Torben Schubert
e-mail: torben.schubert@isi.fraunhofer.de

Empirical Observations on New Public Management to Increase Efficiency in Public Research – Boon or Bane?

Abstract: This paper deals with the influence that New Public Management (NPM) mechanisms have on research performance. Using a large dataset of German university research units, it can be shown that both greater internal hierarchy (especially. "strong chancellors") as well as greater operative flexibility for the researchers themselves increase research performance. Some of the variables, including the presence of research councils, have a positive effect on research efficiency only under some definitions of research output. On the other hand, the introduction of resource accounting systems has a negative impact.

1. Introduction

In many western countries, profound changes have been taking place in the organisation of the public science systems (e.g. De Boer et al., 2007, Meyer, 2007, Frolich, 2005, Smith, 2004). Although the countries started from very different positions, the governance schemes converge towards a pattern that is often described as New Public Management (NPM) (Leisyte and Kizniene, 2006). In contrast to the organisational systems which were common in many continental European countries such as Germany, France, and Austria, the spirit of NPM consists of two pillars (Braun and Merrien, 1999, Schimank, 2007a, 2007b). First, the decision-making competencies of the state authorities are reduced, especially at the purely operative level, leaving much greater steering autonomy to the researchers. Second, the internal hierarchy is strengthened, that is to say, the management authorities (the deans and the university chancellors) gain much greater power over the researchers. Many different mechanisms were developed to deploy the new governance structures in practice. Among these are resource controlling, global budgeting, goal agreements, performance-oriented budgeting by indicators, or performance-oriented payment schemes. In any case, though NPM is *en vogue* among European policy-makers, to my knowledge, there is no empirical proof that the new governance system indeed results in efficiency gains. Indeed, its effectiveness is sometimes implicitly doubted, because it is argued that research is not a routine task and the most empowering setting is that of academic freedom. Anyhow, the latter claim has not been proven either.

Using a unique dataset of a large sample of German research units gathered during a project on "international competitiveness and innovation capacity of universities and research institutions – new forms of governance" funded by the German Research Association (Deutsche Forschungsgemeinschaft, DFG), we will present concise results which point to positive effects of the typical NPM mechanisms on research performance.

2. New Public Management – It is Not Just Hot Air

As already mentioned, NPM is a fashionable concept, especially in politics. Therefore, although it has a large number of academic proponents, it has been criticised that it is merely one more slogan voiced by governmental authorities to organise a silent phasing out of state funding of public research. On the contrary, the proponents argue that the most successful science nations – the UK and USA – have adopted such a system and therefore regard it is beneficial. Despite this heated atmosphere, it is not only legitimate but also necessary to ask what the essence of NPM actually is, and where the postulated efficiency gains should accrue from.

In fact NPM is more than just a slogan. Its theoretical roots date back to two distinct branches of economic theory. The first is that of property rights (see e.g Demsetz, 1967, Buchanan, 1984) which postulates that the separation of property rights (in this context the societal resources devoted to research) and control (in this context the resources controlled by the researcher) leads

to efficiency losses. The second is contract theory (see e.g. Jensen and Meckling, 1976, Holstrom 1979 as well as Bolton and Dewatripont, 2005 as a recent overview), which states that if the principal (in this context the society or the state authorities) cannot observe the effort level of the agent (in this context the researchers), then the agent tends to work too little, where the latter would argue that possible failure has to be accrued to bad luck and not to laziness. This is known as moral hazard. Bearing both lines of argumentation in mind, it is easy to understand the spirit of NPM.

In any case, it is helpful to contrast the pre-NPM system in Germany and New Public Management, because in a certain sense, both are inverted mirror images. This can be seen most easily, by using the 'equalizer'-model of Schimank (2007b). He argues that, in essence, there are 5 governance dimensions, whose specific importance can describe each university system. Therefore, each constellation of the 5 'sliding controllers' ("*Schiebregler*") describes a different university system. Any of these constellations can be thought of as a specific setting on an equaliser, giving a particular tune. The first of the 'sliding controllers' is the level of state regulation ("*staatliche Regulierung*"). This mechanism determines the strength of the governmental influence. The second concerns external control through governmental or societal stakeholders ("*Außensteuerung…durch staatliche oder gesellschaftliche Stakeholder*"), often in the form of research councils (not to be confused with funding councils in the UK), which are similar to a directorate in a company (Mayntz, 2002). The third is academic self-control ("*Akademische Selbststeuerung*") which measures the degree to which the chair holders can decide autonomously. The fourth is internal hierarchical self-control ("*hierarchische Selbststeuerung der Hochschulen*") mapping decision competencies of the deans and chancellors inside the university. The fifth is market control ("*Marktsteuerung*") often induced by the increasing need to acquire external funds.

The old German public science system is characterised by large competencies of the state authorities in regulation of operative management decisions as well as high competencies of the individual chair holders in setting their research and teaching agenda. Cynically, one could remark that this setting may be described as a mixture of command economy and provincial principalities not unlike the state organisation in USSR before 1991. All other governance mechanisms did not play a major role.

NPM, briefly, is exactly the opposite: high degree of control by external stakeholders, of internal hierarchical control, and of market control in conjunction with a low degree of state and chair holder competencies.

By introducing NPM in German universities, which is still an ongoing process, the foremost aim of the state authorities was to increase efficiency in research and teaching.

Four potential sources of efficiency gains from NPM may be identified, which will also guide our empirical design. The first, and most obvious, is to increase operative flexibility of the universities. In theory, this should make resource allocation more efficient, because the decision process is made faster and more problem-oriented (subsidiarity). However, increasing operative

flexibility ceteris paribus reduces accountability, therefore resources may be used in a manner not in accordance with societal goals. Thus, NPM aims at increasing accountability by strengthening internal hierarchical elements, most prominently, the influence of the deans, the chancellors, and the presidents. This second source of efficiency gains might result from a decreased danger of moral hazard (see above). The third is to give the decision-making units feedback and information on their actions. Often this can be achieved by the introduction of internal accounting models. The fourth is to enhance competencies concerning strategic decisions. This once again can be achieved by greater internal hierarchy but also by the newly installed research councils which shall guide, advise, and also decide on items concerning the university as a whole. This might contribute to research efficiency as it sets out an overall agenda for the departments and the university as a whole.

As mentioned above, the goal of this paper is to test for positive impact on research efficiency of university chairs coming from any of these sources. In the next section we will briefly describe the dataset.

3. Empirical Discussion

3.1 Dataset

During February and March 2007 an online survey of German research units at the micro-level was performed. In a lengthy process we were able to identify 1,908 university chairs and corresponding extra-university units from the disciplinary fields of astrophysics, nanotechnology, biotechnology, and economics. Choosing this particular set of disciplines was inspired by a most dissimilar case design (for further discussion see Schmoch and Schubert, 2008). The first divide is along the dimension of natural sciences (astrophysics, bio- and nanotechnology) vs. social sciences (economics). However, the second divide is along the dimension of basic (astrophysics, economics) vs. applied research (bio-, nanotechnology).

Until March we received 473 valid answers (astrophysics: 34, nanotechnology: 201, biotechnology: 136, economics: 102), which implies a participation rate of almost 25%. 140 of the answering units were (public) extra-university belonging to institutions such as the Max Planck or the Fraunhofer Society. The remaining 333 were from universities.

In this paper we will focus on the latter group, mainly for two reasons. First, extra-university units did not have to cope with state authority interference in the past to the same extent as the universities. Therefore, they already had much greater steering autonomy in the past. The very recent policy changes, which are at the focus of this paper, are directed primarily towards the university units. Thus if New Public Management shall be linked to research efficiency, it seems more adequate to restrict the analysis to universities. Second, the extra-university public science sector in Germany is highly fragmented. This heterogeneity might be hard to deal with econometrically.

In this online survey, questions were asked concerning inputs like number of scientists, quality of capital equipment, etc. as well as concerning outputs including publications[1], editorships, cooperation with companies, etc. In addition, questions concerning the governance structures were asked, including the existence of rigid personnel quotas or existence of management accounting systems. Also, we were interested in the perceived power of the deans or the chancellors. By this design we were able to link relationships between inputs and outputs (more specifically efficiency) to governance structures, which is the foremost objective of this paper.

Table 1: Description of the Variables and Summary Statistics

Variable	Shorthand (if used)	Time Period	Unit/Type	Mean	S.D.	Min	Max
Research Outputs							
Fraction of Time Spent on Third Party Research		2004-2005	Percent	69.16	28.91	0	100
Advisory Service for Companies		2004-2005	Count	0.56	1.44	0	11
Co-operations with Companies		2004-2005	Count	2.05	3.46	0	26
Conferred Doctoral Titles		2004-2005	Count	4.25	4.86	0	52
Conferred State Doctoral Titles		2004-2005	Count	0.50	1.07	0	13
Number of Publications in The SCI/SSCI-Database		2004-2006	Count	31.47	40.78	0	320
Number of Citations in The SCI/SSCI-Database		2004-2006	Count	118.82	177.73	0	1359
Research Inputs							
Number of Scientists excluding PhD Students		2005	Count	6.68	8.48	1	77
Number of Scientists including PhD Students		2005	Count	14.47	14.18	1	129
Age of Computers when Replaced		2005	Count	4.54	1.37	2	10
NPM Governance Variables							
Existence of Personnel Quotas	PERSONNEL	2006	Binary	0.80	0.40	0	1
Perceived de facto Influence of the Deans	DEANS	2006	1-5 Likert Scale	3.20	1.02	1	5
Perceived de facto Influence of the Presidents	PRESIDENTS	2006	1-5 Likert Scale	3.65	0.97	1	5
Existence of Goal Agreements	GOAL	2006	Binary	0.30	0.46	0	1
Existence of an Accounting Scheme	ACCOUNT	2006	Binary	0.53	0.50	0	1
Existence of Research Councils	COUNCIL	2006	Binary	0.70	0.46	0	1
Existence of Regular Evaluations	EVAL	2006	Binary	0.39	0.49	0	1

The variables used throughout this paper are described in Table 1. Summary statistics are given. In fact, these are only a snapshot of the 64 variables asked for in total. Without going into detailed analysis, we see that there is great variety among the research groups. For example, the smallest consists of just one person while the largest has a size of 129 scientists including PhD students. The maximum number of publications of the single research institute is 320 and the minimum number is 0. Some groups do not conduct any third party research, while others have a share of 100%. Some interesting facts can be seen from the NPM variables. About 80% of the research groups still have to face rigid personnel quotas. About 40% are regularly evaluated. 53% control their resource movements via some form of accounting model. 70% have research councils at their universities. In any case, important for our purpose is the fact that the variety does note solely relate to structural figures describing the research unit itself, but also to the governance structures they operate under, i.e. while some research units are still subject to the

[1] Apart from the self-reported numbers of publications, for almost all units we determined the persons belonging to the units. So we were able to use bibliometric indicators coming from the Web of Science like SCI/SSCI publications as well as citation rates and international co-publications. In the following we use the bibliometric indicators from the ISI database instead of the self-reported results.

traditional German governance scheme, others operate under almost purely NPM structures. Some will be located in transitional systems. This heterogeneity allows us to estimate the effect of different NPM tools on research performance via cross section data. In the next section we describe the testing methodology adopted, as it guides some of the preliminary discussions.

3.2 Methodology

In order to test for positive effects of NPM on research performance, an estimate or an indicator of research performance must be obtained. This is not as trivial as it may seem. Many authors recognise, at least implicitly, that research performance is multidimensional (e.g. Rousseau and Rousseau, 1997, Nagpaul and Roy, 2003, Warning, 2004, Johnes, 2006). Also, it is known from prior work that research units specialise in very diverse activities (Laredo and Mustar, 2000, Jansen et al., 2007). Therefore using a single indicator (such as publication counts) as a proxy for research performance will not prove useful, because the results will be distorted in favour of those units specialised in the activity covered by the specific indicator use. Instead we calculated FDH scores[2] as a measure of efficiency for different sets of inputs and outputs, in a first step, and tried to test for partial correlations of these scores with NPM variables in a second stage regression.

The FDH estimator (compare Deprins et al., 1984) is similar to DEA estimator (Farrel, 1957, Charnes et al. 1978) of technological efficiency except that convexity of the production set is not imposed. Because we do not know if the production frontiers in research exhibit non-increasing returns to scale, I regard it as safer to use the more flexible FDH efficiencies, which are consistent under any shape of the production frontier.

More formally, letting Ψ be the set of feasible input-output relations, then the FDH estimator of this production set is defined as

$$\widehat{\Psi}_{FDH} = \left\{(x,y) \in \mathbb{R}_{+}^{p+q} \mid y < y_i, x > x_i, \quad (x_i, y_i) \in \chi\right\} \tag{1}$$

where x is p-dimensional input-vector, y is a q-dimensional output-vector and χ is the set of observations. Letting $D_0 = \{i \mid (x_i, y_i) \in \chi, x_i \le x_0, y_i \ge y_0\}$ the set of indices of units dominating unit 0, then it can be shown that the input-oriented FDH efficiency estimator can be calculated as[3]

---

2 The FDH scores for each research unit were calculated with reference only to the own disciplinary peers to account for field-specific differences in the production technology.

3 Note that this is the Farrel-type definition. The software package FEAR for R-statistical einvrionment written by Paul W. Wilson, which I used, claculates the Shephard distance function (compare Shephard, 1970), which is the reciprocal.

$$\hat{\vartheta}(x_0, y_0) = \inf\left\{\vartheta \mid (\vartheta x_0, y_0) \in \widehat{\Psi}_{FDH}\right\} = \min_{i \in D_0} \max_{j=1,\ldots,p}\left(\frac{x_i^j}{x_0^j}\right) \tag{2}$$

Though the two-stage approach of regressing explanatory variables on previously estimated efficiency scores (no matter if DEA or FDH) is very common and intuitive, Simar and Wilson (2007) concisely show that it has been misused in practically all prior treatments, for several reasons. First, the usual asymptotic inference in the second-stage regression is flawed, because the FDH estimates are serially correlated and so are the regression residuals. Second, the very common Tobit specification of the second stage regression, which is motivated by the (estimated) high share of absolutely efficient units, is conceptually misguided, because it does not reflect the properties of the underlying data generating process, which has – by assumption – a continuous distribution (instead of mass-points) at the frontier. Using a Tobit specification rather reflects finite sample problems of the FDH estimator than the characteristics of the density close to the frontier itself. Simar and Wilson (ibid.) argue that the correct model is a truncated regression model. Though the estimates from this model converge at the same low rate as the FDH estimator itself (especially the convergence is not the order of $n^{-1/2}$, which would be that of usual parametric estimation), they are at least consistent. To obtain valid inference they propose complicated bootstrap algorithms in the second-stage truncated regression. Without going into more details, I applied algorithm #1 described in their article. For any of the regressions performed 1,000 bootstrap replications were used. As far as I am aware of the literature, this is the first paper employing the methodology proposed by them.

3.3 Definition of Research Output and Input

As research output is sometimes proxied only by bibliometric indicators, one objective of this survey was to be very broad in the definition of what research output actually is.

All in all, we collected 12 output measures, where 5 belonged to the knowledge and reputation generating section (publications, citations, conference articles, international co-publications, professorial job offers), 3 belonged to interaction with business and governmental bodies (advisory services for companies, cooperation with companies, membership in advisory boards), and the remaining 4 belonged in the maintenance dimension (number of doctoral titles, number of state doctoral theses, editorships, and scholarships).

In prior work (Jansen et al., 2007) we found four types of units with distinct output profiles concerning research, that is to say, research units specialise in certain activities by choosing specific output bundles. The typical units found are those that publish many papers, those that write not as many but highly cited papers, those that engage in graduate teaching, and those that engage in transfer activities such as cooperation with companies or memberships in scientific

advisory boards. We therefore have some indication which output indicators should be included in order to measure output.

However, because universities are pushed to develop unique agendas and profiles (Enders, 2001) this balanced scheme may not provide the only reasonable output definitions. Some university profiles may emphasise graduate teaching while giving not as much weight to publications. Others, in turn, may have a focus on technology transfer to companies.

Because the influence of NPM mechanisms may differ as the notion of output changes, I defined four different output bundles, which are, to a certain extent, inspired by our prior work (Jansen et al., ibid.) and our prior expectation (reputation, transfer, and maintenance). The one linked to that most directly is what we called the balanced set, because for each dimension identified (see above) simply one corresponding variable (or one composite of variable) was picked to cover this dimension. The other sets are derived from this balanced definition simply by including one additional variable from one specific dimension to replace that of another. By this procedure the balanced scheme is biased into a specific direction (e.g. graduate teaching).

Despite the fact that one of our main results (Jansen et al., ibid) was that third party funds are detrimental to publication and basic research activities under certain conditions, this variable was included in the remaining three output sets. This was done not because we believe that it is a good output indicator, but because many university managers, and especially the state authorities, believe that it is. Therefore it is used widely in Germany (Orr et al., 2007, Jäger, et al., 2005, Jäger, 2006 Lescenzsky and Orr, 2004).

a) The balanced scheme

The balanced output set includes the variables publications, fraction of time spent on third party research as a proxy for third party funds, number of advisory services for companies plus number of cooperations with companies, and conferred doctoral plus state doctoral degrees. The balanced scheme is defined as a set of output indicators consisting of the dimensions highlighted in the following three sets without giving special weight to any.

b) The transfer-oriented scheme

The variables included are fraction of time spent on third party research as a proxy for third party funds, number advisory services for companies and number of cooperations with companies as separate dimensions, and conferred doctoral plus state doctoral degrees. This output definition highlights the task of technology transfer to companies.

c) The graduate-teaching-oriented scheme

The variables included are fraction of time spent on third party research as a proxy for third party funds, conferred doctoral and state doctoral degrees as separate dimensions, and number of advisory services for companies plus number of cooperations with companies. By using this set of output indicators, the focus is set on the task of education and qualification.

d) The publication-oriented scheme

The variables included are number of publications, number of citations per publication as a measure for impact, number of advisory services for companies plus number of cooperations with companies, and conferred doctoral plus state doctoral degrees. This output definition is dominated by the task of conducting basic research as measured by bibliometric indicators.

We did not make a distinction concerning inputs. These were defined to be the number of scientists (excluding doctoral or PhD students) and the inverse of the average age of the computers when replaced in that unit as a measure of the quality of capital stock. Unfortunately, the monetary value of the capital stock proved to be largely unknown to the research group leaders. This was the best still objective measure we could think of.

3.4 Ex ante Hypotheses

Referring to the discussion in section 2, four sources of potential efficiency gains were identified. In this section we will revisit this theoretical discussion and operationalise the dimensions with variables from the survey. Additionally, we give the ex ante expectation on the estimated coefficient resulting from the notion of NPM theory.

a) Increasing Operative Flexibility

I use a single variable for this dimension, namely the existence of strict personnel quotas (PERSONNEL – 0 if false and 1 if true). Binding personnel quotas reduce the flexibility of the research units. Bearing in mind that higher input-oriented FDH scores imply lower efficiency, NPM would predict a positive coefficient for the regression model.

In fact, other questions concerning operative flexibility, such as fraction of resources that can be spent freely, were asked. However, it turned out that few research units were familiar with this data. Therefore the item non-response was high and the quality of the answers may be doubted. Therefore, I decided not include these variables.

b) Increasing Short- to Mid-Term Accountability

For this dimension I use three variables. The first two are perceived strength of the de facto exerted influence of the deans (DEANS, 1-5 Likert scale, 1 very low and 5 very high) and perceived strength of the de facto exerted influence of the chancellors and presidents

(PRESIDENTS, 1-5 Likert scale, 1 very low and 5 very high). As NPM says that increasing the power of those in charge implies higher efficiency because losses due to moral hazard are reduced, I expect both coefficients to be negative.

The third concerns existence of goal agreements (GOAL, 0 if false and 1 if true). Goal agreements can be thought of as a treaty between the research unit and usually the deans, in which both parties agree on some objective to be accomplished over some time horizon. Since this is a tool to steer the activities of the research units, once again NPM predicts a positive impact and the coefficient should be negative.

c) Resource Control and Feedback

I use one variable to cover this field, which is the existence of an accounting system monitoring the resource movements of the respective research unit (ACCOUNT, 0 if false and 1 if true). The coefficient is expected to be negative.

d) Increasing Long-Term Strategic Capabilities

Two variables are used to cover this field. The first is the existence of a research council at the university of the research unit (COUNCIL, 0 if false and 1 if true). The second is if the research unit is evaluated regularly (EVAL, 0 if false and 1 if true). EVAL is included in the fourth block because evaluations tend to aim at long-term strategic orientation of the research units. Therefore research councils set an agenda for departments and universities and evaluations – at best – pass on these decisions to the research units. Following NPM both mechanisms should increase research efficiency and the coefficients should be negative.

3.5 Estimation Results

As noted in section 3.2, to test the influence of NPM on research efficiency, a semi-parametric approach was employed where FDH scores are calculated in a first step and in a second step, the NPM variables described in the previous section are regressed on the FDH efficiency scores. As noted in section 3.3, different sets of output variables were used to determine if the influence of NPM fluctuates as the definition of output changes. The results of this estimation can be found in the following table.

Table 2: Influence of NPM Variables on Research Efficiency – Truncated Normal Regression Model

Dependent Variable: Input oriented FDH score				
	balanced	transfer-oriented	teaching-oriented	publication-oriented
CONST	0.8984**	1.9881**	1.4934**	1.1833***
Increasing Operative Flexibility				
PERSONNEL	0.7041***	0.2427	0.4385***	0.5024***
Increasing Accountability				
DEANS	0.0738	-0.0589	-0.0625	-0.0192
PRESIDENTS	-0.2318***	-0.1451*	-0.0781	-0.0917*
GOAL	0.1151	-0.4793***	-0.1790	0.0092
Resource Control and Feedback				
ACCOUNT	0.3557***	0.3330***	0.2208*	0.2768***
Increasing Strategic Capabilities				
COUNCIL	-0.1730	-0.4352***	-0.2473**	-0.0804
EVAL	-0.1451	-0.1156	-0.0110	-0.3461***
n	266	243	243	243

*** significant at 1%-level, ** significant at 5%-level, * significant at 10%-level

We did not include field dummies to account for level effects in efficiency, because we calculated the FDH score for each unit with reference only to the units of the same discipline. Since FDH scores are "normalised" (i.e. one or above) discipline specific level effects should largely be absent.

In the following, we adopt the strategy of offering explanations for the estimation results explicitly only for those that do not follow directly from the discussion on NPM in section 3.4. After each passage we summarise the main conclusions in catch phrases.

We start with some very general observations. It may indeed be ascertained that the influence of the different NPM governance mechanisms differs as the output bundle changes. Some variables (especially COUNCIL, EVAL, and GOAL) are significant only for some output bundles. Some variables (especially PERSONNEL, PRESIDENTS, and ACCOUNT) are significant over (almost) all output bundles, while the variable DEANS is not significant in any. Interesting with the last observation is that the presidents and chancellors in fact have a profoundly positive impact on research efficiency, while we cannot detect any such effect for the deans. Two explanations seem reasonable. First, the deans are elected by those who they ought to govern and after their election period ends, they return to the subgroup of people who are governed. This may lead to situations where the deans remain rather inactive by setting weak incentive schemes (Wigger and Dehm, 2006). Closely related is the more sociological argument that the deans are socialised within the science system and therefore, by ethos, will not push the subordinate chair

holders too hard (de Boer et al., 2006). The presidents, on the contrary, may be far enough away from the units at the basis to take unpopular but also necessary actions.

Conclusion 1: most NPM variables tested have a profound impact on research efficiency. Some of them, especially the existence of rigid personnel quotas, the influence of presidents and chancellors, and the existence of an accounting model have an impact under almost all output definitions.

After a more detailed look at the results, we can state that the variables COUNCIL, EVAL, GOALS, PERSONNEL, and PRESIDENTS, when significant, have the sign predicted by NPM, while the introduction of an accounting tool (ACCOUNT) curiously has a negative impact on research efficiency. In fact, latter result seems counterintuitive. However, three explanations may seem reasonable. First, the additional effort of maintaining an accounting model simply outweighs the positive effect from more reliable information and feedback on management decisions, especially if the research unit is small. Second, many leaders of research units may not be familiar with accounting models and therefore tend to ignore the data from such models. However, if the second explanation is true then, over time, the usefulness of accounting data may grow as the chair holders get accustomed with it. Third, accounting systems cannot simply be carried over to the needs of research units, especially because the cost and activity accounting has to be reduced to a pure cost accounting, since the results from research activities are either hard to price in the absence of a market or are completely intangible (for a discussion see Ambrosy et al., 2007). It might be true that the frictional losses from an adoption of these models originaly developed for companies are too high to make them a useful tool. In any case, we must be careful in interpreting this result. As seen in Table 1, some of the research units in the sample are very large. 6 units were larger than 35 scientists (excluding doctoral students) and many of these had an accounting system. We are quite sure that it would be very unwise for the units to stop controlling resources for obvious reasons. The estimation results are likely to come about because the sample is dominated by small units, where the need for bookkeeping might be limited.

Conclusion 2: except for the introduction of accounting models, which affects efficiency negatively under any output definition, all other NPM variables contribute positively to research efficiency. Strong presidents and chancellors as well no- existence of personnel quotas have a positive impact under three out of four output definitions.

Turning to the variables which are significant only under some output definitions (COUNCIL, EVAL, GOAL), some interesting patterns emerge. First, evaluations affect efficiency only for the publication-oriented scheme. Once again, some ad hoc explanations may be found. On the one hand, it may be possible that evaluations set a special focus on the assessment of research activities, i.e. if they focus on, say teaching, they would affect teaching positively. On the other hand, evaluations tend to have a long-term horizon by creating incentives for setting out a long-erm agenda. To the extent, that publication-related activities cannot be changed from one day to another, this might simply make them a very appropriate tool to foster basic research. Second, goal agreements are significant only for transfer-dominated activities. This is pretty obvious

because goal agreements tend not to be related to publication activities, but instead are tailored towards more routine tasks, such as cooperation with companies. However, the expectation was that this variable would be significant for teaching as well. Third, the existence of research councils has a significant impact on transfer and graduate teaching. This, on the contrary, meets the expectations completely, because major players in the councils are state representatives. Research councils are a modern channel for state authorities to exert their former direct influence. In any case, it is known from the German policy debate that the responsible authorities have a high preference both for education, and therefore graduate teaching, as well as short- to mid-term economic goals, and in following that, for knowledge transfer to companies (Hassink, 1996, compare also Schimank, 1988). The argument implied then is that the estimation results simply reflect the preferences of the research councils.

Conclusion 3: publication-related activities are made more efficient by regular evaluations of the research units. Efficiency in technology transfer tends to be increased both by goal agreements and the existence of research councils, while efficiency in teaching activities is increased by the existence of research councils.

4. Conclusion

It was demonstrated that many NPM instruments have a positive influence on research efficiency. In any case, this depends on the definition of what research activities actually comprise. Evaluations are a reasonable mechanism to enhance publishing activities. Goal agreements and the existence of research councils push the research units in the direction of increased transfer and teaching efficiency. High competencies of the chancellors and presidents as well as the non-existence of strict personnel quotas contribute positively almost every where. We conclude that NPM may exert considerable positive effects on university research if not employed blindly.

Acknowledgement

The research underlying this paper was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) in a project on "performance indicators for research groups" (SCHM 1719/1-2) which is part of a larger research group on "international competitiveness and innovation capacity of universities and research institutions – new forms of governance" (FOR 517).

References

Ambrosy, R.,Heise, S., Kirchhoff-Kestel, S. & Müller-Böling, D. (1997). Integrierte Kostenrechnung: Unterwegs in Richtung zu einem modernen Hochschulmanagement. *Wissenschaftsmanagement*, 4, 204–213.

Bolton, P. & Dewatripont, M. (2005). *Contract Theory*. (The MIT Press, Cambridge, Massachusetts)

Braun, D. & Merrien, F.-X. (Eds.) (1999). *Towards a New Model of Governance for Universities? A Comparative View*, (London)

Buchanan, J. (1984). *Die Grenzen der Freiheit*. (Mohr Siebeck, Tübingen)

Charnes, A., Cooper, W. & Rhodes, E. (1978). Measuring the Inefficiency of Decision Making Units. *European Journal of Operational Research*, 2, 429–444.

de Boer, H.F., Enders, J. & Schimank, U. (2006). On the Way towards New Public Management? The Governance of University Systems in England, the Netherlands, Austria, and Germany, internal communication

de Boer, H.F., Enders, J. & Leisyte, L. (2007). Public Sector Reform in Dutch Higher Education: The organizational Transformation of the University, *Public Administration*, 85 (1), 27–46

Demsetz, H. (1967). Toward a Theory of Property Rights. *American Economic Review*, 57,347–359

Enders, J. (2001). A Chair System in Transition: Appointments, Promotions, and Gate-keeping in German Higher Education. *Higher Education*, 41, 3–25

Farrell, M. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society: Series A*, 120, 253–281.

Frolich, N. (2005): Implementation of new Public Management in Norwegian Universities, *European Journal of Education,* 40 , 223-234

Larédo, P., & Philippe Mustar, (2000). Laboratory activity profiles: An exploratory approach. *Scientometrics*, 47, 515-539.

Leisyte, L. & Kizniene, D. (2006). New Public Management in Lithuania's Higher Education, *Higher Education Policy*, 19, 377-396

Leszczensky, M. & Orr, D. (2004). Staatliche Hochschulfinanzierung durch Indikatorgestützte Mittelverteilung. *Hochschulinformationssystem Working Paper*

Hassink, R. (1996). Technology transfer agencies and regional economic development, *European Planning Studies*, 4, 167-184

Holmstrom, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics*, 10, 74–91

Jansen, D., Wald, A., Franke, K., Schmoch, U. & Schubert, T., (2007): Third Party Research Funding and Performance in Research. On the Effects of Institutional Conditions on Research Performance of Teams, *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59, 125-149

Jäger, M., Leszczensky, M., Orr, D. & Schwarzenberger, A. (2005). Formelgebundene Mittelvergabe und Zielvereinbarungen als Instrumente der Budgetierungan Deutschen Universitäten. *Hochschulinformationssystem Working Paper*

Jäger, M. (2006a). Leistungsbezogene Budgetierung an Deutschen Universitäten. *Wissenschaftsmanagement*, 3, 30–36.

Jensen, M. & Meckling, W. (1976). Theory of the Firm: Managerial Behaviour, Agency Costs, and Ownership Structure. *Journal of Financial Economics*, 3, 305–360.

Mayntz, R. (2002). University Councils: An Institutional Innovation in German Universities. *European Journal of Education*, 37, 21–28.

Meyer, L.H. (2007). Collegial participation in university governance: A case study of institutional change, *Studies in Higher Education*, 32, 225-235

Nagpaul, P.S., Roy, S. (2003): Constructing a multi-objective measure of research performance, *Scientometrics*, 56, 383-402

Orr, D., Jäger, M. & Schwarzenberger, A. (2007): Performance-based funding as an instrument of competition in German higher education, *Journal of Higher Education Policy and Management*, 29, 3-23

Schimank, U. (1988): The contribution of University research to the technological innovation of the German economy: Societal auto-dynamic and political guidance, *Research Policy*, 17, 329-340

Schimank, U. (2007a): Die Governance-Perspektive: Analytisches Potential und anstehende konzeptionelle Fragen, (In T. Brüsemeister et al. (Eds.), *Governance, Schule, Politik. Zwischen Antagonismus und Kooperation*, Wiesbaden), (in press)

Schimank, U. (2007b): Ökonomisierung der Hochschulen – Eine Makro-Meso-Mikro-Perspektive, internal communication

Shephard, R.W. (1970). *Theory of Cost and Production Function*, (Princeton University Press, Princeton, New Jersey)

Simar, L. & Wilson, P.W. (2007): Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, 136, 31-64

Smith, D. (2004): A Decade of Doing Things Differently: Universities and Public-sector Reform in Manitoba, *Canadian Public Administration*, 47, 280-303

Schmoch, U. & Schubert, T. (2008). Are International Co-publications an Indicator for Quality of Scientific Research? *Scientometrics*, 74, forthcoming

Rousseau, S. & Rousseau, R. (1997). Data Envelopment Analysis as a tool for constructing scientometric indicators, *Scientometrics*, 40 1, 45-56

Warning, J. (2004). Performance Differences in German Higher Education: Empirical Analysis of Strategic Groups, *Review of Industrial Organization*, 24, 393-408

Wigger, B.U. & Dehm, R. (2006). A Principal-Agent Approach to a Selfadministered Organization with an Elected Principal. *unpublished*