

An econometric theory of performance bureaucracy

Roger J. Bowden*

The road to hell is paved with good intentions
(trad./anon.)

Abstract

Performance bureaucracies are organizational structures that arise in response to internally or externally derived metrics for mission or goal attainment, often associated with public sector accountability. Measurement and certification are key drivers of the internal political process. This becomes counterproductive when goal setting is exposed to hidden or latent variables reflecting incomplete information or coalface disaffection, so that managerial controls are metrically insufficient. An econometric framework is developed that draws on the semiotics in common use by performance oriented management regimes. Nash cooperative budget solutions among managers lead to overinvestment in measurement and control, the problem of the commons, and adverse performance outcomes.

Key words: Accountability, agency, hidden variables, latent variables, predictive sufficiency, problem of commons, rubrics.

JEL numbers: D73, D02, L32, C01, C71

*Kiwicap Research Ltd, Wellington, New Zealand, emails roger.bowden@kiwicap.co.nz; roger.bowden@uni-ulm.de. The author is grateful to the Institute of Finance, Ulm University, Germany and the Department of Finance at Xiamen University, China for their hospitality and financial support while writing up this paper. Happily, neither of these institutions has yet been materially infected by the concerns it raises.

1. Introduction

Bureaucracy as a socio-economic phenomenon has over the years attracted much discussion and debate from economists, political scientists, sociologists and management analysts. Precise definitions of just what qualifies as a bureaucracy are hard to find, though a working definition might consist in the use of standardised reporting and procedural rules rather than personal judgments, administered by a hierarchic organisational structure. Max Weber's view of an impersonal administrative governance, based on rules as fair processes, probably captures the essence of an idealised bureaucratic regime (see, e.g. Merton (1952)); likewise, Marx's view of state bureaucracy as transitional to mutual cooperation among the workers, grounded again on 'fair rules' (e.g. Draper (1979)). More sceptical themes centre around monopoly and/or lack of market based discipline as to soaring administration costs; complexity, including a plethora of rules rather than reliance on common sense; inability to correct errors and intolerance of dissenting views; agency problems and managerial empire building based on asymmetric information. Contributions that highlight such aspects have been originated by Niskanen (1968, 1994), also Downs (1964), on the costs and distortions of agency; Tullock (1965) on moral adverse selection; Crozier (1964) on the inability to learn from errors, and proneness to internal strategic games; and Von Mises (1944) on rules and empires. Related discussions appear in the literature of the political process, including the ability of legislators to supervise those who administer the legislation. Further discussion of these and other aspects can be found in Breton and Wintrobe (1985), McCubbins *et al* (1987), Sappington (1991), Miller (1992), Lindblom and Woodhouse (1993), Tirole (1994), and Pierre (1995).

In other words, this is a venerable literature. Most authors – and indeed most people – would consider bureaucracy as inherently bad, as something that needs to be corrected, or at best as a more or less inevitable transition process. It is probably also true that most people still conceive of a bureaucracy as embodied in a cumbersome set of procedural rules or as a set of possibly unnecessary proformas requesting information that may never be used, much as in the old German civil service.

Recent years, however, have seen the rise of formidable (and very expensive) bureaucracies that paradoxically, are based on a desire for accountability, in which measurement of one kind or another plays a key role in assessing the effectiveness of the organisation as a whole. They could therefore be called 'performance bureaucracies'¹. Performance criteria can vary in nature. Some are based on more or less numerical scales. An example is the Performance Based Research Fund of the New Zealand Tertiary Education Commission, which allocates

discretionary research funding to universities according to their scores on periodically conducted research evaluations, partly following an earlier UK model. Universities have responded by creating a research bureaucracy of their own, partly funded by internal taxes on research income, with online reporting forms, compulsory courses on the research process, research points and courses for ‘mentoring’, and other exhortatory devices. Performance bureaucracies can also arise from certification requirements, so the performance metrics involved are essentially categorical, though again based on some form of measurement. The Basle Capital Adequacy standards are of this nature, together with the respective national regulatory regimes based on them such as the European Capital Markets Directive. Banks must demonstrate that they meet economic capital ratios against selected asset classes, and have suitable measurement and management regimes in place to assess the risk. Similar schemes are now in place for the insurance industry (Solvency II) and the funds management industry. The Association of Academic School of Business (AACSB) Standards are also of the general certification class, as is ISO 9001 and similar quality assurance regimes for social responsibility, carbon conservation, and other correct attitudes.

Performance bureaucracies formally address concerns of accountability by ultimate stakeholders such as the governments or society in general, especially in contexts where market based discipline either does not apply or where there is an inherent proneness to market failures, such as financial system instability. But like more traditional bureaucracies, the end result is often far from optimal, and in many cases actually destructive of performance. Thus it has been argued that the Basle II capital adequacy regime actually stimulated bonus-seeking bank staff to look for loopholes, notably in connection with the regulatory capital for credit default swaps, leading indirectly to the credit securitisation explosion and the subsequent financial implosion (Bowden and Lorimer 2009). In the educational context, the AACSB quality assurance standards² themselves suggest that resources should be devoted by the Dean’s office in the pursuit of ‘strategic management standards’, while somewhat disingenuously disavowing a ‘bureaucracy of planning and mission creation’. In the event, the practical result is more often than not an academic corporate state, with an explosion of assurance procedures together with appointments to administer them. Indeed, in some Australasian universities the proportion of permanent staff actually doing teaching and research has fallen to less than 40%, overhead costs have soared, and capable research and teaching staff have emigrated to more relaxed pastures elsewhere. Parallel, and in support, of the performance bureaucracy itself is an official managerial language of quality assurance; some of the more relevant bits are codified in section 2.1 of the present paper. The language and the associated constructive content has spread more widely to the general educational literature³, so that schools as well as universities have to

prepare mappings from assurance rubrics to objectives and goals (see section 2.1). Performance bureaucracy is invasive.

Society does need accountability. But a necessary first step in fixing things is a clear understanding as to just how it all can go wrong. The present paper takes as its starting point the quality assurance metric framework, and the political economy of measurement and control constructed on such a basis. A fundamental informational difficulty is that managers design and implement assurance metrics that are not statistically sufficient for the desired goal outcomes; hence ‘metric insufficiency’. Goal outcomes depend upon a set of domain variables, which are descriptions of the operational capability of the organisation. In particular, the domain variables depend upon characteristics and attitudes of the workers themselves – the ‘coalface’. To the extent that performance metrics do not span the domain variables, there will be a set of hidden variables, in control theoretic terms. The hidden variables, including worker ‘buy in’ or the lack of it, mean that the controls based on the metrics can be counterproductive.

Managerial motivations themselves adversely impact on the domain, most notably in the form of key performance objectives that are tied to the exercise of controls rather than to substantive outcomes, leading to overrides or intolerance of staff dissent. In addition, resource allocation games among managerial goal centres will vacuum resources away from the coalface, becoming what is, in effect, the problem of the commons. If this is so, then latent domain components can be activated that may make it very difficult to learn from repeated cycles of measurement and control, especially where the goal measurement is itself only episodic. Finally, there is a symbiotic relationship between performance and traditional brands of bureaucracy. Requirements of the former lead to preoccupation with the prescriptive documentation and procedures of the latter⁴.

The scheme of the rest of the paper is as follows. Section 2 sets the organisational framework for subsequent discussion, using the language of performance bureaucracy. Section 3 turns to the econometric framework, casting the measurement problem as one of metric insufficiency and hidden variables. Section 4 turns to the decision theoretic framework that underpins repeated cycles of measurement and control. This is used in section 5 to set up a Nash cooperative game among senior managers. The solution is analysed in terms of the problem of the commons, latent variable effects, learning failures, and the cost-information trade-off. Section 6 offers some brief concluding remarks.

2. The organisational framework of measurement

Performance bureaucracies are commonly based upon a more or less formal framework of measurement and control. The structure is viewed as a necessary common basis for institutional discourse on how to make management more effective. By the same token, it therefore also a natural vehicle for a review as to how it can go wrong. The present section is a summary outline of this framework.

In common manifestations, the organisational structure of information and performance has a four tier structure, with its own semiotics; terms such as goals, objectives, metrics, mappings, rubrics, and granularity are italicised in what follows.

(Level 1): An institutional welfare or utility function (here denoted u)

This is left unspecified in what follows, except in so far as it is increasing in its arguments. It can be originated from a higher level, such as a board of trustees; by edict from a government department or international organisation; or in the form of a consensual interpretation (according to deputed agents) as to the common will or the common good. Note that the word ‘utility’ does not typically figure in the lexicon; it is used here for convenience and to correspond with general economic theory. The word *mission* is often used to connote something similar, e.g. in the general context of a managerial strategic plan; so institutional utility would be derived from fulfilling its mission statement.

The institutional utility function is a function of:

(Level 2): A vector of n primary *goal* values: $\mathbf{g} = (g_1, g_2, \dots, g_n)$.

Goals are measurable in some way. The measurement could begin and end with the organisation; one could describe this as ‘internally referenced’. In an ‘externally referenced’ system, goals are measurable with respect to some environmental outcome or by some external agency. This could be a market based outcome. Or it could be driven by some higher level administrative entity to whom our organisation reports under accountability provisions. Measurement at two such levels is in fact a common model. Thus Basle-based regulatory regimes measure economic capital as an interim goal but the ultimate goal is to prevent systemic meltdown, which is a market outcome. University administrators mount annual metric exercises for teaching or research effectiveness, but from time to time there are less frequent external references such as university rankings or PBRF research scores. An issue in such cases is the validity of the metrics used for internal goal scoring with respect to the external assessments. In what follows, the measurement process itself is presumed to be stable, meaning that there is no change in measurement criteria. However, one of the risks to be considered is that of a change in an external reference entity and/or its

measurement procedures. The management of each goal is devolved to its own management centre, headed by an executive with responsibility for that area; thus centre or manager G , responsible for goal value g .

Each goal outcome is in turn is regarded by its management as a function of a number of:
(Level 3) *Objectives* $b_j^g; j = 1, 2, \dots, j_g$.

These are intended to represent operational targets.

The achievement or otherwise of objectives is measured at any time by means of a number of *management metrics*; k_j of these are associated with objective j . They can collectively be assembled as:

(Level 4) *Metrics* $((m_{kj}^g; k = 1, 2, \dots, k_j; j = 1, 2, \dots, j_g; g = 1, 2, \dots, n))$.

The metrics m_{kj}^g are functions of the domain variables (see below), taking $X \rightarrow \mathcal{M}$. The metric outcome space \mathcal{M} is left unspecified (elements could be qualitative or categorical in nature) except to the extent that a preference ordering can be defined over the outcome space, so that managers know whether one metric outcome is better than another.

The assemblage of metrics is referred to in the management and education literature as a *rubric*, denoted M in what follows. Metrics, and metric performances, have to be *mapped* into objectives, which are in turn are mapped into goals. A mapping, in this sense, is a mathematical correspondence. Metric *granularity* refers to whether a given metric rubric is finer, more inclusive, or more informative than another. Alternative formal constructs for metric granularity are explored in section 3.4.

For present purposes, goal performance and metric measurements can be regarded as functions of two sorts of variables.

(a) Domain variables $\mathbf{x} = (x_1, x_2, \dots, x_d)$

Domain variables are organisational state variables. They are, or could be, affected in some way by current or feasible controls or other actions of management; hence one could call them managerial domain variables. They could encompass such things as employee quality or attitudes, alternatively equipment or facilities. The space of possible values of the domain state variables is denoted X and it is assumed that a σ -algebra \mathfrak{S}_x exists over this space. It will be assumed that managers do not have comprehensive or detailed knowledge of the domain variables, but attempt to represent their effects in a smaller number of metrics.

(b) Environmental disturbances ε_g for goal G. These are not affected in any way by management actions and can be taken as inherently unpredictable. In particular, they are supposed for present purposes to be statistically independent of the domain variables \mathbf{x} .

The full four tier system requires tensor notation to handle the three dimensions of variation in metrics and goals. However, the essentials can be illustrated with a simpler condensed system in which level 3 is omitted, so that metrics feed directly into goals, or alternatively each goal has just one objective. The metrics associated with goal G can then be assembled into a vector $\mathbf{m}^g(\mathbf{x})$, so that the rubric M is now two dimensional. In the educational management literature it is sometimes referred to as a ‘matrix’, though it need not be rectangular and there appears to be no implicit reference to linear operators on any underlying vector space. However, all the ideas that follow can be extended to the four tier structure with appropriate modifications of notation, sometimes indicated in parentheses as discussion proceeds.

Finally, the metrics can be functions mapping domain variables into a variety of spaces, not all of them continuous or complete; the results could be in terms of some categorical scale, for instance, as indeed might be the domain variables themselves. However, it will often be expositionally convenient to imagine that the metrics are continuous, and at times differentiable, functions on and into complete metric spaces, with the understanding that this is simply to illustrate general properties or directions.

3. Metric sufficiency and insufficiency

If managerial controls are going to be effective, they should be based on the best possible information about the domain variables. The goal values can also be influenced by external environmental influences that might not be predictable; but in an ideal situation, the managerial metrics should capture all the information about goal values that could have been predicted from the domain variables. This is termed metric sufficiency, for it is essentially statistical sufficiency adapted to the present context.

On the other hand, metric insufficiency means that managerial metrics do not span important influences that can drive goal outcomes. To the extent that this does not hold, there is a hidden variable problem (the ‘hiddens’, for brevity). The spanning properties of the hiddens can be formalised in terms of information theory, and related in turn to the granularity of alternative metric rubrics. The first task is to relate such constructs to the statistical theory of sufficiency.

3.1 Metric sufficiency

In what follows it will be assumed that there exists a σ -algebra $\mathfrak{S}_{g,x}$ generated on $\{g\} \otimes X$ together with a probability measure $P_{g,x}$. The development is based on classical sufficiency although the same ideas can also be given an interpretation in terms of Bayesian predictive sufficiency, in which g would be interpreted as a parameter and $P_{g,x}$ as a family of probability measures indexed by g (Furmanczyk and Niemi (1998)). In order to economise on notation, development in what follows is informal, though at times a little extra precision will be necessary.

Let $f(\cdot)$ stand for a generic probability or probability density. A collection of metrics $\mathbf{m}^g(\mathbf{x}) = \{m_k^g(\mathbf{x}); k = 1, 2, \dots, k_g\}$ will be said to be *sufficient* for goal (objective) G if and only if almost surely in \mathbf{x} and g ,

$$(1) \quad f(\mathbf{x} | g, \mathbf{m}^g) = f(\mathbf{x} | \mathbf{m}^g) .$$

A classical corollary is the Fisher-Neyman factorisation theorem $f(\mathbf{x} | g) = h(\mathbf{x})f(\mathbf{m}(\mathbf{x}) | g)$ where $h(\mathbf{x})$ is independent of g . For present purposes, however, a more informative approach is based on prediction, as to whether the existing set of metrics is effective in predicting goal values.

Definition: A set of metrics $\{\mathbf{m}^g\}$ for goal G exhibits predictive sufficiency if:

$$(2) \quad f(g | \mathbf{x}, \mathbf{m}^g) = f(g | \mathbf{m}^g) .$$

A consequence is that metric sufficiency is equivalent to predictive sufficiency. For if condition (1) holds (sufficiency), then

$$f(g | \mathbf{x}, \mathbf{m}^g) = \frac{f(\mathbf{x} | \mathbf{m}^g, g)f(\mathbf{m}^g, g)}{f(\mathbf{x}, \mathbf{m}^g)} = \frac{f(\mathbf{x} | \mathbf{m}^g)f(\mathbf{m}^g, g)}{f(\mathbf{x}, \mathbf{m}^g)} = \frac{f(\mathbf{m}^g, g)}{f(\mathbf{m}^g)} .$$

Conversely, suppose (2) holds (predictive sufficiency). Using Bayes' theorem on both sides followed by $f(\mathbf{x}, \mathbf{m}^g | g) = f(\mathbf{x} | \mathbf{m}^g, g)f(\mathbf{m}^g | g)$ yields expression (1), i.e. statistical sufficiency.

The purport of metric sufficiency is that once the metric outcome values are known, there is no further information about goal or objective values contained among the domain variables \mathbf{x} . It would endow managers with a degree of confidence that correctly designed controls based upon currently observed metric values would have the desired effect in terms of outcome goal values. In this respect, there is an analogy between metric sufficiency and completeness in general economic equilibrium. In a financial market analogy, one would be able to price g in terms of \mathbf{m}^g . In an incomplete market there are alternative prices for g because not all influences are spanned, and some other resolving criterion is necessary.

Remarks:

(a) The same idea can be extended to an institutional welfare function as a whole. If u is any such welfare function, then an assessment rubric is sufficient for \mathbf{x} if

$$f(\mathbf{x} | u, M) = f(\mathbf{x} | M)$$

(b) Metric sufficiency can be local instead of global. Thus let $\mathfrak{S}_{x_0} \subset \mathfrak{S}_x$; for instance, the sub-algebra could be generated by a region $X_0 \subset X$. Then in expression (2), one would replace the implied \mathfrak{S}_x by \mathfrak{S}_{x_0} . Thus it might be that for normal circumstances, the given metrics are sufficient, but controls designed on that basis might in fact drive the domain variables or their distributions to regions where the given metrics are no longer metrically sufficient for the desired goals.

3.2 Metric insufficiency and hidden variables

Metric *insufficiency* is where the available metrics do not span the effects of the domain variables on objectives or goal outcomes. This corresponds to a model in which there are a set of variables hidden from the observer. The structural form relates g to \mathbf{x} , but the manager either cannot identify this function; may not comprehend the full range of the domain variables; or finds it too expensive to measure them. So he or she relies upon a smaller set of metrics \mathbf{m}^g . In the resulting reduced form, metric insufficiency now means that there will be residual effects that are not spanned by the available metrics. It is therefore possible that controls implemented on the basis of observed metric outcomes could have quite unintended effects, because some of the domain variables have not been accounted for or even understood.

Metric insufficiency formally implies that \mathfrak{S}_m must now be supplemented to span the influence of the domain variables on the goals. Sufficiency might now hold with respect to $\mathfrak{S}_m \cup \mathfrak{S}_{\tilde{\mathbf{x}}} \subseteq \mathfrak{S}_x$. To give this a functional interpretation, we will assume that $\mathfrak{S}_{\tilde{\mathbf{x}}}$ is the natural σ -algebra associated with further set of variables $\tilde{\mathbf{x}}$ that are likewise functions of the domain variables, but provide further information about goal or objective values. In what follows, these are often called the hidden variables, as their existence or effect may be unknown or unobservable to the manager. ‘Hidden’ in this context does not necessarily mean the same as ‘latent’, a distinction explored in section 5.2. Goal sufficiency now applies to the extended set of variables $\{\mathbf{m}, \tilde{\mathbf{x}}\}$ which are equivalent in their probability implications for goal values to the complete set of domain variables $\{\mathbf{x}\}$. Note also that the hidden variables are generally specific to goals; thus $\tilde{\mathbf{x}}^g$ for goal G . One would expect that a variable hidden from one goal manager would likewise not be discoverable to another; but it may simply not be useful for the second

goal. Where there is no danger of confusion in the present section, we will often economise on notation by dropping the g superscripts where the reference to goal G is taken as understood i.e. $(\mathbf{m} \sim \mathbf{m}^g, \tilde{\mathbf{x}} \sim \tilde{\mathbf{x}}^g)$. However, in later sections it may be important to refer to the entire collection of hidden variables across different goals, so in more global contexts $\tilde{\mathbf{x}} = \{\tilde{\mathbf{x}}^g; g = 1, 2, \dots, n\}$ means the complete set.

The existence of a further set of variables, additional to the given metrics, is signalled by surprises in the goal values, relative to what the manager might have predicted based on the metrics alone. Ideally, variables that capture these surprises should be separable from the existing metrics, or representations should exist under which this is the case. In general,

$$(3) \quad f(\mathbf{x} | g) \sim f(\mathbf{m}, \tilde{\mathbf{x}} | g) = f(\tilde{\mathbf{x}} | \mathbf{m}, g) f(\mathbf{m} | g)$$

will represent the signalling value of the goal outcome as to the current state of the domain variables. Suppose that we can find a representation for $\tilde{\mathbf{x}}$ such that almost surely in g and \mathbf{x} ,

$$(4) \quad f(\tilde{\mathbf{x}} | \mathbf{m}, g) = f(\tilde{\mathbf{x}} | g),$$

so that

$$(5) \quad f(\mathbf{m}, \tilde{\mathbf{x}} | g) = f(\tilde{\mathbf{x}} | g) f(\mathbf{m} | g).$$

Condition (4) says that $\mathbf{m}, \tilde{\mathbf{x}}$ are conditionally independent of each other, given g , so that the given metrics \mathbf{m} provide no further information about the hiddens, once the value of the goal is known. A representation of the hidden variables exists such that the signalling effect of g can be factorised into independent effects as in expression (5). The converse is also true: if (5) holds then so does (4).

In this case, using representation (5), and Bayes' theorem, the conditional distribution of the goal value g , given \mathbf{x} , can be written as:

$$(6) \quad \frac{f(g | \mathbf{m}, \tilde{\mathbf{x}})}{f(g)} = \frac{f(g | \mathbf{m})}{f(g)} \times \left[\frac{f(g | \tilde{\mathbf{x}})}{f(g)} / \frac{f(\tilde{\mathbf{x}} | \mathbf{m})}{f(\tilde{\mathbf{x}})} \right].$$

Expression (6) says that the predictive value of the domain variables $\mathbf{x} = (\mathbf{m}, \tilde{\mathbf{x}})$ on the goal value decomposes into the predictive content of the metrics and hiddens separately; but following a metric priority rule, we have to discount by the content of the hiddens that could have been predicted from the metrics.

If we take logs and integrate both sides over the joint density $f(g, \mathbf{m}, \tilde{\mathbf{x}})$, expression (6) can be written in differential entropy terms as

$$(7) \quad I_{g \bullet x} = I_{g \bullet m} + I_{g \bullet \tilde{x}} - I_{\tilde{x} \bullet m} .$$

The total conditional information given by \mathbf{x} about g decomposes into the sum of conditional informations due to each of \mathbf{m} and $\tilde{\mathbf{x}}$, less the information given by \mathbf{m} about $\tilde{\mathbf{x}}$. The discount factor $I_{m \bullet \tilde{x}}$ is at a minimum (zero) when the hiddens can be specified to be statistically independent of the metrics, in which case the metrics and hiddens have independent effects that can simply be summed. In certain cases it is always possible to arrange this, as when the metrics are linear functions of underlying normal domain variables. Of course, so far as the manager is concerned, he or she cannot actually observe the hiddens, except via their indirect influence on realised goal values.

Expression (7) can be rewritten as

$$(8) \quad I_{g \bullet x} - I_{g \bullet m} = I_{g \bullet \tilde{x}} - I_{\tilde{x} \bullet m} .$$

This version can be given a control interpretation. The left hand side refers to the error made in designing controls for goal outcomes directed to improving metric values with no regard to possible unintended effects stemming from metric insufficiency. The magnitude of the error (in informational terms) is given by the right hand side; it is maximal when the metrics and hidden variables are statistically independent of each other and the hiddens have a substantial impact on goal values. One could potentially use (8) as a criterion for constructing a basis for the space of hidden variables, for an omniscient observer. First find the function that when added to the existing metric set minimises the informational error with respect to the complete domain. Then proceed sequentially adding such functions, at each step minimising the residual information error with respect to the domain as whole.

Absent any very strong priors as to the interpretation of the hidden variables $\tilde{\mathbf{x}}$, the latter are essentially arbitrary, except in so far as they contribute collectively to the prediction of goal values; their space can potentially be spanned by any number of alternative measure preserving bases⁵. In this respect there are successively stronger conditions: (a) the metrics cannot be separated from the hiddens in any way; (b) conditional predictive independence holds such that the information given by the domain variables factors into that due to metrics and hidden effects; and (c) a representation can be found in which the hiddens and metrics exhibit full statistical independence. In what follows, it will usually be assumed that representations of the hiddens exist under which conditional separation can be accomplished in the sense of (b). Within such a framework, there remain many degrees of freedom in defining the individual hiddens

themselves. But for purposes of subsequent discussion this turns out not to matter, for what is important is their collective effect on the goal values, and this should remain invariant to any particular choice of a basis for \tilde{x} .

Of course, this should not prevent the managers from trying to understand just how the residual impact is generated, which is halfway towards a structural representation of the hidden. The existence of substantial surprises in the revealed goal values should stimulate a reassessment of whether the existing metrics are fit for the intended purpose.

3.3. Metric granularity

A common response to adverse surprises in goal outcomes is to invest more time and money into refining the metrics. The general meaning of metric *granularity* in management jargon has reference to whether one set of metrics is more detailed than another. This would mean that if \mathfrak{S}_{M_a} refers to the natural σ -algebra generated by metric rubric A and \mathfrak{S}_{M_b} that generated by metric rubric B, then rubric A is more ‘granular’ than rubric B if $\mathfrak{S}_{M_b} \subset \mathfrak{S}_{M_a} \subseteq \mathfrak{S}_x$.

On the other hand, asking more questions does not always reveal additional useful information for predicting goal values. Granularity should therefore connote ‘effective’, as well as ‘more detailed’. In this sense, A is more granular than B for goal G if $I_{g \bullet m_a} > I_{g \bullet m_b}$, which in turn means that the amount of residual information left in the hidden is less for A than is the case for B: $\mathfrak{S}_{\tilde{x}_a^g} \subset \mathfrak{S}_{\tilde{x}_b^g}$. The point is important in considering the collective effect of different regulatory arrangements, e.g. along the lines of current US proposals for financial reform. Expanding the number of regulatory reporting centres may yield little or no new information actually useful for controlling systemic risk outcomes, but on the other hand, add considerable direct or indirect compliance costs to the regulatory control. Additional granularity has also been foreshadowed in announcements made in January-March 2009 by the Basle Committee concerning proposed revisions to the Basle II market risk framework.

4. Measurement and control

The effect of the hidden variables is that control outcomes are now uncertain. It is convenient to develop the theory of controls and their outcomes in terms of structural representations for goal values, which have to this point been implicit rather than explicit. In the extended setting, controls can be repeated in the light of metric results, possibly using a fresh set of metrics. But the cycle of measurement and control involves costs, leading into subsequent discussion on resource allocations.

4.1 Structural and reduced form representations

Goal outcomes may be viewed as the result of two influences, the domain variables and the environmental disturbances. The latter are supposed to be independent of the domain variables, stemming from the external environment or other influences beyond the potential control of the managers. They can be represented with a zero mean disturbance ε_g for goal G. Thus for goal G, an underlying structural relationship holds of the form:

$$(9a) \quad g = \varphi^g(x, \varepsilon_g).$$

The domain variables may not all be relevant for any given goal, or the relevance of specific variables may be latent. An alternative representation focuses on just what is relevant for goal G, namely the metrics, the hiddens, and the environmental disturbance specific to that goal. From the viewpoint of the manager of goal centre G, the existence of hidden variables \tilde{x}^g means that the complete domain X is unobservable, i.e. is not measurable with respect to the current set of metrics m^g . However for each goal centre G, there exists a transformation: $X \rightarrow \tilde{X}^g \times \mathcal{M}^g$ that is measure preserving conditional upon any outcome value g . As earlier mentioned, the best situation is where the associated variables can be defined in such a way that they are at least conditionally statistically independent, so that $\mathfrak{S}_{\tilde{x}^g|g} \perp \mathfrak{S}_{m^g|g}$.

The net result is equivalent to a reduced form representation cast in terms of the metrics and the hiddens:

$$(9b) \quad g = \varphi^g(m^g, \tilde{x}^g, \varepsilon_g).$$

Version (9b) will be used in what follows to analyse the impact of metrically based controls upon goal outcomes. Although the hidden variables are now specific to each goal centre, there could well be common effects that affect more than one goal centre.

In a dynamic context, the metrics themselves can be progressively revised as time goes on, with the aim of making them more discriminating or revealing more usable information. If this is to be achieved by asking more detailed questions, or measuring more attributes, the metric rubric is regarded as more ‘granular’, in the sense earlier described. There may be other effects that arise over time, as a result of a repeated cycle of measurement and control. Not all of them are necessarily beneficial.

4.2 Controls and control outcomes

Each management centre has one or more controls, denoted \mathbf{a}^g . These are vector valued objects defined implicitly in terms of their intended effects on observable metrics. The proximal purpose of any proposed control is to improve the metric values, under the supposition that this will in turn improve the goal outcome. This is therefore a metrically based control system. Over time the metrics can themselves be respecified (perhaps to achieve greater granularity), but at any given time the controls are based on the existing metrics.

Thus the controls are *adapted* to the given metrics, in the sense that a control \mathbf{a}^g is a function of observations \mathbf{m}_0^g on the currently (time 0) available metrics, where observations on the latter are in turn functions of the current domain state variables: $\mathbf{m}_0^g = \mathbf{m}^g(\mathbf{x}_0)$. For brevity it will be assumed that metric observations are not disturbed by the environmental variables ϵ_g .

The above corresponds to a devolved system in which the management centre for goal G sets its controls with regard only to its own metrics \mathbf{m}^g . In a fully coordinated system, the control for each management centre would be a function of the entire metric rubric M, over all goal centres. The objective in either case is to improve goal performance, to be achieved in the first instance via an improved value for the metrics, as they currently exist. The collection of controls at any time will be denoted by $A = \{\mathbf{a}^g; g = 1, 2, \dots, n\}$; adapting the management jargon, this could be referred to as the ‘control rubric’⁶.

The effect of the controls, and of their cost, is to drive the domain variables. Given the current domain state \mathbf{x}_0 , the outcome state is represented as:

$$(10) \quad \mathbf{x} = \xi(\mathbf{A}, \boldsymbol{\eta}_x; \mathbf{x}_0)$$

where the further set of variables $\boldsymbol{\eta}_x$ reflect exogenous movements in the domain variables that have no relationship to prior managerial actions (e.g. workers come and go for incidental reasons). For present purposes the $\boldsymbol{\eta}_x$ can be treated as random disturbances.

The set of domain states achievable with strategy rubric \mathbf{A} is denoted by

$$(11) \quad X(\mathbf{A}) = \{\mathbf{x} : \mathbf{x} = \xi(\mathbf{A}, \boldsymbol{\eta}_x; \mathbf{x}_0)\}.$$

Given the current state \mathbf{x}_0 , there is a transition probability distribution function on $X(\mathbf{A})$ induced by $\boldsymbol{\eta}_x$ and the controls, denoted by $F_x(\mathbf{x}; \mathbf{A}, \mathbf{x}_0)$. The resulting outcome for goal centre G is given by the structural representation $g = \varphi^g(\mathbf{x}, \epsilon_g)$, where ϵ_g refers to new (period 1) external

circumstances. In turn, this induces a marginal probability distribution function F_g for the goal outcome g .

The new domain state \mathbf{x} also induces an outcome $\mathbf{m}^g(\mathbf{x})$ value for the existing metrics, which in view of (11) can be written in terms of the controls as

$$(12a) \quad \mathbf{m}^g = \mathbf{v}_g(\mathbf{A}, \boldsymbol{\eta}_x; \mathbf{x}_0).$$

A special case is where the metrics for goal G are unaffected by control settings for other goals, except via the common cost channel as it affects the operational divisions:

$$(12b) \quad \mathbf{m}^g = \mathbf{v}_g(\mathbf{a}^g, C_d, \boldsymbol{\eta}_x; \mathbf{x}_0).$$

Here C_d represents the allocation of financial resources to the operational ‘coalface’; it is considered further below.

If the existing metrics are insufficient to span the new goal value, a new set of unspanned hidden variables $\tilde{\mathbf{x}}^g$ is defined implicitly by the reduced form:

$$(13) \quad g = \varphi^g(\mathbf{m}^g, \tilde{\mathbf{x}}^g, \boldsymbol{\varepsilon}_g).$$

Thus $\tilde{\mathbf{x}}^g = \tilde{\boldsymbol{\xi}}^g(\mathbf{A}, C_d, \boldsymbol{\eta}; \mathbf{x}_0)$, emphasising that managerial controls can generate a new set of hidden variables that will impact on goal value, in ways that are not predictable on the basis of the metrics alone. In fact, the result can be quite adverse if the controls and/or their cost activate latent domain components.

The above constitutes a single representative cycle of measurement and control. Starting with an initial metric system and metric values, management controls are designed and implemented. This generates new domain values as workers react to the controls. In turn, this generates new values, not only of the given metrics, but also of the hiddens. The resulting goal values may not be to the liking of management. So in a second and possibly subsequent cycle, measurement is again carried out, perhaps on a new set of more ‘granular’ metrics; and given the results, a new set of controls implemented.

4.3 *The cost of controls*

To each proposed control \mathbf{a}^g , there is a cost C_g which can be associated with implementing the control itself or in the measurement (metrics) necessary to implement the control and record its results. In general, the more granular the metric rubric, the higher the cost of management information and the higher also the costs of controls designed on its basis. In bank management,

for instance, Basle II was more granular than Basle I, and the announcements of January and March 2009 foreshadow more of the same. Likewise, the NZ PBRF scores earlier quoted required the construction of complete online systems so that staff could enter and edit their progressive publications into their HR space, and managers could execute database tabulations to identify problem areas. The sum of costs over the different goal centres is given by

$$C_v = \sum_{g=1}^n C_g, \text{ where the subscript } v \text{ stands in this context for 'management overhead'.$$

Management costs C_v are of the nature of overhead or indirect costs. In addition, there are direct or operational costs, which refer to the costs of paying and materially supporting line workers; for brevity the latter collectively referred to as the 'coalface'. Direct costs associated with the operations of the coalface are denoted by C_d . Thus if at any time total budget money is fixed as \bar{C} , then

$$(14) \quad C_v + C_d = \bar{C}$$

expresses the tradeoff between management and operational financial resources.

Managerial controls are by no means costless in their impact on the coalface operations. The costs can be reflected reduced dollar funding for operations, in terms of worker morale, or commonly both. In adverse scenarios, worker morale suffers because of compliance costs associated with controls, resentment at the time and resources wasted, or managerial careerism in their advocacy. One could use the monetary cost of management C_v as a proxy for such non-monetary effects, but in view of the constraint (14), this can be absorbed into coalface funding C_d ; the latter thereby encompasses both operational funding limitations and other discouragement effects.

5. Resource allocation and goal outcomes.

In what follows, each manager submits an action plan, which consists of a number of proposed metrics and controls based on the metrics. Bargaining on budgets take place at the senior management team and the outcome will be represented as the solution to a Nash cooperative game. Metric and control implementation may be re-jigged to accord with finally agreed budgets, so this is a budget-driven tatonnement process.

In a fully devolved system, managers bargain over the funding to go to their respective goal centres to support their metric and control activities. The operational coalface has no formal representation but their influence is (hopefully) reflected in acknowledgement by managers that if the coalface is starved of funding this will impact adversely upon their own goal targets.

However, some features are characteristic of the problem of the commons in public good analysis, notably the free rider problem. It is in the interest of each goal manager to maximise the funds available for his or her own administration (metrics, controls) while hoping for more of the burden of direct costing to be at the expense of other divisions. And what may appear minor to the individual manager adds up, over all the goal centres, to a major drain on time and resources for the coalface.

5.1 Nash cooperative allocations of managerial resources

The above considerations suggest a Nash cooperative solution to the devolution bargaining game. Managers have a direct utility function which in the first instance is a function only of their own goal outcomes:

$$(15a) \quad u_g = u^g(g); g = 1, 2, \dots, n.$$

Following Niskanen (1968, 1994), it might also be the case that managers have an additional motive to enlarge their own administrative division, following the general rule that the more staff reporting to the manager, the higher is his or her seniority and the greater his or her own remuneration. In addition, no self respecting manager would wish to be seen to be any less proactive than his peers. The amount spent is taken as a proxy for the degree of managerial activity in originating and administering metrics and controls. Such considerations might justify augmenting the utility function in the form

$$(15b) \quad u_g = u^g(g, C_g); g = 1, 2, \dots, n.$$

In either case it will be expositionally convenient to assume that the utility functions are differentiable in their respective arguments.

Given their understanding of the effects and costs of their metrics and controls, each manager is assumed to have an indirect utility function expressing the expected utility $E[u_g]$ from (15a) or (15b) as a function of the available funding to their centre and to the operational coalface:

$$(16) \quad v_g = v^g(C_g, C_d); g = 1, 2, \dots, n$$

Given a current state and expected utilities $\mathbf{x}_0, \{v_{g0}\}_{g=1, \dots, n}$, the Nash cooperative solutions for the centre budgets and the common budget for the coalface may be obtained as:

$$(17) \quad \arg \max_{\{C_g\}, C_d} \prod_{g=1}^n (v_g - v_{g0})^{\alpha_g} \quad \text{subject to} \quad \sum_{g=1}^n C_g + C_d = \bar{C} : C_g, C_d \geq 0,$$

where the α_g are political power weights with $\sum_{g=1}^n \alpha_g = 1; \alpha_g \geq 0$.

The solutions may be characterised in terms of the elasticities:

$$(18) \quad e_g = \frac{\partial \log v_g}{\partial \log C_g}, g = 1, \dots, n; \quad e_d^g = \frac{\partial \log v_g}{\partial \log C_d}.$$

It takes the form of expenditure shares for the budget centres and the coalface:

$$(19) \quad \frac{C_g}{\bar{C}} = \frac{\alpha_g e_g}{\sum_g \alpha_g (e_g + e_d^g)}; \quad \frac{C_d}{\bar{C}} = \frac{\sum_g \alpha_g e_d^g}{\sum_g \alpha_g (e_g + e_d^g)}.$$

5.2 Solution outcomes

The solution (19) can be used to examine the actual outcomes, which may well be quite different from the anticipated outcomes. Adverse managerial surprises can rise in several ways, for both the metrics and the hiddens are impacted by controls, which are proxied here by cost allocations.

Using expressions (13) and (15a), expand $u(g) = u^g(\mathbf{m}^g, \tilde{\mathbf{x}}^g, \varepsilon)$ around the current value

$u_0^g = u(\mathbf{m}_0^g, \tilde{\mathbf{x}}_0^g, 0)$. For brevity, write

$$\frac{\partial \mathbf{m}_g}{\partial \log C_g} = \beta_g; \quad \frac{\partial \tilde{\mathbf{x}}_g}{\partial \log C_g} = \gamma_g, \text{ with a similar definition for the coalface expenditures}$$

$$\frac{\partial \mathbf{m}_g}{\partial \log C_d} = \beta_d; \quad \frac{\partial \tilde{\mathbf{x}}_g}{\partial \log C_d} = \gamma_d. \text{ In both cases the terms } \beta^*, \gamma^* \text{ can be regarded as functions of } C_g,$$

C_d . Derivatives are evaluated at the current domain point $\tilde{\mathbf{x}}_0^g$.

Then to terms of first order,

$$(20) \quad v_g e_g \approx \beta_g \cdot \frac{\partial u_0}{\partial \mathbf{m}_g} + \gamma_g \cdot \frac{\partial u_0}{\partial \tilde{\mathbf{x}}_g},$$

where the subscript asterisk refers to either manager centre g or the coalface d . Indeterminacy in the choice of the $\tilde{\mathbf{x}}^g$ can be resolved by adapting the elasticities γ_g, γ_d to conform with any measure preserving alternative choice. What is important in the current context is the total effect

$(\gamma_g \cdot \frac{\partial u_0}{\partial \tilde{\mathbf{x}}_g})$, which should remain invariant.

Referring back to expression (19) for the equilibrium shares, suppose that the hiddens are defined in such a way that $\frac{\partial u_0}{\partial \tilde{\mathbf{x}}_g} > \mathbf{0}$; so that in a neighbourhood of the current point \mathbf{x}_0 , the

hiddens are intrinsically beneficial to goal G. However, the effect of spending designed to

improve metrics inhibits the expression of the hiddens, to the extent that

$\gamma_g'(C_g) < 0; \gamma_d'(C_d) > 0$. In this case,

$$\gamma_g' \frac{\partial u_0}{\partial \tilde{x}_g} < 0; \gamma_d' \frac{\partial u_0}{\partial \tilde{x}_g} > 0.$$

The optimal solution is therefore to reduce expenditure on goal centre G and redirect it to the coalface.

But in practice this may not be done. The first difficulty is that by definition the hiddens are either unknown, or at best only partly appreciated as to their nature and effects. A second is that managers might convince themselves that any possible negative effects are more than outweighed by the positive value of metric enhancement (the first right hand term in (20)). A third aspect is that mentioned in connection with specification (15b), as to the direct utility value of spending to goal centre managers. The problem in this respect is that the coalface is not represented in the senior management team. Or if so, the representation is chosen or otherwise motivated to be accommodating, or self selected to be that way because of the costs of complexity in getting involved.

To the extent that the first of the above holds, the existence of possible adverse hidden effects constitutes a source of uncertainty in managerial policy. The goal values that result are inconsistent with those predicted on the basis of metrics. This may be true even where perfect metric predictability exists, i.e. where managers can design and execute controls to achieve exactly the intended result in terms of the metrics.

In standard control theoretic situations it may be possible to resolve such problems over time by means of learning algorithms. Continual repetition would lead to adaptive learning of the impact of controls, and operational identification of the hiddens in the form of the associated error structure in a regression of goal outcomes on metrics. But in the current context there are a number of reasons why adaptive learning algorithms may be ineffective, even where the external environment does not change.

- (a) The cycle of control and results may be limited in frequency, for example if goals are externally assessed only every five years.
- (b) There may be an identification problem where the domain state variables are affected by multiple goal setting: a poor performance in achieving one goal might be blamed on controls in search of a better performance on another. Managers, in other words, can blame each other.

- (c) Latent expression of domain variables. Up to some critical point, the coalface workers may be unaffected by revisions in controls. But at some point the latter become so expensive and intrusive that latent domain attitudes change to frustration or resentment, and goal outcomes are at that point adversely impacted. This is similar to a sudden change of phase in physical systems. Some of the hiddens are latent variables, only expressed as a result of poorly designed controls insufficiently observant of coalface welfare.
- (d) Measurement and control could result in active avoidance behaviour, in which staff or members being controlled seek to take advantage of loopholes in the regulatory regime, or ambiguities in the measurement process. The net result is that goal outcomes can be perverse, certainly relative to the intended effect. As mentioned in the Introduction, the Basle II capital adequacy regime arguably encouraged bank managers in search of enhancement to take advantage of grey areas in directives on economic capital measurement, with disastrous effects for the system being regulated.

5.3 The cost-information trade off and other welfare effects.

If the results from current metrics and controls are a frustrating surprise, a possible managerial strategy is to increase the granularity of the metrics, and hopefully thereby reduce the free dimension or predictive value of the hiddens. However, there is a cost to doing so, expressed as a rise in C_g and thereby a fall in C_d . Even if the exercise was achievable with no additional direct monetary cost, there will typically be indirect costs in terms of coalface motivation, or avoidance activities.

The problem is exacerbated by the problem of the commons. While each manager individually thinks of his or her own expenditure as a relatively minor and affordable diversion from the operational budget, the total effect is in fact much more damaging. Suppose that all goal centres are allocated equal office expenditures so that $C_g = \frac{1}{n} C_v$. Then if $g = g(C_g, C_d)$, and $C_d = \bar{C} - C_v$, it follows that:

$$\frac{dg}{dC_g} = g_1 - ng_2$$

Managerial spending that only looks minor to each manager becomes multiplied up in the coalface commonality, so that in fact each goal performance is exposed to the full amount of the sum collectively diverted for the purposes of management metrics and control.

Other problems may stem from adverse incentive effects of the metrics. Managerial metrics imposed at the level of the individual worker commonly also amount to implied certification. If otherwise productive workers feel that their own personal certification is

incorrect, then they may lose morale, or leave altogether to seek what is in effect a second opinion on their personal merits. This can result in adverse self selection of staff. An example was the personally individualised PBRF research ratings of NZ universities, which was recorded in their HR file. If the employee disagreed with his or her rating, a common result among younger staff was emigration; in effect, to seek a second opinion at institutions overseas, especially if theirs was a more intellectually demanding area or discipline in high demand offshore.

The problem of incorrect metrics in general is that they can induce inappropriate control policies designed to reinforce those metrics. This may pay off well for internal goals, but lead to poor results for external or higher level goals. Thus the payoff for higher PBRF ratings was additional government funding for the university as a whole, and therefore for managerial support. But the consequences in term of international university ratings have been unfavourable. This is an instance in which the institution has some monopoly power⁷ on the output side but faces market forces for its key inputs.

On the other hand, information is also valuable, for without it, managers cannot pursue even minimally adequate supervision and control. Two points can be made in this connection. The first is that managerial metrics may be an unnecessarily formal way to derive information. They are often instituted in response to another common management imperative, namely a demand for reproducible evidence, as distinct from word of mouth. But in fact much of the evidence thus gathered and interpreted is little more than a codification of somebody's opinion, which may or may not be particularly expert⁸. So the costs of the metric process itself are therefore a response to the formalisation of evidence. In that respect, performance bureaucracies inherit the old preoccupations with documentation. If this is seen as necessary for external goal verification (higher level 'meta' goals), the costs are inevitable. But it may not be for internal goal setting. The first issue is therefore whether less formal means of information gathering can be effective. A second point is that internal agency problems may arise in respect of information transmission up the line. Bonus seeking line managers may be apprehensive of passing along accurate information about coalface disaffection, because of a fear that its nature reflects adversely on their own management or ability to inspire the coalface with 'buy in' to goal centre initiatives.

Assuming that the costs of metrics and control are inevitable, there remains the problem of where to stop. Referring back to expressions (18-20), the cost response elasticity e_g will initially be rising, collectively enough to outweigh any adverse effect e_d arising from coalface

expenditure. But as the ‘granularity’ of metrics and control rises, the costs of compliance and complexity rise to the point where further spending on $\{C_g\}$ or its aggregate C_v leads to falling goal values $\{g\}$ and hence overall institutional utility.

Thus at some point, the marginal value of investment in metric granularity and control turns negative. But the optimum point is at best fuzzy because of informational difficulties. An adverse goal outcome in the current measurement round could simply reflect the need for additional metric granularity, and more precisely directed controls, predicated on the results. But in reality could be that coalface has by this stage become fed up with the reporting and compliance together with its cost externalities. In econometric terms, latent variables are activated, that are hidden from the point of view of goal performance.

6. Concluding remarks

Bureaucracy has its own political realities. When times are good, the civil service expands; when times are bad, the budget cuts start and the coalface suffers. When a new government is elected, a new bureaucracy is required to implement its campaign promises, no matter that one of those promises might have been to curb excessive spending on the state itself. When things go wrong with the economy or as now, with the financial system, a resort to the bureaucracy of reporting, regulatory reporting and compliance is the first point of refuge, no matter that the real problem might be one of system design. The end result in all such cases is greater complexity. And one of the costs of complexity is a tendency for workers to opt out from participation and input into the processes of management. To adapt the jargon, ‘stakeholders’ become ‘disengaged’ at a time when their goodwill and input may be badly needed.

The present contribution covers just one aspect of the political economy of bureaucracy, namely the interaction between the problems of measurement, budgetary games, and control. Yet measurement is at the heart of a performance bureaucracy, for a demand for evidence is very much a part of it. It is manifested in a distrust of professional judgment in favour of implied certifications that in reality are no more than someone else’s judgement. The demand for certification itself is a product of the performance bureaucracy, driven by bonus seeking managers and their key performance initiatives.

Nor is market discipline necessarily the right answer, though it might well help in abolishing sources of monopoly that help to sustain public bureaucracies. As the subprime crisis has shown us, the free market can be far from disciplined, though there are those that claim that the crisis itself reflected avoidance activity or loopholes in respect of the Basle II regulatory

regime. In many circumstances there does have to be compliance and control, and some form of informational feedback is necessary. Weberian rules and Tayloristic strategic plans have their place, provided that they are flexible enough to cope with changing times and revised information. What is needed is a rigorous cost benefit analysis as to just why the information is needed; what form it should take; how the information is gathered; what forms of bias or metric insufficiency might result; and if the information is required in pursuit of certification, whether the certification itself is really needed. Performance bureaucracies are high risk games.

References

- Bowden, R.J. (2008) *The Economic State of the Nation*: Wellington NZ: Kiwicap.
- Bowden, R. and D. Lorimer (2009) 'Credit securitisation and system stability', Working paper, Institute of Finance, Ulm University and Kiwicap Research Ltd.
- Breton, A. and Wintrobe, R. (1985), *The Logic of Bureaucratic Conduct*, 2nd. ed., Cambridge: Cambridge University Press,
- Chow, Y.S. and H. Teicher (1988), *Probability Theory: Independence, Interchangeability, Martingales*, Berlin:Springer.
- Crozier, M. (1964). *The Bureaucratic Phenomenon*, London: Tavistock Publications.
- Downs, A. (1964) 'A theory of bureaucracy', Rand Corporation Research Report.
- Draper, H. (1979) *Karl Marx's Theory of Revolution, Volume 1: State and Bureaucracy*. New York: Monthly Review Press.
- Furmanczyk, K. and W. Niemi (1998) Sufficiency in Bayesian models, *Applicationes Mathematicae*, 25,1,113-120.
- Gao, J. (2007) *Nonlinear time series: Semiparametric and non parametric methods*, Monographs on Statistics and Applied Probability, London: Chapman & Hall.
- Huber, J.D. and C.R. Shipan (2002) *Deliberate Discretion: The Institutional Foundations of Bureaucratic Autonomy*, Cambridge: Cambridge University Press.
- Lindblom, C. E. (1983) 'The Science of 'Muddling Through'', pp. 265-271 in: Richard J. Stillman II (ed.) *Public Administration: Concepts and Cases*. Boston: Houghton and Mifflin.
- Lindblom, C. E. and E. J. Woodhouse, *The Policy-Making Process*, Englewood Cliffs, NJ: Prentice Hall, 1993.

- McCubbins, M. D., R. G. Noll, and B. R. Weingast (1987) 'Administrative procedures as instruments of political control', *Journal of Law, Economics and Organization* 3, 2, 243-277.
- Miller, G. J., (1992) *Managerial Dilemmas: The Political Economy of Hierarchy*, Cambridge: Cambridge University Press.
- Niskanen, W. A. (1968), 'Nonmarket Decision Making: the Peculiar Economics of Bureaucracy', *American Economic Review, Papers and Proceedings*, 53, 293-305.
- Niskanen, W. A. (1994), *Bureaucracy and Public Economics*. Aldershot: Edward Elgar.
- Pierre, J. (ed.) *Bureaucracy in the Modern State*. Aldershot: Edward Elgar, 1995.
- Sappington, David E. M. (1991) 'Incentives in Principal-Agent Relationships' *Journal of Economic Perspectives*, 5, 2, 45-66.
- Selznick, P. (1948) 'Foundations of the theory of organization', *American Sociological Review*, 13(1), 25-35. Also in Shafritz, J. M & Ott, J. S. (Eds). *Classics of Organization Theory*. Orlando, FL: Harcourt Brace.
- Selznick, P. (1996) 'Foundations of the Theory of Organizations'
- Tirole, J. (1994) 'The Internal Organization of Government', *Oxford Economic Papers*, 46, 1-29.
- Tullock, G. (1965) *The Politics of Bureaucracy*, Washington, DC: Public Affairs Press..
- Vachris, M.A, (2003) 'Principal-Agent Relationships in the Theory of Bureaucracy' in C. K. Rowley and F. Schneider (ed.) *The Encyclopedia of Public Choice*, New York: Springer.
- Von Mises, L. (1944). *Bureaucracy*, New Haven: Yale University Press.
- Watson, Tony J. (1980). *Sociology, Work and Industry*, Routledge.
- Weber, Max, 'Bureaucracy' in Shafritz - Ott (eds.), *Classics of Organization Theory*. pp. 81-87.
- Weber, M. (1952) 'The Essentials of Bureaucratic Organization: An Ideal-Type Construction', pp. 19-27 in Merton, R. K. (ed.), *Reader in Bureaucracy*, New York: The Free Press, 1952.

Endnotes

¹ The difference is perhaps one of degree rather than kind. In more traditional state bureaucracies, if benefit claimants or tax payers manage to rot the system because the information requested of them is insufficient, then the organisation is not doing its job. So raising more forms to be filled corresponds to correct administrative procedure, and to that extent might possibly be taken to indicate satisfactory performance by the office itself. On the other hand, counting the number or length of different forms to be filled in is unlikely to meet modern demands for measurement validity for organisational performance.

² AACSB ‘Eligibility Procedures and Standards for Business Accreditation’, Revised July 1 2009.

³ The internet is full of such education references; see e.g. http://www.education-world.com/a_curr/curr248.shtml on rubrics for teaching. Similar operating rubrics and mappings have been adopted by university managers in NZ.

⁴ Instructive in this respect are the Committee for European Banking Supervisors (CEBS) Guidelines for external credit certification, intended as the book of words for a ‘common basis application pack’. *Guidelines* amounts to 40 pages of testing procedures, ‘mappings’, and codes of conduct. The final document is dated 20 January 2006, more than a year prior to the subprime meltdown and credit crunch. Back testing, it seems, should be applied to the certifiers as well as the certified.

⁵ Consider the simple deterministic model: $g = x_1 + x_2$; $m = 2x_1 - x_2$. This is a structural interpretation, according to which the goal value depends equally upon the two domain variables x_1, x_2 while the metric depends positively on the first but negatively on the second. One can span g in terms of the x_1, x_2 by adding a new variable of the form $\tilde{x} = \theta_1 x_1 + \theta_2 x_2$, to give a reduced form representation

$g = \frac{1}{2\theta_2 - \theta_1} [(\theta_2 - \theta_1)m + 3\tilde{x}]$. Here m, \tilde{x} are jointly spanning, but it is apparent that the sign of m can

be anything we please, depending on just how the hidden spanning variable \tilde{x} is defined, in this case by choosing the values of θ_1, θ_2 . Some further criteria is needed; for instance that x_1, x_2 are statistically independent, which is equivalent to defining \tilde{x} as the residual in a linear regression of g on m .

⁶ The adaptation is not ‘managerially correct’: management actions are supposed to be consensual and workers are not control objects.

⁷ In NZ the monopoly power derives from the fact that student and research funding are both effectively available only for NZ registered universities, access to which is fixed by the Minister of Education under the provisions of the Education Act 1985.

⁸ Examples in the case of NZ universities are the almost exclusive use of student evaluations to assess teaching expertise; and for PBRF research purposes, success in competitive Marsden Fund applications for funding.