



No. 12-2012

Ivo Bischoff and Frédéric Blaeschke

**Window-Dressing and Lobbying in Performance-
Budgeting: a Model for the Public Sector**

This paper can be downloaded from
<https://www.uni-marburg.de/fb02/makro/forschung/magkspapers>

Coordination: Bernd Hayo • Philipps-University Marburg
Faculty of Business Administration and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Window-Dressing and Lobbying in Performance-Budgeting: a Model for the Public Sector

Ivo Bischoff* and Frédéric Blaeschke[†]

This version: March 1, 2012

Abstract

Performance budgeting schemes in the public sector have to operate with imperfect performance measures. We argue that these imperfections lead to wasteful fund-seeking (window dressing and lobbying) by the administrative units that produce public services. We develop a game-theoretical model to analyse the trade-off between the productivity-enhancing effect of performance budgeting and the social waste it induces. The optimal performance-budgeting scheme crucially depends on the objective functions of administrative units, the available performance signal and the welfare function used. We compare a performance signal based on units' effort to a signal based on their output and show that the former evokes more social waste while the latter amplifies regional inequality. Forgone welfare gains or even welfare losses arise when the government is opportunistic. Our model and its major conclusions apply to a large array of publicly installed contests such as programs of international organisations like the IMF and conditional grant schemes in federalist countries.

Key words: Performance budgeting, rent-seeking, bureaucracy, public-sector efficiency, conditional grants, opportunistic government

JEL: D 7, H 77, H 5, H 11

* Department of Economics, University of Kassel, Nora-Platiel-Strasse 4, 34109 Kassel, E-mail: bischoff@wirtschaft.uni-kassel.de, Tel. ++49 561 8043033, Fax. ++49 561 8042818.

[†] (corresponding author) Department of Economics, University of Kassel, Nora-Platiel-Strasse 4, 34109 Kassel, E-mail: blaeschke@uni-kassel.de, Tel. ++49 561 8043034, Fax. ++49 561 8042818.

1. Introduction

The notion that public sector production is inefficient is widely accepted in the economics and public administration literature. The underlying cause is the asymmetric distribution of information between policy-makers who choose the level of public services and allocate funds and the administrative units that produce these services. Important information, e.g., about the production function for public services is available to the administrative units but not to the policy makers. Administrative units make use of this asymmetry and pursue their own goals rather than public interest (e.g., de Groot, 1988; Moe, 1997, 2006). Since the 1990s, many countries introduced performance budgeting (hereafter PB) to reduce this asymmetry in information and improve public sector efficiency (e.g., OECD, 2007; Robinson, 2007). Under PB, the subordinate units have to provide policy-makers with information about their performance. Next to a number of numerical performance-indicators, the units prepare reports that give a comprehensive account of their activities and performance (e.g., Jordan and Hackbart, 1999; Kelly and Rivenbark, 2003: ch. 5-7; Robinson, 2007a; Lockwood and Porcelli, 2010). This information is expected to improve public sector efficiency because it allows better informed choices. More importantly, policy-makers can allocate funds based on performance, thereby inducing competition among administrative units and making them improve performance (e.g., Robinson, 2007b). At the same time, it is widely acknowledged that the available numerical performance indicators are imperfect in that they capture only part of benefits generated by public services (e.g., Joyce, 1993; Crain and O’Roark, 2004; Robinson, 2007a; Jones and McCaffery, 2010). Therefore, an incentive-compatible PB-scheme cannot rely on numerical indicators only. The largest part of the existing PB-schemes uses additional information from performance reports and other sources (e.g., Kelly and Rivenbark, 2003: ch. 5-7; Gilmour and Lewis, 2006; OECD, 2007: ch. 3).

While the use of this additional information is likely to produce a more comprehensive view of the unit's performance, it raises the administrative costs of PB-schemes: Especially the preparation of performance reports ties up substantial amounts of resources. The resources used there are missing in public service production. More importantly, the use of additional information causes a serious new problem: The rich informational content of performance indicators and especially of performance reports cannot be appreciated in simple allocation rules. Its appreciation requires a non-transparent and complex budget-allocation process in which policy-makers allocate funds in discretionary and often difficult or even arbitrary decisions (e.g., Moynihan, 2005). This in turn provides incentives for the administrative units to engage in socially wasteful fund-seeking: First, they can try to attract funds by sugarcoating the performance report or directing efforts to highly visible measures that make the overall performance look more favorable than it is. These activities will hereafter be called window-dressing.¹ Second, they may lobby policy-makers to favor them in budgetary decisions. Both, window-dressing and lobbying aim at redirecting rents but do not yield any net benefit for the public. Thus, like the lobbying activities of firms and interest groups in the theory of rent-seeking (e.g., Tullock, 1980; Congleton, 2008), resources spent on window-dressing and lobbying represent social waste. Unlike in the rent-seeking models, however, not all activities that potentially attract funds are wasteful: Any effort that increases true performance may attract funds and increase overall welfare at the same time. In other words, there is a trade-off in PB: Distributing funds based on performance indicators and reports increases performance but also increases social waste from window-dressing and lobbying.

Contest theory (e.g., Clark and Riis, 1998) provides the theoretical framework to analyze the benefits of different PB-schemes while rent-seeking theory (e.g., Tullock 1980; Congleton, 2008) can be used to assess the social waste associated with different PB-schemes. So far, the literature lacks a formal model that accounts for both benefits and social waste simultaneously. In this paper, we provide the first theoretical model that analyses this trade-

off.ⁱⁱ The model shows how benefits and social waste depend on the production function for public services, the performance signal and the objective functions of both policy makers and administrative units. The units' performance and fund-seeking effort as well as the policy makers' choice of PB-scheme are explained endogenously in the model. We focus on two questions: First, under which conditions can PB yield net benefits in welfare? Second, if PB is potentially welfare-enhancing, what is the optimal share of performance-based funds?

Our main results can be summarized as follows: First, we compare PB schemes in which every sub-unit receives a performance bonus varying with relative performance to schemes based on a contest for a limited number of "prizes" awarded based on relative performance. We show that they are equivalent with respect to their impact on the units' behavior and overall output. Second, we show that the optimal PB-scheme differs fundamentally for different performance signals. We compare a performance signal driven by the units' effort to produce efficiently (effort-based signal) to a signal that depends on the output that units produce using the basic lump-sum funding (output-based signal). Under the output-related signal, the level of wasteful window-dressing is much lower. It is even possible to resolve the trade-off between productivity gains and wasteful fund-seeking if the government distributes more than 50 percent of all available funds based on performance. The factors that limit the applicability of PB under the effort-based performance signal do not apply. A government that wants to maximize overall output and has the opportunity to choose between performance-signals should use the output-based signal. On the other hand, PB-schemes based on the output-based signal produce higher regional inequality in public services. This inequality increases in the share of performance-based funds. Thus, the dominance of the output-based signal is weakened if welfare is sensitive to regional inequality. Finally, we account for the role of opportunistic policy makers and show that policy makers who maximize political support will not exploit the full potential of PB-schemes. If their aim is to extract rents policy makers misuse PB-schemes and cause welfare losses.

The remainder of the paper is organized as follows: Section 2 outlines the concept of PB and briefly reviews the related literature from economics and public choice. In section 3, we present a game-theoretical model to capture the trade-off described above. Section 4 introduces two major modifications to the model and analyses the implications of different welfare functions for the optimal PB-scheme. Section 5 extends the basic model to assess the welfare implications of PB if policy-makers are opportunistic. In section 6, we discuss the main results of our model and implications for PB and beyond. Section 7 concludes.

2. The concept of PB and related literature

The concept of PB originates from an exchange between scholars from management science, administrative sciences and practitioners in public administration. Since the 1990s, a large number of countries have installed elements of PB (OECD, 2007). Their numerical performance indicators include output measures such as number of students per year for schools and universities, average unemployment duration for employment agencies and performance ratings from surveys among customers of the administrative unit (e.g., Game, 2006). PB-schemes are widespread in the field of higher education where policy makers can compare the performance of different institutions (e.g. state universities) within their jurisdiction and allocate funds based upon relative performance (e.g., Guthrie and Neumann, 2007; Orr et al., 2007; Shin, 2010, Wilkesman and Schmid, 2012). Here, numerical indicators have a substantial influence on budget allocations. Some countries apply similar schemes to local hospitals or cadastre services (OECD, 2007; ch. 3). Despite the importance of numerical indicators in some fields, their overall importance for the allocation of performance-based funds is limited. PB-schemes that rely on a fully-specified allocation formula based on numerical indicators only are the rare exception (OECD, 2007; Robinson, 2007).

Two main arguments stand against allocating funds by a formula that is based on numerical performance-indicators only. The first argument is the heterogeneity in public services: Many

administrative units within one constituency are monopolists for a particular public service. In this case, a meaningful interpretation of their performance reports requires a yardstick from other budgetary years or constituencies (e.g., Gilmour and Levis, 2006). Even if this yardstick is available (e.g. Greiling, 2005; Lockwood and Porcelli, 2011), the allocation of funds across different services requires additional information about political priorities, marginal productivities and reasons for possible difference in performance rankings across services (e.g., Greiling, 2005; OECD, 2007). On the other hand, there are many other public services that are produced by more than one administrative unit within the same jurisdiction. State universities and schools are good examples here. In these cases, comparable performance-measures are readily available. Nevertheless, a formula-based fund allocation rule that assigns funds solely based on numerical performance indicators is not desirable. Here, the second argument applies: Tying funds to imperfect indicators sets incentives for the administrative units to maximize indicator scores rather than performance (e.g., Joyce, 1993; Cragg, 1997; Gilbert and Rocaboy, 2004; Gilmour and Lewis, 2006). To illustrate this point, consider an incentive-scheme that pays teachers by their students' scores on central tests (for instance PISA, baccalaureate or SAT). Here, teachers face incentives to "teach to the test" rather than help students develop their cognitive skills in general. At the same time, they face incentives to neglect their duties to promote social skills, prevent social exclusion or to impact values and social norms and help internalize them. When test results are used as performance indicator only (like in PISA), teachers may even manipulate test scores by asking weak students be absent on the day of the test. To avoid these wrong incentives, efficiency-enhancing PB-schemes must employ other pieces of information next to the numerical indicators to assess the performance and allocate funds.

There is a number of models in the economics and public choice literature that are related to PB. These models take the asymmetric distribution of information between policy-makers and administrative units as starting point. The first strand of literature is theory of bureaucracy (for

a review, see Moe 1997; Wintrobe, 1997). Bureaucrats are assumed to know the public benefits generated by these services as well as the costs of producing them while policy-makers only know the benefits. This asymmetry gives bureaucrats a strong bargaining position in the budget process. Depending on the bureaucrats' objective function, the bargaining process results in inefficiently high output (Niskanen, 1971) and/or x-inefficiencies and managerial slack (Leibenstein, 1966; Migué and Belanger, 1974). Another strand of literature applies the standard principal-agent models to the relationship between policy-makers (principals) and bureaucrats (agents) (for a review, see Moe, 2006).ⁱⁱⁱ In these models, bureaucratic output depends on the bureaucrat's effort and an exogenous factor that is beyond his control. As both are unobservable for the policy-maker, neither output nor effort is contractible. The fact that policy-makers and bureaucrats follow different objective functions causes moral hazard and leads to poor public sector efficiency. One possible remedy is to link the agent's individual pay, career prospects or other benefits to the unit's output and thereby increase the congruence in objectives among principal and agent (e.g. Weingast, 1984; Banks and Sundaram, 1998). Recent studies show that personalized incentive schemes may be harmful because they crowd out the intrinsic motivation of public sector employees (e.g., Frey, 1998; Francois, 2000; Besley and Ghatak, 2005). The literature on contest theory suggests that contests for prizes or extra funds can be an alternative measure to induce efficiency gains (e.g., Clark and Riis, 1998). McCubbins et al. (1987) argue that administrative procedures within the bureaucratic apparatus serve as instruments to control the bureaucrats and thereby reduce the bureaucrats' possibility to exploit their informational advantage. In this context, PB may be seen as a mixture of administrative procedures and incentive schemes for bureaucrats. While the numerical performance indicators and reports are an element of control, performance-based funding initiates a contest between administrative units and sets incentives for them to improve efficiency.

The literature sketched above concentrates on the positive sides of the incentives in PB-schemes but it ignores the fact that real-life PB-schemes based on imperfect performance signals set incentives for the bureaucrats to spend resources on window-dressing and lobbying. There is a large body of literature on the social costs of lobbying and other forms or wasteful rent-seeking (e.g, Tullock, 1980; Congleton, 2008). The rent-seeking literature provides powerful instruments by which we can model the administrative units' competition for performance-based funds. Yet it stresses only the negative aspects of this competition – i.e. the resources wasted in this competition. So far, the trade-off between positive and negative aspects of PB-induced competition for funds has not been analyzed in a formal model. Here our paper comes in. Section 3 develops a game-theoretical model that captures this trade-off. In sections 4 and 5, we introduce a number of modifications to the model.

3. The basic model

Consider a jurisdiction that is divided into N districts. The government of this jurisdiction is in charge of providing citizens in all districts with the public service X . It has delegated the task to N administrative units (hereafter units) – one unit per district. Citizens from district i can only consume the services X_i produced by unit i ($i = 1, \dots, N$). Let X_i be described by the following production function:

$$X_i = \alpha_i(v_i - \lambda_i) \cdot \varepsilon_i \quad (1)$$

Here, α_i is the efficiency parameter ($0 \leq \alpha_i \leq 1$), v_i depicts the gross funds of unit i and λ_i represents the amount of resources spent on preparing numerical performance indicators and writing the performance report (hereafter PB-reporting). Both the efficiency parameter α_i and the resources used for PB-reporting λ_i are controlled by the administrative unit. The random term ε_i ($E(\varepsilon_i) = 1$) captures all factors that influence the unit's output yet are beyond the control of the unit leader. Following the principal-agent literature, we assume that these

factors cause a substantial variance in output and thus neither the efficiency-parameter α_i nor output X_i is contractible. The random term ε_i is unobservable. It is assumed to follow a symmetric distribution within the limits $[1-d, 1+d]$, $d < 1$, and to be uncorrelated with all other determinants of output.

Without PB-scheme, the resources used for PB-reporting $\lambda_i = 0$ and all available resources can be spent on public service production. If a PB-scheme is installed, units have to spend resources on PB-reporting. There is a lower limit λ^{min} for the resources spent on PB-reporting. It comprises of the unavoidable costs of collecting information, preparing the numerical performance indicators and writing an “uninspiring and any old how“ performance report that meets the minimum requirements for such reports. However, the unit can also go beyond λ^{min} and spend additional resources on window-dressing the performance report. All resources spent on PB-reporting are no longer available for public service production and thus reduce output. This holds for the unavoidable λ^{min} as well as for any resources spent beyond λ^{min} .

If a PB-scheme is installed, performance indicators and the performance report provides the government with the imperfect performance signal π_i for unit i . In the basic model, we assume that the jurisdictional government receives the following signal:

$$\pi_i \equiv \pi(\lambda_i, \alpha_i) = \lambda_i^r \alpha_i^{1-r}, \quad r \in [0, 1]. \quad (2)$$

As the level of the efficiency-parameter α_i captures the effort of the unit-leaders to use the resources directed towards public service production efficiently, this signal is hereafter called effort-based signal. It increases in the efficiency-parameter α_i (i.e. in the effort of the unit leaders to produce efficiently) and in the resources that unit i spends on PB-reporting (λ_i). PB-reporting/window-dressing and efficiency cross-fertilize: resources spent on PB-reporting have a larger impact on the overall performance signal if efficiency is high and vice versa. In the basic model, we assume that policy-makers are not susceptible to lobbying. Thus, the

coefficient r is exogenous and measures the unavoidable degree of imperfection in the performance signal. The higher r , the more the performance signal responds to resources used for PB-reporting and thus window-dressing. It can thus be interpreted as a parameter describing the degree of asymmetry in information between policy-makers and unit leaders (next to the dispersion of the random term ε described by d).

Let the unit leaders be risk-neutral and maximize the following objective function.

$$\Theta_i = X_i - a(\alpha_i)^z, \quad a \geq 0, z > 1 \quad (3)$$

The leader of unit i is interested in the output of his unit X_i for altruistic reasons (e.g., Francois, 2000; Besley and Ghatak, 2005), for reasons of prestige or due to career concerns (e.g., Dewatripont et al., 1999). On the other hand, efficiency has opportunity costs because it reduces the possibility to provide benefits for himself and for his staff and requires him to put pressure on his staff to reduce slack. This impairs a good relationship between unit leader and his staff and thereby reduces his utility. By setting $z > 1$, we assume that marginal opportunity costs of efficiency increase in the degree of efficiency.

Let $h[X_i]$ be the welfare that X_i generates among the citizen in district i . We assume $h[X_i] = X_i$. We furthermore assume a benevolent jurisdictional government that maximizes the simple utilitarian welfare function:

$$WF = \sum_{i=1}^N h[X_i] = \sum_{i=1}^N X_i \quad (4)$$

All funds for the production of X are provided by the government. Let the total amount of funds be given by F . For reasons of simplicity, we assume that PB-schemes entail administrative costs on the units' level while their costs on jurisdictional level are negligible. Without PB-scheme, every unit receives funds of F/N regardless of their performance. If the government introduces a PB-scheme, it has to make two decisions: First, it has to set the share

of funds (f) that are distributed based on performance. The remaining fraction ($1-f$) is distributed as block fund of equal size to all N units ($b=(1-f)F/N$). Second, the government has to decide about the number of recipients K . Two regimes can be differentiated. First, it can initiate a contest for a performance prize given to $K < N$ units. Following the theory of rent-seeking and prize-contests (e.g., Tullock, 1980; Clark and Riis, 1996), we assume that all K winners receive the same performance prize ($pp = fF/K$) and we model the process of prize-allocation as if it was a lottery in which K performance-prizes are awarded consecutively.^{iv} Second, the government can provide every unit with a performance bonus pb_i that depends on relative performance such that:

$$pb_i = fF \left(\frac{\pi_i}{\sum_{j=1}^N \pi_j} \right). \quad (5)$$

Hereafter, we denote performance boni by the PB-scheme with $K = N$. Thus, all possible PB-schemes can be described as a combination of f and K .

[insert table 1 about here]

The interaction between the government and its administrative units can be modelled as a sequential game consisting of four stages (see table 1): In stage 1, the government designs the PB-scheme by setting the share of performance-based funds f and the number of recipients K and provides all N units with the block-fund. In stage 2, the unit leaders decide about the efficiency of public service provision α_i and the amount of resources to spend on PB-reporting (λ_i). Given this decision, they use block funds net of PB-reporting expenditures ($v_i - \lambda_i$) to start providing citizens with public services. At this stage, the government receives the performance-signal π_i . In stage 3, the government distributes the remaining funds according to the PB-scheme set in stage 1. Finally, in stage 4, the units that receive performance prizes respectively boni use these additional funds to expand the production of public services.

We assume that the government knows the production function, the function of the performance signal and the unit leaders' objective function. Based on this information, the government can form rational expectations concerning the units' choice of efficiency level, the resources used for PB-reporting and their expected output for all possible PB-schemes and distributions of funds. It receives a unit-specific performance signal π_i in stage 2 of the game and observes the unit-specific output X_i in stage 4. Given the unobservable stochastic component ε_i these two pieces of information (π_i, X_i) are insufficient to calculate the true unit-specific efficiency-level α_i and the resources used for PB-reporting λ_i even after all information is revealed. Thus, the government cannot differentiate between those units that only spent the unavoidable amount λ^{min} of resources on PB-reporting and those units that wasted additional resources on window-dressing. Therefore, it cannot install an incentive scheme that punishes excessive PB-reporting ex post and thus eliminates the incentives to engage in socially wasteful window-dressing. Instead, it has to live with it and choose the PB-scheme, i.e., the combination of f and K , that maximizes welfare. The government solves this game by backward induction. For this purpose, it is necessary to develop the unit's reaction functions, $\alpha_i = \alpha_i(f, K)$ respectively $\lambda_i = \lambda_i(f, K)$. Unit leaders choose efficiency and resources for PB-reporting to maximize expected utility:

$$\max_{\lambda, \alpha} \left\{ E[\Theta_i] : \lambda = 0 \vee \lambda^{min} \leq \lambda \leq b; 0 \leq \alpha \leq 1 \right\} \quad (6)$$

In the absence of PB (i.e. $f = 0$), they do not spend any resources on PB-reporting and choose the efficiency level α_i that maximizes their objective function:

$$\alpha_i^{aut} = \min \left\{ 1; \left[\frac{F-1}{N \cdot az} \right]^{1/(z-1)} \right\} \quad (7)$$

For reasons of simplicity, we assume that all units are identical with respect to their objective function Θ_i and production function $X_i(\cdot)$. Consequently, we assume that $\pi_i = \pi_j \forall i, j$ holds in

the Nash-equilibrium (see Berry, 1993; Clark and Riis, 1996). We hereafter drop the subindex i to save notation. For ease of the argument, we furthermore assume that the lower limit for the resources used for PB-reporting λ^{min} is sufficiently small so there exist an equilibrium in pure strategies. Assuming an inner solution, we arrive at the following Nash-equilibrium solutions for the units' choice when facing a performance-prize scheme (expression (8)) respectively performance-bonus scheme (expression (9)):

$$\lambda_{pp}^* = r \frac{fF}{KN} M \quad \alpha_{pp}^* = \left[\frac{F}{N} \frac{1 + (1-2r)M f/K}{za} \right]^{\frac{1}{z-1}} \quad (8)$$

$$\text{with } M = \left[\frac{K(N-1)}{N} - \sum_j^{K-1} \frac{K-j}{N-j} \right]$$

$$\lambda_{pb}^* = r \frac{fF}{N} \frac{(N-1)}{N} \quad \alpha_{pb}^* = \left[\frac{F}{N} \frac{1 + (1-2r) f(N-1)/N}{za} \right]^{\frac{1}{z-1}} \quad (9)$$

First of all, it is important to note that any PB-scheme only increases efficiency (i.e. yield $\alpha^* > \alpha_i^{aut}$) if $r < 0.5$. For all constellations satisfying this condition, expressions (8) and (9) reveal the essential trade-off: Increasing the share of performance-related funds f improves overall welfare by causing all units to be more efficient ($\partial \alpha^* / \partial f > 0$). At the same time, a concomitant increase in resources spent on PB-reporting ($\partial \lambda^* / \partial f > 0$) reduces welfare. Under performance-prizes, a similar trade-off exists when the government changes K as both efficiency and PB-reporting expenditures decrease in K . For a given share of performance-based funds f , a performance-prize for $K = I$ units leads to the same efficiency parameter and PB-reporting expenditures as performance boni. Regardless of whether prizes or boni are used to incentivize performance, both increase in per-unit funds F/N . The efficiency level α^* decreases in the disutility the unit leaders witnesses as efficiency increases (depicted in z and a), while these parameters are irrelevant for the resources used for PB-reporting (λ^*). The

poorer the quality of the performance signal (i.e. the higher r), the more resources are spent on PB-reporting and the lower the gains in efficiency.

The jurisdictional government maximizes overall welfare by choosing f and K .

$$\text{Max}_{f,K} \left\{ E \left[WF(F, \lambda^*, \alpha^*) \right] \quad \text{s. t. } 0 \leq f \leq 1, \lambda^* = \lambda(f, K), \alpha^* = \alpha(f, K), \lambda^* \leq b, \alpha^* \leq 1 \right\} \quad (10)$$

We arrive at the following inner solution for share of funds f to be used for performance prizes (expression(11)) respectively for performance boni (expression (12)).

$$f_{pp}^* = \frac{K(1-2r) - (z-1)r}{M(1-2r)zr} \quad (11)$$

$$f_{pb}^* = \frac{N(1-2r) - (z-1)r}{(N-1)(1-2r)zr} \quad (12)$$

The optimal share f^* increases in the quality of the performance signal (i.e. decrease with r). The optimal performance-bonus uses the same share of funds as the optimal performance-prize with $K = 1$. The government has a degree of freedom when choosing f and K for performance prizes. The optimal share f_{pp}^* increases with the number of recipients K .

Assuming an inner solution for λ^* , α^* and f^* for all values of K and substituting (11) and (12) in (8) and (9) yields expressions for the level of PB-reporting and efficiency in the social optimum. These are the same for performance bonus and performance prizes:

$$\lambda^* \Big|_{f_{pp}^*, f_{pb}^*} = \frac{F(1-2r) - r(z-1)}{Nz(1-2r)} \quad \alpha^* \Big|_{f_{pp}^*, f_{pb}^*} = \left[\frac{F}{N} \left(\frac{1}{za} + \frac{(1-2r) - r(z-1)}{za(1-2r)} \right) \right]^{1/(z-1)} \quad (13)$$

Thus, all efficient performance-prize-schemes have the same effect on unit behavior and expected output. Furthermore, an efficient boni-scheme yields the same results as any efficient prize-scheme. All optimal PB-schemes yield the same jurisdiction-wide output and

welfare as defined by expression (4). Thus, performance prizes and performance boni are perfect substitutes.

The results above based on the assumption that the maximization problems of unit leaders and of the jurisdictional government yield inner solutions. An inner solution first implies that PB-schemes are welfare-enhancing in principle. Rewriting the numerator in expressions (11) and (12) shows that any PB-scheme only yields welfare gains if $(1+z)r < 1$. This defines a lower limit for f^* . For all cases where $(1+z)r \geq 1$, $f^* = 0$. Similarly, $f^* = 0$ if it is in the self-interest of unit leaders to produce services efficiently (i.e. $\alpha^{aut} = 1$, see expression (7)). Via α^{aut} , the available funds F and parameter a influence the optimal share of performance-based funds f^* indirectly because they determine whether PB-schemes are necessary in the first place.

If PB-schemes are welfare-enhancing in principle, there are the following upper limits for f^* .

$$f_{pp}^* \leq \min \left\{ \frac{K(zaN/F - 1)}{M(1 - 2r)}, \frac{1}{1 + rM/K} \right\}, \quad f_{pb}^* \leq \min \left\{ \frac{N(zaN/F - 1)}{(N - 1)(1 - 2r)}, \frac{N}{N + r(N - 1)} \right\} \quad (14)$$

The first terms in expression (14) account for the fact that, by definition, the efficiency parameter α^* cannot exceed 1. Raising the share of performance-based funds f beyond this limit induces higher PB-reporting expenditures without increasing efficiency. The second terms results from the restriction that the resources spent on PB-reporting cannot exceed the block grant (i.e. $\lambda^* \leq b$). Theoretically, the government could make use of the latter restriction and set very high values for the share of performance-based funds f to raise efficiency and at the same time limit PB-reporting to the size of the block grant. Taking this argument to the limit, it would set a value of f just below 1 and thereby dissolve the trade-off described in expressions (8) and (9). We do not allow for this solution for an obvious reason: If $\lambda = b$, the units have no resources to start production of public services in stage 2 of the game (see table 1). Production in stage 2 is, however, a necessary precondition for the existence of the

performance signal. Thus, PB-schemes that deliberately exploit the λ -restriction destroy their own informational basis.

As $\Delta(K/M)/\Delta K > 0 \quad \forall K=1, \dots, N-1$, the question whether an inner solution exists depends on the value K . Performance boni and performance prizes with a small number of winners K are applicable to a wider range of parameter constellations. In some parameter constellations, the set of optimal PB-schemes consists of performance bonus and performance prizes with a small number of winners K only.

4. Modifications to the basic model

The current section 4 provides three extensions to the basic model. The analyses in the main text are restricted to performance bonus schemes. The results are equivalent for the performance prize schemes as long as welfare is measured by total output (see appendix A).

4.1 Modification 1: Budget-maximizing unit leaders

The basic model used an objective function for the unit leader that is rather optimistic compared to those used in the models following the Niskanen tradition. In this section, we accommodate this tradition and use the following, modified objective function:

$$\Theta_i^N = v_i - \lambda_i - a(\alpha_i)^z, \quad a \geq 0, z > 1 \quad (15)$$

Accordingly, unit leaders are motivated not by output but by their disposable budget ($v_i - \lambda_i$) (e.g., Tullock, 1980; Wintrobe, 1997; Chang and Turnbull, 2002). Other things equal, the benefits the unit leader can extract for himself and his staff increases in the disposable budget.^v The Nash-equilibrium for the units' behavior now reads:

$$\lambda^N = r \frac{fF}{N} \frac{(N-1)}{N} \quad \alpha^N = \left[\frac{F}{N} \frac{(1-r)f}{za} \frac{N-1}{N} \right]^{1/z} \quad (16)$$

The modified objective function yields the same amount of resources spent on PB-reporting λ^* as the one used in the basic model but changes the efficiency-setting behavior. The optimal share of funds f is given by:

$$f^N = \frac{1}{r(z+1)} \frac{N}{N-1} \quad (17)$$

Again assuming an inner solution and substituting expression (17) into (16) yields:

$$\lambda^N \Big|_{f^N} = \frac{1}{(z+1)} \frac{F}{N} \quad \alpha^N \Big|_{f^N} = \left[\frac{F}{N} \frac{(1-r)}{za(z+1)} \right]^{1/z} \quad (18)$$

The optimal PB-regime induces lower efficiency at the same amount of PB-reporting if unit leaders care about disposable funds rather than about output. This result is intuitively clear because increasing benefits and slack cause higher opportunity costs for unit leaders who care about output.

It is important to note that efficiency drops to 0 in the absence of PB (i.e., $\alpha^{aut} = 0$) under the modified objective function. Therefore, PB is always welfare-enhancing even if the performance-signal primarily depends on the PB-reporting expenditures (i.e. r is high). This result is fundamentally different to the conclusion in the basic model according to which PB only leads to welfare gains if $(I+z)r < I$. In fact, if the latter condition holds under the modified objective function (15), expression (17) yields values for the share of performance-based funds of $f^N > I$. In other words, whenever the performance signal has sufficient quality to make PB-schemes welfare-enhancing in the basic model, it is extremely welfare-enhancing in the modified version assumed here. The upper limits for f^N are:

$$f^N \leq \min \left\{ \frac{N^2 az}{(N-1)(1-r)F}, \frac{N}{N+r(N-1)} \right\} \quad (19)$$

4.2 Extension 2: Modified performance-signal

The performance signal used so far contains information concerning the unit's effort to be produce efficiently (at the margin) – blurred by window-dressing. This gives PB-schemes based on this signal high legitimacy, especially when the quality of the performance signal is high (i.e. r is low). The major shortcoming of this signal is that it does not account for the negative impact that PB-reporting (including window-dressing) has on final output. In this section, we assume that policy-makers receive an output-related performance signal based on the output generated by unit i while using the block b in stage 2 of the game:

$$\pi_i^X \equiv \pi(\lambda_i, X_i) = (\lambda_i)^r (\alpha_i (b - \lambda_i) \cdot \varepsilon_{i,S2})^{1-r}, \quad r \in [0, 1]. \quad (20)$$

Here, the random term $\varepsilon_{i,S2}$ captures the impact that factors beyond the control of the unit leaders have on the production in stage 2. We assume that $\varepsilon_{i,S2}$ follows the same distribution as the random term ε_i that disturbs final output in stage 4. The correlation between the two random terms is given by ρ_ε . The output-based performance signal changes the game in a number of ways. We illustrate the changes using the basic model in section 3 as a benchmark. Hence, the assumptions of the basic models apply here. We also assume that the Nash-equilibrium in stage 2 is symmetric in that all unit leaders choose the same efficiency level and the same amount of resources for PB-reporting. This implies that the heterogeneity due to the random term $\varepsilon_{i,S2}$ materializes only after unit leaders have made their decisions.

Assuming an inner solution, the units' choice in equilibrium is given by:

$$\lambda^{\pi X} = \frac{fF(N-1)}{2N^2} + \frac{b}{2} - \sqrt{\left(\frac{fF(N-1)}{2N^2} + \frac{b}{2}\right)^2 - rb \frac{fF(N-1)}{N^2}} \quad (21)$$

$$\alpha^{\pi X} = \left[\frac{F}{N} \cdot \frac{1 - \lambda_{pb}^{\pi X} + (1-r)f(N-1)/N}{za} \right]^{\frac{1}{z-1}} \quad (22)$$

Compared to the effort-based signal, window-dressing has a smaller impact on the expected funds ($\partial \pi_i^X / \partial \lambda_i < \partial \pi_i / \partial \lambda_i$) and thus unit leaders waste less resources on PB-reporting (see figure 1). More importantly, the resources used for PB-reporting do not rise monotonically in the share of performance-based funds f but fall beyond a threshold value $f^{X,t}$. It can be shown that this threshold value $f^{X,t} > 0.5$ for all parameter constellations (proof, see Appendix B). Beyond $f^{X,t}$, a further increase in performance-based funds no longer evokes additional window-dressing yet has the potential to further improve efficiency in public service production (see figure 1). Thus, increasing the share of performance-based funds increases welfare even if efficiency cannot be increased further. Consequently, the strategy that maximizes overall output is to choose the highest possible share of performance-based funds f that still leaves block funds b sufficiently high to receive a performance signal. This strategy is optimal regardless of the values of model parameters $r, z, a, F, N, \rho_\varepsilon$ – provided units do not produce efficiently even without performance-based funds (i.e. $\alpha^{aut} < 1$). Based on the output-based performance signal, PB is applicable even when the unit leaders are reluctant to improve efficiency (i.e., z, a are high) and/or the performance signal is of low quality (i.e., r is high) – constellations in which PB-schemes based on effort-based signals cannot yield welfare gains.

[insert figure 1 about here]

The output-based signal is different to the effort-based one in a second, very important respect: Due to the impact of random factors on output, the new performance signal is heterogeneous among units even if they choose the same efficiency parameter α_i and resources used for PB-reporting λ_i . Substituting expression (20) into expression (5) yields the performance bonus that unit i can expect under the modified performance signal:

$$pb_i = fF \left(\left(\varepsilon_{i,S2} \right)^{1-r} / \sum_{j=1}^N \left(\varepsilon_{j,S2} \right)^{1-r} \right). \quad (23)$$

Unit i 's bonus depends only on factors beyond the control of its leader. The distribution of performance boni is unequal. Inequality becomes larger the better the performance signal is (i.e. the lower r). The impact of this unequal distribution on overall output depends on the correlation ρ_ε . If $\rho_\varepsilon = 0$, expected overall output does not depend on the initial inter-district distribution of performance-based funds. If $\rho_\varepsilon > 0$, the output-based performance signal provides policy makers with an additional piece of information. It indicates which units' final output can expect to be amplified by favorable exogenous circumstances (and which units can expect mitigating effects). Allocating PB-funds according to this information increases overall output. The higher the correlation ρ_ε between random terms, the stronger the impact of PB-funds on overall output. This provides an additional rationale for output-maximizing policy makers to distribute a large share of funds based on performance under the output-based performance signal and to use this signal rather than the effort-based one.

4.3 Modification 3: Welfare-functions sensitive to inter-district inequality

So far, we operate with a simple welfare function that defines jurisdiction-wide welfare as sum of output over all N districts (see expression (4)). It permits an analytical solution of the government's optimization problem and allows for comparative statics. However, it is insensitive to the distribution of output across districts. This can be challenged on grounds of external validity. More importantly, the assumption is likely to drive the results derived so far. In this section, we assume that the jurisdictional government aims at guaranteeing a high level of public services in all districts. Consequently, we use the Rawlsian welfare function:

$$WF^R = \min \{X_i\} \quad (24)$$

Based on this welfare function, we re-evaluate the optimality of different PB-schemes for different performance signals and parameter constellations. The change in welfare function does not change the unit's reaction functions $\alpha_i = \alpha_i(f, K)$ respectively $\lambda_i = \lambda_i(f, K)$ but

changes the results of the government's optimization problem. Three results are noteworthy: First, the optimal bonus scheme dominates all prize schemes because it can achieve the same gains in efficiency at the same level of PB-reporting without inducing inter-district inequality in output. This conclusion holds regardless of the performance-signal used. Second, the PB-scheme that maximizes overall output also maximizes welfare as measured by the Rawlsian welfare function if the effort-based signal is used. The limits for the applicability and the comparative statics for performance boni under the effort-based signal are not sensitive to the welfare function used. This holds because a) the random terms in stage 2 do not influence the performance bonus a certain unit receives and b) inter-district inequality in outcome solely depends on the distribution of random terms in stage 4. Neither the correlation with possible random terms in stage 2 nor the share of performance-based funds influences the degree of inter-district inequality that materializes in the end. Third, the output-based performance signal does not always dominate the effort-based signal under the Rawlsian welfare function. The question which signal yields higher welfare depends on the parameter values. If the output-based signal is used, the parameters describing the distribution in random terms (d, ρ_ε) have a negative impact a) on the performance bonus of the unit i with the most unfavourable random term ($\varepsilon_{i,s2} = 1-d$) and b) on the final output of this unit. As these negative effects increase in the share of performance-based funds, they have the potential to outweigh the positive impact of an output-based performance signal on output. The larger the dispersion in random terms (i.e., the larger d) and the larger the correlation ρ_ε between random terms across stages, the more likely the negative effects limit the share of performance-based funds under an output-based signal. Therefore, the output-based signal is less likely to be superior the larger d and ρ_ε are – other things equal. Figure 2 shows critical combinations of these two parameters for which the superior performance signal changes.

[insert figure 2 about here]

Different lines show critical combinations for different values of the other parameters.^{vi} If the quality of the performance signal is low (i.e. r is high) and unit leaders' are reluctant to increase efficiency (i.e., a and z are high), the effort-based performance signal evokes massive social waste from PB-reporting per unit efficiency gain and PB is limited in its ability to increase output. Under the output-based performance signal, these factors do not limit the ability to increase output. Therefore, the output-based signal is more likely to be superior the larger the parameters r , a and z are – other things equal.^{vii} Consequently, the area where the output-based signal is superior increase in the parameters r , a and z .

5. The role of opportunistic policy-makers

In sections 3 and 4, the jurisdictional government is assumed to be benevolent. We now drop this assumption in favor of assuming opportunistic government. Two forms of opportunism will be differentiated. First, we assume that the jurisdictional government wants to maximize its chance of re-election (e.g., Besley, 2006) and thus maximizes political support Φ :

$$\Phi = \sum_{i=1}^N h(X_i) - \sum_{i=1}^N s\alpha_i^z, \quad s \geq 0 \quad (25)$$

If citizens vote retrospectively, the government can increase its chance of re-election by ensuring an efficient production of public services and thus increasing welfare (e.g., Paldam, 2004). In this respect, it behaves as if it was benevolent. At the same time, however, a high level of efficiency in public service production implies that the administrative staff in the different units cannot enjoy benefits or slack. Through their unions, due to their influence on public opinion and due to the number of voters they represent, public employees have substantial influence on the incumbent's chance of re-election (e.g. Moe, 1997, 2006). Increasing efficiency of public services to attract votes from the general public therefore comes at the cost of losing support among public employees. Once the government accounts for this effect, its objective function becomes similar to that of the unit leaders in the basic

model (see expression (3)). For reasons of simplicity, we assume that the exponent z is the same in both functions, but we allow for a value of $s \neq a$. The larger s , the larger the public employees' impact on political support for the incumbent. Given this new objective function, we arrive at the following inner solutions for the government's optimization problem:

$$f^\phi = \frac{N}{(N-1)} \cdot \frac{(1-s/a)(1-2r) - r(z-1)}{(1-2r)[zr + s/a(1-2r)]} \quad (26)$$

The optimal share of performance-based funds f^ϕ decreases in s and becomes 0 if $s \geq a[1 - r(z-1)/(1-2r)]$. For all parameter constellations that yield inner solutions, $f^\phi < f^*$.

In sum, PB is used to a less than optimal extent. The social costs of opportunism increase with the relative power of the interest groups on the constituency level (i.e in s/a). However, the government will not use PB in a way that reduces welfare compared to a situation without it.

This conclusion changes if the government is interested in extracting rents (McChesney, 1997; Page, 2005). In this case, their objective function reads:

$$\Phi^\lambda = \sum_{i=1}^N \lambda_i \quad (27)$$

To understand the implications of this objective function, we must take a closer look at the interpretation r and λ_i . So far, we assumed that the PB-scheme is designed to yield the best possible performance indicator, i.e the one that operates at the lowest possible value of r . Hence, r represents the unavoidable degree of imperfection resulting from window-dressing that the government cannot identify as such and thus correct for. If the government is interested in extracting rents, its utility increases in the amount of resources that the administrative units transfer back to it through campaign-support and other forms of lobbying. To capture these activities in our model, we use a wider interpretation of λ_i for the remainder of this section. Accordingly, λ_i contains unavoidable expenditures on PB-reporting as well as

additional resources spent on window-dressing and lobbying. Consequently, the performance signal is no longer a performance signal but represents the unit's input in the fund-seeking contest equivalent to the rent-seeking effort in the standard rent-seeking model (Tullock, 1980). A government that wants to extract large amounts of rents will prefer the effort-based performance signal over the output-based one because the former yields higher values for the lobbying effort. As λ^* increases in r , the government can expect more campaign-contributions etc. if the PB-scheme operates at a large value of r . Thus, r becomes a strategic parameter and rent-extracting governments operate the PB-scheme at high values of r . Furthermore, they have an incentive to reserve a large share of performance-based funds f because campaign contributions as a part of λ increase in this share. As a result, a PB-scheme that maximizes extracted rents is highly inefficient. If the government sets $r > 0.5$, units will not increase efficiency beyond α^{aut} at all and the model yields the same implications as the standard rent-seeking model – that is all lobbying expenditures represent social waste.

6. Discussion

6.1 Model results and their implication for PB

The previous sections presented a game-theoretical model that captures the interaction between a jurisdictional government and its subordinate administrative units under PB. We analysed the welfare-effects of PB-schemes that operate with an imperfect performance signal which can be manipulated through window-dressing (and lobbying). First, we find that performance boni and performance prizes (for any number of recipients) are perfect substitutes when welfare in a jurisdiction is measured by overall output across districts. When the welfare function is sensitive to inter-district inequality in output, performance boni dominate performance prizes while the properties of the optimal performance bonus scheme does not depend on the form of the welfare function. Second, we find that the results fundamentally depend on the performance-signal available. Under the effort-based signal, the

optimal PB-scheme depends on the quality of the performance-signal (as expressed in r) and the objective function of the unit leader (as expressed in a , z and the question whether he cares about output or disposable funds). If unit leaders care about output and their preferences for benefits and slack is low and/or available funds are sufficiently high, local service production is efficient even without PB. However, PB-schemes are always welfare-enhancing when unit leaders do not care about output. In all cases, the optimal share of performance-based funds increases in the quality of the performance signal (i.e. decreases in r) and decreases the more reluctant unit leaders are to increase efficiency (i.e. decrease in a and z). Compared to the effort-based signal, PB-schemes that rely on the output-based signal can improve efficiency at substantially lower waste from PB-reporting. PB-schemes with large shares of performance-based funds ($f > 0.5$) can even dissolve the trade-off between productivity-enhancing effects and wasteful fund-seeking. Policy makers maximize output by choosing the highest possible value for f that still leaves basic funding b sufficiently high to get a performance signal. If both signals are available, the PB-scheme that maximizes overall output employs the output-based signal. On the other hand, PB-schemes based on the output-based signal produce substantial inequality in output across districts. If this inequality matters for jurisdictional welfare, the output-based signal no longer dominates the effort-based signal in all constellations. The effort-based signal is superior when the dispersion in random terms is high and the correlation in random terms across stages is high. While only a benevolent government can be expected to exploit the full benefits of PB, opportunistic governments on the jurisdictional level do not necessarily imply that PB-schemes yield welfare-losses. If the government tries to accommodate the public employees' preferences for slack and benefits, they set performance boni and prizes that are too small but nevertheless yield welfare gains. Contrary to that, rent-extracting governments will deliberately design PB-schemes that evoke excessive lobbying and thus reduce welfare.

The analysis above is based on a number of simplifying assumptions. First, we assumed that the unavoidable amount of resources used for PB-reporting λ^{min} is sufficiently small to allow for solutions in pure strategies. This implies that all inner solutions involve some degree of window-dressing. If we account for the fact that λ^{min} may in many cases be substantial, there are a number of additional parameter constellations in which the welfare losses from the unavoidable administrative costs of PB-reporting cannot be outweighed by the gains in efficiency from PB. Under the effort-based performances signal, this applies to parameter constellations for which the optimal share of performance-based funds is small, e.g. because of a poor quality performance signal or a strong reluctance of unit leaders to increase efficiency. Thus, the range of parameter constellations in which PB-schemes are welfare-enhancing shrinks. Under the output-based performance signal, significant fixed costs of PB-reporting are more likely to become binding because the Nash-equilibrium level of PB-reporting is much lower here. They limit the possibility to reduce social waste from window-dressing by increasing the share of performance-based funds. Once λ^{min} becomes binding, welfare gains from increasing this share can only be generated as long as they induce efficiency gains (i.e. change α^*). To loosen the restrictions from λ^{min} and broaden the applicability of PB-schemes, the government can reduce the frequency at which units' performance is assessed. Existing PB-schemes generally initiate comprehensive performance assessments on a multi-annual basis (e.g., OECD, 2007: ch. 2).

Second, we model PB as a one-shot game. In reality, PB-schemes are applied over longer periods of time and involve consecutive performance reports and decisions about the allocation of funds. Thus, PB-schemes can employ information about unit-specific output X_i from previous periods as an additional numerical performance indicator. This makes it costly for the units to spend resources on window-dressing because it reduces expected performance-based funds in the next round. More importantly, increasing efficiency has a

positive influence on expected funds and output in the current and in future budgetary periods. Drawing on the logic on output-based performance signals (see section 4.2), the use of past performance as performance indicator can further increase overall output if the unit-specific random terms ε_i ($=\varepsilon_{iSt}$) are (positively) correlated inter-temporally. In this case, past output indicates which units' production can expect to be amplified by favorable exogenous circumstances (and which units can expect mitigating effects) in upcoming periods. On the other hand, the main limitation of the alternative performance-indicator applies: Given the substantial variance of the random factors in both stages, the resulting interdistrict inequality in output is substantial. Thus, it may not be welfare-enhancing to use past output as a prominent performance indicator. Moreover, positively correlated random terms imply that there are districts in which output is systematically lower even if the administrative units exert the same effort as other units. Allocating less PB-funds to these districts means lowering output in already disadvantaged districts. This destroys the legitimacy of PB-schemes and runs against the principles of fiscal equalization as well as performance budgeting (e.g., Gilmour and Levis, 2006). To comply with these principles and at the same time keep up incentives for units to increase efficiency, the PB-scheme can employ a discriminating fund allocation procedure that – other things equal – grants disadvantaged regions higher boni. A thorough analysis of such PB-schemes requires the solution of games with truly heterogeneous players. This lies beyond the scope of this paper.

6.2 Generalizations

The analysis above yields a number of important insights that reach beyond the context of PB. First, the model applies to vertical grants that international organisations like the IMF or the World Bank give to less developed countries or countries in financial crises. These institutions provide substantial financial support upon application and involving conditionality (e.g., Vaubel, 2006; Dreher, 2009). Support programs generally assume that the

recipients suffer from poor institutional quality (e.g., Hefeker, 2006; Vaubel, 2006). By tying aid to institutional reforms, grantors set incentives to improve institutional quality. At the same time, however, they have to live with the fact that the applications and reports they receive sugarcoat true institutional standards or reforms. If the interest groups in these countries are strong and sugarcoating has a large influence on the signal of institutional quality, these programs induce welfare losses rather than welfare gains. If welfare gains are feasible, the comparative statics of our model apply.

Second, our model opens a new perspective on conditional grants in federalist countries. Conditional grants represent an incentive-scheme for lower-level governments that is similar to the performance prizes or boni (e.g., Boadway et al., 1999; Bessho and Terai, 2011). Local governments in most countries receive block grants depending on population size (and possibly a number of other indicators). In addition, supra-ordinate governments support them with conditional grants for specific projects that improve the quality of local public services. In their application for competitive grants, municipalities have to provide a broad account of their performance in public service production and the value-added by the prospective project. Whenever the demand for project funds exceeds the means reserved for them, some applications have to be turned down and some projects have to be downsized. It is precisely this competition for grants that induces efficiency gains in public service production yet evoke window-dressing and lobbying (Borck and Owings, 2003; Bessho and Terai, 2011).

Inefficiencies in local public service production may result from government failure (e.g., Chernick, 1979; Byrnes and Dollery, 2002) and regional spillovers (e.g., Oates, 1999; Shah, 2006). Government failure may be caused by opportunistic local governments that are interested in extracting benefits (e.g., Belleflamme and Hinriks, 2005) or by regional pressure groups pressing local authorities for favourable yet inefficient solutions in public service production (e.g., Dougan and Keynon, 1988; Austen-Smith, 1997; Lorz, 2001).^{viii} By giving

in to these groups, local governments gain their support yet lose support among voters in the next election. Thus, the objective-function used in the basic model (see expression (3)) captures the situation of a local government striving for re-election. If interest groups are stronger on the local level than on the supra-ordinate level, our results in section 5 indicate that conditional grants can increase net welfare. They can be a valuable instrument to reduce local government failure especially in situations where the preconditions for effective yardstick competition are not met, e.g., because mobility costs are high or citizens lack the necessary information to evaluate the local government's performance (e.g., Besley and Case, 1995; Besley, 2006). If inefficiencies result from regional spillovers, our model does not incorporate all essential features of the inter-fiscal game (e.g., Oates, 1999; Fenge and Wrede, 2007). Nevertheless, its essential lessons apply.

A number of authors argue that the granting governments pursue their own political goals when distributing grants. They may prioritize certain regions to maximize political support (e.g., Grossman, 1994; Worthington and Dollery 1998). If the prioritized recipients know this, they will be less motivated to increase efficiency but also spend less resources on window-dressing. The same holds for those regions that do not expect to be prioritized. Consequently, conditional grants will be less effective but also less harmful – the net effect is undetermined. If, however, the supra-ordinate government is interested in extracting rents, conditional grants are just as harmful as PB-schemes (see section 5).

7. Conclusion

Since its introduction in the 1990s, PB has spread continuously and is now applied for a substantial number of different public services in many countries. The existing PB-schemes assess the performance of administrative units in a complex procedure that uses numerical performance indicators but also draws on information from performance reports and others sources. While this is an adequate way to cope with the limitations of numerical indicators,

the complex assessment procedure evokes socially wasteful fund-seeking. The existing theoretical models focus either on the positive aspects of incentives or on the negative aspect of wasteful fund-seeking. We provide the first model that accounts for both aspects.

We show that the impact of PB and the optimal PB-scheme differs fundamentally for different performance signals. The applicability of PB is limited if policy-makers have to work with an effort-based performance signal. The level of wasteful window-dressing is much lower if policy makers use an output-based performance signal. Therefore PB-schemes based on this signal are more widely applicable and yield higher gains in overall output. However, PB-schemes based on the output-related signal come at the price of substantially higher inter-district inequality in public services. Once policy-makers account for this inequality, the output-based performance signal no longer dominates the effort-based signal. Depending on the constellations, the optimal PB-scheme employs the one or the other signal. In section 5, we show that policy makers who maximize political support will not exploit the full potential of PB-schemes. If their aim is to extract rents, policy makers misuse PB-schemes and cause welfare losses.

There are interesting questions for future research. From a theoretical perspective, the model can be extended to a world with heterogeneous recipient units of PB-funds and the role of targeting. Targeting rules can help governments to prioritize certain (disadvantaged) districts (e.g. Stein 2002). Once window-dressing is taken into account, they can also play an important role in reducing overall window-dressing costs. From an empirical perspective, one challenge lies in the development of an indicator for window-dressing effort. Such an indicator can be used to test hypotheses on the driving factors of window-dressing and help to identify wasteful from less wasteful PB-schemes.

Beyond PB, the main insights from our model can be generalized to contexts where i) international organisations tie financial aid to performance and ii) governments in federalist

countries apply conditional grants to increase efficiency on the subordinate level, especially if inefficiencies result from government failure. Taking an even more general perspective, the logic of the above analysis applies to a very general question: What are the “contests” that a society should install and how should the rules of the contest be designed? While this question is being debated ever since the birth of economics, the modelling approach developed in this paper provides a possible framework for a thorough account of the benefits and costs of contests that operate on imperfect quality signals.

References

- Amazeen, M. (2011). Gap (RED): Social responsibility campaign or window dressing? *Journal of Business Ethics* 99: 167 – 182.
- Austen-Smith, D. (1997). Interest groups: Money, information and influence. in: Mueller (ed.): 296-321.
- Bähr, C. (2008). How does sub-national autonomy affect the effectiveness of structural funds? *Kyklos* 61, 1: 3-18.
- Banks, J. S. and Sundaram, R. K. (1998). Optimal retention in agency problems. *Journal of Economic Theory* 82(2): 293-323.
- Belleflamme, P. and Hindriks, J. (2005). Yardstick competition and political agency problems. *Social Choice and Welfare* 24: 155-169.
- Berry, S. K. (1993). Rent-seeking with multiple winners. *Public Choice* 77: 437-443.
- Besley, T. J. (2006). *Principled agents? The political economy of good government*. Oxford: Oxford University Press.
- Besley, T. J., Case, A. C., 1995. Incumbent behavior: vote seeking, tax setting and yardstick competition. *American Economic Review* 85 (1): 25–45.
- Besley, T. and Ghatak, M. (2005). Competition and Incentives with Motivated Agents. *American Economic Review* 95 (3): 616-636.
- Bessho, S. and Terai K. (2011). Competition for private capital and central grants: the case of Japanese industrial parks. *Economics of Governance* 2: 135-154.
- Boadway, R. Horiba, I. and Raghbendra, J. (1999). The provision of public services by government funded decentralized agencies. *Public Choice* 100: 157-184.
- Borck, R. and Owings, S. (2003). The political economy of intergovernmental grants. *Regional Science and Urban Economics* 33: 139-156.
- Byrnes, J. and Dollery, B. (2002). Local government failure in Australia? An empirical Analysis of New South Wales. *Australian Journal of Public Administration* 61(3): 54-64.
- Chang, C. and Turnbull, G. K. (2002). Bureaucratic behaviour in the local public sector: A revealed preference approach. *Public Choice* 113: 191-209.

- Chernick, H. A. (1979). An economic model of the distribution of project grants. in: Mieszkowski, P. and Oakland, W. H. (eds). *Fiscal federalism and grants-in-aid*. Washington, D.C: The Urban Institute: 81-103.
- Clark, D. J. and Riijs, C. (1996). A multi-winner nested rent-seeking contest. *Public Choice* 87: 177-184.
- Clark, D. J. and Riijs, C. (1998). Competition over more than one prize. *American Economic Review* 88(1): 276-289.
- Congleton, Roger D. (ed.) (2008). *40 years of research on rent seeking*. Berlin etc.: Springer.
- Cragg, M. (1997). Performance incentives in the public sector: Evidence from the Job Training Partnership Act. *Journal of Law Economics, and Organization* 13: 147-168.
- Crain, W. M. and O'Roark, J. B. (2004). The impact of performance-based budgeting on state fiscal performance. *Economics of Governance* 5: 167-186.
- de Groot, H. (1988). Decentralization decisions in bureaucracies as a principal-agent problem. *Journal of Public Economics* 36: 323-337.
- Dewatripont, M., Jewitt, I. D. and Tirole, J. (1999). The economics of career concerns, part I: comparing information structures. *Review of Economic Studies* 66(1): 183-198.
- Dougan, W. R. and Kenyon, D. A. (1988). Pressure groups and public expenditures: The flypaper effect reconsidered. *Economic Inquiry* 26: 159-170.
- Dreher, A. (2009). IMF conditionality: theory and evidence. *Public Choice* 141: 233-267.
- Fenge, R. and Wrede, M. (2007). EU financing and regional policy: Vertical fiscal externalities when capital is mobile. *Finanzarchiv* 63: 457-476.
- Frey, B. S. (1998). *Not just for the money - An economic theory of personal motivation*. Cheltenham. Elgar.
- Francois, P. (2000). 'Public service motivation' as an argument for government provision. *Journal of Public Economics* 78: 275-299.
- Game, C. (2006). Comprehensive performance assessment in English local government. *International Journal of Productivity and Performance Management* 55: 466-479.
- Geys, B., Heinemann, F. and Kalb, A. (2010) Voter involvement, fiscal autonomy and public sector efficiency: Evidence from German municipalities. *European Journal of Political Economy*, 26(2): 265-278.

- Gilbert, G. and Rocaboy, Y. (2004). The central government grant allocation problem in the presence of misrepresentation and cheating. *Economics of Governance* 5: 137-147.
- Gilmour, J. B. and Lewis, D. E. (2006). Does performance budgeting work? An examination of the Office of Management and Budget's PART scores. *Public Administration Review* 66: 742-752.
- Greiling, D. (2005). Performance measurement in the public sector: the German experience. *International Journal of Productivity and Performance Management* 54: 551-567.
- Grossman, P. J., Mavros, P. and Wassmer, R. E. (1999). Public sector technical inefficiencies in large US cities. *Journal of Urban Economics* 46: 278-299.
- Grossman, P. J. (1994). A political theory of intergovernmental grants. *Public Choice* 78: 295-303.
- Guthrie, J. and Neumann, R. (2007). Economic and non-financial performance indicators in universities – The establishment of a performance-driven system for Australian higher education. *Public Management Review* 9: 231-252.
- Hefeker, C. (2006). Project aid or budget aid? The interests of governments and financial institutions. *Review of Development Economics* 10(2): 241-252.
- Jones, L. R. and McCaffery, J. L. (2010). Performance budgeting in the U.S. Federal government: history, status and future implications. *Public Finance and Management* 10: 482-523.
- Jordan, M. M. and Hackbart, M. M. (1999). Performance budgeting and performance funding in the States: a status assessment. *Public Budgeting and Finance*: 68-88.
- Joyce, P. G. (1993). Using performance measures for federal budgeting: proposals and prospects. *Public Budgeting and Finance*, Winter: 3-17.
- Kelly, J. M. and Riverbank, W. C. (2003). Performance budgeting for state and local government. New York, London: M. E. Sharpe.
- Leibenstein, H. (1966). Allocative efficiency vs X-efficiency. *American Economic Review* 56: 392-415.
- Lockwood, B. and Porcelli, F. (2010). Incentives schemes for local government: theory and evidence from comprehensive performance assessment in England. CESifo Working Paper No. 3483.

- Lorz, O. (2001). On the effects of capital mobility on local infrastructure policy and rent-seeking. *Regional Science and Urban Economics* 31: 319–337.
- Li-Wen, L. (2010). Corporate social responsibility in China: window dressing or structural change? *Berkeley Journal of International Law* 28: 64–100.
- Michaels, R. (1988). The design of rent-seeking competitions. *Public Choice* 56: 17-29.
- Migué, J.-L. and Belanger, G. (1974). Towards a general theory of managerial discretion. *Public Choice* 17: 27–43.
- McChesney, F. S. (1997). *Money for nothing: politicians, rent extraction and political extortion*, Cambridge and London: Harvard University Press.
- McCubbins, M. D., Noll, R. G., and Weingast, B. R. (1987). Administrative procedures as instruments of political control. *Journal of Law, Economics, and Organization* 3: 243–77.
- Moe, T. M. (1997). The positive theory of public bureaucracy. in Mueller (ed.): 455-480.
- Moe, T. M. (2006). Political control and the power of the agent. *Journal of Law, Economics, and Organization* 22: 1–29.
- Morey, M. R. and O'Neal, E. S. (2006). Window dressing in bond mutual funds. *Journal of Financial Research* 29: 325–347.
- Moynihan, D. P. (2005). What do we talk about when we talk about performance? Dialogue theory and performance budgeting. *Journal of Public Administration Research and Theory* 16(2): 151-168.
- Mueller, D. C., (ed) (1997). *Perspectives on public choice. A handbook*. Cambridge: Cambridge Univ. Press.
- Niskanen, W. A. (1971). *Bureaucracy and representative government*. Chicago: Aldine-Atherton.
- Oates, Wallace E. (1999). An essay on fiscal federalism. *Journal of Economic Literature* 37: 1120-1149.
- OECD (2007). *Performance budgeting in OECD countries*. OECD Publishing.
- Orr, D., Jaeger, M. and 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.

- Paldam, M. (2004), Are vote and popularity functions economically correct? In Rowley, C. K. and F. Schneider (eds), *The Encyclopedia of Public Choice*, Vol. 1: 49-59. Dordrecht: Kluwer Academic Publishers.
- Page, S. (2005). What's new about the New Public Management? Administrative change in human services. *Public Administration Review* 65: 713-27.
- Robinson, M (ed.) (2007). *Performance budgeting*. New York. Palgrave Macmillan.
- Robinson, M (2007a). Result information. In: Robinson (2007): 26-45.
- Robinson, M (2007b). Making performance budgeting work. In: Robinson (2007): 113-127.
- Schultze, C. (1974). Sorting out the social grant programs: An economist's criteria. *American Economic Review* 64, 2: 181-189.
- Shah, A. (2006). A practitioner's guide to intergovernmental fiscal transfers. World Bank Policy Research Working Papers 4039.
- Shin, J. C. (2010). Impacts of performance-based accountability on institutional performance in the US. *Higher Education* 60: 47-68.
- Stickney, C. P. (1975). Window dressing the interim-earnings report: An empirical assessment for firms initially going public. *Journal of Business* 48: 87 – 97.
- Stein, William E. (2002). Asymmetric rent-seeking with more than two contestants. *Public Choice*, 113(3): 325–36.
- Tullock, G. (1980). The costs of transfers. in: Buchanan, J. M. et al. (eds). *Towards a theory of the rent-seeking society*. College station: Texas A&M University Press: 269-282.
- Vaubel, R. (2006). Principal-agent problems in international organizations. *Review of International Organizations* 1: 125–138.
- Weingast, B. R. (1984). The Congressional-Bureaucratic System: A Principal-Agent Perspective. *Public Choice* 44: 147–191.
- Wilkesmann, U. and Schmid, C. (2011). The impacts of new governance on teaching at German universities. Findings from a national survey. *Higher Education* 61 (forthc.).
- Wintrobe, R. (1997). Modern bureaucratic theory. In: Mueller (ed.): 373-390.
- Worthington, A. C. and Dollery, B. E. (1998). The political determination of intergovernmental grants in Australia. *Public Choice* 94: 299-315.

Appendix A: Optimal performance prize scheme in modification 1 and 2 and section 5

Under modification 1, the units' Nash-equilibrical reaction functions read:

$$\lambda_{pp}^N = r \frac{fF}{KN} M \quad \alpha_{pp}^N = \left[\frac{F}{N} \cdot \frac{(1-r)}{za} \frac{fM}{K} \right]^{\frac{1}{z}}$$

The optimal share of funds f to be used for performance prizes is given by:

$$f_{pp}^N = K \cdot \frac{1}{Mr(z+1)}$$

The limits for an inner solution are given by:

$$f_{pp}^N \leq \min \left\{ \frac{KNaz}{M(1-r)F}, \frac{1}{1+rM/K} \right\}$$

Under modification 2, units choose the following value of α_i and λ_i in equilibrium:

$$\lambda_{pp}^{\pi X} = \frac{fFM}{2KN} + \frac{b}{2} \pm \sqrt{\left(\frac{fFM}{2KN} + \frac{b}{2} \right)^2 - rb \frac{fFM}{KN}} \quad \alpha_{pp}^{\pi X} = \left[\frac{F}{N} \cdot \frac{1 - \lambda_{pp}^{\pi X} + (1-r)M f/K}{za} \right]^{\frac{1}{z-1}}$$

If policy-makers maximize political support (see section 5), the optimal performance prize scheme is described by the following expression:

$$f_{pp}^{\phi} = \frac{K}{M} \cdot \frac{(1-s/a)(1-2r) - r(z-1)}{(1-2r)[zr + s/a(1-2r)]}$$

Appendix B: Proof for $f^{X,t} > 0.5$

Solving the contest for the output based performance signal under the assumption that the random terms are perfectly uncorrelated, we can derive two possible solutions for λ :

$$\lambda_{+/-}^{\pi X} = \frac{fF(N-1)}{2N^2} + \frac{(1-f)F}{2N} \pm \sqrt{\left(\frac{fF(N-1)}{2N^2} + \frac{(1-f)F}{2N} \right)^2 - r \frac{f(1-f)F^2(N-1)}{N^3}}$$

Only the negative square root solution (hereafter: $\lambda^{\pi X}$) is valid because the positive square root solution satisfies the upper limit for λ only for extremely small values of f (if at all).

From the FOC of the units' optimization problem it follows that:

$$\alpha^{\pi X} = \left[\frac{F}{Naz} + \frac{fF}{az} \frac{N-1}{N^2} - \frac{1}{az} \lambda \right]^{\frac{1}{z-1}}$$

$$\frac{\partial \alpha^{\pi X}}{\partial f} = \frac{1}{z-1} \frac{\left[\frac{F}{Naz} + \frac{fF}{az} \frac{N-1}{N^2} - \frac{1}{az} \lambda \right]^{\frac{1}{z-1}}}{\left[\frac{F}{Naz} + \frac{fF}{az} \frac{N-1}{N^2} - \frac{1}{az} \lambda \right]} \left(\frac{F}{az} \frac{N-1}{N^2} - \frac{1}{az} \frac{\partial \lambda}{\partial f} \right)$$

$$\frac{\partial \alpha^{\pi X}}{\partial \lambda} = \frac{1}{z-1} \frac{\left[\frac{F}{Naz} + \frac{fF}{az} \frac{N-1}{N^2} - \frac{1}{az} \lambda \right]^{\frac{1}{z-1}}}{\left[\frac{F}{Naz} + \frac{fF}{az} \frac{N-1}{N^2} - \frac{1}{az} \lambda \right]} \left(-\frac{1}{az} \right)$$

Taking f to its limits we can get $\alpha^{\pi X}|_{f=0} = \left[\frac{F}{Naz} \right]^{\frac{1}{z-1}}$ and $\alpha^{\pi X}|_{f=1} = \left[2 \frac{F}{Naz} - \frac{F}{N^2 az} \right]^{\frac{1}{z-1}}$.

In the permitted interval $f \in [0;1]$ the graph of $\lambda^{\pi X}$ is inversely U-shaped and has exactly two roots at $f_1^0 = 0$ and $f_2^0 = 1$. The first derivative of $\lambda^{\pi X}$ with respect to f reads:

$$\frac{\partial \lambda^{\pi X}}{\partial f} = \frac{F}{2N} \left[-\frac{1}{N} - 0.5 \frac{1}{\sqrt{A}} \left(2f \left(\frac{N-1}{N} \right)^2 - 2(1-f) + (2-4f-4r+8rf) \frac{N-1}{N} \right) \right]$$

$$\text{where } A \equiv f^2 \frac{(N-1)^2}{N^2} + (1-f)^2 + 2f(1-f) \frac{N-1}{N} - 4rf(1-f) \frac{N-1}{N}$$

Evaluating expression $\lambda^{\pi X}$ at $f=0$ and $f=1$ yields:

$$\left. \frac{\partial \lambda^{\pi X}}{\partial f} \right|_{f=0} = r \frac{F}{N} \frac{N-1}{N} > 0 \quad \left. \frac{\partial \lambda^{\pi X}}{\partial f} \right|_{f=1} = -r \frac{F}{N} < 0$$

Evaluating expression $\lambda^{\pi X}$ at $f = 0.5$ yields:

$$\left. \frac{\partial \lambda^{\pi X}}{\partial f} \right|_{f=0.5} = \frac{F}{2N} \left[-\frac{1}{N} + 0.5 \frac{1}{\sqrt{A|_{f=0.5}}} \left(1 - \left(\frac{N-1}{N} \right)^2 \right) \right]$$

where
$$A|_{f=0.5} = \frac{1}{4} \frac{(N-1)^2}{N^2} + \frac{1}{4} + \left(\frac{1}{2} - r \right) \frac{N-1}{N} = \frac{4N^2 - 4N + 1}{4N^2} - r \frac{N-1}{N}$$

We show that $\left. \frac{\partial \lambda^{\pi X}}{\partial f} \right|_{f=0.5} > 0$ by proving that $\left. \frac{\partial \lambda^{\pi X}}{\partial f} \right|_{f=0.5} \leq 0$ is impossible for all relevant

parameter constellations:

$$\frac{F}{2N} \left[-\frac{1}{N} + 0.5 \frac{1}{\sqrt{\frac{4N^2 - 4N + 1}{4N^2} - r \frac{N-1}{N}}} \left(1 - \left(\frac{N-1}{N} \right)^2 \right) \right] \leq 0$$

$$\Leftrightarrow \left(\frac{4N^2 - 4N + 1}{4N^2} \right) \leq \left(\frac{4N^2 - 4N + 1}{4N^2} - r \frac{N-1}{N} \right) \Leftrightarrow r \leq 0$$

This result contradicts the assumption of the positive range of $0 < r \leq 1$. The only valid case would be a perfect performance signal ($r = 0$), which is irrelevant for our analysis. We may

conclude that $\left. \frac{\partial \lambda^{\pi X}}{\partial f} \right|_{f=0.5} > 0$. Thus, the value of f that maximizes λ lies between 0.5 and 1.

ENDNOTES

ⁱ In general the term window dressing is used in the sense of “making something appear better than it is”. The accounting literature uses it to refer to activities such as accounting tricks to make balance sheets and business performance reports appear better (e.g. Stickney 1975). But the term is also used in a variety of fields where economic appearance or labels may play a prominent role such as investment banking (e.g. Morey and O’Neill 2006), structural policy (e.g. Li-Wen 2010) or corporate social responsibility (e.g. Amazeen 2011).

ⁱⁱ Michaels (1988) models a rent-seeking contest with multiple inputs. However, all his inputs cause social waste. The aim of his paper was not to model the trade-off that we focus on here but to show that relative prices determine the mix of different forms or rent-seeking effort.

ⁱⁱⁱ A number of papers discuss similar matters in the context of federal structures in which the supra-ordinate government delegates task to local jurisdictions.

^{iv} The probability p_i that unit i wins in one of the K rounds is given by (Clark and Riis, 1996):

$$p_i = p_i^1 + \sum_{j=1}^{k-1} \left[\prod_{s=1}^j (1-p_i^s) p_i^{j+1} \right]$$

Here, p_i^1 stands for the probability that unit i wins the first prize, $(1-p_i^j)$ represents unit i ’s probability of not having won one of the previous prizes and p_i^j represents unit i ’s probability to win in round $j > 1$. We assume that all probabilities strictly depend on relative performance such that $p_i^1 = \pi_i / \sum_{v=1}^N \pi_v$ and p_i^j is given by the ratio of π_i and the sum over the performance signals of all units still in the contest (Clark and Riis, 1996).

^v Note that assuming that gross funds v_i rather than disposable funds $(v_i - \lambda_i)$ would automatically imply that units use the all block funds for window-dressing and thus overall outcome and welfare would be zero for all N-K units that do not receive a prize in the end. Thus, either $f=0$ or $f=1$ are optimal from the policy-makers point of view. We decided to restrict the analysis in the main text of the paper to the more interesting case where unit managers are interested in disposable funds.

^{vi} As the optimal share of performance-based funds under the output-based signal cannot be derived analytically, it is impossible to define the conditions under which one or the other performance signal is superior in general terms. Figure 3 is thus generated in simulations. We assume a symmetric triangular distribution of all random terms. The critical values were generated using 50 randomly created samples for the random term vectors for every single combination of $[\varepsilon_H - \varepsilon_L]$ and ρ_ε . (Simulation routines and results are available with the author).

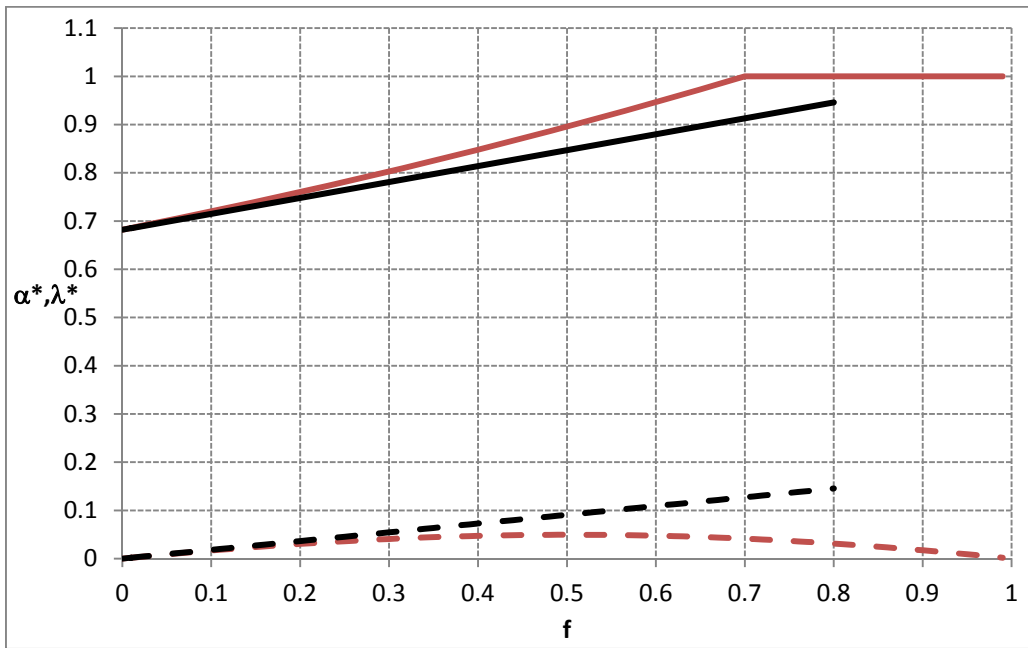
^{vii} The primary impact of the ration F/N is on α^{out} and applies to both signals in a similar way.

^{viii} Supra-ordinate governments and institutions frequently justify conditional grants by arguing that they increase welfare within the recipient region itself. For example, the EU cohesion programs assume that economically weak regions may suffer from poor institutional quality (e.g., Bähr, 2008). A number of papers recognize the justification (e.g., Schultze, 1974; Chernick, 1979; Byrnes and Dollery, 2002). The empirical literature has come up with evidence for inefficiencies (e.g., Grossman et al., 1999; Geys et al., 2010).

Table 1: Game structure

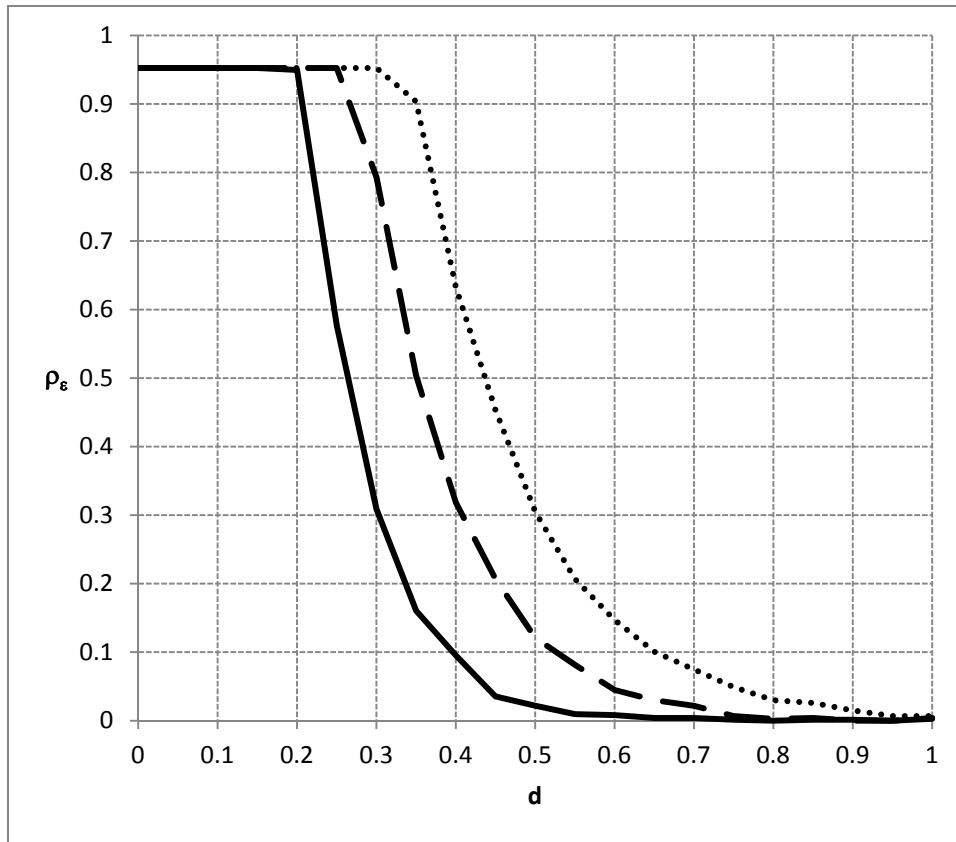
Stage	Activity	Information
1	Policy-makers sets f and K and allocate the block fund to the units.	f, K
2	The units choose λ_i (window-dressing and lobbying) and α_i (efficiency level) given f and K and start the production of X using $b_i - \lambda_i$ and α_i . Policy-makers observe the performance signal π_i	π_i
3	Policy-makers observe distribute the performance-based funds according to the PB-scheme.	
4	The units receiving additional funds use these to expand production.	WF, X_i

Figure 1: Efficiency and PB-reporting as a function of f for both performance signals



The solid lines represent the Nash-equilibrical efficiency-parameter (α^*) for different values of f , the dashed line represent the Nash-equilibrical resources used for PB-reporting (λ^*) for an arbitrary parameter constellation. Black (red) lines represent the units' reaction function under the effort-based (output-based) performance signal. The equilibrical solution for the effort-based signal are limited by the upper limit for PB-reporting expenditures $\lambda \leq b$. Therefore f cannot exceed 0.8.

Figure 2: Effort-based versus output-based performance signal under a Rawlsian welfare function: critical values for d and ρ_ε



For all combinations (d, ρ_ε) on the left of resp. below the lines, the output-based signal is superior to effort-based model. The opposite is true for all combinations to the right resp. above the lines. The exogeneous parameter values used in this figure are: $N = 32$, $F = 15$, $z = 1.7$, $a = 0.55$, $r = 0.075$ (solid line), $r = 0.1$ (dashed line) resp. $r = 0.125$ (dotted line).