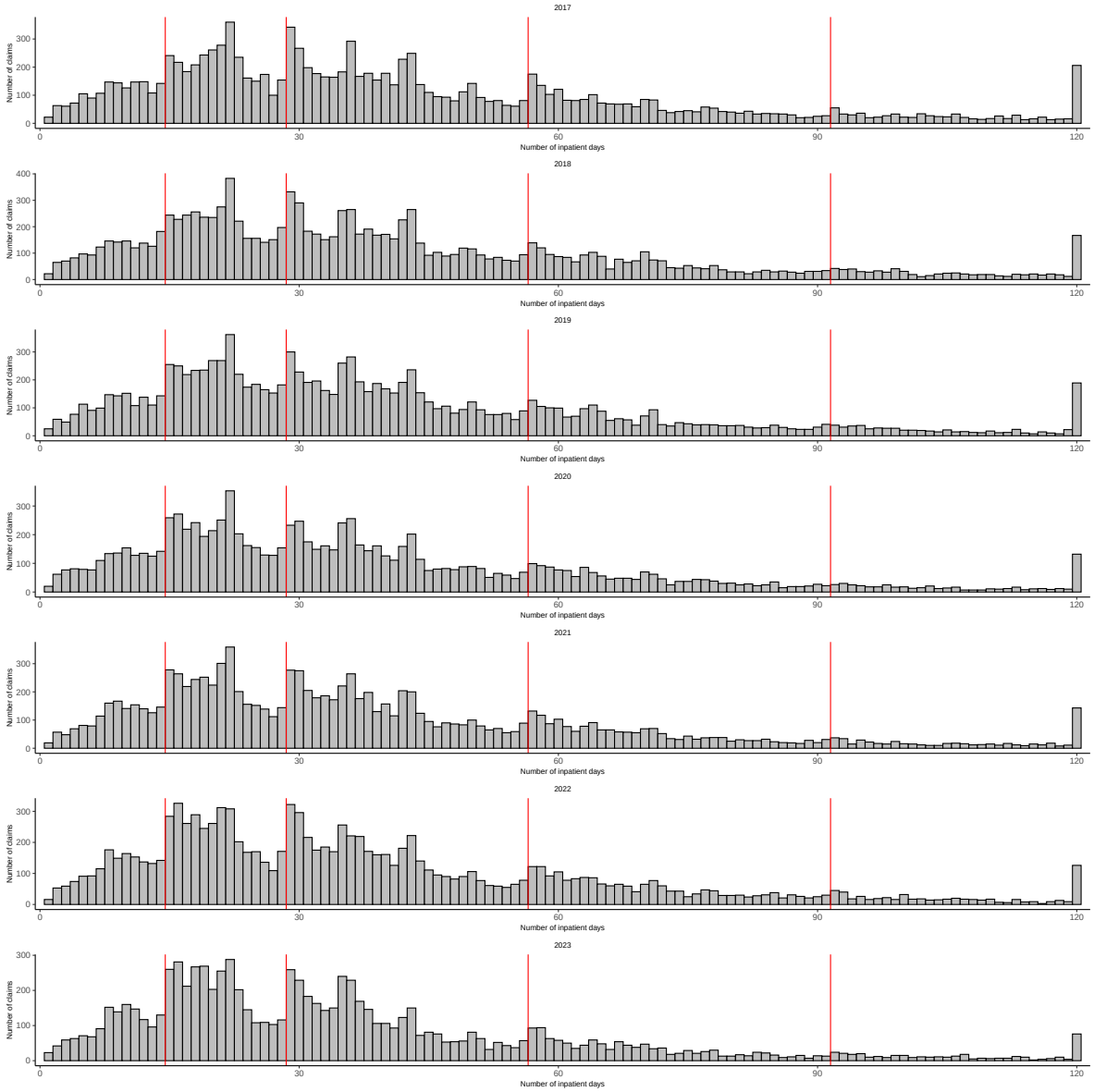
In model specification (5), the claims are aggregated over the years for each GR provider. This choice is disputable when substantial deviations exist between the distributions of claims for the different years. Therefore, to check whether the aggregation of claims over the years is reasonable, Figure 6.4

shows the distributions of the length of stay for the last seven years (2017–2023) separately.

There are no significant differences in the shape of the distributions visible over the years. Even in the COVID-19 years 2020 and 2021, in which the number of GR patients was considerably lower, the distribution of the length of stay in GR was not significantly different from the years unaffected by the COVID-19 pandemic.

Figure 6.4

Distribution of the length of stay of geriatric rehabilitation claims at CZ belonging to clinical trajectories for the different years (2017–2023)

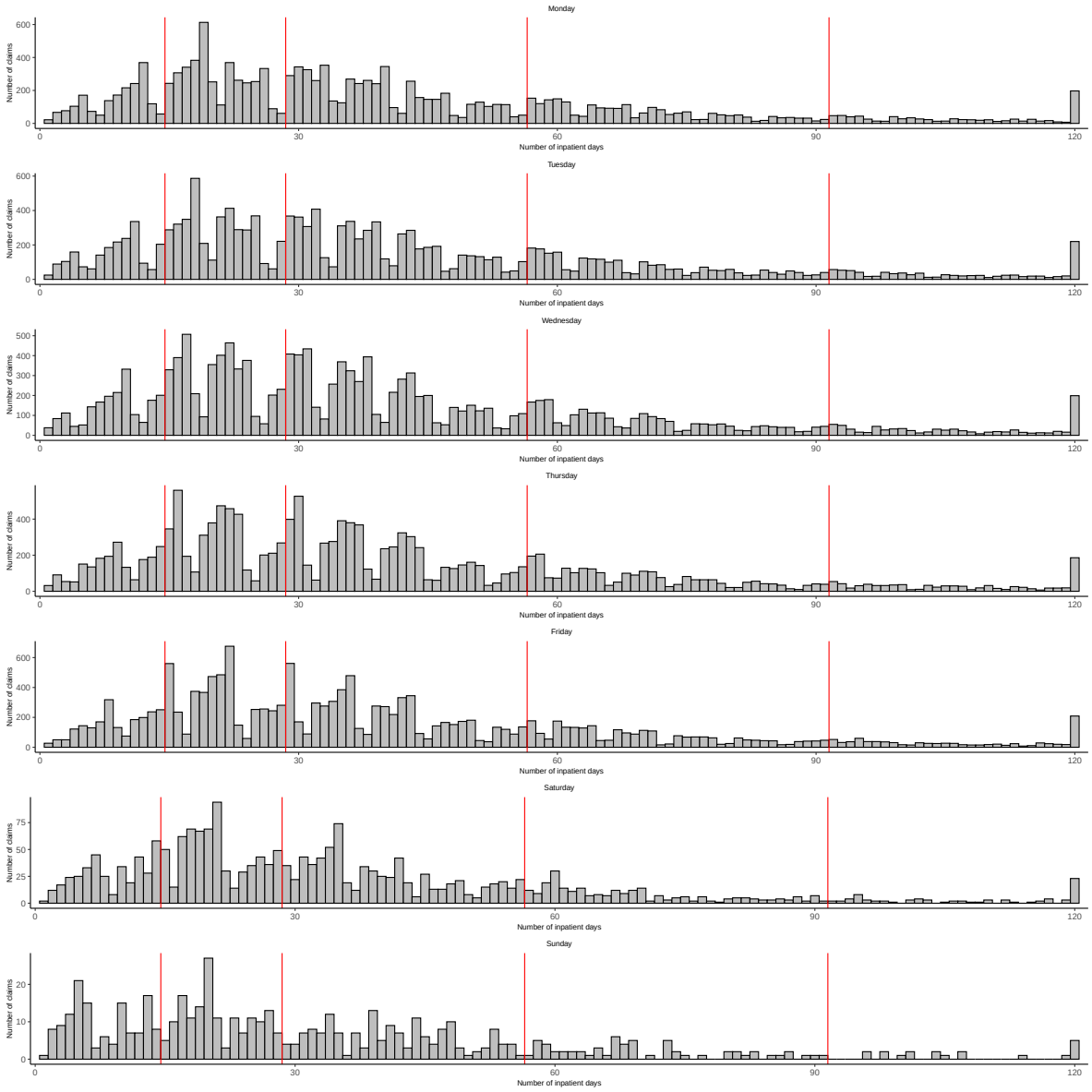


6.2.2 Admission day

The distributions of the length of stay of GR claims at CZ for each admission day in the week are shown in Figure 6.5. The first bin after each threshold represents the same day as the admission day. The weekly peaks can be observed by the relatively high peaks every seven days. Furthermore, the weekend effect can be recognised by the high peaks followed by the drop in claims in the two days immediately after.

Figure 6.5

Distribution of the length of stay of geriatric rehabilitation claims at CZ for each admission day separately of clinical trajectories in the years 2017–2023



6.2.3 Estimation results

On the level of a GR provider, the strength of the behavioural distortion is regressed on the fraction of years in which the budget ceiling is exceeded, the share of female patients, the mean age of the patient group and the share of patient admission on the different days in the week, as described in Section 5.2. The corresponding least squares estimates of the coefficients are shown in Table 6.1 for $R = 2$. This choice for R seems most logical when looking at Figure 6.3 because the peaks after the thresholds mainly consist of two bins corresponding to two days. The sample consists of all 167 GR providers with, on average, 89 claims to determine the strength of the behavioural distortion for $R = 2$.

Table 6.1

Least squares estimates for the strength of behavioural response with $R = 2$

Variable	$\hat{\beta}$	std. error	t-value	p-value
Intercept	-2.081	2.029	-1.025	0.307
Budget ceiling	0.240	0.105	2.286	0.024*
Female	0.117	0.364	0.321	0.749
Age	-0.016	0.014	-1.163	0.247
Monday admission	2.876	1.847	1.557	0.121
Tuesday admission	2.677	1.814	1.476	0.142
Wednesday admission	4.906	1.917	2.561	0.011*
Thursday admission	3.513	1.749	2.008	0.046*
Friday admission	3.815	1.886	2.023	0.045*
Saturday admission	2.740	2.354	1.164	0.246
R-squared		0.103		
Adjusted R-squared		0.050		
F-statistic		1.961 (0.047*)		

Note. In total, 167 geriatric rehabilitation providers are included.
*$p<0.05$.

The estimated coefficient for the budget ceiling variable is 0.240, which is statistically significant at a 5% level. The positive sign indicates that a stronger behavioural distortion of a GR provider is associated with a larger fraction of years in which the budget ceiling is exceeded. This finding could suggest that GR providers who are likely to exceed the budget ceiling may strengthen their strategic discharge behaviour. However, it could also be that GR providers who tactically exploit the financial incentives are more likely to exceed the budget ceiling. In the sensitivity analysis, the estimated coefficient remains statistically significant at a 5% level for $R = 3$ but not for $R = 1$, as shown in Table D.2 and Table D.1 in Appendix D, respectively. A possible reason for the insignificant estimate for $R = 1$ is that only, on average, 48 GR claims are left for each GR provider to determine the strength of behavioural distortion, which may be too few to result in a reliable metric.

The estimated coefficients for the share of patient admissions on Wednesday, Thursday and Friday are 4.906, 3.513 and 3.815, respectively. These are all statistically significant at a 5% level. Since the thresholds are located at whole weeks, it implies for $R = 2$ that if patients are admitted on, for example, Thursday, they have a relatively high chance of getting discharged on Thursday or Friday instead of Tuesday or Wednesday. A similar reasoning applies if Wednesday or Friday is the admission day. Hence, these estimates show that the weekend effect matters. For $R = 1$ and $R = 3$, the estimated coefficients are no longer statistically significant at a 5% level. An explanation for the insignificant estimates for $R = 1$ could be the low number of claims to determine the strength of behavioural distortion, while the reason for $R = 3$ could be that there is always overlap with the weekend when three days before and after a threshold are considered, which could cancel out the weekend effect.

The R-squared and adjusted R-squared, as shown in Table 6.1, are equal to relatively low values of 0.103 and 0.050, respectively. So, the independent variables do not explain much of the variation in the dependent variable. However, the overall model is still statistically significant at a 5% level with an F-statistic of 1.961. To summarise, the model provides useful information, but there might be an issue with missing variables or the sample size to estimate the underlying relationship accurately. After all, only four types of variables are included in (5) and 167 GR providers are considered, each having only 89 claims on average in the $[-R, R]$ windows around the thresholds.

6.3 Insights from the qualitative assessment

6.3.1 Re-evaluation of the incentives criterion

Table 6.2 shows the scores, which can take a value ranging from $--$ to $++$, that the Dutch Healthcare Authority gave to the modular payment system with registration per hour (NZa, 2024a). The highest possible score ($++$) is obtained for three of the five criteria. The second criterion received a relatively low score ($-$) because the direct patient-related time must be registered per five minutes. The third criterion got a medium score ($+/-$) with the explanation that a modular payment system contains a production incentive for overtreatment but no incentive for undertreatment or risk selection. Yet, the Dutch Healthcare Authority questions how strong the financial incentive for overtreatment will be given the labour shortages and the increasing demand for elderly care due to the ageing population.

In my opinion, the third criterion is not assessed correctly. While the Dutch Healthcare Authority assumes that large demand for care and pressing labour shortages will limit the production incentive for overtreatment, there are currently no indications that this scenario will become reality soon. Contrarily,

Table 6.2

Score overview of modular payment system

	Score
Criterion 1: Easy to scale up and down	++
Criterion 2: Low administrative burden	−
Criterion 3: Right incentives	+/−
Criterion 4: Limited patients	++
Criterion 5: Simplicity	++

Note. Source: Dutch Healthcare Authority (NZa, 2024a).

there is — from a demand perspective — no waiting list at all for GR, and there is — from a supply perspective — enough personnel available to provide the multidisciplinary treatment. There might be a lack of nursing staff, but as Figure 6.3 suggests that GR providers are currently gaming the payment system, this possible shortage seems not urgent yet because there would otherwise be no room to let patients stay longer than necessary. Furthermore, when the ongoing experiments of an immediate ambulant GR trajectory (so, without a prior inpatient stay) turn out to be successful and get structurally reimbursed, fewer patients are expected to stay in GR facilities in the coming years.

In the scenario of no waiting lists and enough personnel available, there is no reason to assume that the production incentive for overtreatment in the new modular payment system will be modest. Figure 6.3 indicates that healthcare providers are currently gaming the payment system with respect to the length of stay in GR. The new modular payment system will make it even more rewarding to game the system because it is profitable to prolong treatment endlessly with constant marginal revenues that are higher than constant marginal costs. In other words, every treatment hour or inpatient day yields the same amount of profit, regardless of how long the treatment has been going on. Moreover, there is no gain for GR providers to admit new GR patients at the expense of already admitted GR patients because an hour more or a day longer at the beginning of the treatment process yields as much as an hour more or a day longer at the end of a treatment process. This has, contrary to what the Dutch Healthcare Authority aims for, a negative impact on the flow in the chain.

In the long run, the ageing population and growing personnel shortages (especially in nursing staff) will inevitably have an enormous impact on GR. The question is, however, when this will happen. As long as there are no concrete indications that this will happen in the coming years, it seems too early to implement the new modular payment system because it is plausible that there will be more gaming then. Therefore, it would be better to wait, or to introduce a modular payment system with decreasing tariffs to limit the incentive for overtreatment.

6.3.2 Possible adjustments to the current payment system

The scores for the current payment system in GR based on DTCs, as given by the Dutch Healthcare Authority, are shown in Table 6.3. Compared to the scores of the modular payment system as shown in Table 6.2, the current payment system in GR scores worse for every criterion, except for an equal score for the third criterion.

Table 6.3

Score overview of payment system based on DTCs

	Score
Criterion 1: Easy to scale up and down	$+/-$
Criterion 2: Low administrative burden	$--$
Criterion 3: Right incentives	$+/-$
Criterion 4: Limited patients	$+$
Criterion 5: Simplicity	$+ \sim +/-$

Note. Source: Dutch Healthcare Authority (NZa, 2024a).

Adjusting the DTC payment system in GR was not part of the multi-criteria analysis performed by the Dutch Healthcare Authority. Only the DTC system in its current form was evaluated (NZa, 2024a). However, with certain improvements to better meet the criteria, a change of payment system may not be needed at all. A big advantage of keeping the current system is that GR providers do not have to change their ICT systems, which would save a lot of money. With this in mind, I propose several improvements to the current GR payment system for all criteria, except for the fourth one, because this criterion largely overlaps with the first.

The first criterion is inherently disadvantageous for the DTC payment system because DTCs are, per definition, bundles of care. Scaling up and down the care does not automatically result in a different bundle. This issue can be alleviated by setting more thresholds. However, this would come at the cost of increasing the incentives for gaming. That is, the more thresholds you set, the more you move to the modular payment system in which it is beneficial to extend the treatment duration endlessly. Another solution would be to move the current thresholds to different locations. As shown in Figure 6.1 and Figure 6.2, the thresholds for the treatment intensity are often located after the peaks in claims. Besides, the locations of the thresholds for the length of stay originate from the old frequency distribution before the transition to the Health Insurance Act, but the mean length of stay decreased by seven days since then (Bouwstra et al, 2017). Therefore, moving the thresholds to the left would result in better alignment with the current distribution of claims.

Regarding the second criterion, the DTC payment system received the lowest possible score ($--$) from the Dutch Healthcare Authority because direct and indirect patient-related time must be registered. The solution is simple, though. Like the modular payment system, the DTC payment system can only reimburse the direct patient-related time and discount the indirect patient-related time in the reimbursement for the length of stay.

The third criterion got an average score ($+/-$) because the Dutch Healthcare Authority argues that, although there is only a limited production incentive for overtreatment, there is an incentive for strategic discharge behaviour and no incentive for rapid flow in the chain. In contrast to what the Dutch Healthcare Authority states, there is an incentive for rapid flow as the average marginal revenue decreases for the current DTC payment system, while this stays constant in the modular payment system. Hence, it is not advantageous in the current DTC payment system to continue treatment indefinitely, while it is in the modular payment system. Moreover, the strategic discharge behaviour can possibly be reduced by not setting the thresholds for the length of stay exactly at whole weeks. As shown in Figure 6.3, it seems that physicians think in weeks of treatment rather than in days. To break this pattern, thresholds can be set at, for example, every ten days.

The fourth criterion is particularly relevant for the main diagnosis group 'amputations' in GR, which is the smallest group with a treatment share of only 2.5% (see Table C.1 in Appendix C). The distribution of the length of stay for this group deviates from the other main diagnosis groups as there are no significant behavioural distortions visible (see Figure E.1 in Appendix E). However, this does not imply that the current reimbursement scheme is inappropriate. The thresholds are set around equal distances in the entire period of 1–120 days. Hence, no adjustments are needed from my perspective. Potentially, thresholds can be moved or additional thresholds can be set, similar to the argumentation belonging to the first criterion.

The fifth criterion got a score between $+$ and $+/-$. The Dutch Healthcare Authority argues that claims in the DTC payment system generally do not correspond to the exact provided treatment because bundling of care takes place within DTCs. As a result, there is heterogeneity within the claims as the provided treatment varies per subgroup for the treatment intensity and the length of stay. This issue is hard to solve even when more thresholds are implemented because DTCs remain bundles of care. Introducing additional DTCs results in less heterogeneity but makes the system more complex. However, another solution to simplify the DTC payment system would be to reduce the number of components by half by no longer distinguishing between CVA and other diagnoses.

7 Conclusion and discussion

Stepwise tariffs in the reimbursement scheme of GR may incentivise physicians to game the payment system. Analysis of the GR claims at CZ from the years 2017–2023 indicates that gaming behaviour is present regarding the length of stay but not regarding the treatment intensity. Furthermore, regression results on a GR provider level show that stronger behavioural distortions are correlated with larger fractions of years in which the budget ceiling is exceeded. Also, a larger share of patient admissions on Wednesday, Thursday and Friday are associated with stronger behavioural responses. Finally, a qualitative assessment of the proposed new payment system, which abolishes the stepwise tariffs and introduces constant tariffs, points out that the efficiency of healthcare provision in GR will likely deteriorate as the financial incentives for overproduction increase.

Finding strategic discharge behaviour with respect to the length of stay in GR is in line with previous related research. Only Pott et al. (2021) found no evidence for strategic discharge behaviour, but they examined healthcare providers' responses to discontinuous decreasing tariffs, while all other related studies investigated discontinuous increasing tariffs. As discharging patients earlier as opposed to later does not likely affect patients' benefits in the same way (i.e. an early discharge is more harmful to patients than being discharged later than is optimal), this could explain why physicians' responses differ. In contrast, the absence of strategic discharge behaviour regarding the treatment intensity is somewhat surprising based on what one would expect from economic theory and the existing literature. This thesis mentions five possible explanations for the absence but does not investigate these further. In other words, this thesis shows *how* similar financial incentives affect physicians' behaviour regarding the length of stay and the treatment intensity differently, but it remains up to future research to reveal *why* this is the case. I recommend conducting interviews with healthcare professionals and the management board of healthcare institutions to obtain deeper insights into this issue.

This thesis is the first to provide evidence of the positive correlation between exceeding the budget ceiling and the strength of behavioural distortion on a GR provider level. This information may help policymakers and health insurers to better understand how the budget ceiling relates to the strategic discharge behaviour of GR providers. However, the obtained insights do not provide any insights on cause and effect. Hence, it remains unclear whether a stronger response to discontinuous tariffs leads to a higher chance of exceeding the budget ceiling or vice versa. Further research could perform experiments to estimate the underlying causal effects.

The estimated coefficients in the multiple linear regression analysis were highly sensitive to the choice of the window with presumed excess and missing mass. An explanation for this sensitivity could be that, on average, only 48, 89 and 126 GR claims were available per GR provider to determine the strength of behavioural distortion when looking at one, two or three days before and after each threshold, respectively. These relatively low numbers of included GR claims may also cause the low value for the (adjusted) R-squared and the high standard errors. Future research should try to collect more claims data to determine the strength of responses more reliably. Including more years is, in that regard, not an option because health insurers are only allowed to store data for seven years. However, data from multiple health insurers could be combined to obtain a dataset with more GR claims.

As many healthcare systems around the world move away from fee-for-service payment systems to bundled care payment systems, much of the literature is focused on this direction of change. The proposed transition for the payment system in GR encompasses exactly the other way around. As a result, no relevant literature was identified, and it was only possible to assess the proposed changes using insights from economic theory and existing GR claims data. However, it would be interesting to compare the outcomes of the two payment systems quantitatively. Therefore, I would advise the Dutch Healthcare Authority to closely monitor the performance of GR providers in the coming years. Starting in 2025, the transition period of five years begins, in which an increasing number of GR providers will adopt the modular payment system. The two payment systems will exist next to each other in these five years, making a good comparison possible. The monitoring should preferably continue after the national implementation of the modular payment system as it is not unlikely that GR providers try to obtain a similar production output in the transition period and start taking advantage of the changed incentives after the abolition of the current payment system in GR.

What the exact future for GR will look like is unknown, but the Dutch Healthcare Authority would be wise to think about different scenarios when the modular payment system does not turn out as expected. In that regard, I recommend two options: 1) introduce decreasing (instead of constant) tariffs for the modules to limit the financial incentive for overtreatment, and 2) maintain the current payment system in GR but adapt it to reduce the existing behavioural distortions regarding the length of stay. For the second option, it is essential that the thresholds are moved from whole weeks since the GR claims data suggest that physicians think in weeks of treatment rather than in days. In any case, both options will contribute positively to the efficiency of care provision.

References

ABF Research. (2023, April 20). *Arbeidsmarktprognoses zorg en welzijn vernieuwd [Updated labour market forecasts for healthcare and welfare]*. https://abfresearch.nl/publicaties/arbeidsmarktprognoses-zorg-en-welzijn-vernieuwd/

ActiZ. (2019). *Geriatrische revalidatiezorg infographic [Geriatric rehabilitation infographic]*. https://www.studiogrz.nl/wp-content/uploads/2019/03/ActiZ_GRZ-Infographic.pdf

ActiZ. (2022, June 27). *De seinen op groen voor revalidatie in de thuissituatie [Green light for rehabilitation at home]*. https://www.actiz.nl/de-seinen-op-groen-voor-revalidatie-de-thuissituatie

Bachas, P., Kondylis, F., & Loeser, J. (2021, March 8). *Ready, set, bunch!* https://blogs.worldbank.org/impactevaluations/ready-set-bunch

Bouwstra, H., Wattel, L. M., de Groot, A. J., Smalbrugge, M., & Hertogh, C. M. (2017). The influence of activity-based funding on treatment intensity and length of stay of geriatric rehabilitation patients. *Journal of the American Medical Directors Association*, *18*(6), 549.e15–549.e22. https://doi.org/10.1016/j.jamda.2017.02.003

Card, D., Lee, D. S., Pei, Z., & Weber, A. (2015). Inference on causal effects in a generalized regression kink design. *Econometrica*, *83*(6), 2453–2483. https://doi.org/10.3982/ECTA11224

Centraal Bureau voor de Statistiek. (2023, December 15). *Prognose: Bijna 18 miljoen inwoners, 19 miljoen in 2037 verwacht [Forecast: Nearly 18 million residents, 19 million expected in 2037]*. https://www.cbs.nl/nl-nl/nieuws/2023/50/prognose-bijna-18-miljoen-inwoners-19-miljoen-in-2037-verwacht

Centraal Bureau voor de Statistiek StatLine. (2023a, December 15). *Prognose bevolking; Kerncijfers, 2023–2070 [Population forecast; Key figures, 2023–2070]*. https://opendata.cbs.nl/#/CBS/nl/dataset/85742NED/table?dl=9CBFC

Centraal Bureau voor de Statistiek StatLine. (2023b, June 14). *Ziekenhuisopnamen en -patiënten; Diagnose-indeling VTV [Hospital admissions and patients; Diagnostic classification VTV]*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84067NED/table?dl=A181F

Centraal Bureau voor de Statistiek StatLine. (2023c, July 7). *Zorguitgaven; kerncijfers [Healthcare expenditure; Key figures]*. https://opendata.cbs.nl/#/CBS/nl/dataset/84047NED/table?dl=A2787

CZ. (n.d.). *Privacy statement*. https://www.cz.nl/privacy#Bewaard

CZ. (2023). *ELV, GRZ en GZSP: Zorginkoopbeleid 2024 [ELV, GR and GZSP: Healthcare purchasing policy 2024]*. https://zorginkoopbeleid.cz.nl/grz-elv-gzsp/download-inkoopdocument

Douven, R., Remmerswaal, M., & Mosca, I. (2015). Unintended effects of reimbursement schedules in mental health care. *Journal of Health Economics*, *42*, 139–150. https://doi.org/10.1016/j.jhealeco.2015.03.008

Einav, L., Finkelstein, A., & Mahoney, N. (2018). Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica*, *86*(6), 2161–2219. https://doi.org/10.3982/ECTA15022

Eliason, P. J., Grieco, P. L., McDevitt, R. C., & Roberts, J. W. (2018). Strategic patient discharge: The case of long-term care hospitals. *American Economic Review*, *108*(11), 3232–3265. https://doi.org/10.1257/aer.20170092

Ellis, R. P., & McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of Health Economics*, *5*(2), 129–151. https://doi.org/10.1016/0167-6296(86)90002-0

Gaspar, K., Croes, R., Mikkers, M., & Koolman, X. (2023). Length of hospital stays and financial incentives: Evidence from Dutch rehabilitation centers. *The European Journal of Health Economics*, 1–11. https://doi.org/10.1007/s10198-023-01615-5

Gaspar, K., & Koolman, X. (2022). Provider responses to discontinuous tariffs: Evidence from Dutch rehabilitation care. *International Journal of Health Economics and Management*, *22*(3), 333–354. https://doi.org/10.1007/s10754-021-09322-5

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, *8*, 435–464. https://doi.org/10.1146/annurev-economics-080315-015234

Nederlandse Zorgautoriteit. (2012). *Productstructuur geriatrische revalidatiezorg 2013 [Product structure geriatric rehabilitation 2013]*. https://puc.overheid.nl/nza/doc/PUC_2938_22/1/

Nederlandse Zorgautoriteit. (2024a). *Advies nieuwe bekostiging tijdelijk verblijf [Advice on new financing for temporary stay]*. https://puc.overheid.nl/nza/doc/PUC_758153_22/

Nederlandse Zorgautoriteit. (2024b). *DBC-zorgproducten [DTC care products]*. https://www.opendisdata.nl/msz/zorgproduct

Nederlandse Zorgautoriteit. (2024c). *Kerncijfers en analyses bij advies nieuwe bekostiging tijdelijk verblijf [Key figures and analyses for advice on new financing for temporary stay]*. https://puc.overheid.nl/nza/doc/PUC_758153_22/1/#content_bijlagen

Parliamentary Documents II. (n.d.). *Toekomst AWBZ [Future AWBZ]*. https://zoek.officielebekendmakingen.nl/kst-30597-184.html

Pletscher, M. (2016). Marginal revenue and length of stay in inpatient psychiatry. *The European Journal of Health Economics*, *17*, 897–910. https://doi.org/10.1007/s10198-015-0735-4

Pott, C., Stargardt, T., Schneider, U., & Frey, S. (2021). Do discontinuities in marginal reimbursement affect inpatient psychiatric care in Germany? *The European Journal of Health Economics*, *22*, 101–114. https://doi.org/10.1007/s10198-020-01241-5

Rijksoverheid. (2023). *Niets doen is geen optie. IBO Ouderenzorg [Doing nothing is not an option. IBO Elderly care]*. https://open.overheid.nl/documenten/b0d11071-3141-4b48-8c81-4e5b372427b3/file

SiRM. (2023). *Geriatrische revalidatiezorg in zwaar weer: Analyse financiële situatie grz 2019–2022 [Geriatric rehabilitation in dire straits: Analysis financial situation gr 2019–2022]*. https://www.actiz.nl/sites/default/files/2023-10/GRZ_in_zwaar_weer-rapport_oktober_2023_1.pdf

Studio GRZ. (2014). *Handleiding registratie geriatrische revalidatie [Registration manual geriatric rehabilitation]*. https://www.studiogrz.nl/wp-content/uploads/2021/06/Handleiding-Geriatrische-Revalidatie-v20141113823.pdf

Vektis. (2023). *Verzekerden in beeld 2023: Inzicht in het overstapseizoen [Insured persons in the picture 2023: Insight into the transition season]*. https://www.vektis.nl/uploads/Publicaties/Zorgthermometer/Zorgthermometer%20Verzekerden%20in%20Beeld%202023.pdf

Vektis. (2021, July 23). *Factsheet geriatrische revalidatiezorg [Fact sheet geriatric rehabilitation]*. https://www.vektis.nl/intelligence/publicaties/factsheet-geriatrische-revalidatiezorg

Verenso. (2020). *Toegang geriatrische revalidatiezorg zonder voorafgaande opname in het ziekenhuis [Access geriatric rehabilitation without prior admission to the hospital]*. https://www.verenso.nl/_asset/_public/Thema-en-projecten/GRZ/201203_Leidraad_GRZ-toegang-versie-1-1_DEF.pdf

Zorginstituut Nederland. (n.d.). *Geriatrische revalidatiezorg (Zvw) [Geriatric rehabilitation (Health Insurance Act)]*. https://www.zorginstituutnederland.nl/Verzekerde+zorg/geriatrische-revalidatiezorg-zvw

Zorginstituut Nederland. (2020). *Analyses zorgproducten geriatrische revalidatiezorg [Analyses of care products for geriatric rehabilitation]*. https://www.zorgcijfersdatabank.nl/binaries/content/assets/zorgcijfersdatabank/publicaties-en-factsheets/memo-grz-analyses-zin.pdf

Zorginstituut Nederland. (2023, October 25). *Totale zorgkosten Zorgverzekeringswet geriatrische revalidatiezorg [Total healthcare costs Health Insurance Act geriatric rehabilitation]*. https://www.zorgcijfersdatabank.nl/databank?infotype=zvw&label=00-totaal&geg_zvw=jjverdiepnew23&geg_wlz=jjaarNEW&meta_tabel=kosten&tabel=B_kost&item=211

Zorgverzekeraars Nederland. (n.d.). *Zorgverzekeraars [Health insurers]*. https://www.zn.nl/zorgverzekeraars/

Appendix A List of all diagnosis-treatment combinations in geriatric rehabilitation

Table A.1

Overview of all diagnosis treatment combinations in geriatric rehabilitation

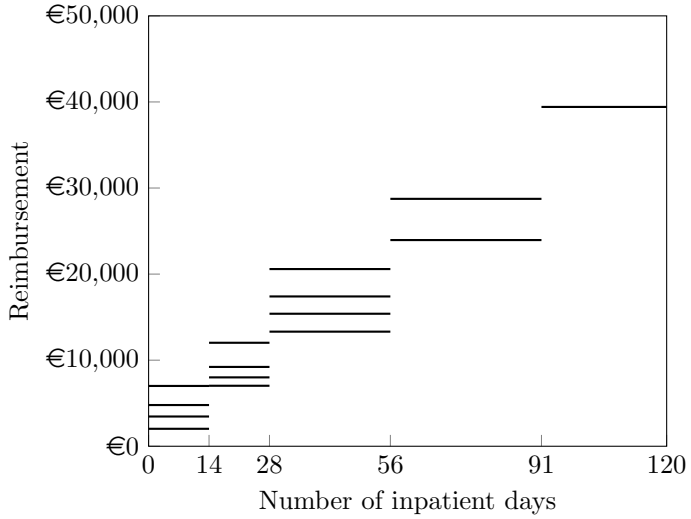
Code	Maximum price[a]	Trajectory	Diagnosis	Length of stay[b]	Treatment intensity[c]
998418022	14,280	Clinical	CVA	29-56	0-39
998418026	13,314	Clinical	Other	29-56	0-27
998418029	7,384	Clinical	CVA	15-28	0-20
998418032	7,027	Clinical	Other	15-28	0-14
998418034	2,173	Clinical	CVA	1-14	0-7
998418036	2,028	Clinical	Other	1-14	0-5
998418043	25,671	Clinical	CVA	57-91	0-75
998418045	24,084	Clinical	CVA	29-56	>84
998418046	19,405	Clinical	CVA	29-56	52-84
998418047	16,892	Clinical	CVA	29-56	39-52
998418048	13,309	Clinical	CVA	15-28	>58
998418049	10,328	Clinical	CVA	15-28	26-58
998418050	8,706	Clinical	CVA	15-28	20-26
998418051	9,544	Clinical	CVA	1-14	>45
998418052	5,303	Clinical	CVA	1-14	13-45
998418053	3,849	Clinical	CVA	1-14	7-13
998418060	23,954	Clinical	Other	57-91	0-52
998418062	20,585	Clinical	Other	29-56	>59
998418063	17,403	Clinical	Other	29-56	36-59
998418064	15,394	Clinical	Other	29-56	27-36
998418065	12,022	Clinical	Other	15-28	>41
998418066	9,215	Clinical	Other	15-28	18-41
998418067	8,002	Clinical	Other	15-28	14-18
998418068	7,005	Clinical	Other	1-14	>32
998418069	4,787	Clinical	Other	1-14	9-32
998418070	3,455	Clinical	Other	1-14	5-9
998418071	42,522	Clinical	CVA	>91	-
998418072	32,185	Clinical	CVA	57-91	>75
998418073	39,420	Clinical	Other	>91	-
998418074	28,742	Clinical	Other	57-91	>52
998418075	7,506	Ambulatory	CVA	-	>32
998418076	2,443	Ambulatory	CVA	-	5-32
998418077	375	Ambulatory	CVA	-	0-5
998148079	6,065	Ambulatory	Other	-	>23
998418080	1,447	Ambulatory	Other	-	4-23
998418081	336	Ambulatory	Other	-	0-4

Note. Source: Dutch Healthcare Authority (NZa, 2024b).
[a]In prices 2023 (euro).
[b]In number of days.
[c]In number of treatment hours.

Appendix B Reimbursement scheme of clinical other trajectories in geriatric rehabilitation

Figure B.1

Reimbursement scheme of clinical trajectories belonging to other diagnoses in geriatric rehabilitation in prices 2023



Note. The horizontal lines represent different treatment intensities.

Appendix C Geriatric rehabilitation claims at CZ broken down by main diagnosis group

Table C.1

Overview of main diagnosis groups of geriatric rehabilitation claims belonging to clinical trajectories at CZ

Main diagnosis group	Year							
	2017	2018	2019	2020	2021	2022	2023[a]	2017–2023[a] (%)
CVA	2004	1989	1974	1828	1869	1822	1434	12,920 (16.8)
Elective	1678	1602	1451	1000	1053	1200	967	8951 (11.6)
Trauma	3448	3454	3483	3127	3359	3583	2707	23,161 (30.1)
Amputations	303	259	304	295	226	316	228	1931 (2.5)
Other	4334	4649	4374	3956	4460	4536	3630	29,939 (38.9)
Total	11,767	11,953	11,586	10,206	10,967	11,457	8966	76,902(100.0)

[a]Not complete yet.

Appendix D Multiple linear regressions results for $R = 1$ and $R = 3$

Table D.1

Least squares estimates for the strength of behavioural response with $R = 1$

Variable	$\hat{\beta}$	std. error	t-value	p-value
Intercept	0.817	2.628	0.311	0.756
Budget ceiling	0.215	0.136	1.580	0.116
Female	0.291	0.471	0.616	0.539
Age	-0.027	0.018	-1.532	0.128
Monday admission	1.142	2.391	0.478	0.634
Tuesday admission	1.939	2.349	0.826	0.410
Wednesday admission	1.043	2.480	0.421	0.675
Thursday admission	1.211	2.265	0.535	0.594
Friday admission	1.774	2.442	0.726	0.469
Saturday admission	-0.279	3.048	-0.092	0.927
R-squared		0.067		
Adjusted R-squared		0.013		
F-statistic		1.237 (0.276)		

Note. In total, 167 geriatric rehabilitation providers are included.

Table D.2

Least squares estimates for the strength of behavioural response with $R = 3$

Variable	$\hat{\beta}$	std. error	t-value	p-value
Intercept	-0.884	1.881	-0.470	0.639
Budget ceiling	0.235	0.097	2.421	0.017*
Female	0.083	0.337	0.247	0.805
Age	-0.012	0.013	-0.915	0.362
Monday admission	1.417	1.711	0.828	0.409
Tuesday admission	1.279	1.681	0.761	0.448
Wednesday admission	3.038	1.775	1.711	0.089
Thursday admission	2.163	1.621	1.334	0.184
Friday admission	2.114	1.748	1.210	0.228
Saturday admission	1.171	2.181	0.537	0.592
R-squared		0.081		
Adjusted R-squared		0.027		
F-statistic		1.508 (0.150)		

Note. In total, 167 geriatric rehabilitation providers are included.
*$p < 0.05$.

Appendix E Distributions of the length of stay for the main diagnosis groups

Figure E.1

Distribution of the length of stay of geriatric rehabilitation claims at CZ belonging to clinical trajectories from the years 2017–2023 for the different main diagnosis groups