



Improving study success and diversity in Dutch higher education using performance agreements

Ben Jongbloed¹  · Frans Kaiser¹ · Don F. Westerheijden¹

Received: 30 April 2019 / Accepted: 3 December 2019 / Published online: 24 December 2019
© The Author(s) 2019

Abstract

More and more governments have started to introduce elements of performance in the funding mechanisms for their higher education institutions. An example is a performance agreement: a contract signed between the funding authority and an individual higher education provider. In the Netherlands, a policy experiment involving performance agreements was concluded in 2016. We analyse whether the agreements actually have helped achieve the goals of improving student completion rates, educational quality and increasing the diversity in educational offerings. We present some indicators relating to these goals and discuss what can be learned from the performance agreements experiment in the Netherlands.

Keywords Tertiary education · Public funding · Higher education policy · Performance based funding · New public management · Diversity

Introduction

The models for funding public higher education institutions (hereafter: universities) vary enormously across countries. In line with the new public management-inspired idea (Ferlie et al. 1996) that funds should flow to institutions where performance is manifest, many countries have implemented performance-based funding of some sort. In performance-based funding, the recurrent public budget (or *core grant*) that the university receives is to a lesser or greater extent made dependent on a set of performance measures or output criteria. One way of doing this is to include performance indicators in the funding formulas that determine the core grant per institution. Another option is for the funding authorities to make an agreement with the universities on delivering a particular set of services in the period ahead. In the latter case, a performance contract is negotiated where the university is rewarded for delivering on its

✉ Ben Jongbloed
b.w.a.jongbloed@utwente.nl

¹ Center for Higher Education Policy Studies (CHEPS), University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

strategic plan (Jongbloed and Vossensteyn 2016). According to De Boer et al. (2015), the most frequently used performance indicators used in the funding formulas of OECD countries are the number of degree completions (Bachelor, Master and PhD), the ECTS credits earned by students and measures of research performance (e.g. research quality, publication output, competitive research revenues generated).

The Netherlands is no exception to this, and for quite a number of years the Dutch funding authorities have employed a performance-driven funding formula that determines the core grants of research universities and universities of applied sciences. In recent years, performance agreements were added to the Dutch higher education funding system. This was done in a policy experiment carried out over a five-year period (2012–2016). After this period, the policy was evaluated in order to see whether it should be continued in the future. This article discusses the experiment; how it was set up, and what impact it has had on the performance of Dutch universities in terms of increasing student performance, educational quality and diversity in programme offerings and research. Finally, we will discuss the lessons we may draw from the experiment.

Performance agreements

Funding formulas normally include a mix of input and output measures that reflect broad dimensions of an institution's current activities (e.g. student enrolments across its programmes) and/or its performance (e.g. degree completions, research output). The measures apply equally to all universities and relate to realisations, that is to the recent past. This implies that funding formulas are backward looking (i.e. *ex post*) funding mechanisms. A performance agreement is forward looking. It provides *ex ante* funding, because it includes the goals that a university intends to achieve in the coming period. In its plan, the university specifies the performance that it expects to deliver. In return, the university receives its core grant or a part thereof. The ambitions included in the agreement will reflect the particular mission of the university – its context and its strengths. While the university's goals will have to be in line with the overall national objectives set for higher education, the university normally will have room to also make some institution-specific choices in terms of goals, ambition levels and measures taken to achieve the goals. The performance agreement can include a financial penalty or sanction of some sort if at the end of the contract period it turns out that the objectives have not been achieved.

Performance agreements are not solely meant to strengthen performance but also have other aims (De Boer et al. 2015). One goal is to encourage universities to strategically position themselves through their choice of educational offerings and research focus. This is also known as institutional differentiation, or profiling. Another goal is to improve the strategic dialogue between the government and the universities, making it richer content-wise and less oriented on compliance. Yet another goal is to inform policy-makers and the public at large about the universities' performance, thus improving accountability and transparency (Jongbloed et al. 2018b). The agreements therefore address one of the risks of formula funding, namely that all universities will respond to the formula's indicators in the same way and produce more homogeneity instead of more diversity in the system (Codling and Meek 2006).

Compared to funding formulas, a system of performance agreements leaves more room for universities to specify their individual performance dimensions and have these connected to

financial rewards. Performance agreements therefore can handle situations where universities have multiple objectives and – within some nationally-set boundaries – can set their own target levels, given their particular mission and strengths. To monitor the universities' progress in meeting their agreements, the funding authorities exercise some form of oversight. In the Dutch case, the Minister of Education installed an independent *Review Committee* to oversee the agreements.

Elsewhere (Jongbloed and Vossensteyn 2016) we have shown some of the characteristics of the various performance agreements that are in place in Austria, Denmark, Finland, Germany, Ireland and some other countries. In some of these countries (e.g. Finland, Ireland) the agreements are linked to (a part of) the university's core grant, whereas in others (e.g. Denmark, Germany), the agreements are a steering instrument for the government, next to the funding formula. The Netherlands belongs to the first group of countries: the performance agreement constitutes on average 7% of the education component in a university's core grant. Next to the education component, the core grant also includes a research component. It needs to be stressed that the performance agreements are superimposed on a funding formula that – from the early 1990s onwards – is performance-based. Through this formula, some 20% of the university's education component in its core grant is based on degrees. In addition, 40% of the (separate) research component in the universities' core grant is also based on (BA, MA, PhD) degrees. This implies that, on average, a quarter (for universities) to a third (for universities of applied sciences) of the core grants in the Netherlands is based on performance measures.

The Dutch experiment: arrangements and goals

The Netherlands has a binary system of higher education, which means there are two types of programmes: research-oriented education, traditionally offered by research universities, and professional higher education, offered by universities of applied sciences (UASs). University programmes differ not only in focus, but also in access requirements, length and degree nomenclature.

There are 18 research universities in the Netherlands, including one Open University, and 38 universities of applied sciences. The universities of applied sciences have more of a regional function and focus in particular on their education mission, although in recent years they also have started to strengthen their practice-based research, partly thanks to dedicated public funds for research and research-oriented staff positions.

In 2009, the Minister of Education installed a committee – the Committee on the Future Sustainability of the Higher Education System (named *Veerman Committee*, after its chair) – to look at performance and diversity in Dutch higher education. The committee regarded the binary distinction as valuable and practical. It stated that eliminating binary divides would risk institutions competing with each other for the same students, making them more alike. Most importantly, in its advisory report (Veerman Committee 2010) the committee called for a threefold differentiation in higher education:

- (1) a differentiation in institutional types (research universities and universities of applied sciences);
- (2) a differentiation between institutions of the same type (i.e. institutions choosing their own profile);
- (3) a differentiation in the range of programmes offered.

The third dimension of differentiation translates into the range of education programmes offered in response to the increased heterogeneity of the student population. Diversity in programme offerings is seen as one of the major factors associated with the positive performance of higher education systems (van Vught 2008). Diversity is associated with the need to offer access to higher education to students with different educational backgrounds, allowing students smoother transfers to other programmes. This contributes to making students successfully complete their programmes. Dutch universities of applied sciences in particular have to contend with heterogeneity in terms of the educational preparation and the orientation of their students. In contrast, student intake in research universities is more homogeneous in terms of educational background.

The Veerman Committee was positive about the overall quality of Dutch higher education. The outcomes of the Dutch quality assurance and accreditation systems show that the generic quality is good. However, student satisfaction surveys show that, while across the board students take a positive view of the quality of higher education, there are some weaknesses in terms of teaching logistics and the degree to which highly skilled and motivated students are challenged during their programmes. In addition, the committee felt there are weaknesses related to the high level of student drop-out and completion rates. It also pointed at the relatively long time-to-degree and low levels of success for ethnic minority students compared to native Dutch students. The committee recommended that the quality and diversity of Dutch higher education had to increase, in particular when it comes to raising study success for students, and offering programmes that meet the needs of, on the one hand, non-Western ethnic minority students and, on the other, highly motivated and talented students.

As a result of the recommendations of the Veerman Committee, performance agreements were introduced in 2012. The agreements were signed between the Education Ministry and each individual university. They were formulated both in terms of quantitative indicators and qualitative ambitions. The agreements aimed at the following goals:

- Improving the quality of education in universities and universities of applied sciences in terms of, among other things, measures of students' success and other indicators of quality;
- Enhancing programme differentiation within and between universities, encouraging universities to exhibit clearer education profiles and focused research areas. This should produce a higher level of diversity in the higher education system;
- Strengthening the focus of universities on their valorisation function (i.e. knowledge exchange, research commercialization, promoting entrepreneurship).

For the period 2013–2016, 7% of the education component in the institutions' core grant (annually, on average, EUR 135 million for the research universities and EUR 175 for the universities of applied sciences sector) was tied to performance agreements. The remainder of the core grant of universities continued to be based primarily on the funding formula described above. A *Review Committee* consisting of five independent higher education experts was installed by the minister of Education in 2011, with the remit to oversee the performance agreements. The committee's task was to develop criteria for assessing the agreements, monitor each institution's progress in realizing its ambitions during the contract period, and, at the end of the period (i.e. in the year 2016), make a recommendation to the minister about whether the goals in the agreement had been met or not. If a university did not achieve its agreed goals it risked losing part of its core grant for the years ahead. It should be mentioned that the performance

agreement arrangements were set up as a policy experiment. Depending on an external evaluation, the future of the performance agreements experiment was to be determined.

For their performance agreements the universities agreed with the ministry to make use of seven mandatory indicators to state their ambitions with respect to improving student success and educational quality. The indicators used for this were: student completion (bachelor students only), student drop-out rates in Year 1, share of Year 1 students switching to other programmes, the number of students in honours programmes (aimed at students selected on the basis of their talents and motivation), student satisfaction scores, teaching intensity (i.e. the number of student contact hours per week in the first year of degree programmes), academic staff qualifications (e.g. the share of academic staff holding a university teaching qualification), and the share of overheads (indirect costs). Two of these performance indicators, completion rates and drop-out rates, received most of the attention - during the annual monitoring by the committee and at the end of the performance agreement period. It is to these indicators in particular that we will pay attention. The universities' ambitions with respect to increasing programme diversity and institutional profiling were stated in more qualitative terms, relating to topics such as starting new degree programmes and phasing out old ones, introducing student mentoring programmes, setting up research centres, engaging in partnerships with local business, et cetera.

The Dutch experiment: increased study success?

In order to learn about the impact the performance agreements have had on the performance of Dutch universities we will first look at the outcomes with respect to student success, one of the key areas related to educational quality. We will base our discussion on the reports prepared by the Review Committee and the underlying data collected by us as part of our work for this committee.

In term of the results achieved by the universities over the period of the performance agreements, we focus on two performance indicators only: degree completion and drop-out. This is, firstly, because of limits set to the length of this article, but also because completion and drop-out were the indicators regarded by those involved in the performance agreements as the key indicators. The definitions of the two indicators are as follows:

- Completion rate: the proportion of full-time bachelor's students who, after the first year of study, re-enrol at the same university and who earn a bachelor's degree at that same university in the standard time to degree plus 1 year;
- Drop-out rate: the proportion of the total number of full-time bachelor's students (only first-year students) who, after 1 year, are no longer enrolled in the same university.

For the two key student success indicators, Table 1 presents averages for research universities and universities of applied sciences. It shows the initial situation (at the start of the performance agreement), the ambition levels chosen by the institutions, and the levels that were achieved (realisations). For the 13 research universities that we have data for, we distinguish three subsets: the large comprehensive universities (3 in total), the technical and agricultural universities (4 in total) and the remaining 6 universities (mid-sized comprehensive and other institutions). In the universities of applied sciences sector, we distinguish seven subsets, based

Table 1 Completion rates and drop-out rates: overview of realisations, ambitions and initial situation for research universities and universities of applied sciences (per type of institution)

	n	Completion rate (%)			Drop-out rate (%)		
		Initial (year 2011)	Ambition (year 2015)	Realisation (year 2015)	Initial (year 2011)	Ambition (year 2015)	Realisation (year 2015)
Research universities ¹							
Large comprehensive (intake >3000 students)	3	62.9	72.3	74.9	17.7	17.0	14.3
Technical/agricultural	4	42.3	61.3	68.2	18.5	21.8	15.0
Other/Mid-sized comprehensive	6	66.4	73.3	78.0	18.2	16.5	15.7
Overall	13	60.2	70.1	73.9	17.4	18.1	14.7
Universities of applied sciences							
Arts	7	69.2	71.8	75.8	23.0	23.5	19.4
Agriculture	3	75.1	79.3	65.8	27.6	23.5	23.9
Teacher training	6	78.5	76.5	75.5	30.3	28.7	28.9
Large (intake >3000 students)	8	64.3	68.1	59.8	27.3	27.3	27.0
Mid-sized to large (intake between 1000 and 3000 students)	5	62.5	66.3	56.1	28.3	28.7	29.9
Small to mid-sized (intake between 500 and 1000 students)	4	70.9	72.0	62.1	28.5	28.1	28.4
Small (intake <500 students)	3	75.0	71.7	73.1	22.1	21.0	20.7
Overall	36	69.9	71.6	66.9	26.8	26.2	25.6

¹ Excludes the Open University, the three theological universities and the University of Humanistic Studies

Source: Based on Review Committee (2017a)

on the scope (specialised versus broad/comprehensive) and the size (from small to large) of the institutions.

The table illustrates that the research universities booked substantial results in terms of reducing drop-out and increasing completion rates. The average completion rates in research universities increased from on average 60% to 74%, and drop-out rates in the first year of degree programmes declined from 17% to 15%. The sharpest rise in completion can be observed among the four technical research universities: from an average of 42% to 68%. For many research universities the 2015 completion rates equalled or exceeded the ambition set for 2015. At two of the four universities that fell short, the 2015 completion rates were close to the target values.

Figure 1 pictures the trajectories of the completion rate (on the vertical axis) and the drop-out rate (on the horizontal axis) for the individual universities over the period 2011–2015. The background colours in Fig. 1 indicate the ‘preferred quadrant’: a low drop-out and a high completion rate (top-left corner – green) are preferred over a high drop-out and a low completion rate (bottom-right – in red). In the period 2012–2016, all three types of research universities moved towards higher completion rates and lower drop-out rates.

In the universities of applied sciences sector, the average completion rate fell from approximately 70% to 67% (Table 1). However, drop-out was pushed back slightly, from 27% to 26%. A relatively large number of universities failed to realise their ambitions. We do not show the pattern for individual institutions, but a quadrant picture like Fig. 1 would primarily have shown a movement towards lower completion and higher drop-out. Large differences between the various types of universities of applied sciences can be observed (Table 1). Only the specialised fine arts colleges (7 in total) managed to raise completion rates and lower drop-out. Most of the other types of universities of applied sciences saw their completion rates drop to percentages that were not just lower than ambition levels, but also lower than the starting position. In quite a few cases, there was a persistent downward trend in the numbers, that only appeared to take a turn for the better in the final year of the performance agreements.

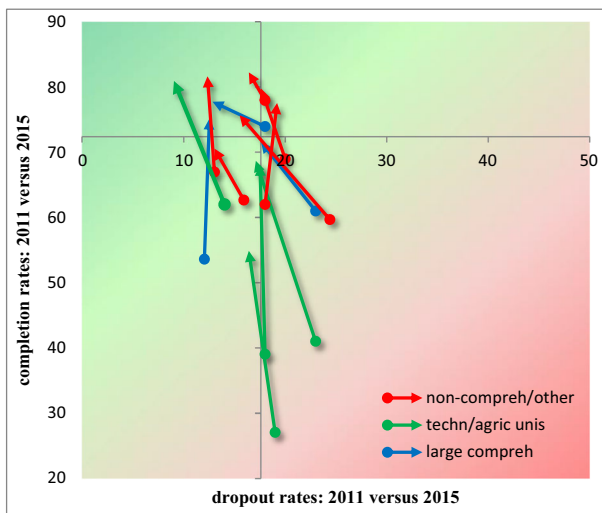


Fig. 1 Completion rate and drop-out; initial situation (year 2011) and realisation (year 2015); research universities

At the end of 2016, the Review Committee made assessments of each institution's performance in the light of the institution's ambitions and the information presented to the committee through the institution's annual reports and a meeting with the committee (Review Committee 2017a). The committee looked at all seven indicators, although completion and drop-out received most of the attention, because for the other five indicators results were mostly positive.

First of all, the committee made a positive assessment of the performance agreements of the research universities. For the universities of applied sciences its assessment was less positive. The disappointing results for these institutions with regard to student completion can in part be attributed to the trade-offs that were made between access, quality and completion – the three classic goals in higher education. Quality was interpreted by these institutions mostly in terms of meeting accreditation standards, which partly relate to the pedagogical model, the counselling and supervision offered to students, and the quality of the students' thesis work. The trade-off between the three goals was most strongly manifested in the large universities of applied sciences that have a highly diverse student population. In handling the quality standards that were placed upon the universities of applied sciences by the accreditation agency, many UASs did not want to prioritise completion rates over quality standards. In addition, the UAS sector felt an obligation to continue to provide access opportunities to students that, from an academic point of view, might be somewhat less-prepared compared to others. The trade-offs, therefore, were made in favour of quality and access and, consequently, at the expense of completion rates. Nevertheless, the UASs did manage to produce scores in the student satisfaction surveys (another indicator in the performance agreements) that showed no evidence of a decline in students' appreciation of their programme.

Taking all this evidence into account, the Review Committee in its advice to the minister concluded that only six UASs had not achieved their performance agreements, despite their efforts to increase study success and to improve other areas of performance. The minister decided to impose a financial penalty on these six institutions, but decided to only apply half of the envisaged penalty in appreciation of the initiatives the universities of applied sciences had taken with regard to improving quality and study success.

In the autumn of 2016, when the performance agreements experiment was concluded, most of the attention in the popular press and among the relevant stakeholders in the higher education sector was given to the institutions' realised values for the seven performance indicators. This was because the financial sanctions attached to the performance agreements were very much tied to whether an institution had met its agreed ambition levels on these indicators. Representatives of the universities argued that indicators like completion and drop-out rates touch upon areas that are difficult to control by the university – much less than the indicators related to the number of student contact hours (the 'teaching intensity' indicator), the share of academic staff holding a university teaching qualification or overhead shares. Because Dutch higher education is based on open entry, the universities of applied sciences stated they had very few opportunities to influence the quality of the student cohort that commences a higher education career. And compared to research universities, universities of applied sciences cater for a much more challenging student population, enrolling relatively more students with non-Western backgrounds.

The focus on indicators and quantitative targets is very much in line with a new public management type of approach to governance in higher education, where financial consequences are tied to performance targets (Ferlie et al. 1996; Hood 2007). However, in a strict version of new public management the targets would be imposed from the top, with little room

or acknowledgement for the professionals at the ‘shop-floor level’. In the case of the Dutch performance agreements, target levels were set by the universities themselves in the light of their own strengths and weaknesses. This meant that in formulating their performance agreement, the universities had the opportunity to define their ambitions, given the composition of their student population. Looking back at the results of the performance agreement experiment, it turned out that some universities of applied sciences may have set their ambitions too high and overestimated their opportunities to influence student success. The instruments they employed to raise student completion were either not effective enough or required more time to make an impact. In this respect, the performance agreements provided an important learning opportunity for the universities.

The Dutch experiment: encouraging diversity?

While the contents and results of the performance agreements in many ways stress measurable results, other valuable and often qualitative aspects of the universities’ portfolio were also part of the agreements.

Part of the budget tied to performance agreements (on average: two-sevenths; the remainder was tied to the seven quantitative indicators) was awarded to universities in the form of competitive funds. This selective budget was awarded in proportion to the quality of a university’s performance agreement plans for programme differentiation and research concentration. Assessing the quality of the plans by means of scoring them on three criteria (level of ambition, alignment with policy agendas and feasibility) was part of the tasks of the Review Committee. The universities that in 2012 had submitted the best plans received relatively more selective funding than universities with a mediocre proposal.

The Dutch performance agreements essentially were about making higher education better aligned to the needs of society and creating a higher education system that offers increased quality and diversity. Recognising that a uniform policy tends to create uniform reactions, the performance agreements were seen as the way to create diversity in terms of the universities’ degree programmes and research focus. The question is whether there is evidence that diversity increased in the period during which the performance agreements were in place. We will only focus here on educational diversity, disregarding diversity in research.

Diversity is a prominent theme in higher education (Birnbaum 1983; Marginson 2017) and science and technology policy (Nowotny et al. 2001). Diversity is held to be important because it is seen as a means to enhancing rigour and creativity, offering flexibility in the face of uncertain future progress, and promoting learning across programmes (Stirling 2007). Diversity may act as a ‘resource pool’ in providing flexibility and resilience. More broadly, institutional and technological diversity are seen as stimuli for innovation and productivity. Diversity is a property of a *system* (e.g. the higher education system), rather than of its individual elements (e.g. universities). The concept of diversity, however, is a multi-faceted notion that combines many aspects. Stirling defines diversity as a combination of three basic properties – *variety*, *balance* and *disparity* (Stirling 2007).

Variety is the number of categories (types) into which system elements are apportioned. All else being equal, the greater the variety, the greater the diversity. Obviously, a crucial issue here is to resolve the categories used. Distinguishing additional (sub-)categories however makes this a rather tricky issue.

Balance is the answer to the question: ‘how much of each type of thing do we have?’ (Stirling 2007, p. 709). Balance is perfect when each category is equally represented in the population. It is considered that, all else being equal, the more even the balance, the greater the diversity.

Disparity is the answer to the question: ‘how different from each other are the types of thing that we have?’ (Stirling 2007, p. 709). Disparity goes beyond variety and balance by accounting for the nature of the categorization. All else being equal, the more disparate are the represented elements, the greater the diversity.

Diversity is a combination of these three basic properties. However, one needs to recognise that each property constitutes the other two. Variety and balance, for instance, cannot be characterized without first considering disparity. The diversity of a system (e.g. the contents of degree programmes in higher education) can only be assessed when its elements (i.e. programmes in this case) have been grouped into categories (e.g. disciplinary areas). Once this categorization has been done, *variety* corresponds to the number of categories; *balance* to the way the elements are spread among categories (e.g. the number of new students embarking on every category of programme); *disparity* to the degree of difference between the categories.

In order to analyse educational diversity, the Review Committee was confronted with the need to categorize universities and their degree programmes. For this, it felt that simply referring to research universities and universities of applied sciences – the two categories of higher education institutions – was not enough. The committee operationalised diversity by looking at the level and range of programmes offered by universities (e.g. two-year associate degrees, bachelor’s degrees, master’s degrees, broad-based bachelor programmes, two-year research master’s, selective honours programmes, and professional master’s programmes offered by universities of applied sciences). It employed a categorization in terms of the programmatic scope of the university, combined with the size of the institution in terms of the student intake numbers for the bachelor phase (see Table 1).

As part of its analysis of institutional differentiation the committee analysed three partly overlapping features of a university’s educational profile: (1) the range of programmes offered by a university, to see whether or not an institution is broadening the scope of its programmes by covering more disciplinary areas, (2) to what extent a university is focusing on particular programmes within that programme range, and (3) the market share of the programmes provided by the university. In this article we cannot possibly cover all aspects related to these profiling dimensions (see Review Committee 2017a for more on this topic). Therefore, we will focus on the second and third aspect only: focus and market share. ‘Focus’ is one aspect of the disparity in the system. It touches on the question of whether universities differ in terms of the emphasis they give to particular disciplinary areas – both in their education and in their research activity. To analyse this in a quantitative way, information on student intakes in the institution’s respective degree programmes may be used.

With respect to education activity, the distribution of new entrants across the programmes within a university indicates the institution’s focus areas within the range of programmes on offer. Programmes are only ‘counted’ in the analysis of diversity if they have students; and the more new entrants a programme has (i.e. the bigger its share – within the institution, or nationally), the bigger the presence of the programme. The committee quantified the equality in this distribution by means of an inequality coefficient. The Gini inequality index used for this is an indicator of balance (or evenness). A Gini coefficient of zero expresses perfect evenness, while a Gini coefficient of 1 (or 100%) expresses extreme inequality. The more

unbalanced the distribution, as reflected in a higher Gini coefficient, the more sharply the focus areas will stand out and the clearer the institutional profile. The development of educational focus areas is not necessarily an indication of increasing diversity, because two different universities could choose to focus on the same areas or themes, thus reducing diversity (DiMaggio and Powell 1983). For students, the presence of focus areas can make a university stand out more clearly in the higher education landscape.

Focus areas do not result solely from an institutional profiling strategy; they also evolve as a result of fluctuations in new entrants' interest in specific programmes. This is an aspect that can be covered by a concentration or market share index, as reflected in the Herfindahl-Hirschmann Index (HHI). The HHI is often used to measure industrial concentration in a market. This indicator is defined as follows:

$$HHI = \sum_{i=1}^n MA_i * MA_i$$

where MA_i is the market share of institution i and n the number of institutions.

In the particular version of the HHI used here, the squared market shares across all programmes offered by an individual university are summed to arrive at the HHI for the individual university. If more of its students are enrolled in programmes where the institution only has a small market share, the institution's aggregated market share is relatively low. If a lot of its students enrol in programmes where it has a large market share, its aggregated market share will be high. The aggregated market share can be interpreted as a measure of concentration and therefore as another profiling feature for a university.

From the analysis of the Gini coefficients from 2006 onwards, we observe that student intake in the majority of research universities is spread out increasingly more evenly across the bachelor's programmes on offer. This trend largely has continued during the period of the performance agreements. This was even more the case for student intake in the master's programmes. This can be interpreted as a tendency towards fewer focus areas, and the Review Committee interpreted this as indicating less diversity (Review Committee 2017a). In the universities of applied sciences sector, most institutions show a more even spread – more *balance* – of students across bachelor's programmes, giving no indication of a strengthening of particular focus areas. However, when looking at the offer of master's programmes in the universities of applied sciences sector, there were more institutions moving to more clearly visible focus areas during the period of the performance agreements.

In the period up to 2011, most research universities saw a decline in their market shares (thus, a decline in their HHI) for their bachelor's programmes on offer. After the introduction of the performance agreements in 2011, however, the picture changed: nine out of 17 research universities saw a relative rise in intake in bachelor's programmes with a large market share. This indicates more diversity. For most research universities, the market share indicator for their master's programmes declined over the period up to the year 2015, especially during the period of the performance agreements. In the universities of applied sciences sector, the number of bachelor's and master's programmes with a large market share increased up to 2011. After this year (i.e. during the period of the performance agreements), however, this trend subsided. There was no visible growth anymore in the institutions' market share for bachelor's or master's programmes.

A more integrated assessment of whether the Dutch universities have worked on establishing a more clear educational profile for their institution can be made by combining the results

of the analyses of focus areas (by means of Gini coefficients) and those of market shares (HHI index) in educational provision. We have done this by placing the development of the Gini and HHI indices for the 3 years 2006, 2011 and 2015 in a quadrant graph (see Fig. 2). If a university saw its market share (horizontal axis) rise, while at the same time the inequality across the degree programmes (vertical axis) in its educational offer grew, then the university distinguishes itself more from other universities. In Fig. 2 this would then be shown as a movement towards the top-right corner. A movement in the direction of an educationally less distinct university would be shown as an arrow that points towards the bottom-left corner.

Figure 2 shows the result for the provision of master's level programmes by research universities, where the changes over time are relatively larger and clearer than for areas like bachelor's programmes. The arrows in Fig. 2 clearly illustrate that research universities tend to get a less clear profile over time. For the universities of applied sciences, that mostly offer bachelor's degree programmes (not shown in the picture), one cannot detect a clear movement to one or the other quadrant, meaning there is no indication of an increased profiling of institutions – meaning less disparity, i.e. less diversity. Looking back at the performance agreements in the Netherlands, we conclude that the results related to the objective of increasing diversity are rather mixed.

Conclusion: lessons and the way ahead

Overlooking the outcomes of the performance agreements experiment in the Netherlands, the evidence shows that the research universities managed to increase quality and completion in education, while the universities of applied sciences sector experienced several problems in achieving the wished for completion rates. In terms of the diversity goal, results are inconclusive. When it comes to diversity in educational programmes, there is no clear sign of

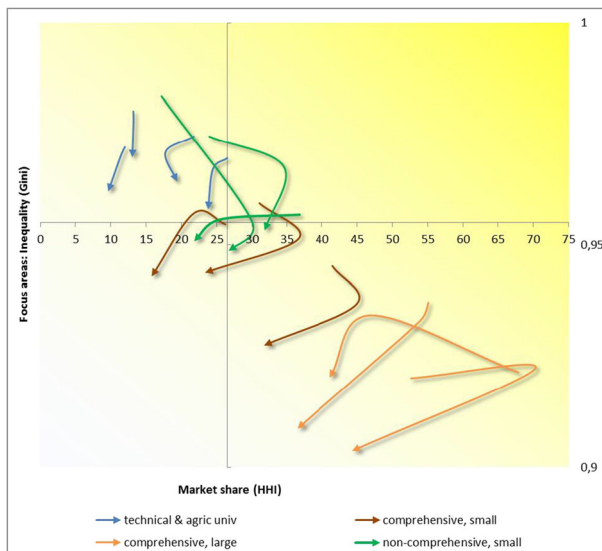


Fig. 2 The provision of Master's programmes by research universities: Indicators for Focus areas and Market shares (from 2006 via 2011 to 2015)

institutional differentiation: most institutions exhibit an increasingly equal spread of educational activity over their programmes – and this was also the case for research.

This may produce the impression that performance agreements have not achieved a lot. However, in evaluations of the policy experiment published in 2017 by three different committees, the conclusions were much more positive. First, the Review Committee itself produced an evaluation report (Review Committee 2017b). Second, the association of universities of applied sciences ordered an evaluation (Slob et al. 2017). Third, the Minister of Education ordered an independent committee to evaluate the experiment and make recommendations for a future system of performance agreements (Evaluatiecommissie Prestatiebesteding Hoger Onderwijs 2017). The three committees agreed on many issues. On the positive side, they concluded that the performance agreements had contributed to the following outcomes:

- putting the improvement of students' study success more prominently on the institutions' agendas;
- intensification of the debate about the drivers of study success (both among universities and within universities);
- more attention for the profiling (differentiation, focus areas) of universities;
- improvement of the dialogue between stakeholders in higher education (executive boards of universities, ministry, department heads, associations of universities, Review Committee, representatives of business and community), including the possibility for universities to share their 'story behind the numbers' with the Review Committee;
- increased transparency and accountability, thanks to the setting of targets and the use of indicators.

Less positive were universities and student associations about:

- the decline of university autonomy due to the setting of national targets and use of mandatory indicators;
- the additional bureaucracy and administrative cost due to the emphasis on indicators;
- the financial penalty associated with the non-achievement of goals;
- the choice and definition of indicators, which in some cases contributed to unintended effects (e.g. an over-emphasis on quantitative outcomes instead of qualitative achievements);
- the lack of time available for a well-considered construction of the procedures surrounding the experiment;
- that the experiment was managed largely by stakeholders (executive boards, managers, ministry, national committees and organisations) that were quite distant from the 'shop floor level', with a small role only for students in this process.

Nevertheless, in the evaluations of the performance agreements by the three evaluation commissions the need was reaffirmed for incorporating a performance-oriented component in the funding mechanism for universities. The then Minister of Education expressed her intention to continue with some form of performance agreements, but was keen to stress that the agreements should ultimately be about the quality of higher education and quantitative targets should not receive priority over qualitative ones.

On the topic of potential financial sanctions tied to quality agreements there was less agreement. On the one hand the Review Committee in its evaluation concluded that attaching

financial consequences to agreements fosters their effectiveness. It argued that both the international literature and the Dutch experiment have shown that agreements are taken more seriously by all the parties and have greater impact if financial consequences are attached (Review Committee 2017b). Elsewhere there was a preference for rewarding universities that fully delivered on the performance agreement, but not punishing universities financially if they had not met their agreement. The rectors' associations showed little enthusiasm for performance agreements and stated that universities should always have autonomy to decide on their ambitions in dialogue with their internal and external stakeholders and be accountable to those same stakeholders. The term used here is *horizontal accountability* (Jongbloed et al. 2008), meaning universities primarily report to those agents that are not their hierarchical superiors (e.g. ministry) but to students, regional stakeholders and professional organisations. Universities' executives clearly prefer horizontal accountability over a vertical type of steering.

Now that the evaluations of the performance agreements experiment have been published and a new coalition government is in place from October 2017 onwards, the decision has been made to continue the agreements under the label of Quality Agreements (Ministry of Education 2018). The Quality Agreements will only concern educational quality and are no longer also about research. It was agreed there will be mild financial consequences attached to the agreements and less steering by the government in the process. Indicators will play a role, but their role is determined by the university itself. The agreements therefore are more horizontal. There also will no longer be an independent expert committee reviewing the agreements. Instead, the assessment is placed with an existing organization (i.e. the national accreditation agency) that will integrate the monitoring of quality agreements in its regular assessment of the institution's educational quality. Universities are expected to discuss progress in their internal decision-making bodies, giving a bigger role to student representatives.

What the Quality Agreements will bring is still unclear. But what is clear is that the agreements have lost whatever they included in terms of new public management ingredients. That ambitions are to be agreed in close dialogue with the universities' relevant (local) stakeholders implies that the agreements will develop more into a steering instrument that fits the public value management paradigm (Stoker 2006; Jongbloed et al. 2018a).

Whether performance agreements, or indeed performance-based funding formulas, matter for the performance of higher education is a question that cannot be answered on the basis of the Dutch experiment with performance agreements alone. Although the Review Committee claims that the agreements were indeed effective, causality is difficult to prove. Nevertheless, many countries continue to employ performance-based funding mechanisms, but many do so without necessarily having evaluated how effective the approach actually is (Jongbloed and Vossensteyn 2016). Policy evaluations on the impact of performance-based funding are rare. On the one hand there is some evidence that formula-based performance funding has failed to increase degree completions (Hillman et al. 2015). On the other, there is some scattered evidence that points to its benefits (Claeys-Kulik and Estermann 2015; De Boer et al. 2015). The Dutch performance agreements can therefore be seen as an experiment on the way towards providing some of that evidence and using it to inform the design of future funding mechanisms for higher education.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or

exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bimbaum, R. (1983). *Maintaining diversity in higher education*. San Francisco: Jossey-Bass.
- Claeys-Kulik, A. L., & Estermann, T. (2015). *DEFINE thematic report: Performance-based funding of universities in Europe*. Brussels: European University Association.
- Codling, A., & Meek, L. V. (2006). Twelve propositions on diversity in higher education. *Higher Education Management and Policy*, 18(3), 1–24.
- De Boer, H., Jongbloed, B., et al. (2015). *Performance-based funding and performance agreements in fourteen higher education systems*. Report for the Ministry of Education, Culture and Science. The Hague: Ministry of Education, Culture and Science.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.
- Evaluatiecommissie Prestatiebekostiging Hoger Onderwijs. (2017). *Van Afvinken naar Aanvonken*. Den Haag: Ministerie van OCW.
- Ferlie, E., Ashburner, L., Fitzgerald, L., & Pettigrew, A. (1996). *The new public management in action*. Oxford: Oxford University Press.
- Hillman, N. W., Tandberg, D. A., & Fryar, A. H. (2015). Evaluating the impacts of “new” performance funding in higher education. *Educational Evaluation and Policy Analysis*, 20(10), 1–19.
- Hood, C. (2007). Intellectual obsolescence and intellectual makeovers: reflections on the tools of government after two decades. *Governance*, 20(1), 127–144.
- Jongbloed, B. W. A., & Vossensteyn, J. J. (2016). University funding and student funding: international comparisons. *Oxford Review of Economic Policy*, 32(4), 576–595.
- Jongbloed, B., Enders, J., & Salerno, C. (2008). Higher education and its communities: Interconnections, interdependencies and a research agenda. *Higher Education*, 56(3), 303–324.
- Jongbloed, B., Kaiser, F., van Vught, F., & Westerheijden, D. F. (2018a). Performance agreements in higher education: A new approach to higher education funding. In A. Curaj, L. Deca, & R. Pricopie (Eds.), *European higher education area: The impact of past and future policies* (pp. 671–687). Dordrecht: Springer.
- Jongbloed, B., Vossensteyn, H., van Vught, F., & Westerheijden, D. F. (2018b). Transparency tools for higher education. In E. Hazelkorn, H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 211–228). Edward Elgar.
- Marginson, S. (2017). Horizontal diversity in higher education systems. *Does the growth of participation enhance or diminish it?* Paper for the CGHE seminar 6 July 2017, London: UCL Institute of Education / University College London.
- Ministry of Education (2018). *Aanbieding sectorakkoorden hogescholen en universiteiten*. (Letter to Parliament). Brief aan de Tweede Kamer, 9 April 2018. The Hague: Ministry of Education.
- Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. London: Polity Press.
- Review Committee / Higher Education and Research Review Committee (2017a). *System report 2016. Fourth annual monitoring report on the progress of the profiling and quality improvement process in higher education and research*. The Hague: Review Committee.
- Review Committee / Higher Education and Research Review Committee. (2017b). *Prestatieafspraken: Het Vervolgproces na 2016. Advies en Zelfevaluatie*. The Hague: Review Committee.
- Slob, A., Jeene, B. G., Rouwhorst, Y. T. M., Theisens, H. C., & van Welie, E. A. A. M. (2017). *Kwaliteit door Dialoog. Eindrapport van de commissie prestatieafspraken hbo*. Den Haag: Vereniging Hogescholen.
- Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15), 707–719.
- Stoker, G. (2006). Public value management: a new narrative for networked governance? *The American Review of Public Administration*, 36(1), 41–57.
- van Vught, F. (2008). Mission diversity and reputation in higher education. *Higher Education Policy*, 21(2), 151–174.
- Veerman Committee (2010). *Threefold differentiation. For the sake of quality and diversity in higher education. Recommendations of the Committee on the Future Sustainability of the Dutch Higher Education System*. The Hague: Ministry of Education.