COGITATIO

Article

Keeping One's Shiny Mercedes in the Garage: Why Higher Education Quantification Never Really Took Off in Germany

Maarten Hillebrandt

Faculty of Law, University of Helsinki, 00014 Helsinki, Finland; E-Mail: maarten.hillebrandt@helsinki.fi

Abstract
The cybernetic dream of regulatory 'dashboard control' has taken off in the German higher education system. Both government regulators and university managers are engaged in the creation of waves of increasingly fine-grained quantitative data. Yet a wide range of recent case studies of the German higher education sector attest that in spite of this 'datafication' frenzy, the impact of the collected data mass on regulatory and managerial decision-making capacities seems to have remained relatively limited. This article explores why, in spite of the considerable investment in quantitative data infrastructures in the German higher education sector, this did not result in significant overt analytical capacity building. It explores three hypotheses: 1) a legal hypothesis according to which quantification is curbed by legal protections under the *Rechtsstaat*; 2) a dysfunctionality hypothesis which holds that decision makers reject quantification as a flawed and impracticable pursuit; and 3) an egalitarian federalism hypothesis which argues that Germany's federal states seek to prevent commensurability to avoid comparison and competition. The article finds that, in spite of its inconspicuousness, quantification indeed does inform various central decision-making processes. However, different legal, political, and relational factors prompt decision makers to engage in a hybrid, tempered and, overall, untransparent application of numerical data.

Keywords
administrative capacity; data; education; Germany; quantification

1. The Shiniest Car: Unused in the Garage?

The cybernetic dream of regulatory 'dashboard control' has taken off in higher education systems across the world (de Boer, Enders, & Schimank, 2008; Espeland & Sauder, 2016; Hood, James, Peters, & Scott, 2004, Chapter 3). Germany is no exception to this trend. Over the past three decades, the sector has engaged in the creation of waves of increasingly fine-grained data, which often takes on a numerical form (Franzen, 2018; Huber & Hillebrandt, 2019; Kleimann, 2016). Since about 10 years ago, university administrations, too, have begun to expand their performance data regarding teaching and research activities. The construction of information systems, quality assurance mechanisms, and core data sets

is thus the talk of the town in German higher education (Biesenbender & Hornbostel, 2016; Seyfried & Pohlenz, 2017).

Surprisingly however, the impact of the collected data mass seems to have remained relatively limited in German higher education. Indeed, the initiation of numerical data collection projects with steering potential, followed by the relative disregard of that data in decision making processes appears to be a recurrent feature. To be sure, quantitative data are used for purposes of steering and reaching decisions on such issues as budgeting, research funding, hiring, and programme accreditation (Huber & Hillebrandt, 2019; Kleimann & Hückstädt, 2018; Leibner, 2017; Oberschelp, 2017). However, the impact of such quantification is largely displaced by

pre-existing arrangements. This even appears to be the case when such data are allegedly included in decisional procedures. Most German states' performance-based budgeting policy, for example, is based on only a few indicators, and capped to maximal year-on-year fluctuations, making the metric relatively inconsequential (Oberschelp, 2017, p. 109; Schimank, 2009, p. 134).

Some forms of quantification appear to have barely caught on. Higher education regulators and managers, for example, generally consider article-level metrics, which have mushroomed over the past decade, too context-specific and complex for decision-making frameworks (Franzen, 2018). Meanwhile, faculty appointment committees rather orient themselves on the intangible notion of research reputation (Kleimann & Hückstädt, 2018). (Re)accreditation procedures for study programmes, in turn, rely foremost on qualitative peer review (Schneijderberg & Steinhardt, 2018), and data collection activities to improve teaching quality are met with mistrust by university teachers (Seyfried & Pohlenz, 2017, p. 98). Finally, university rankings, produced by newspapers and specialised organisations, are almost wholly ignored by regulators and university managers alike (Leibner, 2017, p. 48).

The above brief overview, though impressionistic, offers a first indication that, even when universities and government regulators are quite interested in the establishment of number-based data infrastructures (Prenzel & Lange, 2017, p. 18), the impact of such data on decision making appears to be faltering. Decision making is decoupled from data evidence, objects appear difficult to quantify, or comparison between objects is considered problematic and is thus resisted. This article explores why the 'shiny Mercedes' of elaborate quantitative data in the German higher education sector has not resulted in analytical capacity building to the extent that it did elsewhere.

2. Connecting Quantification and Analytical Capacity

Quantification has for some time been associated with managerialism in the higher education sector and beyond. However, in the case of Germany, the link between quantification and analytical capacity, which is a precondition of managerialism, appears to be weak. This section develops a number of hypotheses as to why this might be the case.

2.1. Analytical Capacity as a Precondition of Managerialism

The past decades have seen the rise of the new public management in higher education sectors in Western Europe and beyond (de Boer et al., 2008). The new public management paradigm is widely perceived to be underpinned by a neoliberal programme for the public sector, in particular higher education, in which efficiency, accountability and a consumer orientation moved to the

foreground as central guiding principles (Power, 1997, p. 43; Shore & Wright, 2000, p. 60). The diagnosis of neoliberal managerialism sparked a debate about the extent to which higher education sectors in different countries, not least in Germany, actually follow its central tenets (de Boer et al., 2008; Hood et al., 2004; Schimank, 2005). This article takes a step back from this discussion, to consider a *necessary precondition* for a managerialist outlook in higher education: the collection of and systematic engagement with performance and quality information for steering universities, or *analytical capacity* (Head, 2008, p. 2; Wollmann, 1989, p. 233). The process by which *numerical* indicators are used to effect impact on managerial or regulatory decision making (de Boer et al., 2008, pp. 37–40; Parrado, 2014, p. 88; Power, 1997, pp. 98–104), has been defined as *quantification* (Huber & Hillebrandt, 2019; for an alternative, more encompassing definition of quantification, see Espeland & Sauder, 2016, p. 21). The relation between quantification and analytical capacity may vary in intensity from one context to another. In some instances, the two may be tightly coupled to the extent that numerical indicators determine decisional outcomes. In these cases, we might say that the managerialist dream of 'governance by numbers' is fully realised (Huber & Hillebrandt, 2019, p. 260).

Various criteria however need to be fulfilled before we can speak of truly impactful analytical capacity (see Figure 1). Crucially, a *data infrastructure* must be in place. Data infrastructures typically require experts who collect and compile analytical data in a systematic manner (Parrado, 2014). Moreover, to speak of governance by numbers, that data needs to be predominantly *cast in quantitative terms*.

Quantitative data infrastructures can only enhance analytical capacity when they come about in a *clear decision-making context*. Governing by numbers goes beyond the mere observation and consideration of data. Instead, the indicators influence decisions in a direct way. The focus of analytically oriented decision makers is on the question: "What happens if we change these settings?" (Head, 2008, p. 1). As such, there must be at least a *plausible linkage between the collected data and the decisional output*. The more numerical representations alter decisional outcomes, the stronger the analytical capacity that quantification provides (Parrado, 2014, p. 89). In other words, "What matters is that moment when numbers oust judgment, or at least marginalise it or limit its operation to specific domains" (Kurunmäki, Mennicken, & Miller, 2016, p. 395). In practice however, decision makers' attitude towards data often appears more driven by circumstantial conditions than by the computing capabilities to which they have access as such (Head, 2008, p. 9; Parrado, 2014, p. 93).

Previous research found that while German federal government departments took active steps to enhance their data processing capacities, the connection between the resultant knowledge base and decision making remained weak (Wollmann, 1989, pp. 261–262). This study
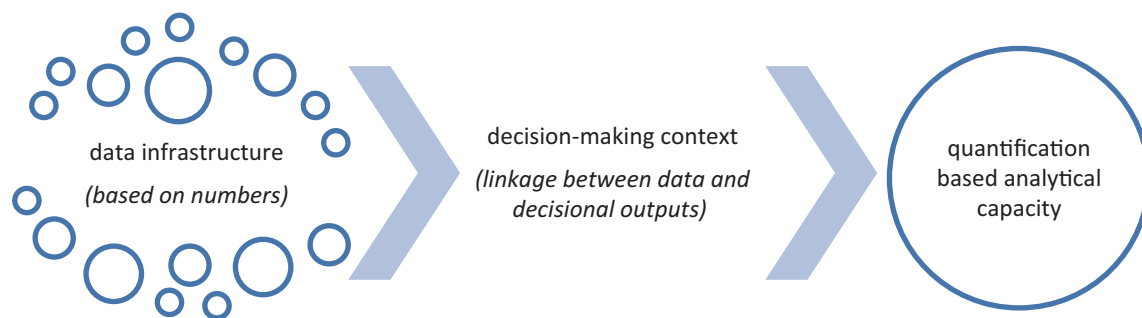
Figure 1. Governing by numbers: Quantification-based analytical capacity.

seeks to test and update these findings in the context of the considerably more digitised (Parrado, 2014, p. 86), decentralised (Schimank, 2005) and autonomised (Kleimann, 2016) German higher education sector.

2.2. Obstacles to Quantification-Based Analytical Capacity

The picture that emerges from recent literature sketches German higher education as a system that has, over the past years, produced a 'shiny new Mercedes' worth of new data infrastructures. However, the Mercedes appears mostly to stand in the garage. The collected data hardly make up a decisive factor in managerial and regulatory decision making. If German higher education decision makers have the numbers, why do they not let them govern? In order to explain this picture, this article considers three alternative hypotheses derived from different literatures.

According to the first explanation, here described as the *legal hypothesis*, the general desire among regulators and managers to rely on quantification is significantly curbed by existing legal arrangements, leading to a 'driving ban.' This view derives from the comparative policy literature, which describes the German decision-making context as a corporatist system built on a legalistic order, as opposed to the more liberal marketist Anglo-Saxon or social-democratic Scandinavian systems (Esping-Andersen, 2009). Although a direct comparison of the analytical capacity engendered by quantification in different higher education systems goes beyond the scope of this article, and the treatment of this hypothesis must therefore remain exploratory, it is here noted that this view is well-anchored in the higher education literature, which typically characterises German higher education as a system with strong academic rights entrenchment (Hood et al., 2004; Hüther & Krücken, 2013, p. 316; Seyfried & Pohlenz, 2017, pp. 99–100; see also de Boer et al., 2008, for further elaboration of this theme). This is believed to create a relatively resilient and reform-resistant governance system, undermining advances of governance by numbers.

A second (and alternative) explanation is the *dysfunctionality hypothesis*. According to this hypothesis, regulators and managers recognise the general value of quan-

titative data, but consider them unfit for incorporation in decision-making instruments. Negative experiences or general scepticism lead to the gradual side-tracking of indicators in the steering system, giving them a 'flat tire.' This hypothesis derives from a long-standing critical rationalist strand in the public management literature, which argues that management instruments—including data infrastructures—when unwieldy, must either be repaired or decoupled from decision making (Bouckaert & Balk, 1991; Van Thiel & Leeuw, 2002). Governing by numbers has been associated a host of dysfunctionalities associated with technical, definitional, and behavioural deficits (Oberschelp, 2017, pp. 112–118; Van Thiel & Leeuw, 2002, p. 270). The literature on (higher education) quantification has long recognised this problematic side to quantification-based steering, highlighting various potential unforeseen and perverse aspects in a variety of contexts (Espeland & Sauder, 2016; Huber & Hillebrandt, 2019; Strathern, 2000). It is less clear how German regulators and managers relate to such commonly reported observations.

The third hypothesis, that of *egalitarian federalism*, is rooted in the literature on federal politics (Lijphart, 1984/2010). This literature identifies a competitive dynamic both among the affiliated component states and between the decentral and the central government layers. It builds on the observation that German higher education makes up a devolved policy area (Dobbins & Knill, 2016, p. 72), and as such effectively functions as a governance quilt of sixteen independent and loosely cooperating systems (Hüther & Krücken, 2013, p. 308). Local differences in preference and political outlook lead to the absence of a coherent overarching steering logic, akin to a car with 'scattered parts.' The absence of a powerful co-ordinating centre (the federal government) means that decisions towards convergence take place on a lowest common denominator basis. It may be expected that this protective structure hinders efforts at coordinated quantification, as the latter opens new conflict lines regarding decisional autonomy.

The three hypotheses are shown in Table 1. Beyond these hypotheses, it also needs to be established whether previous scattered observations of the decoupling of analytical capacity from decision making can actually be generalised across the system. It is for exam-

Table 1. 'The shiny Mercedes': Why higher education quantification never really took off.

Hypothesis	Description	Situation
Legal hypothesis	'driving ban'	Influence of quantification is curbed by (constitutional) legal protections
Dysfunctionality hypothesis	'flat tire'	Failure of quantification leads to its gradual side-tracking
Egalitarian federalism hypothesis	'scattered parts'	Disconnect between state steering instruments, marginalising quantification

ple imaginable that the different hypotheses interact, or only apply under specific circumstances. For this reason, the hypotheses are treated in an exploratory manner, as scenarios that aid theorisation.

3. Research Approach

In order to investigate the perceived dynamics around different quantitative data infrastructures, the article considers the managerial and regulatory situation of universities in three German federal states, whereby the focus lies on decision makers and knowledge creators and brokers. Taken together, the selected states (whose identity is suppressed for reasons of respondent anonymity) are broadly representative of the available variety in German higher education sector along several variables, including number of public universities, expenditure (Deutschland in Zahlen, 2017), professor-student ratios (Destatis, 2018, p. 26), and political-ideological legacies (Lanzendorf & Pasternack, 2009; Schimank, 2009). Between them, the selected states enrol a sizeable proportion of all German students (Federal Statistical Office, 2018).

The analysis builds on data collected for a European comparative study on quantification in the public sector, namely the Open Research Area project entitled 'Quantification, Administrative Capacity and Democracy,' funded in Germany by the German Research Foundation (project number 627097). Part of the 'Quantification, Administrative Capacity and Democracy' project work consisted of the creation of a register of current quantification-based regulatory instruments and related policy documents, as well as interviews with various actors in the higher education sector. Regulatory instruments were inventoried on the basis of a review of relevant academic publications in German higher education journals, reports, policy documents, and websites produced by the central actors in the German higher education sector, including those produced by various authorities such as the science and education ministries of the selected states and the federal government, relevant state-affiliated agencies and institutes such as the statistics offices of the selected states and the federal government, the German Council of Science and Humanities, the German Research Foundation, the Accreditation Council, and the German Centre for Higher Education Research and Science Studies, and finally, the

central administrations of the selected universities. The interviews in turn served to establish which quantification instruments are most salient from a regulatory and managerial perspective, and to certify that no instruments were overlooked. As quantitative data infrastructures presented by (commercial) third parties were found to have a negligible bearing on regulatory and managerial decision making, they are not treated in detail in the analysis (see however Ringel, Brankovic, & Werron, 2020; Krüger, 2020).

For this article, 15 interviews with 19 university managers and state regulators are included. The managers interviewed (12) occupy positions at central and faculty level in six universities (4 in state A, 1 each in states B and C) that vary in terms of size (in student numbers), age (founding year), wealth (in terms of budget per student and third-stream funding), and teaching pressure (students per professor). The selected regulators (7) occupy positions in the science and education ministries, or in directly related agencies. An overview of interviewees can be found at the end of this article. The analysis consisted of three steps. First, a baseline inventory of quantification-based data infrastructures at the university, state, and federal levels was conducted. Second, the background and functioning of these infrastructures were explored with interviewees, to identify to what extent and in what ways these infrastructures are coupled to decisional processes of steering and management. The third and final step consisted of establishing the presence or absence of the obstacles as formulated in the three hypotheses, both at university and state level. The following section will consider the findings more closely, whereby the focus lies on the third step of the analysis (identifying the presence/absence of evidence of the three hypotheses). References in this section remain limited to quotations (details on interview-based evidence can be obtained from the author).

The analysis also has a number of shortcomings. First, the single-country focus means that the study cannot verify whether identified explanations are specific to the German case, or whether they are more widely generalisable. Further (comparative) case studies would be required to corroborate this aspect. Second, due to space constraints, the reported analysis can only summarily engage with regulatory differences per level and pertaining to functional differentiation of activity (teaching/research). However, thorough efforts are made to

clarify the scope of empirical claims. Third, although efforts were made to include a wide variety of local conditions through a most-different cases and within-cases selection, the data derived from the interviews may not be fully representative of all public universities and states in the German higher education context.

4. Quantification-Based Analytical Capacity: The Invisible Mercedes

A baseline inventory of data infrastructures corroborates the starting position of this article that the German higher education sector is indeed undergoing a process of 'datafication' (Table 2). This development is nearly universally confirmed by the respondents.

At the same time, many of the same respondents argue that direct governance by numbers remains a marginal phenomenon, suggesting that data infrastructures are indeed decoupled from decision making. Respondents identify performance-based budgeting as the major exception to this trend. However, the financial impact and administrative burden of performance-based budgeting are everywhere limited (see also Section 4.2). The downside of explicitly decoupling data analyses from regulation is that indicators remain inconsequential and obtain a largely rhetorical character:

It is always the best for all, if it continues the way it is. Good, you could say, then we also don't need to collect new numbers, then we actually don't need to do anything. But I'll just say, the influence of the media on politics is getting stronger. It is true that a science minister today gets asked more and more about various subjects, and then of course he wants to have

Table 2. The 'datafication' of the German higher education sector.

Locus	Type of data infrastructure	
	Teaching	Research
Faculties	Indicator dashboards (levels of advancement differ strongly)*	
		Performance contracts with faculties or professors (some quantification)
University administrations	Central data analysis departments with business intelligence systems, financial health data and data warehouses**	
		Performance contracts with faculties or professors (some quantification)
State ministries		Performance contracts with universities (some quantification)
	Performance-based university budgeting system relying on quantitative indicators	
	Data visualisations to enhance transparency regarding performance	
	Instruments for charting universities' student absorption capacities and average operational teaching costs per discipline area	
Federal Science Council		'Research rating': distribution data on different indicators
		'Core Data Set': promoting standardisation of various definitions for different reporting purposes***
State and federal statistics offices	Various mandatorily reportable data per state, discipline or semester	
(Semi-) public research bodies	Graduate surveys at national and state level (partially quantified)	

Notes: * This partially depends on the extent to which discipline areas are already accustomed to working with quantitative data; ** I.e., data concerning research projects (e.g., external funding attracted and data from the German Research Foundation Funding Atlas), concerning publications (journal impact and citations, often divided by the number of professors), and data concerning teaching modules (including students attracted, number of credits handed out, student satisfaction); *** Most universities are currently aligning their internal data accordingly.

something to answer. (Respondent 2, translation by the author)

In summary, the inventory of quantification instruments confirms that number-based data infrastructures are indeed present, growing, and actively promoted in key areas of German higher education governance, but are hardly coupled directly to decisional outcomes. Decision makers tend to predominantly use number-based data infrastructures as a way of 'getting grip' on a managerial context. What holds them from taking the final leap, by letting the numbers decide for them? The following analysis suggest that the answer to this question lies in a number of specific legal, administrative, and federal political conditioning factors.

4.1. Legal Rights and Obligations: Acting in the Shadow of Legalistic Culture

According to the legal hypothesis, quantification-based analytical capacity is incongruent with legal protections guaranteed by the German *Rechtsstaat*. Respondents, both university managers and government regulators, frequently asserted that they would like to instigate data-driven decision making, but are legally prevented from doing so. The force of legal constraints is underlined by a series of significant (constitutional) court interventions over the past decades setting out the rules of engagement in higher education. In one state, for example, the competitive element between universities was removed from the performance-based budgeting system after criticism from its Court of Auditors. Another legal constraint on the use of quantitative data for strategic purposes is data protection law, which in most cases prohibits the dissemination of individually attributable data:

> We can't even say how high the study discontinuation rate is….Because we cannot verify study careers since, for reasons of data protection, such data are not collected. (Respondent 2, translation by the author)

Clearly, a strict interpretation of the notion of data protection sets limits on the possibility of equity-enhancing quantification going in the direction of affirmative action to ensure *de facto* enrolment and graduation of ethnic minorities, akin to policies developed in other countries, notably England. There, an Office for Fair Access keeps explicit track of the diversity of student populations within specific universities for regulatory purposes (United Kingdom Department of Education and Skills, 2004).

The constitutional protection of academic autonomy is also strong. In particular, tenured professors enjoy extensive protections from interference. In practice, this makes it difficult to motivate them to provide performance data. A weaker variety of the autonomy argument applies to the university organisation as a whole. In some cases, universities refuse to deliver requested data, because they are unconvinced about the knowledge creators' methodology. In regulatory relations where they cannot avoid quantification, universities will lobby hard for the adoption of the least consequential form. One area in which they have been unsuccessful, is that of calculating capacity for creating study places. Here, court action to protect citizens' legal right to higher education has created a *numerus clausus* system building on an elaborate formula for absorption capacity. In other choices, for example human resources, university management is constrained by detailed regulation protecting professional prerogatives.

To a large extent however, reluctance to rely on overt quantification to arrive at managerial or regulatory decisions is related to *semi*-legal argumentation. When driving rules are heavily regulated, it may seem more opportune to leave the car inside altogether. In many cases managers or regulators would be free to engage in data-driven decision making, they seek to avoid confrontations, which in turn strengthens their interlocutors' negotiating position. No law, for example, prevents state ministries from drawing a far more rigid connection between quantified performance and university budgets. However, consensualism requires the downplaying of performance transparency to avoid embarrassment for 'losing' universities. University managers could propose sharper internal financial distribution formulas. However, their position of 'leader among peers' means that university managers generally steer clear of openly confrontational quantification:

> I think, one should always bear in mind the specificity that we are not in the open economy, where you can harshly say: "You have not reached your target, tomorrow you can find yourself another place to work." I am very happy…that I am analysing it externally. (Respondent 9, translation by the author)

The legal hypothesis is thus far from a generalisable feature. The law is rather invoked by academics or universities to highlight specific prerogatives, than to resist quantification per se. However, when too conspicuous, quantification is deemed to intervene undesirably in the existing system of interdependencies of the higher education system. The reliance on data infrastructure is thus accorded a relatively invisible place in the background.

4.2. Dysfunctional Data: Inducing Caution, but Not Rejection

The dysfunctionality hypothesis presents an alternative explanation: Decision makers' believe that data-driven decision making does not work. Indeed, regulators are well aware of the negative side effects of governing by numbers, and there is strong reluctance to couple quantitatively measured performance to significant budgetary decisions:

As a university, I can't from one day to the next lose €10 million….Most of the money is locked into wages and salaries. In existing contracts. They have to be honoured as well. And [the buildings] need to be heated, and the lights have to be on, and then there's hardly anything left. Then you can cut a few research assistants and…some project budget, but that's not all too much. (Respondent 2, translation by the author)

This perception does not threaten the perpetuation of the performance-based budget model, which functions unobtrusively in the background (see Section 4, introduction). The already limited impact of the model is further constrained by additional demand-based teaching quality and pay-per-student funding schemes introduced to address the pressing financial situation in the sector. Both schemes are largely input-oriented without regard for output.

Quantification-based analytical capacity entails various difficulties related to technical aspects. Required data may be unreliable or unavailable in a standardised, longitudinal form. In some cases, the desired data are altogether absent:

The coalition agreement of the [current state] government…stated that the drop-out rate should be reduced by 20 percent….We read that the next day, and said: "How should that work? We don't even know what the *current* drop-out rate is." (*laughs*) (Respondent 3, translation by the author)

A further technical aspect is the feasibility of data infrastructure building. For effective steering, decision makers require up-to-date numbers. However, in many cases data collection is a resource-intensive and time-consuming affair. Beyond technical aspects, there is the definitional question what indicators are required for particular policy objectives. Available data may not fit goals or be irrelevant for some academic disciplines. For this reason, most budgetary allocation models, whether imposed on or within universities, avoid complex and controversial data. Decision makers struggle with this problem and arrive at different solutions in different states:

We always found this discussion difficult. I observed that in the other states as well: with research, they are always quick [with operationalisations]….That's not controversial. In the area of teaching, there is a discussion in our state as well as in others: what is actually the right indicator? The question is, when I count graduates, do I actually capture anything that represents teaching quality? (Respondent 14, translation by the author)

In terms of behavioural aspects of quantification, it is widely believed that too much transparency (e.g., through ranking) leads to inequitable outcomes and must remain limited. This stands in stark contrast to ex-

periences in many countries, including the United States, where private-party rankings hold strong sway over the sector (Espeland & Sauder, 2016). While such rankings also exist in Germany, decision makers largely refuse to make prestige and resource allocation dependent on them. At the same time, both regulators and managers point out that risks can be hedged successfully, for example by capping maximal budgetary fluctuations, or by incorporating discretionary elements. Particularly university managers see quantification as a hybrid form amenable to experimentation (see also Huber, 2020). As such, a distinction can be made between data-driven and data-*informed* decision making. While the former makes decisions explicitly dependent on particular data pictures, be it through the setting of thresholds, minimum values, classifications, or more formulaic, even algorithmic, numerical composites, the latter merely uses such pictures as an inspiration that is expressly non-committal and open-ended.

In sum, the findings do not convincingly support the dysfunctionality hypothesis. Decision makers show an awareness of the potential dysfunction of quantification-based analytical capacity, but this does not stop them from using numbers to inform their decision making. Aware of the risks involved, they strive for practical feasibility, modesty of expectations, and hybridity. Even with a flat tire, the Mercedes might get you further than going by foot, albeit perhaps rather via backroads than the highway. This promotes an untransparent and somewhat detached incorporation of key indicators into traditional decision-making processes.

4.3. Resistance to National Commensuration: Transparency Limits Engagement

The egalitarian federalism hypothesis focusses particularly on the attitude of state regulators, and applies only to German higher education inasmuch as it is regarded as a single system. The broad support for different national data initiatives shows that the states are willing to cooperate actively to make standardisation and numerical comparison more feasible, albeit reluctantly. The research rating was mainly thought of as a counterweight to international rankings that were questioned on grounds of methodology and their fit with the German higher education landscape. Indeed, for a variety of reasons, Germany consistently undershoots its expected target based on sector size and budget in these rankings.

Still, coordinated national 'datafication' efforts remain modest. One regulator in a more forward state speculated that the kind of performance visualisation tool that existed in his state could never be mainstreamed nationally:

We are after all a large state, we have a whole row of universities here that can also be compared….We would very much like to do this for the entire federal republic, but we happen to have this federalism

and every state knows best in this area, so in that sense...we're not going to get that. (Respondent 3, translation by the author)

Some evidence suggests that when national quantification efforts create too much visibility, this diminishes support. When the pilot study of the research rating exercise performed in a limited number of discipline areas disappeared 'in the drawer,' observers speculated this might have been due to the undesirable transparency that this created about states' university performance. While there was not much to be gained from good performance, widespread under-performance would most likely trigger negative media attention and sour relations between individual states and their universities. Such political sensitivities play out differently in other federal states, such as the United Kingdom, where higher education policies are devolved, but where the vast majority of universities are based in England, where the introduction of various transparency-enhancing regulatory instruments in a top–down fashion has been going on for over two decades now.

One result of the limited enthusiasm for direct state comparison is the tendency of states to disown national quantification processes, by emphasising universities' role as stakeholders in the process:

In the context of the pilot phase, the states did not refer the findings back to themselves....Even to the contrary, if I remember correctly, in the [institutional] process, they even retrospectively said: "We as states actually have no opinion at all about whether we need the research rating or not. We'll hear what the universities have to say about it." (Respondent 8, translation by the author)

Beyond state efforts, comparison of state higher education sectors also emerges from other knowledge brokers. The Funding Atlas published by the German Research Foundation, however focuses throughout on universities and disciplines. This is illustrated by the fact that, seemingly deliberately, not a single table breaks findings down by state (German Research Foundation, 2018). This differs from the Federal Statistics Office, which in its regular publications offers state comparisons on a wide range of indicators, from on-time graduations to the academic staff-to-student ratio or expenditure per student and professor (Destatis, 2018), or the German Rectors' Conference, which compiles data on e.g., the diversity of study programmes on offer and *numerus clausus* courses per state (German Rectors' Conference, 2018).

The findings suggest that while the states endorse diplomacy-like cooperation towards opening up the German university landscape in numbers and indicators, they withdraw when things get too transparent. Beyond federal programmes, the contribution of knowledge creators and brokers such as university platforms, the German Research Foundation, and the Federal Statistics Office to the creation of a quantified competitive field remains limited and removed from state or federal decision-making processes. Perhaps the German states prefer to regard higher education quantification as a Mercedes garage workshop: harmless when left to tooling fans, potentially harmful when the cars are actually taken out for a drive.

5. Conclusion

Regulators and managers in the German higher education system have a 'shiny Mercedes' worth of number-based data infrastructures at their disposal, but rarely couple this data to decision making in the form of quantification-based analytical capacity. In 2009, the German sociologist Schimank stated that "policy learning implies a willingness to learn, [which] is, whether on the professors' side or the state's side, still hardly a given" (Schimank, 2009, p. 136, translation by the author). Ten years later, this picture seems to have shifted, at least as regards state regulators and university managers. Both groups show a broad interest in 'learning from the numbers,' and letting the new insights inform their decision making. However, they do so in a manner that remains tied into institutionalised traditional decision-making arrangements. As a result, quantification-based analytical capacity functions in a largely tempered, hybrid, and untransparent manner that has little to do with algorithmic understandings of direct 'governance by numbers.'

The three different hypotheses discussed in the analysis further clarify the manner in which legal arrangements, administrative prudence, and federal politics constrain certain manifestations of quantification-based decision making. While legal arrangements and the German *Rechtsstaat* squarely protect prerogatives such as academic and institutional autonomy, this in itself does not prohibit regulatory or managerial decision making through quantification. For example, university management typically enters professorial appointment or salary negotiations with a quantified overview of the person's performance and achievements. That said, in many decisional contexts a degree of administrative prudence and consensualism prevents regulators and managers from engaging in forms of quantification that are considered too disruptive, for example by creating undesired transparency or antagonising particular actors. The German states, for example, endorse data-informed cooperation only to the extent that quality differences are not shown too transparently, to avoid stirring undesired competition.

In the limited number of cases where steering is directly indicator-based, decision makers shun data complexity, generally prioritising simplicity, reliability and feasibility over multidimensionality, completeness and precision. A prime example are performance-based budgeting systems, which function on the basis of only a few indicators and whose potential effects are capped in advance. The role played by data infrastructures in in-

forming decision making in the background is harder to capture. Particularly managers inform themselves with a wide variety of data that differ significantly from one university to the next.

It thus emerges that data infrastructures are in fact present in a variety of central decision-making settings. Yet, where quantification is most consequential, it is also most watered down by discussion forums and/or expert evaluations. As such it functions to a large extent under the radar, in the sense that the strength of its influence is not easily demonstrable. To outer appearances, the 'shiny Mercedes' of analytical capacity appears to stay mostly in the garage. In reality, regulators and managers across Germany might take it out for an inconspicuous drive more often than has previously been thought. Further research may help uncover the structural consequences of the growth of quantification in German higher education governance, and the extent to which it has encroached on traditional academic self-governance, while national comparative research could demonstrate whether the German Mercedes experience is unique, or rather similar to that of the Rolls-Royces and Lamborghinis of other higher education systems.

Acknowledgments

Conflict of Interests

The author declares no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Biesenbender, S., & Hornbostel, S. (2016). The research core dataset for the German science system: Developing standards for an integrated management of research information. *Scientometrics*, *108*(1), 401–412.

Bouckaert, G., & Balk, W. (1991). Public productivity measurement: Diseases and cures. *Public Productivity and Management Review*, *15*(2), 229–235.

de Boer, H., Enders, J., & Schimank, U. (2008). Comparing higher education governance systems in four European countries. In N. C. Soguel & P. Jaccard (Eds.), *Governance and performance of education systems* (pp. 35–54). Wiesbaden: Springer.

Destatis. (2018). Hochschulen auf einem Blick [Universities in overview]. *Destatis*. Retrieved from https://www.destatis.de/Migration/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/BroschuereHochschulenBlick.html

Deutschland in Zahlen. (2017). Staatliche Hochschulausgaben je Studierenden: In Euro [State university expenditure per student: In Euros]. *Deutschland in Zahlen*. Retrieved from www.deutschlandinzahlen.de/tab/bundeslaender/bildung/bildungsausgaben/staatliche-hochschulausgaben-je-studierenden

Dobbins, M., & Knill, C. (2017). Higher education governance in France, Germany, and Italy: Change and variation in the impact of transnational soft governance. *Politics and Society*, *36*(1), 67–88.

Espeland, W. N., & Sauder, M. (2016). *Engines of anxiety: Academic rankings, reputation, and accountability*. New York, NY: Sage.

Esping-Andersen, G. (2009). *The three worlds of welfare capitalism*. Cambridge: Polity Press.

Federal Statistical Office. (2018). Studierende Wintersemester 2017/2018 [Enrolled students, winter semester 2017/2018]. *Federal Statistical Office*. Retrieved from www.statistikportal.de/de/studierende

Franzen, M. (2018). Die digitale Transformation der Wissenschaft [The digital transformation of science]. *Beiträge zur Hochschulforschung*, *40*(4), 8–28.

German Rectors' Conference. (2018). Statistische Daten zu Studienangeboten an Hochschulen in Deutschland [Statistical data regarding programme offer at universities in Germany]. *German Rectors' Conference*. Retrieved from https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-03-Studium/02-03-01-Studium-Studienreform/HRK_Statistik_BA_MA_UEbrige_WiSe_2019_20_finale_internet.pdf

German Research Foundation. (2018). *Förderatlas 2018: Kennzahlen zur öffentlich Finanzierten Forschung in Deutschland* [Research Atlas 2018: Indicators of publically funded research in Germany]. Bonn: German Research Foundation. Retrieved from https://www.dfg.de/sites/foerderatlas2018/download/dfg_foerderatlas_2018.pdf

Head, B. (2008). Three lenses of evidence-based policy. *Australian Journal of Public Administration*, *67*(1), 1–11.

Hood, C., James, O., Peters, B. G., & Scott, C. (Eds.). (2004). *Controlling modern government: Variety, commonality and change*. Cheltenham: Edward Elgar Publishing.

Huber, M. (2020). Steered by numbers: How quantification differentiates the reform of a German university. *Politics and Governance*, *8*(2), 26–35.

Huber, M., & Hillebrandt, M. (2019). "Pay for promise" in higher education: The influence of NPM on resource allocation in German universities. *Historical Social Research*, *44*(2), 247–269.

Hüther, O., & Krücken, G. (2013). Hierarchy and power: A

conceptual analysis with particular reference to new public management reforms in German universities. *European Journal of Higher Education*, *3*(4), 307–323.

Kleimann, B. (2016). *Universitätsorganisation und präsidiale Leitung: Führungspraktiken in einer multiplen Hybridenorganisation* [University organisation and rectoral management: Leadership practices in a multi-hybridities organisation]. Wiesbaden: Springer.

Kleimann, B., & Hückstädt, M. (2018). Auswahlkriterien in Berufungsverfahren: Universitäten und Fachhochschulen im Vergleich [Selection criteria in appointment procedures: Universities and politechnic colleges compared]. *Beiträge zur Hochschulforschung*, *40*(2), 20–46.

Krüger, A. K. (2020). Quantification 2.0? Bibliometric infrastructures in academic evaluation. *Politics and Governance*, *8*(2), 58–67.

Kurunmäki, L., Mennicken, A., & Miller, P. (2016). Quantifying, economising, and marketising: Democratising the social sphere? *Sociologie du Travail*, *58*, 390–402.

Lanzendorf, U., & Pasternack, P. (2009). Hochschulpolitik im Ländervergleich [A state comparison of university policies]. In J. Bogumil & R. G. Heinze (Eds.), *Neue Steuerung von Hochschulen: Eine Zwischenbilanz* [The new steering of universities: Taking stock] (pp. 13–28). Berlin: Edition Sigma.

Leibner, T. (2017). University governance and rankings: The ambivalent role of rankings for autonomous accountability and competition. *Beiträge zur Hochschulforschung*, *39*(3/4), 30–51.

Lijphart, A. (2010). *Patterns of democracy*. New Haven, CT: Yale University Press. (Original work published 1984)

Oberschelp, A. (2017). Das Fächerrating des Landes Hessen: Ausgestaltung, Einsatz und Leistungsfähigkeit eines Instruments im Rahmen der Landeshochschulsteuerung [The disciplines rating in the state of Hessen: Design, implementation and performance of an instrument in the context of state university regulation]. *Beiträge zur Hochschulforschung*, *39*(1), 104–129.

Parrado, S. (2014). Analytical capacity. In M. Lodge & K. Wegrich (Eds.), *The problem-solving capacity of the modern state: Governance challenges and administrative capacities* (pp. 86–104). Oxford: Oxford University Press.

Power, M. (1997). *The audit society: Rituals of verification*. Oxford: Oxford University Press.

Prenzel, M., & Lange, S. (2017). Evidenzbasierte Governance von Organisationen in Forschung und Lehre: Erwartungen an die Wissenschafts und Hochschulforschung [Evidence-based governance of organisations in research and teaching: Expectations of science studies and higher education governance]. *Beiträge zur Hochschulforschung*, *39*(1), 10–20.

Ringel, L., Brankovic, J., & Werron, T. (2020). The organizational engine of rankings: Connecting "new" and "old" institutionalism. *Politics and Governance*, *8*(2), 36–47.

Schimank, U. (2005). 'New public management' and the academic profession: Reflections on the German situation. *Minerva*, *43*, 361–376.

Schimank, U. (2009). Governance-Reformen nationaler Hochschulsysteme: Deutschland in internationaler Perspektive [Governance reforms of national higher education systems: Germany in international perspective]. In J. Bogumil & R. G. Heinze (Eds.), *Neue Steuerung von Hochshulen: Eine Zwischenbilanz* [The new steering of universities: Taking stock] (pp. 123–137). Berlin: Edition Sigma.

Schneijderberg, C., & Steinhardt, I. (2018). Accreditation of X qualities instead of quality X: A normative analysis of criteria of the German higher education accreditation regime. *Higher Education Policy*, *32*(1), 5–28.

Seyfried, M., & Pohlenz, P. (2017). Zwischen Wunsch und Wirklichkeit: Qualitätsmanagement als weiches Disziplinierungsinstrument [Between wish and reality: Quality management as soft disciplining instrument]. *Beiträge zur Hochschulforschung*, *39*(3/4), 96–115.

Shore, C., & Wright, S. (2000). Coercive accountability: The rise of audit cultures in higher education. In M. Strathern (Ed.), *Audit cultures: Anthropological studies in accountability, ethics, and the academy* (pp. 57–89). London: Routledge.

Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics, and the academy*. London: Routledge.

United Kingdom Department of Education and Skills. (2004). Guidance letter from Secretary of State for Education and Skills. *United Kingdom Department of Education and Skills*. Retrieved from www.offa.org.uk/about/background

Van Thiel, S., & Leeuw, F. (2002). The performance paradox in the public sector. *Public Performance and Management Review*, *25*(3), 267–281.

Wollmann, H. (1989). Policy analysis in West Germany's federal government: A case of unfinished governmental and administrative modernization? *Governance*, *2*(3), 233–266.

About the Author

Maarten Hillebrandt is Postdoctoral Researcher at the Erik Castrén Institute, University of Helsinki, where he is involved in a research project on the transparency of EU decision making. Previously, he participated in the international comparative ORA-funded research project 'Quantification, Administrative Capacity and Democracy' (2016–2019), in which researchers from the London School of Economics, Paris MinesTech, Bielefeld University, Helmut Schmidt University Hamburg, and Leiden University were co-collaborators. Maarten specialises in EU institutional law and politics, government transparency policy, and quantification.