

The longer-run performance effects of agencification: theory and evidence from Québec agencies

AIDAN R. VINING

Beedie School of Business, Simon Fraser University, Canada
E-mail: vining@sfu.ca

CLAUDE LAURIN

HEC Montréal, Canada
E-mail: claude.laurin@hec.ca

DAVID WEIMER

Robert M. La Follette School of Public Affairs, University of Wisconsin-Madison, USA
E-mail: weimer@lafollette.wisc.edu

Abstract: Although governments worldwide are increasingly choosing to deliver services through organisations with greater autonomy than traditional bureaus, the implicit assumption that such agencification contributes to long-run efficiency remains largely untested. Agencification gives agency managers more autonomy and access to incentive mechanisms that lead to greater efficiency if they are not offset by inefficiencies resulting from managerial discretion. We test the hypothesis that agencification improves efficiency by examining the longer-run performance of 13 agencies in the province of Québec, Canada over approximately 10 years. We find that these agencies experienced long-term productivity gains, but that these gains reached a plateau over the time period studied. In addition, we describe changes in several measures of performance. A survey of the managers of these agencies indicates that they perceive agencification as having a substantive impact, but worry about the sustainability of autonomy and their capacity to show continued gains in measured performance over time.

Key words: agencification, corporatisation, efficiency, new public management, productivity, quango

Introduction

Over the last decade, there has been a global trend away from the use of traditional governmental bureaus to deliver services. The most common feature of these reforms has been an increase in the autonomy of agencies

and the introduction of more “business-like” organisational structures, language, accounting procedures and incentive systems. A variety of labels have been used to describe this restructuring process, including “agencification”, “corporatization” and even “quangocratization”. (Hereafter, for convenience, we describe *both* the process and the outcome as agencification.¹) The OECD (2005, 114–115) estimates that such agencies now account for 50% or more of public expenditures and employment in some countries. For example, in Sweden, such agencies now dominate government service provision (Jacobsson and Sundström 2007). The trend does not appear to be abating.

The canonical feature of agencification is that governments retain formal ownership and politicians retain hierarchical control of nonetheless more autonomous organisations. As it remains a creature of government, the agency’s primary organisational goal remains something other than profit maximisation. Therefore, it is distinct from privatisation where the formal organisational goal becomes profit maximisation.² Agencification, however, does usually involve more emphasis on both operational efficiency and enhanced revenue generation (Bertelli 2006, 232). Furthermore, the strategic direction and day-to-day management are often entrusted to managers that are not part of the civil service. As a result of these characteristics, such agencies are somewhat awkwardly balanced between government and market (Greve 1999; Bertelli 2006). What effect does this tension have on their performance? Most importantly, what is the *long-run* impact of agencification on performance?

Although there is empirical evidence suggesting that agencification does produce short-term beneficial impacts, that evidence does not address the longer-run impact question because it is based on cross-sectional data or very brief time series. We address this gap by examining the performance of 13 agencies in the province of Québec, Canada. We examine multiple dimensions of performance for these agencies using a variety of measures over approximately a 10-year period following agencification.

To summarise our empirical results, we find that these agencies showed *sustained* improvement in a number of performance measures. Looking most closely at labour productivity, we find that it increased annually following agencification, although at a declining rate, such that performance plateaued

¹ Henceforth, also, when we use the term “agency”, we mean an entity that has been subject to the process of agencification.

² Some governments have created separate agencies as a precursor to privatisation, because some form of it is usually required before privatisation can take place. Many governments have adopted agencification as a preferred governance mechanism and do not regard it as a precursor to privatisation (Bilodeau et al. 2007; Lægread and Verhoest 2010).

with all gains realised within about 10 years. These findings are substantively and statistically significant.

What is agencification?

There are many ways to categorise governmental organisational designs, or more specifically, the architecture of government ownership and oversight. An initial problem in analysing agencification is the absence of a bright theoretical line separating the resulting agencies from other institutional forms. These agencies are almost never as autonomous as state-owned enterprises (SOEs), but they are usually more autonomous than traditional line bureaus that perform similar tasks (Vining 2011). Lægneid and Verhoest (2010, 4) exemplify this balancing act in their definition of agencification. On the one hand, they argue that an agency is “a structurally disaggregated body, formally separated from the ministry, which carries out public tasks at a national level on a permanent basis, is staffed by public servants, is financed mainly by the state budget, and is subject to public legal procedures”. On the other hand, they also argue that agencies “are not totally independent, because executives normally have ultimate political responsibility for their activities”. Other research identifies a similar tension (Richards and Smith 2006; Verschuere 2007). Lægneid and Verhoest’s definition is quite restrictive in view of the observed evolutionary diversity among agencies (Florio and Fecher 2011), especially cross-nationally. For example, many agencies are not financed “mainly by the state budget” because they have legislative authority to collect fees that at least cover their operating costs (Bertelli 2006; Vining 2011). In some countries, the managers and workers may be “public servants”, but they are not part of the government civil service. In addition, at least in nations with federal systems, many of these agencies operate at the sub-national level.

The efficiency hypothesis

A number of management scholars, applied economists and management consultants have posited that the major rationale for agencification is to improve “efficiency”. The expectation of improvement in efficiency is largely based on new public management (NPM) ideas. The core prescription that relates to agencification is “establishing an operating environment for a government organization which replicates the internal and external conditions of successful private enterprises” (Nicholls 1989, 27). These conditions usually include giving agencies narrower task domains, clearer

mandates, greater access to higher-powered incentives and more autonomy from political interference (Walsh 1995; Kettl 2005; Pollitt et al. 2007).³

Critics have retorted that the “be more like private enterprises” metaphor is simplistic; this has been famously summarised in the aphorism that private sector firms and public sector agencies are “fundamentally alike in all unimportant respects” (Sayre 1958, 102; see also, e.g. Wilson 1989; Hargrove and Glidewell 1990, but see Boyne 2002). Other critics have more specifically questioned the intellectual coherence of the NPM (Hood 1990; Goldfinch and Wallis 2010; Reddy et al. 2011). They raise a fundamental dialectic. On the one hand, NPM appears to be optimistic about the *intrinsic* motivation of public servants, as is public interest theory, because it seeks to “set managers free” to achieve public goals. Yet, on the other hand, NPM scholars also emphasise higher-powered *extrinsic* rewards and punishments, unlike public interest theory. In this regard, NPM appears to be heavily influenced by principal-agent theory (Dixit 2002; Burgess and Ratto 2003; Vining and Weimer 2005). NPM has also been criticised on the more practical grounds that the relative importance of the various recipe ingredients is unclear. Læg Reid and Verhoest (2010, 1), for example, focus on narrower agency task domains: “New Public Management (NPM) assumes that task specialization results in efficiency gains”. Other scholars focus on the potential benefits of agency and managerial autonomy in allowing managers to pursue the public interest freed from political interference (Pendlebury and Karbhari 1998; Talbot 2004). There is empirical evidence that agencification, in its various forms, does actually result in greater autonomy (Verhoest et al. 2004; Læg Reid et al. 2005; Verhoest 2005; Painter et al. 2010).

From a principal-agent *theory* perspective, however, performance improvements from structural reform must fundamentally flow from either reduced moral hazard (hidden action) or reduced adverse selection (hidden information) or some combination of the two (Dixit 2002; Burgess and Ratto 2003; Bertelli 2006). In practice, it is usually not possible to distinguish precisely between moral hazard and adverse selection in their causal contribution to “agency loss”. However, both can, in theory, be reduced by better alignment between managerial rewards and specific performance outcomes – in other words, a shift to more high-powered incentives in

³ There are, of course, other potential rationales for the creation of agencies. In our view, however, they do not plausibly predict efficiency improvement. It is beyond the scope of this analysis to discuss these hypotheses. The authors can provide interested readers with a table that summarises what they consider to be plausible inferences regarding the performance of three of these rationales: a “political commitment” hypothesis, an “institutional isomorphism” hypothesis and a “political economy” hypothesis.

conjunction with narrower task domains that accommodate more measurable mandates (Frant 1996; Hyndman and Eden 2002). In addition, agencies enjoy some formal insulation from political actors. In turn, this also increases managerial autonomy within the agency, particularly the autonomy of the chief executive (CE) who now functions more like a private sector CE officer. Autonomy could improve agency performance by allowing the CE and other managers to act in the public interest, that is, to create greater public value with less political interference. From a principal-agent perspective, however, the problem is that greater autonomy also affords managers greater freedom to engage in inefficient self-interested behaviour (even if only “the quiet life”).

In sum, agencification – clearer incentives and targets for the CE and employees, increased transparency around managerial behaviour and performance and greater pressure on CEs to deliver results – should lead to improved performance if it is not overwhelmed by the countervailing effects of more self-interested behaviour flowing from increased agency autonomy. The net effect of these opposing forces on performance is theoretically ambiguous, as summarised in Figure 1 (adapted from Bilodeau et al. 2007) by the question mark (?). If the net impact is positive, we would expect the rest of the flow in Figure 1 to unfold as in the lower part of the figure and ultimately result in improved performance.

The efficiency hypothesis clearly has both positive and normative content. In its positive, managerial NPM guise (which is of most interest for thinking about performance effects), the hypothesis posits that performance will improve following agencification, as viewed from the principal’s perspective. In addition, from most normative perspectives, it is hard to argue with more output, lower cost or more output at a lower cost. Therefore, the empirical evidence on the performance effects of agencification is crucial. As we discuss below, there *is* some emerging evidence of changed behaviour and improved performance in the short run.

Yet, an important question (holding constant for the moment the problem of appropriate performance measurement) is whether agency managers will continue to deliver services more efficiently over a longer time horizon. Harding and Preker (2000, 19) observe that (although without presenting empirical evidence): “In some cases performance has improved. But often these improvements have not been sustained”. Their finding is consistent with scepticism that public managers are still ultimately subject to political control rather than to market discipline. Thus, here the concern is that not enough has *fundamentally* changed because agencification does not “depoliticize decisionmaking in a sustainable way” (Harding and Preker 2000, 19). Why might this be the case?

First, consider principals. It takes considerable discipline for political principals to continue to monitor performance and the metrics that

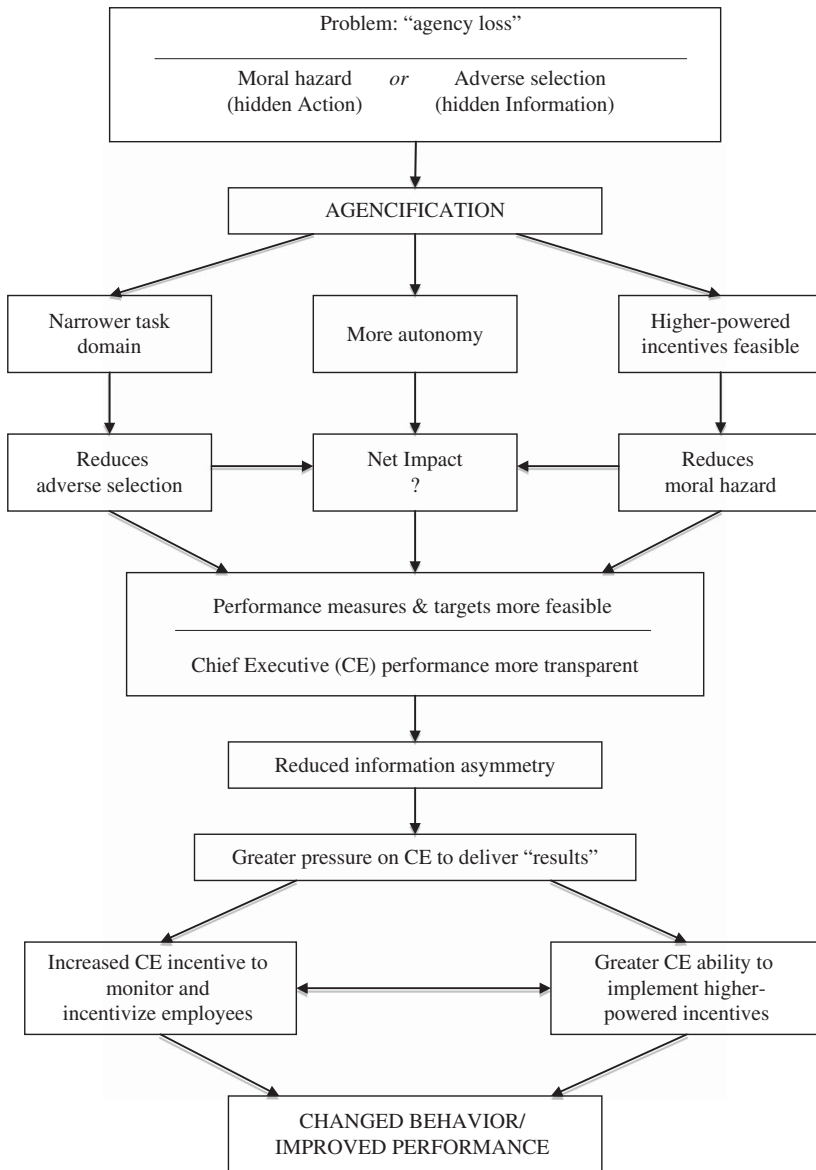


Figure 1 Summary of information and incentive changes with agencification.

appropriately incentivise performance. Although some politicians and senior civil servants may care about efficiency and effectiveness, eventually the attention of these monitors will shift to more pressing political concerns

(Koop 2011). In addition, even initially altruistic politicians are likely to demand other outputs when efficiency conflicts with their electoral needs. This version of the null hypothesis posits that, in the long run, the principal-agent problem is more likely to reside with (political) principals than with (managerial) agents.

Second, consider agents. The motivation of agents subject to high-powered incentives is problematic in the public sector. Burgess and Ratto (2003, 298) summarise these reasons well:

This is due to aspects like multi-tasking, multiple principals, the difficulty of defining and measuring output, and the issue of the intrinsic motivation of workers. In these circumstances the theory predicts that low-powered incentive schemes are optimal and task assignment and work organisation become crucial in promoting better performance and may sometimes be substitutes for high-powered financial incentives.

These caveats suggest that agencification might produce a short-term improvement in performance that fades gradually after the change of status: a “Hawthorne-like” effect (Mayo 1933; Levitt and List 2009). This fade might emerge over time as politicians distance themselves from agencies that engage in activities that could cost votes. Agencification could be useful, because it provides politicians with the appearance of distance and absence of control, while retaining the option to exercise control covertly or even explicitly in a politically salient emergency. For instance, governments frequently create gambling agencies – for example, the Georgia Lottery Corporation and the British Columbia Lottery Corporation (Jensen 2003). These agencies are attractive because they are usually a major source of revenue for government, but most politicians are nervous about the potential accusations of promoting immoral behaviour and hypocrisy – fostering gambling addiction, family dysfunction, indebtedness and the like. Similarly, tolling authorities (and other revenue-generating agencies with some degree of monopoly power) are often structured as stand-alone agencies. Politicians can protest that they oppose the rate increases, lament that they have no power to stop them and pocket the revenues.

To foreshadow the empirical analysis, it is important to emphasise that the appropriate performance metrics for agencies with monopoly pricing power are both complex and controversial. Agencification is likely to provide managers with more discretion to exercise any monopoly power that the agency possesses. However, success at pricing above social marginal cost is not the best way to judge the social value of these agencies (Vining 2011). Thus, measures that capture this, such as revenue increases, expenditure increases or revenue-expenditure margin increases, are normatively suspect. Nonetheless, some governments creating agencies are likely to

regard such revenue increases and related metric changes as one form of performance improvement. In addition, agencies with both monopoly power and enhanced autonomy may well fall prey to greater technical or X-inefficiency (Leibenstein 1966). An increase in this kind of inefficiency would not be captured by revenue, expenditure or revenue-expenditure margin changes. Indeed, increases in revenue, expenditures or margins might enable this kind of inefficiency. To address this issue, cost and productivity measures are required. In sum, for those agencies with some degree of monopoly power, an assessment of performance improvement will be dependent on the appropriateness of the specific metrics. As we discuss in detail later, our strategy is to use an extensive set of measures across the agencies that provide a comprehensive picture of the behaviour and performance changes following their creation. Most importantly, we explicitly examine productivity change, as this is socially valuable under all plausible circumstances. Having provided direct evidence on this normatively unambiguous measure, one can remain agnostic, or reject, a change on any other particular individual metric as really representing improvement (Boyne 2003a).

As we explain further below, our primary hypothesis is that agencification improves performance as measured by productivity. In addition, we hypothesise that performance improvements continue for many years following agencification. However, we also test a secondary hypothesis that performance gains decline over time until a performance plateau is reached.

Existing (and related) empirical evidence on performance effects

The global empirical evidence on the performance effects of agencification is quite limited. Therefore, it is useful to also examine the empirical evidence on performance change in two closely related contexts: (1) changes in bureau autonomy that do not amount to agencification and (2) changes in bureau or agency incentives (and motivation). The relevance of evidence of performance change following changes in autonomy or incentives is clear: agencification increases formal autonomy and provides more opportunities to employ more high-powered, or at least more extrinsic, incentives.

Autonomy and performance

A manifest purpose of many government reorganisations is to increase the autonomy of line bureaus from political influence and hierarchical bureaucratic control. However, evidence on the *performance* effects of autonomy is not extensive. Lewis (2003, 158) has shown that, at the federal level of the United States (US) government, certain kinds of autonomy

(insulation) improve the chances that an agency will survive over time; there is also evidence that politicians delegate autonomy to agencies to insulate desired policies when political uncertainty is high (Lavertu 2013). Concerning more conventional measures of performance, Moynihan and Pandey (2006, 122) review the extant empirical evidence and conclude: “There is some evidence that clear goals and bureaucratic autonomy are clear predictors of public sector performance”. However, also based on a review of the evidence, Rainey and Steinbauer (1999, 16) argue that the evidence suggests a non-linear relationship: “Government agencies will be more effective when they have higher levels of autonomy, but not extremely high levels of autonomy”.

There are only a small number of empirical studies that examine the impact of autonomy on line bureaus (Poister et al. 2011). Wolf (1993) analysed 44 federal agencies in terms of the impact of autonomy (among other variables) on performance. In aggregate, he found that higher levels of political autonomy had a positive effect on performance. Moynihan and Pandey (2006) examined a large sample of information system managers in state-level primary health and human services agencies and concluded that a higher level of “managerial authority” is associated with better performance. Langbein (2009) used survey data from US government employees to examine the relationship between employee perceptions of discretion and productivity. Although her results show a complex, contingent relationship between these variables, her overall conclusion is: “there appears to be a trade-off between accountability (to the executive) and productivity: executive political controls reduce both discretion (presumably raising accountability) and productivity (106)”.

The case study evidence on the insulation of agencies and programmes from *political* influence also generally concludes that increased autonomy results in better performance, although the performance measures have been quite variable (Borins 1998; Carpenter 2001; Barzelay and Campbell 2003; Katyal 2006; Stephenson 2008). In sum, there is *some* evidence that autonomy improves performance and little evidence that it worsens performance.

(Higher-powered) incentives and performance

A contentious issue in public management concerns the effect that individual and group incentives have on public sector performance (for reviews, see Moynihan 2008; Perry et al. 2009; Heinrich and Courty 2010; Langbein 2010). Theory suggests that “pay for performance” (PFP) is likely to be efficacious primarily where output is easy to measure (Lazear 2000; Lazear and Shaw 2007). However, as Langbein (2010, 12) warns, “in complex modern organizations, tasks are complex, and team-based; individual

performance is hard to observe and hard to link to firm profits...when performance is hard to measure, PFP is not even used in the competitive private sector, where efficiency is the unintended by-product of self-interested actors". The crucial question is whether extrinsic rewards "crowd out" intrinsic rewards (Frey and Jegen 2001; Le Grand 2003; Perry and Hondegem 2008). Suffice it to say that "it remains quite open what to expect regarding the relationship between performance and pay" (Binderkrantz and Christensen 2011, 33).

In the US, early empirical research centred on the Job Training Partnership Act of 1982 (Heckman et al. 1996; Courty and Marschke 1997). These studies generally found considerable "gaming" of what were, initially, very blunt incentive schemes. Heinrich (2007) examines the "high performance bonuses" introduced by the US federal government and the incentive effects of the Workforce Investment Act, the largest US employment and training programme. She concludes that the incentive system was so badly designed that any positive motivating effects would have been very surprising. Heinrich (2007, 281) concludes, "the results of the theoretical and empirical investigation suggest that high performance bonus systems are more likely to encourage misrepresentation of performance and other strategic behaviors than to recognize and motivate exceptional performance and performance improvements".

There have also been a number of European empirical studies. Walker and Boyne (2006) provide tentative evidence that (non-financial) incentives (and greater goal clarity) did result in improved performance in local governments in the United Kingdom. However, multiple reforms were introduced simultaneously, complicating the interpretation of specific causality. Kelman and Friedman (2009) studied the introduction of hospital-level financial incentives intended to reduce wait times. They report dramatic improvements in wait times and no evidence of dysfunctional effort substitution or gaming (but see Bevan and Hood 2006 and Bevan 2010). Burgess et al. (2010) study the performance effects of the introduction of both team-based and individual incentives to tax collection in the United Kingdom. They find that team-based incentives raised productivity (even though the teams were quite large – over a hundred members) and that individual incentives raised both the "tax-yield" and productivity. The main source of productivity improvement came from reassignment of more efficient employees to the incentivised tasks. Binderkrantz and Christensen (2011) study the impact of executive performance contracts on over 60 Danish public agencies using 2000, 2005 and 2008 data. (As they only purport to study the effect of PFP on performance, we do not treat this as an "agencification and performance" study per se). They find no substantive impact on performance as a result of these incentives.

In sum, the empirical findings suggest that extrinsic incentives, especially financial incentives, are quite variable in their impact. The one constant appears to be gaming by employees (i.e. strategic responses). This raises the possibility that more sophisticated incentive designs might produce better outcomes. It is also worth emphasising that, in almost all public sector contexts, financial incentives have been low relative to base salary, so that the price effect has been small (Weibel et al. 2010). There is some evidence that intrinsic incentives do work.

Agencification and performance

Obviously, the most relevant evidence directly examines the effect of agencification on performance and it is very limited (Talbot 2004, 105; see Verhoest and Lægread 2010 for a review; see Pollitt and Dan 2011 for a meta-analysis of empirical evidence). Boyne (2003b) specifically notes the lack of (time-series) evidence, and Pollitt and Dan (2011, 32–33) emphasise the paucity of empirical studies that convincingly address productivity change.

In an early study, Shirley (1999) generally found that “agencification” improved performance in developing countries. However, it is unclear what exactly agencification means, as her study covers a number of disparate countries; in many cases, the changes would appear to be better described as SOE reform. Several studies that examine what is described as agencification in China have the same problem (Aivazian et al. 2005; Xu et al. 2005). Brewer (2004) analysed quite highly aggregated evidence from 25 OECD countries. His work supports the idea that a portfolio of “agencification-like” reforms do improve agency performance. Bilodeau et al. (2007) studied 11 agencies at both the federal level in Canada and at the provincial level in Québec and found that the change to an agency form did improve performance for the three years following agencification. Quenneville et al. (2010, 158) analysed 16 agencies in Québec over a 5-year period and concluded that “average annual financial performance across agencies clearly improved over this period”. Cambini et al. (2011) find that (two different versions of) agencification of Italian bus entities resulted in a reduction of production costs. In aggregate, the evidence does suggest improvement, although only based on short time frames.

Nelson and Nikolakis (2012) provide the only extant study of longer-term performance of which we are aware. They assess the agencification of six Australian state-level forest SOEs over the period 1989 to 2007. The changes increased managerial autonomy and encouraged clarity in the definition of goals, and can therefore be considered agencification, albeit from a starting point of much greater political independence than a

traditional bureau. Nonetheless, they report improved performance and profitability over the period.

Likely performance effects of agencification

We conclude from the literature review that: (1) the *long-run* effects of agencification are largely unknown, although there is some minimal evidence showing some performance improvement; (2) there is some evidence showing *short run* improvement from agencification; (3) the evidence suggests that some degree of autonomy improves performance; (4) the evidence on incentives is mixed, but higher-powered extrinsic incentives do not improve performance unless well-designed; and (5) higher-powered intrinsic incentives appear to have some positive impact, although the evidence is not very strong.

In view of this weak and mixed evidence, our specific hypotheses relating to performance change are tentative. However, we conclude that our primary maintained hypothesis is that agencification improves performance, primarily as measured by productivity. In addition, we hypothesise (H1) that annual performance gains continue for many years following agencification. We base this hypothesis on the idea that, while the CE could implement some productivity-improving or cost-cutting changes quite quickly, other operational or “cultural” changes could take a number of years to fully implement. However, we also test a secondary hypothesis (H2): the magnitudes of any incremental performance improvements decline over time. Even if performance does improve for years following privatisation, there is no strong theoretical rationale to expect that these performance improvements will continue to accumulate indefinitely. Indeed, a more plausible assumption is diminishing marginal returns to agencification: after some years, the effects of agencification will “peter out” completely.

Sample selection, data and research design

We examine the consequences of agencification in the province of Québec, Canada over an extended period of time. The final sample consists of 13 agencies for which long-term time series data are available. Appendix 1 summarises the basic information on these 13 agencies, including the name, *de facto* year of creation, last year of data availability, number of full-time equivalent employees (FTEs), agency annual expenditures as of 2008 and a summary of the agency’s main programme or services.⁴ These data were collected from their annual reports up to 2009–2010. In examining

⁴ For an account of Quebec’s process leading to the creation of agencies, see Chapter 1 in Mazouz and Leclerc (2008).

performance change, an important issue is potential bias in the sample selection. We first address this issue and then outline our data and research design.

Sample selection and potential bias

An initial screen of agencification candidates suggested 21 entities created in 2001 or before that date. They basically fell into two categories. First, between 1995 and 2000, during “an experimentation phase” (Mazouz and Tremblay 2006), the Quebec government created five “Autonomous Service Units”. The *Public Administration Act* (2000) formalised their status as autonomous agencies. Second, after the passage of this Act, the Quebec government created 16 more agencies that were required to negotiate and commit to an “accountability and performance convention” (*Convention de Performance et d’Imputabilité*) with an oversight Ministry. We attempted to collect data on all 21 of these entities. However, we could not include three agencies, because they did not disaggregate their performance data from those of their oversight Ministries (*Centre Québécois d’Inspection des Aliments et de Santé Animale*, *Contrôle Routier* and *Géologie Québec*). We could find no clear statement as to why their accounts were handled in this way. In addition, we could not construct a plausible primary measure of output for three entities (*Centre d’Expertise Hydrique*, *Forêt Québec* and *la Régie du Logement*). Finally, we had to drop two more agencies: *Sécurité du Revenu*, because it merged with another agency in 2006, and the *Centre de Signalisation du Québec*, because it was privatised in 2007. As a result, their time series was truncated and we could not validate their data.

Therefore, we include all 13 agencies with sufficient data. Thus, we eliminate some obvious sources of bias. There are, however, some remaining areas of concern. A review of the agencies’ mandates shows that most share the feature that, following agencification, they would have fairly narrow task domains, although we found no clear statement that this was a specific rationale for their agencification. Similarly, consistent with a narrowing task domain, most of the selected entities would have a single primary tangible output that was reasonably easy to measure. However, this was not universally the case, as we had to exclude three agencies because of lack of relevant output data.

Therefore, our sample is biased to the extent that it mostly includes agencies that share the two features of narrow task domain and a primary tangible and measurable output. Our theoretical discussion around Figure 1 suggests that agencies with these features would be most amenable to performance improvement. Furthermore, our reading of government

documents suggests that the Quebec government was concerned with improved efficiency, as it implemented a formal results-based management system during the same period. Thus, our sample is not random in one sense – it probably is biased towards agencies that are more likely to improve performance than would be a randomly selected sample of government line bureaus. Thus, our findings apply most directly to bureaus, or separable parts of bureaus, with narrow task domains. Parenthetically, we also note that this may well turn out to be a common problem in empirical agencification studies, because this kind of change limits the use of straightforward “before-after” comparisons.

Data and research design

The starting point for our analysis is financial data that all agencies disclose in their annual reports. These data include annual revenues (where applicable) and expenditures and the number of FTEs. For this first component, we also identify a measure for the primary output of each agency.⁵ Appendix 1 also reports our selected measures of output. Using the financial statement data, we compute a number of ratio measures that assess the agencies’ financial performance, the average cost of their outputs and their overall rates of labour productivity over about a 10-year time period.

The financial data underlying each measure needs to be adjusted to take inflation into account. The use of nominal dollars (especially over an extended period of years) would overstate the magnitude of change. We convert revenues and expenditures for all agencies for all years to 2003 (Canadian) dollars using the Consumer Price Index – all items (CPI). The CPI is available from the Statistics Canada website (www.statcan.gc.ca).

As each agency produces different outputs, our output measures for agencies are in different units of analysis. For statistical description and analysis to be meaningful, we normalise these data using a standard methodology (see Boardman et al. 2002, 147). Specifically, for each agency in year t , the normalised measure is equal to the ratio of the year t observation to the value of the measure in the base year. It is important to note that, using this normalisation method, each agency is given *equal* weight in the sample (similarly, see Bilodeau et al. 2007) regardless of size. Thus, this approach normalises for size discrepancies as well as for measurement in

⁵ In all cases, the main output was reported for all periods. Only three agencies reported other output measures over the entire period: BIA (one other measure – correlation coefficient of 0.87 with the main output), CGER (one other measure – correlation coefficient of 0.82 with the main output) and EQ (two other measures – correlation coefficients of 0.96 and 0.86 with the main output). Thus, the only other measures that the agencies viewed as sufficiently important to regularly report were highly correlated with their main measures.

different output units. This is the appropriate approach for an examination of the effect of agencification *as an organisational form*. We use only the normalised data in further analysis. We then calculate the annual change in the normalised values for the relevant measures for each agency.

In the first component of our research design, we test hypothesis H1 by computing confidence intervals around the geometric means of normalised values for each standard annual report performance measure. If the confidence interval excludes 1.00, then we conclude that the performance indicator has significantly changed over the period of study. This provides a rough first test of the efficiency hypothesis H1 as well as an overview of other potentially relevant outcomes of agencification. The second component seeks to assess the impact of agencification on performance by estimating a panel regression model of the measures computed from the financial statements of each agency. The regression model allows us to estimate the pattern of performance change over time and explicitly test the efficiency hypotheses, H1 and H2. The third component reports results of a survey that we sent to agency personnel following our initial statistical analysis. In addition to providing a check of the data, it allowed us to elicit executives' views about the role of autonomy in achieving performance improvements.

Results

Overview of financial performance

Table 1 shows the average change in the standard agency measures over the entire study period. It reports the end-of-period geometric mean (appropriate in view of normalised data) of the normalised performance measures and tests whether it is significantly different from one with a probability of type I error set at a 5% level. The changes (increases) for revenues, output and labour productivity are all statistically significant at the 5% level, while the improvement (increase) in the revenue-expenditure margin is statistically significant at the 10% level. We regard the *increases* in total output and labour productivity as *unambiguously* representing performance improvement, as neither measure is subject to the exercise of monopoly power concerns discussed earlier. Both total output and labour productivity increase over the entire study period. We regard the *decrease* in average cost as also *unambiguously* representing performance improvement, especially in a context of generally rising expenditures and revenues. These aggregate results are consistent with the efficiency hypothesis H1.

Table 1 is also informative about the subset of agencies that generated revenues. All of these agencies generated an average annual increase in revenues during the entire study period. The average annual increase in

Table 1. Annual change in performance measures normalised to base year (base year: agency creation year)

		BIA	CARRA	CCQ ¹	CEAEQ ²	CGER	CPF	LSJML	RC	SAG ²	AFE	CR	EQ	RRQ	Average
Normalised revenue change	End of period	1.568	1.917	1.969	7.044	3.307	3.662	4.541	1.115	1.098	-	-	-	-	2.42*
	Annual growth (%)	6.88	7.75	5.88	24.09	12.94	12.85	18.05	2.33	1.99	-	-	-	-	10.30
Normalised expenditure change	End of period	1.205	1.940	1.169	1.044	3.512	1.498	1.186	1.313	1.327	2.128	0.687	0.902	1.268	1.36*
	Annual growth (%)	3.99	7.90	1.30	0.47	14.01	3.61	1.66	3.59	3.36	10.55	-4.54	-1.63	2.78	3.62
Normalised revenue/expenditure change	End of period	1.301	0.988	1.684	6.751	0.942	2.444	3.830	0.849	0.828	-	-	-	-	1.65
	Annual growth (%)	4.54	-0.12	4.64	23.15	-0.52	9.28	16.19	-0.62	-1.27	-	-	-	-	6.14
Normalised FTE change	End of period	1.241	1.215	1.109	0.948	1.371	1.364	1.135	1.059	0.886	1.437	0.692	0.848	1.115	1.090
	Annual growth (%)	2.73	2.24	1.28	-0.42	2.99	2.73	1.33	0.85	-1.70	5.59	-4.36	-2.38	1.24	0.93
Normalised output change	End of period	1.978	1.624	1.675	1.424	1.840	2.831	1.011	1.822	1.024	0.991	0.625	1.561	1.339	1.42*
	Annual growth (%)	10.95	5.54	4.02	3.68	6.01	8.52	0.98	7.74	1.95	-0.06	-5.55	8.94	3.30	4.31
Normalised average cost change	End of period	0.609	1.195	0.698	0.733	1.908	0.543	1.172	0.721	1.296	2.147	1.098	0.578	0.946	0.95
	Annual growth (%)	-2.93	2.23	-2.23	-2.49	6.92	-4.35	2.47	-2.47	3.99	10.68	1.37	-7.97	-0.49	0.36
Normalised productivity change	End of period	1.594	1.336	1.210	1.224	1.343	2.076	0.891	1.721	1.506	0.690	2.528	1.841	1.201	1.40*
	Annual growth (%)	8.13	3.34	2.35	2.74	3.03	6.19	0.08	7.57	8.01	-3.64	14.36	11.61	2.09	5.07

Notes: ¹FTE change calculated from 2001 to 2010 owing to data limitation.

²FTE change calculated from 2001 to the end of period (2010) owing to data limitation.

BIA = Bureau des infractions et amendes; CARRA = Commission administrative des régimes de retraite et d'assurances; CCQ = Centre de conservation du Québec; CEAEQ = Centre d'expertise en analyse environnementale du Québec; CGER = Centre de gestion de l'équipement roulant; CPF = Centre de perception fiscale; LSJML = Laboratoire de sciences judiciaires et de médecine légale; RC = Régie du cinéma; SAG = Service aérien gouvernemental; AFE = Aide financière aux études; CR = Centre de recouvrement; EQ = Emploi Québec; RRQ = Régie des rentes du Québec; FTE = full-time equivalent employees.

*Significantly different from 1.000 at the 0.05 level.

revenue is quite substantial – more than 10%. Table 1 demonstrates that annual revenues increased at a faster rate than did expenditures (3.6% on average). Revenues also increased faster than output (~4.3% on average). The individual agency results also point to the conclusion that performance on the measures tended to improve over time. Across other dimensions of performance, however, the results demonstrate some variability. For those agencies that do not generate revenues, results generally show performance improvement. However, AFE experienced a substantial increase in expenditures and FTEs, along with a decrease in output and productivity, at least when including its unaudited data for the first two years of its operation.

Assessment of performance (efficiency hypotheses)

The central empirical question is whether or not agencification affects long-run productivity. To answer this question, we analyse the data for the 13 agencies over the multi-year period. The results of these panel regressions appear in Table 2.

The dependent variable for each model is the normalised productivity by agency and year. The primary independent variable is the time since the beginning of the agencification in years.⁶ Because of the normalisation of productivity, the coefficient of time can be interpreted as an annual percentage rate of growth. Model 1 provides the basic analysis. In addition to time since agencification, it includes fixed effects for both firms and calendar years. It also includes the square of time to allow for the possibility that the change in productivity is itself changing over time. It thus provides a way of testing hypothesis H2.

The coefficients of both time since agencification and its square in Model 1 are statistically significant. Substantively, the effects of time and its square must be interpreted together. Specifically, productivity gains appear greatest immediately after agencification and decline over time until falling to 0 in about 11 years ($\partial \text{Productivity} / \partial \text{Time} = 0.088 - 0.0078 \times \text{Time}$). Consequently, consistent with hypothesis H2, it appears that there are substantial and fairly long-lasting productivity gains following agencification, but that these gains eventually dissipate so that productivity plateaus.

Model 2 adds the variable FTE, the number of full-time employees in thousands. All else equal, larger agencies show smaller gains in productivity. However, some caution is required because, as the productivity involves division of the agency output measure by FTE, the negative relation may

⁶ Estimating the models with the annual difference in productivity as the dependent variable yields a comparable result: annual efficiency gains decline towards 0 over approximately 10 years.

Table 2. Panel regression analysis of productivity: 13 agencies for up to 13 years

	Model 1 (Full)	Model 2 (Size)	Model 3 (Pre-2008)	Model 4 (GDP)
Time (since agencification)	0.088* (0.028)	0.096* (0.019)	0.11* (0.029)	0.13* (0.039)
Time squared	-0.0039* (0.0015)	-0.0035* (0.0012)	-0.0059* (0.0019)	-0.0034* (0.0013)
Real GDP (billions dollars)	-	-	-	-0.0084 (0.0045)
FTEs (thousands)	-	-0.88* (0.27)	-0.83* (0.28)	-0.88* (0.25)
Constant	0.95* (0.15)	1.2* (0.21)	1.2* (0.21)	3.3* (1.1)
Calendar-year fixed effects	Yes	Yes	Yes	No
Agency fixed effects	Yes	Yes	Yes	Yes
Observations	133	133	115	133
F or Wald statistic	3.03*	3.76*	2.49*	13.46*
R ² (within)	0.30	0.37	0.27	0.33

Notes: GDP = gross provincial product; FTEs = full-time equivalent employees.
Two-tailed significance tests: *p < 0.05.

represent a correlation purely by construction.⁷ Nonetheless, the time pattern of productivity appears similar to that found in Model 1: large initial gains gradually decline towards 0. In this model, the gains reach 0 somewhat later – at about 14 years.

The last two years of the study period saw a worldwide recession. To investigate whether the declining pattern of improvements was an artefact of the recession, we re-estimate the model excluding the observations for 2008 and 2009. Model 3 presents the results from re-estimation of Model 2. The curvilinear relationship continues to hold. We find similar results (not shown) in replicating Model 1.

To assess whether the results are being driven by economic activity in general, we replaced the calendar-year fixed effects with the annual real gross provincial product in billions of dollars. Its coefficient was not statistically significantly different from 0 and the pattern of productivity gains did not change.

In summary, the results of the analysis are consistent with both the hypothesis that annual gains in productivity persist over an extended period (H1) and the hypothesis that the magnitude of these annual gains decreases over time (H2).

Although we are confident that these agencies improved their performance following agencification, our research design does not allow us to rule out the possibility that these agencies would have improved similarly even if they had not been agencified. To investigate this possibility, we would have needed data covering a comparable period for a set of comparable agencies that were not agencified. If these comparable agencies showed similar performance improvements, then the improvements we observed in the agencified agencies would likely be more attributable to (unobserved) broader institutional reform than to agencification. Unfortunately, because of the lack of comparable data from other Québec public bureaus, we could not construct a plausible comparison group of these bureaus against which to compare the results of the agencies (Mazouz and Tremblay 2006). However, we are able to compare the change in the FTEs of the 13 agencies to that of Québec's overall public sector FTEs over approximately the same time period. This comparison shows that the Québec public sector FTEs increased over the whole period, while the agencies in our sample showed virtually no change in overall FTEs. Indeed, they showed substantial declines in the latter part of the study period. Over the period 2000 through 2009, total annual average public sector full-time equivalent employment in Québec *increased* by 26% while it *decreased* by

⁷ FTE has a statistically insignificant coefficient when the dependent variable is first differences; the effects of time since agencification remain unchanged.

56% for the 13 agencies we studied. Thus, the productivity of employees in bureaus would have had to increase substantially to produce the same productivity gains as realised by the agencies. Specifically, output by bureaus would have had to have gone up about 76% over the period to match the average 40% increase in productivity achieved by the agencies – an implausibly large number.⁸

As a further way to investigate whether the agency gains might have occurred without agencification, we compared the employment data of the agencies to that of the three largest SOEs in Quebec, including the electric utility (*Hydro-Québec*, HQ), the lottery corporation (*Loto-Québec*, LQ) and the enterprise responsible for selling alcoholic beverages (*Société des Alcools du Québec*, SAQ). These SOEs are also autonomous but have a clear mandate to generate income for the government.⁹ These SOEs employ, respectively, more than 20,000, 6,000 and 5,000 FTEs. Although their goals are complex (and beyond our scope to discuss), they do have incentives to monitor employee productivity. If there were an underlying trend towards reducing FTEs to improve productivity in the Quebec public sector, these three SOEs would almost certainly reflect it. Our analysis of these three SOEs, however, found that, between 2000 and 2009, the aggregate number of FTEs increased by more than 17% (11.5% for HQ, 13.9% for LQ and 54.4% for SAQ). Furthermore, using revenue per employee as a measure of productivity, we found that productivity actually declined: by 15.2% for HQ, 17.7% for LQ and 7.1% for SAQ. In summary, the SOEs saw reductions in productivity, and the public sector as a whole would have had implausibly large output gains to show comparable increases in productivity to the agencies. Consequently, we conclude that there was no secular trend in increasing productivity and that the improvements in performance that we measure are most plausibly attributable to agencification.

Survey results

Using the performance measures data, we summarised each agency's performance in a report that we gave to each CE. We asked the CEs to confirm the accuracy of the data, answer a brief questionnaire regarding their agency's results and provide their overall responses to the results and, more generally, the impact of agencification.¹⁰ CEs in 11 of the 13 agencies

⁸ As a rough calculation: $[(q_1/1.26)/(q_0/1)] = 1.40$ implies $q_1 = 1.76q_0$, or a 76% increase.

⁹ In Quebec, SOEs have more autonomy than agencies, including a separate board of directors.

¹⁰ These reports and the questionnaire are available from the authors upon request (in French).

responded to the survey.¹¹ Our exchanges with each agencies' personnel provided us with an opportunity to make a number of adjustments to the data and, in one case, to expand the database.¹² Table 3 summarises these survey results.

The survey asked six questions. For each one, the CEs were asked to mark their agreement with a statement on a five-point Likert scale and also to provide open-ended comments on those statements. Question 1 asked respondents if they thought our results accurately portrayed the performance of their agency. A total of 7 out of 11 respondents thought that the results correctly portrayed their agency's long-term performance. However, a number of respondents argued that the results required some context if they were to be useful in assessing their overall performance accurately. Question 2 asked about our selection of the agency's primary output measure. In a number of cases, the CEs expressed some uneasiness at our selections and commented that they would have preferred that we examine all outputs. However, as already noted, the only additional measures that they consistently asked for were highly correlated with the primary measure. Therefore, use of multiple measures would not have materially altered our findings. Question 3 asked about their managerial "autonomy" on three dimensions: budget surplus management (question 3a), human resource management (question 3b) and other aspects of managerial autonomy (question 3c). A majority of the CEs responded that their autonomy to manage budget surpluses and human resources was too restricted. Their comments suggest that they perceived their autonomy had decreased over time as the Québec Treasury Council imposed both salary and hiring freezes. Question 4 asked about their ability to mandate more high-powered "incentives". A number of the respondents argued that these kinds of incentives were not appropriate for their agencies, either for executives or for employees. Question 5 asked them about the role of performance measures in agency management and control. Almost without exception, the respondents' concurred that those performance measures laid out in annual reports are actually used to assess the performance of their agency. Question 6 asked about the use of other performance measures beyond those in the annual report. A majority of the CEs said that they do use other measures, and some provided examples of these measures.

¹¹ AFE and RC did not respond to the survey. Both said they did not have the time or resources to check the data.

¹² Although CPF has not published an annual report since 2007, agency managers did provide access to internal data for 2007–2009. As the information is not publicly available, however, our analysis did not use these data.

Table 3. Summary of agency survey responses

Questions (paraphrased from French)	Average score (Scale 1 to 5)	Per cent agreeing (4 or 5)
Performance assessment		
Q1: Do results accurately portray agency performance?	3.3	64
Q2: Choice of primary output appropriate?	3.6	64
Satisfaction with agencification		
Q3a: Do you have sufficient budget autonomy?	2.5	30
Q3b: Do you have sufficient human resource autonomy?	2.1	9
Q3c: Do you have sufficient autonomy over other domains?	3.8	64
Q4: Authority to mandate high-powered incentives?	1.7	18
Usefulness of performance indicators		
Q5: Are reported performance measures actually used?	4.3	91
Q6: Are other performance measures used?	4.1	82

Note: A total of 11 responses to all questions except Q3a, which had 10 responses.

In aggregate, and consistent with our empirical findings, the survey results suggest that the implementation of NPM-like reform *has* had a substantive impact on these agencies. Yet, managers do worry about the sustainability of autonomy and about the reductionism inherent in numeric measurement of goal achievement. Despite the performance improvements we document, agency managers are somewhat skeptical of the sustainability and depth of reform *over time* (similarly, see Mazouz and Tremblay 2006). Managers appear to fear the presence of a Hawthorne effect.

Discussion and conclusion

The increase in total output, the decrease in average cost and especially the increase in labour productivity suggest performance improvement from agencification, at least in the Québec institutional environment. Our formal test of the efficiency hypothesis focusing on labour productivity indeed shows an initial increase in productivity following agencification. However, the annual gains appear to decline towards 0 over time such that the agencies reach a productivity plateau. Thus, we conclude that, despite some scholarly scepticism about the sustainability of performance improvement (which is to some extent shared by Quebec agency CEs!), agencification *does* result in longer-run improved performance in these agencies, though these performance improvements peak after about a decade. Hence, our conclusion is that the “?” in Figure 1 represents a net positive effect on performance.

A number of explanations can be offered for the long-run plateau in performance. There may indeed be a sort of “Hawthorne effect” such that both the initial increase in the oversight of managers and their performance decline over time. It may also be that the managers’ perceptions of declining autonomy are accurate, so that their opportunities to drive productivity gains decline. An alternative explanation, however, is simple declining marginal returns: over time, managers take actions that exhaust performance improvement. Of course, these explanations are not mutually exclusive. In any event, sorting them out would require long term, detailed case studies of agency behaviour.

There are five caveats to this conclusion. First, we do not have an ideal experimental design – governments simply do not create pairs of bureaus and then agencify one of them. Before–after comparison is rarely possible because of scope and incentive changes that occur at the time of agencification. In spite of these design limitations, we are confident that the agencified agencies showed productivity gains. We have to rely on indirect evidence to conclude that agencification *caused* the gains because we do not have a close comparison group during the agencification period. Further, our results do not imply that agencification will produce productivity gains in *any* bureau. Rather, the results apply most appropriately to bureaus like those agencified in Québec, that is, bureaus with relatively narrow task domains and a measurable primary output.

Second, one can always argue about what the “long run” really is when assessing performance. Certainly, in the long run, we are all dead. In this sense, all life is a Hawthorne effect. The CEs’ perceptions that their autonomy has decreased over time may foreshadow a reversion to the previous status quo performance. Obviously, tracking agency performance over even longer-term time frames would be informative and should be a major focus of future research. Nonetheless, the performance of this set of agencies appears to have improved over a long enough period to support the claim of long-run productivity improvement.

A third caveat is whether agencification is enabling the exercise of monopoly pricing power at the same time that it is generating performance improvement. The fact that aggregate agency revenues grew more quickly than did aggregate primary outputs suggests that this is a legitimate concern. It is not likely, however, to concern political principals, or their central budgetary control agents, who will almost universally regard these particular changes as representing performance improvement. Nonetheless, this issue merits further research. It also suggests the potential value of a “regulatory constitution” that mandates a goal of social welfare maximisation and social marginal cost pricing. The chances of implementing such a constitution are slim, however, because revenue-hungry governments care about more than social welfare.

A fourth caveat is the “black box” nature of the agencification recipe (Pollitt and Dan 2011). In view of the variability in recipes, we do not know which ingredients are critical. We also do not know whether agencification permits or fosters other organisational and human resource changes, such as better employee recruitment, that might have led to better performance over time (Huff 2011). Indeed, we cannot be sure if there *are* critical ingredients; rather, it may simply be: “somebody is finally paying attention to us!” In other words, is there an extended Hawthorne effect?

Fifth, a final caveat relates to the year-to-year inconsistency in the additional customised performance measures revealed by the agencies. Only a few agency-designed measures were used continuously. This inconsistency means that it is almost impossible for either government performance auditors or external researchers to effectively monitor any changes in these kind of customised measures, including the quality of services. More attention to the definitions and continued use of these performance measures would improve both internal and external accountability.

Acknowledgements

The authors would like to thank Marie-Ève Quenneville, Simon Kelly and Dominique Hamel for outstanding research assistance. Laurin acknowledges financial support from the Centre sur La Productivité et La Prospérité. Vining thanks the Social Science Research Council of Canada (SSHRC) for its invaluable financial support.

References

- Aivazian V. A., Ge Y. and Qiu J. (2005) Can Corporatization Improve the Performance of State-Owned Enterprises Even Without Privatization? *Journal of Corporate Finance* 11(5): 791–808.
- Barzelay M. and Campbell C. (2003) *Preparing for the Future: Strategic Planning in the U.S. Air Force*. Washington, DC: Brookings.
- Bertelli A. M. (2006) Delegating to the Quango: Ex Ante and Ex Post Ministerial Constraints. *Governance: An International Journal of Policy, Administration, and Institutions* 19(2): 229–249.
- Bevan G. (2010) Performance Measurement of “Knights” and “Knaves”: Differences in Approaches and Impacts in British Countries after Devolution. *Journal of Comparative Policy Analysis: Research and Practice* 12(1–2): 33–56.
- Bevan G. and Hood C. (2006) What’s Measured Is What Matters: Targets and Gaming in the English Public Health System. *Public Administration* 84(3): 517–538.
- Bilodeau N., Laurin C. and Vining A. R. (2007) Choice of Organizational Form Makes a Real Difference: The Impact of Corporatization on Government Agencies in Canada. *Journal of Public Administration Research and Theory* 17(1): 119–147.

- Binderkrantz A. S. and Christensen J. G. (2011) Agency Performance and Executive Pay in Government: An Empirical Test. *Journal of Public Administration Research and Theory* 22(1): 31–54.
- Boardman A. E., Laurin C. and Vining A. R. (2002) Privatization in Canada: Operating and Stock Price Performance with International Comparisons. *Canadian Journal of Administrative Sciences* 19(2): 137–154.
- Borins S. F. (1998) *Innovating with Integrity: How Local Heroes are Transforming American Government*. Washington, DC: Georgetown University Press.
- Boyne G. A. (2002) Public and Private Management: What is the Difference? *Journal of Management Studies* 39(1): 97–122.
- (2003a) What is Public Service Improvement? *Public Administration* 81(2): 211–227.
- (2003b) Sources of Public Service Improvement: A Critical Review and Research Agenda. *Journal of Public Administration Research and Theory* 13(3): 367–394.
- Brewer G. A. (2004) Does Administrative Reform Improve Bureaucratic Performance? A Cross-Country Empirical Analysis. *Public Finance and Management* 4(3): 399–428.
- Burgess S., Propper C., Ratto M., von Hinke S., Scholder K. and Tominey E. (2010) Smarter Task Assignment or Greater Effort: The Impact of Incentives on Team Performance. *The Economic Journal* 120(547): 968–989.
- Burgess S. and Ratto M. (2003) The Role of Incentives in the Public Sector: Issues and Evidence. *Oxford Review of Economic Policy* 19(2): 285–300.
- Cambini C., Filippini M., Piacenza M. and Vannoni D. (2011) Corporatisation and Firm Performance: Evidence from Publicly-Provided Local Utilities. *Review of Law and Economics* 7(1): 195–217.
- Carpenter D. P. (2001) *The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862–1928*. Princeton, NJ: Princeton University Press.
- Courty P. and Marschke G. R. (1997) Measuring Government Performance: Lessons from a Federal Job Training Program. *American Economic Review Papers and Proceedings* 87(2): 383–388.
- Dixit A. K. (2002) Incentives and Organizations in the Public Sector: An Interpretative Review. *Journal of Human Resources* 37(4): 696–727.
- Florio M. and Fecher F. (2011) The Future of Public Enterprises: Contributions to a New Discourse. *Annals of Public and Cooperative Economics* 82(4): 361–373.
- Frant H. L. (1996) High-Powered and Low-Powered Incentives in the Public Sector. *Journal of Public Administration Research and Theory* 6(3): 365–392.
- Frey B. S. and Jegen R. (2001) Motivation Crowding Theory: A Survey of Empirical Evidence. *Journal of Economic Surveys* 15(5): 589–611.
- Goldfinch S. and Wallis J. (2010) Two Myths of Convergence in Public Management Reform. *Public Administration* 88(4): 1099–1115.
- Greve C. (1999) Quangos in Denmark and Scandinavia: Trends, Problems, and Perspectives. In Flinders M. V. and Smith M. J. (eds.), *Quangos, Accountability, and Reform: The Politics of Quasi-Government*. Basingstoke, UK: Palgrave Macmillan, 83–108.
- Harding A. and Preker A. S. (2000) *Understanding Organizational Reforms: The Corporatization of Public Hospitals. Health, Nutrition and Population Discussion Paper*. Washington, DC: The World Bank.
- Hargrove E. C. and Glidewell J. C. (eds.) (1990) *Impossible Jobs in Public Management*. Lawrence, KA: University Press of Kansas.
- Heckman J. J., Smith J. A. and Taber C. (1996) What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance in the JTPA Program. In Libecap G. (ed.), *Studies in Bureaucratic Behavior*. San Francisco, CA: JAI Press.

- Heinrich C. J. (2007) False or Fitting Recognition? The Use of High Performance Bonuses in Motivating Organizational Achievements. *Journal of Policy Analysis and Management* 26(2): 281–304.
- Heinrich C. J. and Courty P. (2010) Incentives and Their Dynamics in Public Sector Performance Management Systems. *Journal of Policy Analysis and Management* 29(1): 183–208.
- Hood C. C. (1990) De-Sir Humphreyfying the Westminster Model of Bureaucracy: A New Style of Governance. *Governance: An International Journal of Policy, Administration, and Institutions* 3(2): 205–214.
- Huff R. F. (2011) Measuring Performance in US Municipalities: Do Personnel Policies Predict System Level Outcomes? *Journal of Comparative Policy Analysis: Research and Practice* 13(1): 11–33.
- Hyndman N. S. and Eden R. (2002) Executive Agencies, Performance Targets and External Reporting. *Public Money & Management* 22(3): 17–24.
- Jacobsson B. and Sundström G. (2007) Governing State Agencies: Transformations in the Swedish Administrative Model. ECPR Conference, Pisa, Italy, 6–8 September.
- Jensen J. L. (2003) Policy Diffusion Through Institutional Legitimization: State Lotteries. *Journal of Public Administration Research and Theory* 13(4): 521–541.
- Katyal N. K. (2006) Internal Separation of Powers: Checking Today's Most Dangerous Branch from Within. *Yale Law Journal* 115(9): 2314–2349.
- Kelman S. and Friedman J. N. (2009) Performance Improvement and Performance Dysfunction: An Empirical Examination of the Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service. *Journal of Public Administration Research and Theory* 19(4): 917–946.
- Kettl D. F. (2005) 2000 *The Global Public Management Revolution*. Washington, DC: Brookings Institution Press.
- Koop C. (2011) Explaining the Accountability of Independent Agencies: The Importance of Political Saliency. *Journal of Public Policy* 31(2): 209–234.
- Læg Reid P. and Verhoest K. (2010) Introduction: Reforming Public Sector Organizations. In Læg Reid P. and Verhoest K. (eds.), *Governance of Public Sector Organizations: Autonomy, Control and Performance*. Basingstoke, UK: Palgrave Macmillan, 1–18.
- Læg Reid P., Roness P. G. and Rubecksen K. (2005) Autonomy and Control in the Norwegian Civil Service: Does Agency Form Matter? Working Paper No. 4, Stein Rokkan Centre for Social Studies, Bergen, Norway.
- Langbein L. I. (2009) Controlling Federal Agencies: The Contingent Impact of External Controls on Worker Discretion and Productivity. *International Public Management Journal* 12(1): 85–115.
- (2010) Economics, Public Service Motivation, and Pay for Performance: Complements or Substitutes? *International Public Management Journal* 13(1): 9–23.
- Lavertu S. (2013) Issue-Specific Political Uncertainty and Policy Insulation in US Federal Agencies. *Journal of Law, Economics, and Organization* 29(1): 145–177.
- Lazear E. P. (2000) The Power of Incentives. *American Economic Review: Papers and Proceedings* 90(2): 410–414.
- Lazear E. P. and Shaw K. L. (2007) Personnel Economics: The Economist's View of Human Resources. *Journal of Economic Perspectives* 21(4): 91–114.
- Le Grand J. (2003) *Motivation, Agency and Public Policy: Of Knights and Knaves, Pawns and Queens*. Oxford: Oxford University Press.
- Leibenstein H. (1966) Allocative Efficiency vs. X-Efficiency. *American Economic Review* 56(3): 392–415.

- Levitt S. D. and List J. A. (2009) Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments. NBER Working Paper No. 15016, NBER, Cambridge, Mass.
- Lewis D.E. (2003) *Presidents and the Politics of Agency Design: Political Insulation in the United States Government Bureaucracy, 1946–1997*. Stanford, CA: Stanford University Press.
- Mayo E. (1933) *The Human Problems of an Industrialized Civilization*. New York: The MacMillan Company.
- Mazouz B. and Tremblay B. (2006) Toward a Postbureaucratic Model of Governance: How Institutional Commitment is Challenging Québec's Administration. *Public Administration Review* 66(2): 263–273.
- Mazouz B. and Leclerc J. (2008) *La Gestion Intégrée par Résultats*. Québec: Les Presses de l'Université du Québec.
- Moynihan D. P. (2008) The Normative Model in Decline? Public Service Motivation in the Age of Governance. In Perry L. and Hondeghem A. (eds.), *Motivation in Public Management*. Oxford: Oxford University Press.
- Moynihan D. P. and Pandey S. K. (2006) Creating Desirable Organizational Characteristics. *Public Management Review* 8(1): 119–140.
- Nelson H. W. and Nikolakis W. (2012) How Does Corporatization Improve the Performance of Government Agencies? Lessons from the Restructuring of State-Owned Forest Agencies in Australia. *Public Management Journal* 15(3): 364–391.
- Nicholls D. (1989) Corporatisation. *Economic Papers* 8(3): 27–33.
- OECD (2005) *Modernizing Government: The Way Forward*. Paris: OECD Publishing.
- Painter G. D., Burns J. P. and Yee W.-H. (2010) Explaining Autonomy in Public Agencies: The Case of Hong Kong. In Lægveid P. and Verhoest K. (eds.), *Governance of Public Sector Organizations: Autonomy, Control and Performance*. Basingstoke, UK: Palgrave Macmillan.
- Pendlebury M. and Karbhari Y. (1998) Resource Accounting and Executive Agencies. *Public Money & Management* 18(2): 29–33.
- Perry J. L. and Hondeghem A. (2008) Building Theory and Evidence about Public Service Motivation. *International Public Management Journal* 11(1): 3–12.
- Perry J. L., Engbers T. A. and Jun S. Y. (2009) Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence. *Public Administration Review* 69(1): 39–51.
- Poister T. H., Pitts D. W. and Edwards L. H. (2011) Strategic Management Research in the Public Sector: A Review, Synthesis, and Future Directions. *The American Review of Public Administration* 40(5): 522–545.
- Pollitt C., Van Thiel S. and Homburg V. (eds.) (2007) *New Public Management in Europe*. Basingstoke, UK: Palgrave Macmillan.
- Pollitt C. and Dan S. (2011) The Impact of the New Public Management in Europe: A Meta-Analysis. COCOPS, European Commission. http://www.cocops.eu/wp-content/uploads/2012/03/WP1_Deliverable1_Meta-analysis_Final.pdf (accessed 20 February 2014).
- Quenneville M.-È., Laurin C. and Thibodeau N. (2010) The Long Run Performance of Decentralized Agencies in Québec. In Lægveid P. and Verhoest K. (eds.), *Governance of Public Sector Organizations: Autonomy, Control and Performance*. Basingstoke, UK: Palgrave Macmillan, 157–176.
- Rainey H. G. and Steinbauer P. (1999) Galloping Elephants: Developing Elements of a Theory of Effective Government Organization. *Journal of Public Administration Research and Theory* 9(1): 1–32.
- Reddy K., Locke S. and Scrimgeour F. (2011) Improving Performance in New Zealand's Public Corporations: The Effect of Governance Practices. *Governance: An International Journal of Policy, Administration, and Institutions* 24(3): 517–556.

- Richards D. and Smith M. (2006) The Tension of Political Control and Administrative Autonomy: From NPM to a Reconstituted Westminster Model. In Christensen T. and Lægheid P. (eds.), *Autonomy and Regulation. Coping with Agencies in the Modern State*. Cheltenham, UK: Edward Elgar, 181–200.
- Sayre W. S. (1958) Premises of Public Administration: Past and Emerging. *Public Administration Review* 18(2): 102–105.
- Shirley M. M. (1999) Bureaucrats in Business: The Role of Privatization versus Corporatization in State-Owned Enterprise Reform. *World Development* 27(1): 115–136.
- Stephenson M. C. (2008) Optimal Political Control of the Bureaucracy. *Michigan Law Review* 107(1): 53–110.
- Talbot C. (2004) Executive Agencies: Have They Improved Management in Government? *Public Money and Management* 24(2): 104–112.
- Verhoest K. (2005) Effects of Autonomy, Performance Contracting, and Competition on the Performance of a Public Agency: A Case Study. *Policy Studies Journal* 33(2): 235–258.
- Verhoest K., Peters G., Bouckaert G. and Verschuere B. (2004) The Study of Organisational Autonomy: A Conceptual Overview. *Public Administration and Development* 24(2): 101–118.
- Verhoest K. and Lægheid P. (2010) Organizing Public Sector Agencies: Challenges and Reflections. In Lægheid P. and Verhoest K. (eds.), *Governance of Public Sector Organizations: Autonomy, Control and Performance*. Basingstoke, UK: Palgrave Macmillan, 276–297.
- Verschuere B. (2007) The Autonomy-Control Balance in Flemish Arm's Length Public Agencies. *Public Management Review* 9(1): 107–133.
- Vining A. R. (2011) Public Agency External Analysis Using a Modified 'Five Forces' Framework. *International Public Management Journal* 14(1): 63–105.
- Vining A. R. and Weimer D. L. (2005) Economic Perspectives on Public Organizations. In Pollit C., Lynn L. and Ferlie E. (eds.), *Oxford Handbook of Public Management*. Oxford: Oxford University Press, 209–233.
- Walker R. M. and Boyne G. A. (2006) Public Management Reform and Organizational Performance: An Empirical Assessment of the U.K. Labour Government's Public Service Improvement Strategy. *Journal of Policy Analysis and Management* 29(1): 183–208.
- Walsh K. (1995) *Public Services and Market Mechanisms: Competition, Contracting and the New Public Management*. New York: St. Martin's Press.
- Weibel A., Rost K. and Osterloh M. (2010) Pay for Performance in the Public Sector – Benefits and (Hidden) Costs. *Journal of Public Administration Research and Theory* 20(2): 387–412.
- Wilson J. Q. (1989) *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.
- Wolf P. J. (1993) A Case Survey of Bureaucratic Effectiveness in U.S. Cabinet Agencies: Preliminary Agencies. *Journal of Public Administration Research and Theory* 3(2): 161–181.
- Xu L. C., Zhu T. and Lin Y.-M. (2005) Political Control, Agency Problems and Ownership Reform. *Economics of Transition* 13(1): 1–24.

Appendix 1: Summary description of agencies

	Year of creation	Last year of data	Output	FTE (total)	Annual budget of expenditures (in millions)	Purpose
Aide financière aux études (AFE)	1997	2007–2008	Number of applications for loans and bursaries	344	19.50	Manages financial aid programmes available to students
Bureau des infractions et amendes (BIA)	2001	2009–2010	Number of statements and offence reports received, and the number of judgments issued (penal and criminal) with a fine	141	9.39	Ensures follow-up of infraction reports and recovers fines and legal costs that were not paid by the deadline
Commission administrative des régimes de retraite et d'assurances (CARRA)	1995	2009–2010	Amounts paid in retirement pension	441	40.40	Administers pension plans payed by the Québec government to former employees
Centre de conservation du Québec (CCQ)	1995	2009–2010	Number of hours needed for the production at CCQ	30	2.18	Provides restoration services to museums and other institutions engaged in conservation
Centre d'expertise en analyse environnementale du Québec (CEAEQ)	1997	2009–2010	Number of hours of service provided to the Ministry	107	10.45	Provides a wide array of environmental analyses
Centre de gestion de l'équipement roulant (CGER)	1997	2009–2010	Number of vehicles rented	410	53.48	Operates and maintains the vehicle and motorised equipment fleet owned by the government of Québec
Centre de perception fiscale (CPF)	1995	2005–2006 (1)	Amounts of revenues recovered	1,079	80.50	Recovers fiscal and other debts owed to Revenue Québec

Appendix 1: (Continued)

	Year of creation	Last year of data	Output	FTE (total)	Annual budget of expenditures (in millions)	Purpose
Centre de recouvrement (CR)	2001	2008–2009	Amounts recovered	238	12.70	Recovers amounts owed to the “Ministère de la Santé et de la Solidarité Sociale”
Emploi Québec (EQ)	1998	2004–2005	Number of new participants under assistance by Emploi-Québec	4,484	292.70	Provide assistance with various employment related services
Laboratoire de sciences judiciaires et de médecine légale (LSJML)	1996	2009–2010	Number of expert reports produced	137	9.72	Supports police or legal investigations by providing expertise in forensic medicine
Régie du cinema (RC)	2001	2009–2010	Number of documents (permits, certificates and visas) issued by the Régie	43	5.37	Manages the motion picture rating system
Régie des rentes du Québec (RRQ)	1997	2009–2010	Number of pension recipients	1,121	144.90	Administers the Québec Pension Plan and family benefits programme
Service aérien gouvernemental (SAG)	1995	2008–2009	Number of flight hours	162	57.77	Manages and operates government owned aircrafts

Note: In 2007, CPF stopped presenting individual results. The agency provided us unaudited results up to 2008–2009. FTE = full-time equivalent employees.